ENHANCING HIGH-GAIN-OBSERVER PERFORMANCE IN THE PRESENCE OF MEASUREMENT NOISE

By

Alexis A. Ball

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Electrical Engineering

2011

ABSTRACT

ENHANCING HIGH-GAIN-OBSERVER PERFORMANCE IN THE PRESENCE OF MEASUREMENT NOISE

By

Alexis A. Ball

High-gain observers are a prevalent and an important topic in state estimation and output feedback control of nonlinear systems. In the absence of measurement noise, this technique robustly estimates the derivatives of the output while achieving fast convergence. Moreover, for a sufficiently fast observer and a globally bounded controller, the high-gain observer is able to recover the system performance achieved under state feedback control.

However, in the presence of measurement noise, a tradeoff exists between the measurement noise sensitivity and the speed of state reconstruction. As the observer gain is increased, the bandwidth of the observer is extended. As the bandwidth increases, the high-gain observer asymptotically approaches the behavior of a differentiator, exacerbating the presence of measurement noise.

This dissertation addresses the challenging performance issues that arise when implementing high-gain observers in the presence of measurement noise. In particular, we focus on the tradeoff between fast state reconstruction, minimizing the bound on the steady-state estimation error, and rejecting the model uncertainty. The observer design and analysis is approached through three major thrust areas: observer structure, tracking performance and filtering.

To my family; for their endless

support and encouragement.

ACKNOWLEDGMENTS

I am indebted to my advisor, Professor Hassan Khalil, for molding me into the researcher I am today. Professor Khalil's expectations for excellence are uncompromising, and his patience boundless. I feel incredibly fortunate to have matured scholarly under his tutelage. Thank you to my committee members Professor Ranjan Mukherjee, Professor Xiaobo Tan and Professor Ning Xi for their insightful comments that improved the quality of this dissertation.

The many thought provoking discussions that occurred with my colleagues at other institutions would not have been possible without the support of the National Science Foundation, the Alliance for Graduate Education and the Professoriate, the Graduate School at Michigan State University, the College of Engineering at Michigan State University and Dr. Barbara O'Kelly. I simply cannot thank Dr. O'Kelly enough for coming to my rescue on several occasions.

I would like to thank my colleagues and friends for the helpful discussions, free food and distractions. Most of all, thank you to my parents for believing in me and my abilities. Their exaggerations of my superhuman research prowess stimulates me to be that mythical heroine. Last, but certainly not least, thank you to my beloved Zahar who inspires me to pursue the impossible.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

High-gain observers are an important tool in state estimation and output feedback control. Some of the earlier research performed in the spirit of high-gain observers can be viewed in [23] and [22]; see also [37], [44] and [31] for recent results. Yet, even in early works such as [22], it was noted that noise in the system sensors can cause a noticeable (and undesirable) effect on the system dynamics. Thus, the focus of this dissertation is to analyze and address the issues associated with high-gain observer performance degradation in the presence of measurement noise.

1.1 Feedback Control

Before describing the high-gain observer form, it is important to lay the foundation for the types of systems that are considered in this body of work. Namely, the nonlinear

system

$$\dot{z} = \psi(x, z, \varsigma, u) \tag{1.1}$$

$$\dot{x} = Ax + B\phi(x, z, \varsigma, u) \tag{1.2}$$

$$y = Cx \tag{1.3}$$

$$w = \Theta(x, z, \varsigma) \ , \tag{1.4}$$

where $z \in \mathbb{R}^l$ and $x \in \mathbb{R}^n$ are the system states, $y \in \mathbb{R}$ and $w \in \mathbb{R}^s$ are the measured outputs, $u \in \mathbb{R}$ is the control input and $\varsigma(t) \in \mathbb{R}^p$ represents the exogenous signals. To create a more realistic problem, the function $\phi(x, z, \varsigma, u)$ is assumed to be unknown; however, this does not excluded problems where the system model may be known. The system matrices take the form

$$A = \begin{bmatrix} 0 & 1 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdot & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \text{ and}$$

$$C = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \end{bmatrix} \ ,$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$ and $C \in \mathbb{R}^{1 \times n}$. Logically, the system is assumed to have dimension $n \geq 2$. In the event that the dimension condition is not met, the construction of an observer is unnecessary; the measured output y is simply used.

The structure for the model (1.1)-(1.4) includes mechanical systems, electromechanical systems and systems that can be placed in the normal form satisfying the conditions of input-output linearization [30]. Some examples of mechanical and electromechanical systems where the displacements are measured, but not their deriva-

tives can be found in [3, 28, 33] for an induction motor, a rotational/translational actuator and a smart material application, respectively. Furthermore, the additional measurement w may not be needed in every model. In this case, (1.4) can be removed from the system representation. Yet, many models utilize the extra output. For instance, consider a system in which the dynamics are extended by adding integrators; see [30]. Furthermore, the classic example of the magnetically suspended ball is modeled such that the ball position and current are measurement outputs. The position fits the chain of integrators form (in x), whereas the current becomes the state variable w. Hence, many relevant and interesting systems are encapsulated in the types of models of interest in this dissertation.

Assumption 1.1:

- $\varsigma(t)$ is continuously differential and bounded;
- $\varsigma(t) \in \mathcal{D} \subset \mathbb{R}^p$, where $\mathcal{D}$ is compact;
- ϕ, ψ, and Θ are locally Lipschitz in their arguments, uniformly in ς, over the domain of interest; that is, for each compact subset of (x, z, u) in the domain of interest, the functions satisfy the Lipschitz inequality with a Lipschitz constant independent of ς for all $\varsigma \in \mathcal{D}$.

The static state feedback controller takes the following form

$$u = \gamma(x, w, \varsigma) \tag{1.5}$$

and is designed to meet the desired performance objectives. In practice, the controller (1.5) cannot be implemented as written. Recall that only the first state is accessible. Therefore, the controller cannot require values for any state beyond x_1 and remain useful. Given the full state measurement is not available, an alternative method

is necessary to obtain the desired state information. One option is to construct an observer that will estimate the system states from the available measurement y, which leads to a dynamic output feedback controller of the form

$$u = \gamma(\hat{x}, w, \varsigma) \ , \tag{1.6}$$

where the state x is replaced by the estimate $\hat{x}$. For the class of systems defined by (1.1)-(1.4), we consider the high-gain observer

$$\dot{\hat{x}} = A\hat{x} + B\phi_0(\hat{x}, w, \varsigma, u) + h(y - \hat{x}_1) \ . \tag{1.7}$$

Typically, the gain function is defined as

$$h(y - \hat{x}_1) = H(y - \hat{x}_1) \ , \tag{1.8}$$

where

$$H = \begin{bmatrix} \dfrac{\alpha_1}{\varepsilon} & \dfrac{\alpha_2}{\varepsilon^2} & \cdots & \cdots & \dfrac{\alpha_n}{\varepsilon^n} \end{bmatrix}^T \ . \tag{1.9}$$

The α_i's are designed such that the roots of

$$s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1} s + \alpha_n = 0 \tag{1.10}$$

have negative real parts. The function ϕ_0 is locally Lipschitz and a known nominal model of ϕ, which initially appears in (1.2). However, this is not the end of the story. In general, the separation principle does not hold uniformly for nonlinear systems; namely, maintaining the ability to design the state feedback controller independently from the observer, with the result being a stable closed-loop system. The first separation principle for high-gain observers was reported in [50], followed by a more comprehensive theorem in [6]. The stipulation is that the state feedback controller

be globally bounded and the observer parameter ε be chosen sufficiently small. If the state feedback controller is designed to achieve global stabilization, then it can be shown that the output feedback controller utilizing the high-gain observer in (1.7) realizes semiglobal stabilization.

Another important detail is the nature of the stability properties. Let the closed-loop system (1.1)-(1.4) under the state feedback controller (1.5) be denoted as

$$\dot{\chi} = f_r(\chi, \varsigma) \ , \tag{1.11}$$

where

$$\chi = \begin{bmatrix} x \\ z \end{bmatrix} \in \mathbb{R}^N \text{ and } f_r(\chi, \varsigma) = \begin{bmatrix} Ax + B\phi(x, z, \varsigma, u) \\ \psi(x, z, \varsigma, u) \end{bmatrix}.$$

Instead of stabilizing an equilibrium point of the system, the problem is posed as rendering a certain compact set positively invariant and asymptotically attractive. The allure of formulating the results in this fashion, is that the separation principle is no longer limited to stabilization of an equilibrium point [5]. Some examples include servomechanisms [25, 29] and finite time convergence to a set [20]. The regulation problem for servomechanisms, for example, requires that the trajectories of the system reach an invariant manifold where the tracking error is zero, usually called the zero-error manifold. Instead of regulating the system about an equilibrium point, we desire to study the dynamics on the manifold. Therefore, the set that we wish to render positively invariant and asymptotically attractive is the zero-error manifold.

As stated in [4], uniform asymptotic stability with respect to a set $\mathcal{A}$, uniformly in ς, stipulates the following:

- Uniform Stability - For each $\epsilon > 0$ there is a $\delta = \delta(\epsilon)$ such that

$$|\chi(t_0)|_{\mathcal{A}} \leq \delta \Rightarrow |\chi(t)|_{\mathcal{A}} < \epsilon, \forall\ t \geq t_0 \geq 0, \forall\ \varsigma(t) \in \mathcal{D}.$$

- Uniform Attraction - There is a constant $c > 0$, independent of t_0 and $\varsigma(t)$, and for each $\epsilon > 0$ there is $T = T(\epsilon)$ such that

$$|\chi(t)|_{\mathcal{A}} < \epsilon, \forall\, t \geq t_0 + T, \forall\, |\chi(t_0)|_{\mathcal{A}} < c, \forall\, \varsigma(t) \in \mathcal{D}.$$

The expression $|\chi|_{\mathcal{A}} = \inf_{\nu \in \mathcal{A}} \|\chi - \nu\|$ is the distance with respect to $\mathcal{A}$. Furthermore, the system is said to be globally uniformly asymptotically stable with respect to $\mathcal{A}$, if the uniform stability property holds with a class $\mathcal{K}_\infty$ function δ, and the uniform attraction property holds for any $r > 0$ with $T = T(\epsilon, r)$. By extending the definition of stability to a compact positively invariant set instead of just an equilibrium point, a wider variety of problem formulations can be encapsulated in the above setup. Clearly, we can still address stabilization of the origin by defining $\mathcal{A} = 0$. However, we can just as easily structure the control objective as a regulation or tracking problem. For additional examples of control problems that can be formulated as stabilization with respect to a set, see [4].

Assumption 1.2:

- The closed-loop system (1.11) is globally uniformly asymptotically stable with respect to a compact positively invariant set $\mathcal{A}$, uniformly in ς;

- $\phi(x, z, \varsigma, u)$ is zero in $\mathcal{A}$, uniformly in ς.

To uncover some of the interesting properties inherent to high-gain observers, consider the scaled estimation error

$$\eta_i = \frac{x_i - \hat{x}_i}{\varepsilon^{n-i}}\ . \tag{1.12}$$

For a second-order system the scaled estimation errors are

$$\eta_1 = \frac{x_1 - \hat{x}_1}{\varepsilon} \tag{1.13}$$

$$\eta_2 = x_2 - \hat{x}_2 \ , \tag{1.14}$$

which satisfy the singularly perturbed equation

$$\varepsilon\dot{\eta}_1 = -\alpha_1\eta_1 + \eta_2 \tag{1.15}$$

$$\varepsilon\dot{\eta}_2 = -\alpha_2\eta_1 + \varepsilon\delta(x, z, \varsigma, u) \ , \tag{1.16}$$

where $\delta = \phi - \phi_0$. As the value of ε is decreased, the effect of δ in (1.16) is diminished. Hence, high-gain observers have the ability to reject the error due to modeling uncertainty as ε approaches zero. The smaller the value of ε, the faster the time-scale of the observer relative to the plant (or system in x). This difference in time-scale leads to the possibility of peaking behavior. If there is any difference in the initial conditions between the state x_1 and the estimate $\hat{x}_1$, then the initial condition of η_1 will be $\mathcal{O}(1/\varepsilon)$. This peaking phenomenon is a by-product of the observer gain structure in (1.9) and leads to a term of the form

$$\frac{a}{\varepsilon}\exp(-at/\varepsilon) \tag{1.17}$$

in the transient response of the solution to (1.15)-(1.16), where $a > 0$. Given the term will decay rapidly, the effects will be seen primarily in the transient response. However, as ε tends to zero, (1.17) approaches the behavior of an impulse function, where its amplitude peaks at a value $\mathcal{O}(1/\varepsilon)$. Before the impulse-like behavior of the term can subside, this exponential mode has the ability to not only induce an unacceptable transient response, but destabilize the closed-loop nonlinear system. The destabilizing effect of peaking interacting with the nonlinear feedback control was

first observed in [21]. Furthermore, the solution of saturating the control and/or the state estimates outside a compact region of interest to achieve a globally bounded controller, and designing the nominal function $\hat{\phi}$ to be globally bounded in the estimates, protects the system plant from the destabilizing behavior during the peaking period. Notice that the peaking period shrinks to zero as ε tends to zero. Moreover, the system trajectories under output feedback come arbitrarily close to the trajectories under state feedback as the value of ε approaches zero. This amounts to recovering the system performance in addition to the stability properties of the system under state feedback control; see [6].

1.2 High-Gain Observers and Measurement Noise

Return to the singularly perturbed representation of the derivative of the scaled estimation errors in (1.15)-(1.16). By adding noise into the system measurement such that

$$y = x_1 + v$$

becomes the output and v is the additive measurement noise, (1.15)-(1.16) is altered in the following manner

$$\varepsilon\dot{\eta}_1 = -\alpha_1\eta_1 + \eta_2 - (\alpha_1/\varepsilon)v \tag{1.18}$$

$$\varepsilon\dot{\eta}_2 = -\alpha_2\eta_1 + \varepsilon\delta(x, z, \varsigma, u) - (\alpha_2/\varepsilon)v \ , \tag{1.19}$$

where reducing the amount of error in the estimation is no longer as simple as decreasing ε. Unlike the system without measurement noise there exists a tradeoff between the steady-state errors due to the model uncertainty, captured in the function δ, and the measurement noise v. In [2], the norm on the state vector η and, ultimately, the

estimation error satisfies the inequality

$$\|x(t) - \hat{x}(t)\| \leq c_1\varepsilon + c_2\frac{\mu}{\varepsilon^{n-1}}, \forall t \geq T \tag{1.20}$$

for an n-dimensional system with positive constants c_1, c_2 and T. The measurement noise v is assumed to be bounded by the positive constant μ. Furthermore, another tradeoff exists between the speed of state recovery and the accuracy of that estimate in steady-state. Moreover, it is crucial that the observer be sufficiently faster than the dynamics of the plant in order to ensure recovery of the state feedback controller performance. Therefore, choosing smaller values of ε results in better rejection of the modeling uncertainty, faster reconstruction of the system states and recovery of the performance under state feedback control. However, the presence of measurement noise prevents ε from being chosen arbitrarily small. Hence, the work in this dissertation seeks to further analyze the effects of measurement noise, while quantifying and reducing the manifestations of the tradeoff in the system states.

1.3 Organization

The purpose of this dissertation is to tackle the challenging performance issues that arise when implementing high-gain observers in the presence of measurement noise. Thus, the work herein approaches observer design and analysis in the presence of measurement noise through three major thrust areas: observer structure, tracking performance and filtering. The divisions addressing those areas are briefly summarized below.

The first attempt at minimizing the tradeoffs present in the high-gain observer are done through manipulating the observer gain structure. The gain is designed as a function of the estimation error in the first state. Therefore, the gain function

responds to the value of this estimation error, such that the observer experiences a larger gain during the transient period and a lower gain afterwards. By designing the function in this fashion, the closed-loop system is able to obtain reasonably fast state estimation and attenuate a larger portion of the measurement noise in steady-state. One may interpret this result as minimizing the classic tradeoff of speed versus accuracy present in observer design.

However, the effect of measurement noise on the tracking error is less significant than on the estimation error. Simulation observations suggest that pushing the observer gain too large can noticeably compromise the estimation error and, ultimately, the system performance. Yet, such issues are not as apparent for control problems that are formulated in the tracking framework. The effect that the measurement noise has on the system tracking error is analyzed for linear systems and a class of nonlinear systems.

In addition to augmenting the observer form with nonlinearities, the observer performance can potentially be improved by filtering out the measurement noise before feeding the output to the observer. Typically, a lowpass filter is used to remove the noise from signals in the feedback loop. However, depending on the order of the filter, unacceptable phase lag can be introduced with the potential for destabilizing the system. In the interest of providing an alternative to the classic lowpass filter, the feasibility of wavelets for denoising is studied in Chapter 4. The idea of using wavelets to remove noise is not a new one, however, incorporating them into the feedback loop for online denoising is a recent development. Historically, all of the signal is available when the denoising algorithm is applied. Naturally, all of the past and present signal values will not be available in the feedback loop. This is just one of the complications introduced by attempting to denoise the output signal online. Overall, the investigation is carried out using wavelets to design various pre-filters, while comparing the results to the lowpass filter. The final chapter speculates on possible future work and

provides some concluding remarks.

Chapter 2

Nonlinear-Gain High-Gain Observers

2.1 Introduction

High-gain observers have developed into an important topic in state estimation and output feedback control of nonlinear systems, beginning with papers such as [21] and [22]. In the absence of measurement noise, this technique robustly estimates the derivatives of the output while achieving fast convergence [21]. Moreover, for a sufficiently high observer gain and a globally bounded controller, the high-gain observer is able to recover the system performance achieved with the state feedback control. Refer to [31] for a survey on high-gain observers.

However, observer theory reveals that a tradeoff exists between the measurement noise sensitivity and the speed of state reconstruction [38]. As the observer gain is increased, the bandwidth of the observer is extended. As the bandwidth increases the high-gain observer asymptotically approaches the behavior of a differentiator, exacerbating the presence of measurement noise. The authors of [43], in the context of discrete-time models, exploited this knowledge by designing a switched filter

composed of two linear filters (one for the transient response and the other for the steady-state response); the value of the estimation error determines which filter is active. The idea is to use a large filter gain (increasing the filter bandwidth) during the transient behavior to elicit a fast recovery of the state estimates. The filter with the smaller gain is active once the estimation error has reached a steady-state threshold, reducing the filter bandwidth and preventing a large magnification of the measurement noise. In [52], the authors seek to minimize the effect quantization error has on shaft encoder measurements by introducing a dead-zone nonlinearity into the state estimation scheme. The dead-zone nonlinearity is used to alternate "smoothly" between varying filter bandwidths to initially achieve fast state estimation and, ultimately, minimize the quantization error.

Recently, others have addressed the issue of measurement noise and observers in [8, 45, 47]. The work in [47] investigates a high-gain observer with a sign-indefinite gain adaptation for systems with potentially nonlocal Lipschitz functions and noisy output. However, this approach often leads to highly oscillatory, although bounded, behavior in the state estimates. In [8], the authors propose an extended Kalman filter that utilizes an adaptive high-gain parameter to achieve noise rejection and global convergence of the estimated state when careful tuning is exercised. More generally, the work in [45] analyzes observers with improved transient performance for both linear and nonlinear systems; the effects of measurement noise are briefly considered for one high-gain observer design.

The effect of measurement noise in high-gain observers has been studied in [2, 53]. It is shown in [2] that the steady-state estimation error has a component due to modeling uncertainty, which can be attenuated by increasing the gain. Furthermore, the error has a component due to measurement noise that is amplified by increasing the gain. This tradeoff constrains the observer gain, which reduces the observer's ability to quickly reconstruct the states. In [2], the authors construct an observer

to diminish the manifestations of the tradeoff in the system states. A switched-gain observer is proposed in [2] to force a large gain during the transient period for fast state reconstruction, and allow for a smaller gain once the states are satisfactorily estimated to reduce the effect of noise on the steady-state performance. However, a number of complications are generally associated with a switched system. The time in which the gains are switched, trigger threshold, and system peaking are all design issues that must be addressed. Both from an analysis and design/implementation prospective, using a switched observer can be tedious.

The purpose of this chapter is to construct high-gain observers containing a nonlinear gain that takes the form of a piecewise linear function with two or three distinct linear regions. The regions are chosen to correspond to the desired transient and steady-state responses, respectively. By constructing the observer gain in this manner, we can achieve fast state estimation and reduced steady-state error. Furthermore, the observer is devised such that the behavior of the innovation process can be controlled separately from the other estimation errors. This is accomplished by assigning one fast eigenvalue with the remaining eigenvalues chosen relatively slow. Without this key step, the stability analysis for the proposed observers is unattainable. The analysis focuses on the proposed high-gain observer and the closed-loop system dynamics. The discussion concludes with a simulation comparing the system performance under the linear, nonlinear, and switched high-gain observer designs. In particular, it is demonstrated that the nonlinear-gain observers are sufficient in obtaining the desired estimation error dynamics, while reducing the implementation complexity necessary with the switched gain observer.

2.2 Problem Formulation and System Description

Consider the nonlinear system

$$\dot{z} = \psi(x, z, \varsigma, u) \tag{2.1}$$

$$\dot{x} = Ax + B\phi(x, z, \varsigma, u) \tag{2.2}$$

$$y = Cx + v \tag{2.3}$$

$$w = \Theta(x, z, \varsigma) \ , \tag{2.4}$$

where $z \in \mathbb{R}^l$ and $x \in \mathbb{R}^n$ are the system states, $y \in \mathbb{R}$ and $w \in \mathbb{R}^s$ are the measured outputs, $u \in \mathbb{R}$ is the control input, $\varsigma(t) \in \mathbb{R}^p$ represents the exogenous signals and $v(t) \in \mathbb{R}$ is the measurement noise. The function $\phi(x, z, \varsigma, u)$ may not be known. We do not explicitly define noise in the output w, given the purpose of this work is to study how measurement noise directly enters the high-gain observer from y, and mitigate the resulting effects. The triple (A, B, C) represents a chain of n integrators, where it is assumed that $n \geq 2$. Possible sources for the model (2.1)-(2.4) include mechanical systems, electromechanical systems and systems that can be placed in the normal form satisfying the conditions of input-output linearization.

Assumption 2.1:

- $\varsigma(t)$ is continuously differential and bounded;
- $\varsigma(t) \in \mathcal{D} \subset \mathbb{R}^p$, where $\mathcal{D}$ is compact;
- $v(t)$ is a measurable function of t and bounded, where the bound is defined as $|v(t)| \leq \mu$;
- ϕ, ψ, and Θ are locally Lipschitz in their arguments, uniformly in ς, over the

domain of interest; that is, for each compact subset of (x, z, u) in the domain of interest, the functions satisfy the Lipschitz inequality with a Lipschitz constant independent of ς for all $\varsigma \in \mathcal{D}$.

The state feedback controller takes the following form

$$\dot{\theta} = \Gamma(\theta, x, w, \varsigma) \tag{2.5}$$

$$u = \gamma(\theta, x, w, \varsigma) \tag{2.6}$$

and meets the requirements listed in Assumption 2.2.

Assumption 2.2:

- Γ and γ are locally Lipschitz functions in their arguments, uniformly in ς, over the domain of interest;
- Γ and γ are globally bounded functions of x. The necessity for this assumption is detailed in Chapter 1.

Let the closed-loop system (2.1)-(2.4) under the state feedback controller (2.5)-(2.6) be denoted as

$$\dot{\chi} = f_r(\chi, \varsigma) \ , \tag{2.7}$$

where

$$\chi = \begin{bmatrix} x \\ z \\ \theta \end{bmatrix} \in \mathbb{R}^N \text{ and } f_r(\chi, \varsigma) = \begin{bmatrix} Ax + B\phi(x, z, \varsigma, \gamma) \\ \psi(x, z, \varsigma, \gamma) \\ \Gamma(\theta, x, w, \varsigma) \end{bmatrix} .$$

Assumption 2.3:

- The closed-loop system (2.7) is globally uniformly asymptotically stable with respect to a compact positively invariant set $\mathcal{A}$, uniformly in ς;
- $\phi(x, z, \varsigma, \gamma)$ is zero in $\mathcal{A}$, uniformly in ς.

Instead of stabilizing an equilibrium point of the system, the problem is posed as rendering a certain compact set positively invariant and asymptotically attractive. The allure of formulating the results in this fashion, is that the separation principle is no longer limited to stabilization of an equilibrium point [5]; further details are provided in Chapter 1. As stated in [4], uniform asymptotic stability with respect to a set $\mathcal{A}$, uniformly in ς, stipulates the following:

- Uniform Stability - For each $\epsilon > 0$ there is a $\delta = \delta(\epsilon)$ such that

$$|\chi(t_0)|_{\mathcal{A}} \leq \delta \Rightarrow |\chi(t)|_{\mathcal{A}} < \epsilon, \forall\, t \geq t_0 \geq 0, \forall\, \varsigma(t) \in \mathcal{D}.$$

- Uniform Attraction - There is a constant $c > 0$, independent of t_0 and $\varsigma(t)$, and for each $\epsilon > 0$ there is $T = T(\epsilon)$ such that

$$|\chi(t)|_{\mathcal{A}} < \epsilon, \forall\, t \geq t_0 + T, \forall\, |\chi(t_0)|_{\mathcal{A}} < c, \forall\, \varsigma(t) \in \mathcal{D}.$$

The expression $|\chi|_{\mathcal{A}} = \inf_{\nu \in \mathcal{A}} \|\chi - \nu\|$ is the distance with respect to $\mathcal{A}$. Furthermore, the system is said to be globally uniformly asymptotically stable with respect to $\mathcal{A}$, if the uniform stability property holds with a class $\mathcal{K}_\infty$ function δ, and the uniform attraction property holds for any $r > 0$ with $T = T(\epsilon, r)$. By extending the definition of stability to a compact positively invariant set instead of just an equilibrium point, a wider variety of problem formulations can be encapsulated in the above setup. Clearly, we can still address stabilization of the origin by defining

$\mathcal{A} = 0$. However, we can just as easily structure the control objective as a regulation or tracking problem. For examples of control problems that can be formulated as stabilization with respect to a set, see [4] and Chapter 1.

2.3 Observer Dynamics

The intuition behind the nonlinear observer gain is the following:

- Achieve the desired (fast) state reconstruction with ε_1 without sacrificing the steady-state performance;
- Reduce the steady-state estimation error with ε_2 while maintaining an acceptable rate of convergence in the estimates.

For a visual of the two-piece nonlinear-gain, see Figure 2.1. However, when examining the two-piece structure, it appears as if the slope (through the origin) g_1 is not equivalent to the slope g_1 in the linear-gain observer, as shown in Figure 2.2. Hence, it may be prudent to also investigate an observer constructed with a three-piece nonlinear gain; see Figure 2.3.

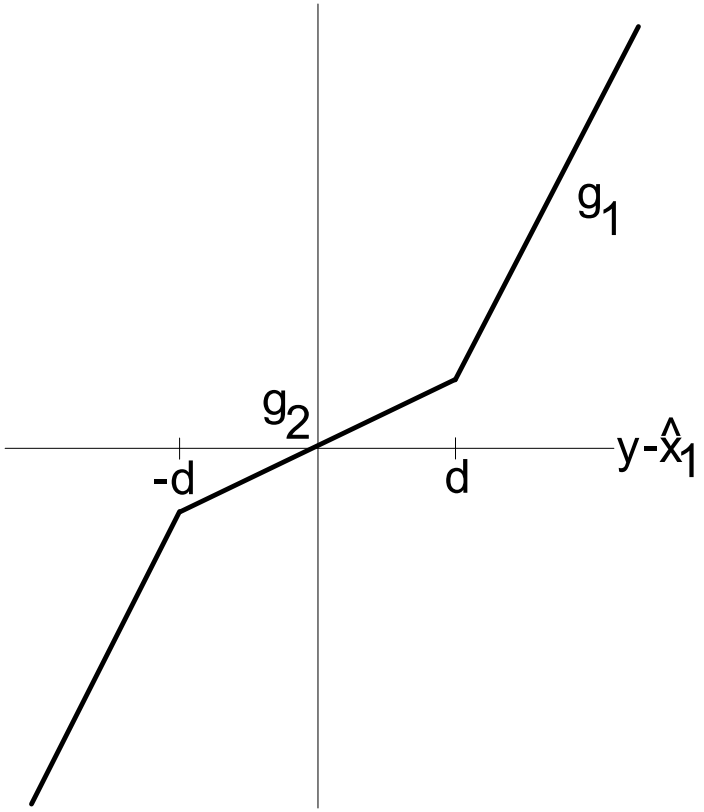

Figure 2.1: Plot of the two-piece nonlinear-gain function.

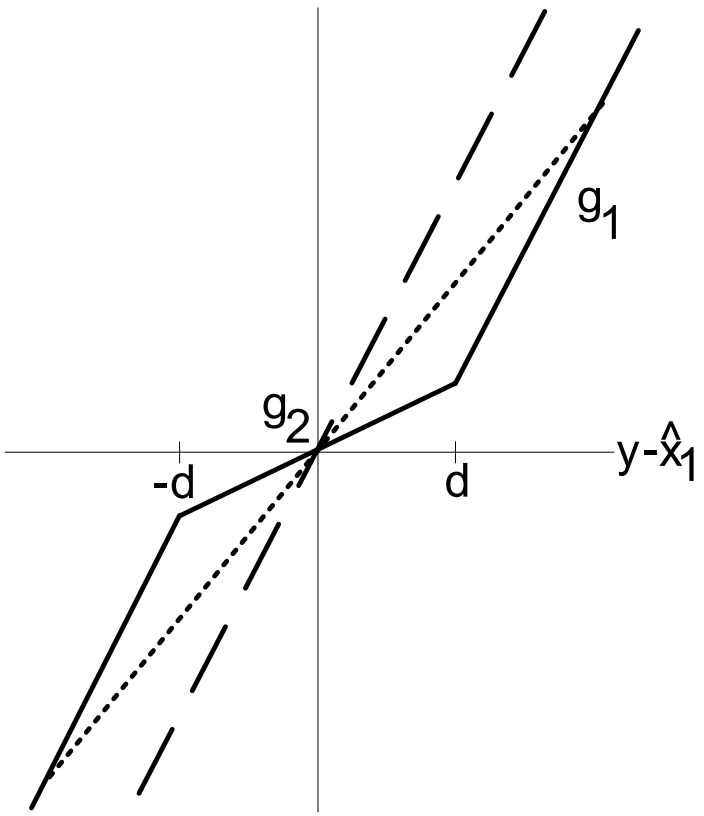

Figure 2.2: Plot of the two-piece nonlinear-gain function compared with the linear-gain (g_1).

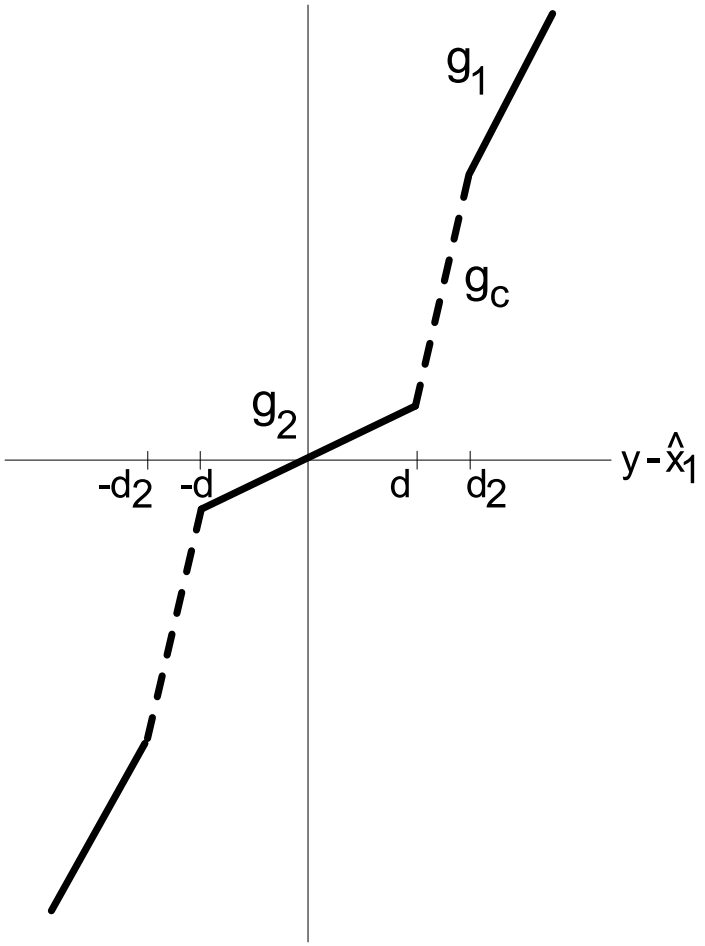

Figure 2.3: Plot of the three-piece nonlinear-gain function.

The high-gain observer is defined as

$$\dot{\hat{x}} = A\hat{x} + B\phi_0(\hat{x}, w, \varsigma, u) + h(y - \hat{x}_1) \ , \tag{2.8}$$

where the nonlinear gain is

$$h_i(y - \hat{x}_1) = \alpha_i \left[g_1^i(y - \hat{x}_1) + d(g_2^i - g_1^i)\text{sat}\left(\frac{y - \hat{x}_1}{d}\right)\right] \tag{2.9}$$

for the two-piece structure and

$$h_i(y-\hat{x}_1) = \alpha_i \left[g_1^i(y-\hat{x}_1) + d(g_2^i - g_{ci})\text{sat}\left(\frac{y-\hat{x}_1}{d}\right) + d_2(g_{ci} - g_1^i)\text{sat}\left(\frac{y-\hat{x}_1}{d_2}\right)\right] \tag{2.10}$$

for the three-piece version. The representations of the piecewise linear functions shown in (2.9) and (2.10) are derived using Proposition 1 of [24], which states that a piecewise linear function can be written as a summation of a linear function with multiple saturation functions. The function "sat" denotes the saturation function defined as

$$\text{sat}(e) = \begin{cases} e, & \text{if } |e| \leq 1 \\ \text{sign}(e), & \text{if } |e| > 1 \end{cases} . \tag{2.11}$$

The expression for g_{ci} is given as

$$g_{ci} = \frac{d_2 g_1^i - d_1 g_2^i}{d_2 - d} ,$$

where the observer gains are defined as

$$g_1 = 1/\varepsilon_1 \text{ and } g_2 = 1/\varepsilon_2, \text{ where } \varepsilon_1 < \varepsilon_2$$

and chosen to correspond to the desired transient and steady-state responses, respectively. Both ε_1 and ε_2 are small positive parameters. The parameter d is defined such that the observer gain is the smaller value g_2 for $|x_1 - \hat{x}_1| \leq d$ and $d_2 > d > \mu$. The function ϕ_0 is a nominal model of ϕ.

Assumption 2.4:

- ϕ_0 is locally Lipschitz in its arguments, uniformly in ς, over the domain of interest;

- ϕ_0 is globally bounded in x and zero in $\mathcal{A}$.

The α_i's are designed such that the roots of

$$s^n + \alpha_1 s^{n-1} + \cdots + \alpha_{n-1} s + \alpha_n = 0 \tag{2.12}$$

are real and negative, with one fast root and $(n-1)$ slow real roots. In this case, (2.12) is written as

$$(s^{n-1} + \beta_1 s^{n-2} + \cdots + \beta_{n-2} s + \beta_{n-1})(s + \lambda) = 0 \ , \tag{2.13}$$

where the first polynomial is Hurwitz with $\mathcal{O}(1)$ real roots and $\lambda \gg 1$. Relating (2.12) to (2.13), it can be seen that $\alpha_1 = \lambda + \beta_1$, $\alpha_i = \beta_{i-1}\lambda + \beta_i \ \forall \ 1 < i < n$ and $\alpha_n = \beta_{n-1}\lambda$. The output feedback controller is obtained by replacing x in (2.5)-(2.6) with $\hat{x}$.

2.4 Closed-Loop System Analysis

For the closed-loop system analysis, the observer dynamics are replaced by the equivalent dynamics of the scaled estimation error

$$\eta = D(\varepsilon_1)(x - \hat{x}) \ , \tag{2.14}$$

where $D(\varepsilon_1) = \text{diag}[1, \varepsilon_1, \cdots, \varepsilon_1^{n-1}]$. The closed-loop system under the output feedback controller can be written as

$$\begin{aligned}\dot{\chi} &= f(\chi, \varsigma, D^{-1}(\varepsilon_1)\eta) \\ &= \begin{bmatrix} Ax + B\phi(x, z, \varsigma, \gamma(\theta, x - D^{-1}(\varepsilon_1)\eta, w, \varepsilon)) \\ \psi(x, z, \varsigma, \gamma(\theta, x - D^{-1}(\varepsilon_1)\eta, w, \varepsilon)) \\ \Gamma(\theta, x - D^{-1}(\varepsilon_1)\eta, w, \varepsilon) \end{bmatrix}\end{aligned} \tag{2.15}$$

$$\varepsilon_1\dot{\eta} = A_0\eta + B_0 v + \varepsilon_1^n B\delta(\chi, \varsigma, w, D^{-1}(\varepsilon_1)\eta) + \bar{h}\delta_1(\eta_1 + v)\ , \tag{2.16}$$

where

$$A_0 = \begin{bmatrix} -\alpha_1 & 1 & \cdots & \cdots & 0 \\ -\alpha_2 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ -\alpha_{n-1} & \cdots & \cdots & 0 & 1 \\ -\alpha_n & \cdots & \cdots & \cdots & 0 \end{bmatrix}, B_0 = \begin{bmatrix} -\alpha_1 \\ -\alpha_2 \\ \vdots \\ -\alpha_{n-1} \\ -\alpha_n \end{bmatrix},$$

$$\bar{h}_i = \alpha_i \left[1 - \left(\frac{\varepsilon_1}{\varepsilon_2}\right)^i\right]$$

and

$$\delta_1(\eta_1 + v) = \begin{cases} \text{sat}\left(\frac{\eta_1+v}{d}\right) d, & \text{two-piece gain} \\ \frac{dd_2}{d_2-d}\left(\text{sat}\left(\frac{\eta_1+v}{d}\right) - \text{sat}\left(\frac{\eta_1+v}{d_2}\right)\right), & \text{three-piece gain} \end{cases}$$

where the value of δ_1 depends on the form of the nonlinear gain. Yet, regardless the structure of the gain, $|\delta_1| \leq d$. The function $\delta(\chi, \varsigma, w, D^{-1}(\varepsilon_1)\eta)$ is de-

fined as $\phi(x, z, \varsigma, \gamma(\theta, \hat{x}, w, \varsigma)) - \phi_0(\hat{x}, w, \varsigma, \gamma(\theta, \hat{x}, w, \varsigma))$, and the matrix A_0 is Hurwitz. The equations (2.15)-(2.16) resemble a model appearing in the standard singularly perturbed form, as shown in [35]. The primary difference between this system and the standard form is the presence of the negative powers of ε_1 in the term $D^{-1}(\varepsilon_1)\eta$. However, δ is a globally bounded function in $\hat{x}$, implying that it is also globally bounded in $D^{-1}(\varepsilon_1)\eta$. This property allows us to extend the analysis associated with standard singularly perturbed systems to the case involving (2.15)-(2.16). The slow dynamics of (2.15) can be approximated by defining $\varepsilon_1/\varepsilon_2 \triangleq \varepsilon_f$, setting $\varepsilon_2 = 0$ and keeping $\varepsilon_f \neq 0$, which yields $\eta = 0$. This reduces (2.15) to the closed-loop system (2.7) under the state feedback controller (2.5)-(2.6). Moreover, the system is globally uniformly asymptotically stable with respect to the compact positively invariant set $\mathcal{A}$. Then, according to a converse Lyapunov theorem in [39], there exists a smooth Lyapunov function $V(\chi)$, two class $\mathcal{K}_\infty$ functions U_1 and U_2, and a class $\mathcal{K}$ function U_3 such that

$$U_1(\|\chi\|_{\mathcal{A}}) \leq V(\chi) \leq U_2(\|\chi\|_{\mathcal{A}}) \tag{2.17}$$

$$\frac{\partial V}{\partial \chi} f(\chi, \varsigma, 0) \leq -U_3(\|\chi\|_{\mathcal{A}}) \tag{2.18}$$

for all $\varsigma \in \mathcal{D}$.

Theorem 2.1: Let Assumptions 2.1 through 2.4 hold and consider the closed-loop system (2.7) with the observer (2.8). Moreover, let $\mathcal{M}$ be any compact set in $\mathbb{R}^N$ and $\mathcal{N}$ be any compact subset of $\mathbb{R}^n$, where $\chi(t_0) \in \mathcal{M}$ and $\hat{x}(t_0) \in \mathcal{N}$. Then, given the positive constant $\varepsilon_f = \varepsilon_1/\varepsilon_2 < 1$, there is a positive constant λ^* such that for $\lambda > \lambda^*$, the following properties hold:

- There exist positive constants μ^* and c_a such that for $\mu < \mu^*$ there is a constant $\varepsilon_a = \varepsilon_a(\mu) > c_a \mu^{1/n}$ with $\lim_{\mu \to 0} \varepsilon_a(\mu) = \varepsilon_2^* > 0$, such

that for each $\varepsilon_2 \in (c_a\mu^{1/n}, \varepsilon_a]$ the trajectories of the closed-loop system are bounded for all $t \geq 0$.

- There exists $\mu_2^* > 0$ and a class $\mathcal{K}$ function ρ_1 such that for every $\mu < \mu_2^*$ and every $\Upsilon_1 > \rho_1(\mu)$, there are constants $T_\Upsilon = T_\Upsilon(\Upsilon_1) \geq 0$ and $\varepsilon_b = \varepsilon_b(\mu, \Upsilon_1) > c_a\mu^{1/n}$, with $\lim_{\mu\to 0} \varepsilon_b(\mu, \Upsilon_1) = \varepsilon_b^*(\Upsilon_1) > 0$, such that for each $\varepsilon_2 \in (c_a\mu^{1/n}, \varepsilon_b]$

$$\max\{|\chi(t)|_{\mathcal{A}}, \|x(t) - \hat{x}(t)\|\} \leq \Upsilon_1, \forall t \geq T_\Upsilon \ . \tag{2.19}$$

- There exist $\mu_3^* > 0$ and a class $\mathcal{K}$ function ρ_2 such that for every $\mu < \mu_3^*$ and every $\Upsilon_2 > \rho_2(\mu)$, there is a constant $\varepsilon_c = \varepsilon_c(\mu, \Upsilon_2) > c_a\mu^{1/n}$, with $\lim_{\mu\to 0} \varepsilon_c(\mu, \Upsilon_2) = \varepsilon_c^*(\Upsilon_2) > 0$, such that for each $\varepsilon_2 \in (c_a\mu^{1/n}, \varepsilon_c]$

$$\|\chi(t) - \chi_r(t)\| \leq \Upsilon_2, \forall t \geq t_0 \tag{2.20}$$

where $\chi_r(t)$ is the solution of (2.7) with $\chi_r(t_0) = \chi(t_0)$.

The last two bullet items are similar to Theorem 1 of [2].

Remark 2.1: The procedure for choosing the design parameters ε_1, ε_2 and d is fairly simple. Initially, the observer gain ε_1 should be chosen to correspond to the desired transient behavior in the estimation error dynamics; the approximate results can be seen by running the chosen value in a linear observer. Without violating the condition on the ratio of the ε's, the value of the parameter ε_2 is set to achieve the desired steady-state estimation error dynamics. Finally, the value of d should be chosen as small as possible, while respecting the lower bound $d > \mu$, imposed by the measurement noise. By choosing the value of d small, the nonlinear observer is able to maximize the use of ε_1 during the transient phase. In the case of the three-piece

nonlinear gain, the additional parameter d_2 should be chosen as close as possible to d.

Proof: In order to place the set $\mathcal{M}$ in the interior of $\Omega_c = \{V(\chi) \leq c\} \subset \mathbb{R}^N$, choose $c > \max_{\chi \in \mathcal{M}} V(\chi)$. The set Ω_c is compact for any choice of c. We have already established that δ is a globally bounded function in $D^{-1}(\varepsilon_1)\eta$. Therefore, there is a constant $L_\delta > 0$, independent of ε_1, such that $\|\delta\| \leq L_\delta$ for all $\chi \in \Omega_c$ and $\eta \in \mathbb{R}^n$.

Consider the fast equation (2.16) for $\chi \in \Omega_c$. This equation possesses both slow and fast variables due to the choice of the eigenvalues shown in (2.13). To transform (2.16) into the singularly perturbed form, A_0 and B_0 are represented as $A_0 = A_{01}\lambda + A_{02}$ and $B_0 = B_{01}\lambda + B_{02}$. The procedure from [35] is used with the change of coordinates

$$\begin{bmatrix} \zeta \\ \eta_1 \end{bmatrix} = T\eta \,, \tag{2.21}$$

where

$$T = \begin{bmatrix} Y \\ Z \end{bmatrix} = \begin{bmatrix} -\beta_1 & 1 & 0 & \cdots & 0 \\ -\beta_2 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ -\beta_{n-1} & 0 & \cdots & \cdots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

and $T^{-1} = \begin{bmatrix} M & N \end{bmatrix}$ with $M \in \mathbb{R}^{n\times(n-1)}$ and $Y \in \mathbb{R}^{(n-1)\times n}$. Applying this change of coordinates to (2.16) yields

$$\varepsilon_1 \dot{\zeta} = YA_{02}M\zeta + YB_{02}v + \varepsilon_1^n YB\delta + Y\bar{h}\delta_1 \tag{2.22}$$

$$\varepsilon_1 \dot{\eta}_1 = -\lambda\eta_1 + \zeta_1 - (\lambda+\beta_1)v + (\lambda+\beta_1)\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right)\delta_1\ , \tag{2.23}$$

where

$$Y\bar{h} = \begin{bmatrix} -\beta_1 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ -\beta_{n-1} & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} (\lambda+\beta_1)\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right) \\ (\lambda\beta_1+\beta_2)\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right) \\ \vdots \\ \lambda\beta_{n-1}\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^n\right) \end{bmatrix} = \lambda\frac{\varepsilon_1}{\varepsilon_2}a + b,$$

$$a = \begin{bmatrix} \beta_1\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right) \\ \beta_2\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right) \\ \vdots \\ \beta_{n-1}\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^{n-1}\right) \end{bmatrix},$$

$$b = \begin{bmatrix} -\beta_1^2\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right) + \beta_2\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right) \\ -\beta_1\beta_2\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right) + \beta_3\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^3\right) \\ \vdots \\ -\beta_{n-1}\beta_1\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right) \end{bmatrix},$$

$$YA_{01} = 0, YB_{01} = 0, ZA_{01}M = 0, ZA_{01}N = -1,$$

$$ZA_{02}M = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \in \mathbb{R}^{1\times(n-1)},$$

$A_{02}N = 0$, and $YA_{02}M$ is by design a Hurwitz matrix. The solution matrix P to the Lyapunov equation $PYA_{02}M + (YA_{02}M)^T P = -I$ is symmetric and positive definite. Let the Lyapunov function candidate for (2.22) be chosen as $W_1 = \zeta^T P \zeta$. It can be shown that

$$\dot{W}_1 \leq -\frac{1}{\varepsilon_1}\|\zeta\|^2 + \frac{2}{\varepsilon_1}\|\zeta\| \left(\|PYB_{02}\|\mu + \varepsilon_1^n \|PYB\|L_\delta \right. \\ \left. + \|P\| \left(\frac{\varepsilon_1}{\varepsilon_2}\lambda\|a\| + \|b\|\right) d\right). \tag{2.24}$$

It follows that

$$\dot{W}_1 \leq -\frac{1}{2\varepsilon_1\|P\|} W_1, \forall\, W_1 \geq L_W \tag{2.25}$$

for

$$L_W = 4\|PYB_{02}\|\mu + \varepsilon_1^n 4\|PYB\|L_\delta \\ + 4\|P\| \left(\frac{\varepsilon_1}{\varepsilon_2}\lambda\|a\| + \|b\|\right) d. \tag{2.26}$$

Thus, the ζ states are bounded. The bound (2.26) is reached within the interval $[t_0,\, t_0 + T_1(\varepsilon_1)]$, where

$$T_1(\varepsilon_1) = 4\|P\|\varepsilon_1 \ln\left(\frac{k_{w_1}^2}{\sigma_1 \varepsilon_1^n}\right) \to 0 \text{ as } \varepsilon_1 \to 0 \tag{2.27}$$

and $k_{w_1} > 0$ is a constant independent of ε_1 and ε_2, and
$\sigma_1 = \sqrt{\|P\|}\|PYB\|L_\delta$.

Meanwhile, the next step is to show that the trajectories of (2.23) reach a positively invariant strip defined as

$$\{|x_1 - \hat{x}_1| \leq L\}\,, \tag{2.28}$$

where $0 < L < d$. Using the Lyapunov function $W_2 = \frac{1}{2}\eta_1^2$, it can be shown that

$$\dot{W}_2 = -\frac{\lambda}{\varepsilon_2}\eta_1(\eta_1+v) - \lambda\left(\frac{1}{\varepsilon_1}-\frac{1}{\varepsilon_2}\right)\eta_1(\eta_1+v-\delta_1) + \eta_1\left[\beta_1\left(\frac{1}{\varepsilon_1}-\frac{1}{\varepsilon_2}\right)\delta_1 - \frac{1}{\varepsilon_1}\beta_1 v + \frac{1}{\varepsilon_1}\zeta_1\right]. \tag{2.29}$$

It follows from $d > |v|$ that $\text{sign}(\eta_1) = \text{sign}(\eta_1 + v)$ whenever $|\eta_1 + v| \geq d$ and

$$-\lambda\left(\frac{1}{\varepsilon_1}-\frac{1}{\varepsilon_2}\right)\eta_1(\eta_1+v-\delta_1) \leq 0.$$

To ensure that the condition in (2.28) is satisfied, we require a tighter bound on the state ζ_1 than what is provided in (2.25). Thus, the ultimate bound on $|\zeta_1|$ is denoted as

$$c_0 + k_1\mu + k_2\varepsilon_1^n L_\delta + k_3\frac{\varepsilon_1}{\varepsilon_2}\lambda d + k_4 d\,, \tag{2.30}$$

where c_0, a constant due to initial conditions, can be made arbitrarily small for t large enough; k_1, k_2, k_3, and k_4 are positive constants. Using the ultimate bound in (2.30), after $t_0 + T_1(\varepsilon_1)$,

$$\dot{W}_2 \leq -\frac{(1-\theta)\lambda}{\varepsilon_2}|\eta_1|^2, \forall\ |\eta_1| \geq U \tag{2.31}$$

for

$$U = \frac{\mu}{\theta} + \frac{\varepsilon_2}{\varepsilon_1}\left(\frac{c_1\mu + c_2 d + c_0 + k_2\varepsilon_1^n L_\delta}{\lambda\theta}\right) + \frac{k_3 d}{\theta}\,, \tag{2.32}$$

where $c_1 = \beta_1 + k_1$, $c_2 = \beta_1\left(1 - \frac{\varepsilon_1}{\varepsilon_2}\right) + k_4$ and $\varepsilon_1 \leq 1$. The parameter $\theta \in (0,1)$ will be chosen in a later step.

To ensure that (2.28) implies

$$|x_1 - \hat{x}_1 + v| \leq d\,, \tag{2.33}$$

d is chosen as $d > L + \mu$. In order for the strip (2.28) to be positively invariant, we require the choice of L to be greater than U. This requirement leads to the inequality $U < L < d - \mu$. We rewrite the foregoing inequality as $U + \mu < d$, and revisit U defined in (2.32). Recall that the terms in U containing μ in the numerator and/or λ in the denominator can be made sufficiently small. However, the third term, $k_3 d/\theta$, does not immediately appear to be necessarily small. To ensure that the inequality $U + \mu < d$ is satisfied, the constant k_3 needs to be handled with care. In particular, we need $k_3/\theta < 1$ to avoid violating the foregoing inequality. According to Lemma 1 of [53], the constant k_3 arises from the response of the system

$$\varepsilon_1 \dot{\bar{\zeta}} = Y A_{02} M \bar{\zeta} + a\lambda \frac{\varepsilon_1}{\varepsilon_2}$$

and is given by

$$k_3 = \int_0^\infty |\bar{C} e^{Y A_{02} M t} a| dt \tag{2.34}$$

for $\bar{C} = [1\ 0\ \cdots\ 0]$. Define

$$G(s) = \bar{C}(sI - Y A_{02} M)^{-1} a\ . \tag{2.35}$$

Then, by selecting the poles and zeros of $G(s)$ to be real and distinct such that

$$G(s) = K \left(\prod_{i=1}^{\bar{m}} \frac{s + \bar{z}_i}{s + \bar{p}_i} \right) \left(\prod_{j=\bar{m}+1}^{\bar{n}} \frac{1}{s + \bar{p}_i} \right),$$

where $\bar{m} \leq \bar{n}$, $\bar{z}_i > \bar{p}_i$ for $i = 1, ...\bar{m}$, the impulse response of (2.34) is nonnegative; see [26] for the proof. This assumption can always be satisfied by an appropriate

choice of $\beta_1, \cdots, \beta_{n-1}$ and $\varepsilon_1/\varepsilon_2$; see Appendix A for examples. Hence,

$$k_3 = \int_0^\infty \bar{C} e^{YA_{02}Mt} a dt = G(s)\Big|_{s=0} \tag{2.36}$$

where solving for k_3 leads to

$$YA_{02}M = \begin{bmatrix} -\beta_1 & 1 & 0 & \dots & 0 \\ -\beta_2 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ -\beta_{n-2} & 0 & \dots & \dots & 1 \\ -\beta_{n-1} & 0 & \dots & \dots & 0 \end{bmatrix},$$

$$\bar{C}(sI - YA_{02}M)^{-1}a = \frac{\beta_1\left(1 - \frac{\varepsilon_1}{\varepsilon_2}\right)s^{n-2} + \cdots + \beta_{n-1}\left(1 - \left(\frac{\varepsilon_1}{\varepsilon_2}\right)^{n-1}\right)}{s^{n-1} + \beta_1 s^{n-2} + \cdots + \beta_{n-1}}$$

and $k_3 = 1 - \left(\dfrac{\varepsilon_1}{\varepsilon_2}\right)^{n-1}$. Therefore,

$$U + \mu = \frac{\mu}{\theta} + \mu + \frac{c_1\mu + c_2 d + c_0 + k_2\varepsilon_1^n L_\delta}{\varepsilon_f \lambda\theta} + \frac{(1 - \varepsilon_f^{n-1})d}{\theta}.$$

Recall that μ must be small enough and λ chosen large enough to ensure that $U + \mu < d$. Let

$$\mu_1^* = \frac{(1-k)\varepsilon_f^{n-1} d}{2 - k\varepsilon_f^{n-1}},$$

$$\lambda^* = \frac{c_1\mu_1^* + c_2 d + k_2 L_\delta}{\varepsilon_f^n[(1-k)d + k\mu_1^*] - 2\varepsilon_f\mu_1^*}$$

and require that $\mu \in (0, \mu_1^*)$, $\lambda > \lambda^*$ and $k \in (0,1)$. Substituting the upper bound μ_1^* and the lower bound λ^* into the expression for $U + \mu$ leads to

$$\begin{aligned} U + \mu &< \frac{\mu_1^*}{\theta} + \mu_1^* + \frac{1}{\theta}[\varepsilon_f^{n-1}(1-k)d + \varepsilon_f^{n-1}k\mu_1^* - 2\mu_1^*] + \frac{(1-\varepsilon_f^{n-1})d}{\theta} \\ &= \frac{1}{\theta}[(1+\theta+\varepsilon_f^{n-1}k-2)\mu_1^* + (\varepsilon_f^{n-1}(1-k)+1-\varepsilon_f^{n-1})d] \\ &= \frac{1}{\theta}[(1+\theta+\varepsilon_f^{n-1}k-2)\mu_1^* + (1-\varepsilon_f^{n-1}k)d]. \end{aligned}$$

By choosing $\theta = 1 - k\varepsilon_f^{n-1}$ and $\varepsilon_1 \leq 1$, the condition $U + \mu < d$ is satisfied.

We will now show that all trajectories reach the strip (2.28) in finite time. The following inequality originates from (2.31)

$$\dot{W}_2 \leq -\lambda \frac{2k\varepsilon_f^n}{\varepsilon_1} W_2, \forall\, W_2 \geq \frac{1}{2}L^2. \tag{2.37}$$

Therefore, the set

$$\Sigma_2 = \{W_2 \leq \frac{1}{2}L^2\} = \{|\eta_1| \leq L\} \tag{2.38}$$

is positively invariant. If $\eta_1(t_0)$ is outside of Σ_2, then from (2.37)

$$W_2(\eta_1(t)) \leq W_2(\eta_1(t_0))\exp\left(-\lambda \frac{2k\varepsilon_f^n}{\varepsilon_1}(t-t_0)\right). \tag{2.39}$$

From the scaling equation (2.14), it can be seen that whenever $x(t_0)$ and $\hat{x}(t_0)$ are bounded, there exists a constant $k_{w_2} > 0$, independent of ε_1 and ε_2, such that $W_2(\eta_1(t_0)) \leq k_{w_2}^2$. From (2.38) and (2.39), it can be seen that η_1 reaches the set Σ_2 within the time interval $[t_0,\, t_0 + T_2(\varepsilon_1)]$, where

$$T_2(\varepsilon_1) = \frac{\varepsilon_1}{\lambda k \varepsilon_f^n} \ln \left(\frac{k_{w_2}}{\sqrt{\frac{1}{2}}L} \right) \to 0 \text{ as } \varepsilon_1 \to 0. \tag{2.40}$$

At this point, η_1 is inside the strip (2.28) and cannot leave the strip for all future time. Inside the strip, the parameter ε_2 is driving the dynamics of the high-gain observer, not ε_1. Therefore, it is appropriate to alter the scaling equation (2.14) to obtain

$$\xi = D(\varepsilon_2)(x - \hat{x}) \ , \tag{2.41}$$

where $D(\varepsilon_2) = \text{diag}[1, \varepsilon_2, \cdots, \varepsilon_2^{n-1}]$. Then, the error dynamics become

$$\varepsilon_2 \dot{\xi} = A_0 \xi + \varepsilon_2^n B \delta + B_0 v \ , \tag{2.42}$$

which is valid for trajectories inside the strip. Take the Lyapunov function candidate as $W_3 = \xi^T S \xi$, where S is the positive definite symmetric solution to the Lyapunov equation $SA_0 + A_0^T S = -I$. Then, it can be shown that

$$\dot{W}_3 \leq -\frac{1}{2\varepsilon_2 \|S\|} W_3, \forall \ W_3 \geq \left(\sigma_2 \varepsilon_2^n + \sigma_3 \mu\right)^2 \ , \tag{2.43}$$

where $\sigma_2 = 4\|SB\| L_\delta \sqrt{\|S\|}$ and $\sigma_3 = 4\|SB_0\| \sqrt{\|S\|}$. Therefore, the set

$$\Sigma_3 = \{W_3 \leq \left(\sigma_2 \varepsilon_2^n + \sigma_3 \mu\right)^2\} \tag{2.44}$$

is positively invariant. For $\xi(t_0)$ outside Σ_3, it can be seen from (2.43) that

$$W_3(\xi(t)) \leq W_3(\xi(t_0)) \text{exp} \left(\frac{-(t - t_0)}{2\varepsilon_2 \|S\|} \right). \tag{2.45}$$

From the scaling equation in (2.41), for bounded $x(t_0)$, $\hat{x}(t_0)$ and ε_2, there exists a constant $k_{w3} > 0$, independent of ε_2, such that $W_3(\xi(t_0)) \leq k_{w3}^2$. It follows from (2.44) and (2.45) that ξ reaches the set Σ_3 within the time interval $[t_0, t_0 + T_3(\varepsilon_2)]$, where

$$T_3(\varepsilon_2) = 4\varepsilon_2 \|S\| \ln\left(\frac{k_{w3}}{\sigma_2 \varepsilon_2^n}\right) \to 0 \text{ as } \varepsilon_2 \to 0. \tag{2.46}$$

Inside the set Σ_3

$$\|x(t) - \hat{x}(t)\| = \|D^{-1}(\varepsilon_2)\xi(t)\| \leq \varepsilon_2\gamma_1 + \frac{\mu}{\varepsilon_2^{n-1}}\gamma_2 \triangleq F_r(\varepsilon_2, \mu) \tag{2.47}$$

for $\gamma_1 = \sigma_2/\sqrt{\lambda_{min}(S)}$ and $\gamma_2 = \sigma_3/\sqrt{\lambda_{min}(S)}$. Therefore, all of the trajectories are traveling towards the positively invariant set $\Sigma = \{\Sigma_2 \cap \Sigma_3\}$. Moreover, the set Σ is reached within the time interval

$$T(\varepsilon_1, \varepsilon_2) = T_1(\varepsilon_1) + T_2(\varepsilon_1) + T_3(\varepsilon_2) \to 0 \text{ as } \varepsilon_2 \to 0\ , \tag{2.48}$$

where the times are defined in (2.27), (2.40) and (2.46). Recall that $\varepsilon_1 < \varepsilon_2$, meaning that decreasing ε_2 will eventually correspond to a decrease in ε_1. Hence, only the reduction of ε_2 is explicitly listed in (2.48).

We will now address the remaining arguments (including the slow vector χ) for the boundness of all trajectories, ultimate boundness where the trajectories come close to the set $\mathcal{A} \times \{x - \hat{x} = 0\}$ and closeness of trajectories. A similar approach for the slow states χ can be found in [2].

For $(\chi, \eta) \in \Omega_c \times \Sigma$ the vector χ can be represented as

$$\dot{\chi} = f(\chi, \varsigma, D^{-1}(\varepsilon_2)\xi). \tag{2.49}$$

The function f is globally bounded in $D^{-1}(\varepsilon_2)\xi$, because it is globally bounded in

$\hat{x}$ and the term $D^{-1}(\varepsilon_2)\xi$ results from substituting $x - D^{-1}(\varepsilon_2)\xi$ for $\hat{x}$. Hence, there exists a positive constant k_f, independent of ε_2, such that

$$\|f(\chi, \varsigma, D^{-1}(\varepsilon_2)\xi)\| \leq k_f. \tag{2.50}$$

The following lemma originally appeared in [2] and is included here for convenience; see [2] for the proof.

Lemma 1: The function $F_r(\varepsilon_2, \mu)$ has the following properties for $\varepsilon_2 > 0$ and $\mu \geq 0$. First, $F_r(\varepsilon_2, \mu)$ has a global minimum at

$$\varepsilon_2 = [(n-1)\gamma_2\mu/\gamma_1]^{(1/n)} \triangleq c_a\mu^{1/n}$$

and

$$\min_{\varepsilon_2>0} F_r(\varepsilon_2, \mu) = (c_1c_a + \gamma_2/c_a^{n-1})\mu^{1/n} \triangleq k_a\mu^{1/n} \ .$$

For $\varepsilon_2 > c_a\mu^{1/n}$, $F_r(\varepsilon_2, \mu)$ is a strictly increasing function of ε_2 and $F_r(\varepsilon_2, \mu) \leq k_b\varepsilon_2$, where $k_b = \gamma_1 + \gamma_2/c_a^n$. Then, given $k_r > 0$, for every $\mu \in [0, (k_r/k_a)^n)$ there exist $\varepsilon_m = \varepsilon_m(\mu, k_r) \geq 0$ and $\varepsilon_M = \varepsilon_M(\mu, k_r) > c_a\mu^{1/n}$, with

$$\varepsilon_m \leq \min\{c_a\mu^{1/n}, (\mu\gamma_2 n/k_r)^{1/(n-1)}\}$$

and

$$\lim_{\mu\to 0} \varepsilon_M(\mu, k_r) = k_r/\gamma_1 \ ,$$

such that $F_r(\varepsilon_2, \mu) \leq k_r$ for all $\varepsilon_2 \in (\varepsilon_m, \varepsilon_M]$.

With $\eta = 0$, the output feedback expression in (2.15) reduces to the state feedback representation shown in (2.7). Then, for $\varepsilon_2 \in (c_a\mu^{1/n}, \varepsilon_2^*)$ and $\mu < \mu_1^*$,

there is a positive constant L_1, independent of ε_2, such that the following Lipschitz condition is satisfied

$$\|f(\chi, \varsigma, D^{-1}(\varepsilon_2)\xi) - f(\chi, \varsigma, 0)\| \leq L_1 \|D^{-1}(\varepsilon_2)\xi\| \tag{2.51}$$

for all $(\chi, \eta) \in \Omega_c \times \Sigma$, where the lower bound on ε_2 is chosen according to Lemma 1. Taking the derivative of the smooth Lyapunov function $V(\chi)$ yields

$$\begin{aligned}
\dot{V} &= \frac{\partial V}{\partial \chi} f(\chi, \varsigma, D^{-1}(\varepsilon_2)\xi) + \frac{\partial V}{\partial \chi} f(\chi, \varsigma, 0) - \frac{\partial V}{\partial \chi} f(\chi, \varsigma, 0) \\
&= \frac{\partial V}{\partial \chi} f(\chi, \varsigma, 0) + \frac{\partial V}{\partial \chi} \Big[f(\chi, \varsigma, D^{-1}(\varepsilon_2)\xi) - f(\chi, \varsigma, 0) \Big] \\
&\leq -U_3(\chi) + \left\| \frac{\partial V}{\partial \chi} \right\| \|f(\chi, \varsigma, D^{-1}(\varepsilon_2)\xi) - f(\chi, \varsigma, 0)\|.
\end{aligned}$$

Let L_2 be an upper bound on $\|\partial V / \partial \chi\|$ over Ω_c. Then,

$$\dot{V} \leq -U_3(\chi) + L_2 L_1 \|x - \hat{x}\| \triangleq -U_3(\chi) + L_r F_r(\varepsilon_2, \mu) \tag{2.52}$$

for all $(\chi, \eta) \in \Omega_c \times \Sigma$. Let $L_3 = (1/L_r) \min_{\chi \in \partial\Omega_c} U_3(\chi)$ and apply Lemma 1 with $k_r = L_3$ and set $\mu_4^* = (k_r/k_a)^n$. Then, for $\mu < \mu_4^*$ and $\varepsilon_2 \in (c_a \mu^{1/n}, \varepsilon_M]$, we have $\dot{V} \leq 0$ for all $(\chi, \eta) \in \partial\Omega_c \times \Sigma$. From the previous analysis, we have that $\dot{W}_3 \leq 0$ for all $(\chi, \eta) \in \Omega_c \times \partial\Sigma$. Therefore, the set $\Omega_c \times \Sigma$ is positively invariant, which implies that the set $\Omega_c \times \Sigma$ is positively invariant.

Given $\chi(t_0)$ is in the interior of Ω_c, we have

$$\chi(t) - \chi(t_0) = \int_{t_0}^{t} f(\chi(\tau), u(\tau)) d\tau.$$

Considering that f is continuous and its arguments are bounded, $\|f(\chi, u)\| \leq k_1$

for all $\chi(t) \in \Omega_c$, where k_1 is independent of ε_2 . Hence,

$$\|\chi(t) - \chi(t_0)\| \leq k_1(t - t_0) \tag{2.53}$$

as long as $\chi(t) \in \Omega_c$. Therefore, there exists a finite time T^*, independent of ε_2, such that $\chi(t) \in \Omega_c$ for all $t \in [t_0, t_0 + T^*]$. The previous analysis showed that η enters the set Σ during the finite time period $[t_0, t_0 + T(\varepsilon_1, \varepsilon_2)]$, where $T(\varepsilon_1, \varepsilon_2) \to 0$ as $\varepsilon_2 \to 0$. Thus, there exists ε_u such that for all $0 < \varepsilon_2 \leq \varepsilon_u$, $T(\varepsilon_1, \varepsilon_2) \leq T^*$. Thus, choosing μ_5^* small enough that $c_a(\mu_5^*)^{1/n} < \varepsilon_u$ and setting $\mu^* = \min\{\mu_1^*, \mu_4^*, \mu_5^*\}$ and $\varepsilon_a = \min\{\varepsilon_M, \varepsilon_u\}$, we conclude that for $0 \leq \mu < \mu^*$ and $c_a\mu^{1/n} < \varepsilon_2 \leq \varepsilon_a$, the trajectory (χ, η) enters the set $\Omega_c \times \Omega$ during the finite time period $[t_0, t_0 + T(\varepsilon_1, \varepsilon_2)]$ and stays there for the remainder of time. Prior to entering the set, $\chi(t)$ and $\eta(t)$ are bounded by (2.53) and (2.45), respectively. Therefore, the closed-loop trajectories are bounded.

Ultimate Boundness: The proof for ultimate boundness utilizes the dynamics of the ξ vector inside the strip (2.28).

Apply Lemma 1 to the expressions (2.47) and (2.52) by setting $k_r = \tau$, where $\tau \in (k_a\mu^{1/n}, L_3]$ and set $\varepsilon_b = min\{\varepsilon_a, \varepsilon_M\}$; the equations (2.47) and (2.52) are satisfied for all $t \geq t_0 + T^*$. Thus, for $\varepsilon_2 \in (c_a\mu^{1/n}, \varepsilon_b]$ the function $F_r(\varepsilon_2, \mu) \leq \tau$. Then, from (2.52)

$$\dot{V} \leq -U_3(\chi) + L_r\tau = -\frac{1}{2}U_3(\chi) - \frac{1}{2}U_3(\chi) + L_r\tau.$$

This results in

$$\dot{V} \leq -\frac{1}{2}U_3(\chi)$$

for all $\chi \notin \{U_3(\chi) \leq 2L_r\tau\}$. Given $U_3(\chi)$ is positive definite and continuous, there is a positive constant $\tau^* < L_3$ such that the set $U_3(\chi) \leq 2L_r\tau$ is compact

for $\tau \leq \tau^*$. Let $c_o(\tau) = \max_{U_3 \leq 2L_r\tau}\{V(\chi)\}$, then $c_o(\tau)$ is a nondecreasing function that tends to zero as $\tau \to 0$. Choose a class $\mathcal{K}$ function $\varphi(\tau)$ such that $\varphi \geq c_o(\tau)$. Then, for $\varphi(\tau) \leq V(\chi) \leq c$

$$\dot{V} \leq -\frac{1}{2}U_3(\chi).$$

Therefore, there exists a time $T_\varphi = T_\varphi(\tau) \geq 0$ such that

$$\dot{V}(\chi(t)) \leq \varphi(\tau)$$

for all $t \geq t_0 + T^* + T_\varphi(\tau)$. It follows from (2.17) that

$$|\chi(t)|_{\mathcal{A}} \leq U_1^{-1}(\varphi(\tau)) \triangleq \rho_a(\tau)$$

for all $t \geq t_0 + T^* + T_\varphi(\tau)$. Moreover,

$$\max\{|\chi(t)|_{\mathcal{A}}, \|x(t) - \hat{x}(t)\|\} \leq \max\{\tau, \rho_a(\tau)\} \triangleq \rho_b(\tau)$$

where ρ_b is a class $\mathcal{K}_\infty$ function. Define ρ_1 as $\rho_1(\mu) = \rho_b(k_a\mu^{1/n})$ with $\mu < \mu_2^* = \min\{\mu^*, (\tau^*/k_a)^n\}$, where ρ_1 is a class $\mathcal{K}_\infty$. Given $\Upsilon_1 > \rho_1(\mu)$, take $\tau = \min\{\tau^*, \rho_b^{-1}(\Upsilon_1)\}$ and set $T_\Upsilon = t_0 + T^* + T_\varphi(\tau)$ to achieve the inequality in (2.19).

Closeness of Trajectories: Using the fact that the closed-loop system under state feedback is uniformly asymptotically stable with respect to the set $\mathcal{A}$ and the inequality in (2.19), given $\Upsilon_2 > 2\rho_1(\mu)$, there exists a time $T_{\Upsilon_2} = T_{\Upsilon_2}(\Upsilon_2) > 0$, independent of ε_2, such that

$$|\chi(t)|_{\mathcal{A}} \leq \frac{\Upsilon_2}{2} \tag{2.54}$$

and

$$|\chi_r(t)|_{\mathcal{A}} \leq \frac{\Upsilon_2}{2} \tag{2.55}$$

for all $t \geq T_{\Upsilon_2}$ for $\varepsilon_2 \in (c_a \mu^{1/n}, \varepsilon_b]$. Then,

$$\|\chi(t) - \chi_r(t)\| \leq \|\chi(t) - x\| + \|\chi_r(t) - x\| \tag{2.56}$$

for all $t \geq T_{\Upsilon_2}$ and $x \in \mathcal{A}$. Then, we take the infimum of the right-hand side of (2.56), over all $x \in \mathcal{A}$, and substitute the values in (2.54) and (2.55). This results in the following

$$\|\chi(t) - \chi_r(t)\| \leq |\chi(t)|_{\mathcal{A}} + |\chi_r(t)|_{\mathcal{A}} \leq \Upsilon_2 \tag{2.57}$$

for all $t \geq T_{\Upsilon_2}$. Furthermore, we can see from (2.53) that

$$\|\chi(t) - \chi(t_0)\| \leq k_1(t - t_0) \tag{2.58}$$

and similarly

$$\|\chi_r(t) - \chi(t_0)\| \leq k_1(t - t_0) \tag{2.59}$$

for all $t \in [t_0, t_0 + T(\varepsilon_1, \varepsilon_2)]$. Therefore,

$$\|\chi(t) - \chi_r(t)\| \leq 2k_1 T(\varepsilon_1, \varepsilon_2) \tag{2.60}$$

for all $t \in [t_0, t_0 + T(\varepsilon_1, \varepsilon_2)]$. Viewing the closed-loop system under output feedback as a perturbation of the closed-loop system under state feedback and applying Theorem 3.4 of [30] over the interval $[t_0 + T(\varepsilon_1, \varepsilon_2), T_{\Upsilon_2}]$ gives the following

$$\|\chi(t) - \chi_r(t)\| \leq 2k_1 c_3 T(\varepsilon_1, \varepsilon_2) + c_4 F_r(\varepsilon_2, \mu) \tag{2.61}$$

for some constants $c_3 \geq 1$ and $c_4 > 0$, independent from ε_1 and ε_2. It can be seen from (2.60) and (2.61) that the time for which (2.61) is valid can be simplified to $[t_0, T_{\Upsilon_2}]$. Furthermore, it can be verified that $T(\varepsilon_1, \varepsilon_2)$ is a class $\mathcal{K}$ function of ε_2 for $\varepsilon_2 \leq (1/e)(k_{w3}/\sigma_2)^{1/n} \triangleq \bar{\varepsilon}_2$. The inequality $c_a(\mu_6^*)^{1/n} < \bar{\varepsilon}_2$ can be satisfied for a sufficiently small μ_6^*. Then, for each $\mu < \mu_6^*$ the following statement holds

$$\begin{aligned} \min_{\varepsilon_2 \in (c_a \mu^{1/n}, \bar{\varepsilon}_2]} & \{2k_1 c_3 T(\varepsilon_1, \varepsilon_2) + c_4 F_r(\varepsilon_2, \mu)\} \\ & = 2k_1 c_3 T(\varepsilon_f c_a \mu^{1/n}, c_a \mu^{1/n}) + c_4 k_a \mu^{1/n} \\ & \triangleq \rho_3(\mu). \end{aligned} \tag{2.62}$$

Moreover, it can be shown that $\rho_3(\mu)$ is a class $\mathcal{K}$ function. For each $\Upsilon_2 > \rho_3(\mu)$, there exists $\bar{\varepsilon}_c = \bar{\varepsilon}_c(\mu, \Upsilon_2) > c_a \mu^{1/n}$ with the $\lim_{\mu \to 0} \bar{\varepsilon}_c(\mu, \Upsilon_2) = \bar{\varepsilon}_c^*(\Upsilon_2) > 0$ such that for all $\varepsilon_2 \in (c_a \mu^{1/n}, \bar{\varepsilon}_c]$

$$2k_1 c_3 T(\varepsilon_1, \varepsilon_2) + c_4 F_r(\varepsilon_2, \mu) \leq \Upsilon_2 \tag{2.63}$$

is satisfied. The inequality in (2.20) can be found by taking $\rho_2(\mu) = \max\{2\rho_1(\mu), \rho_3(\mu)\}$, $\mu_3^* = \min\{\mu_2^*, \mu_6^*\}$ and $\varepsilon_c = \min\{\varepsilon_b, \bar{\varepsilon}_c, \bar{\varepsilon}_2\}$, in combination with (2.57), (2.61) with the simplified time and (2.63).

2.5 Simulation: Field Controlled DC Motor

The system under consideration is a field controlled DC motor [30], where it is desired that the shaft angular velocity track a reference signal as shown in Figure 2.4. The

system is represented as

$$\dot{x}_1 = x_2 \tag{2.64}$$

$$\dot{x}_2 = \phi(x, z, u) \tag{2.65}$$

$$\dot{z} = \psi(x, z, u) \tag{2.66}$$

$$y = x_1 + v \tag{2.67}$$

$$w = z \ , \tag{2.68}$$

where x_1 is the rotor position, x_2 the rotor angular velocity and z the armature current. Notice that the available measurements are the rotor position and the current. Moreover, the current measurement is not an output resulting from the chain of integrators.

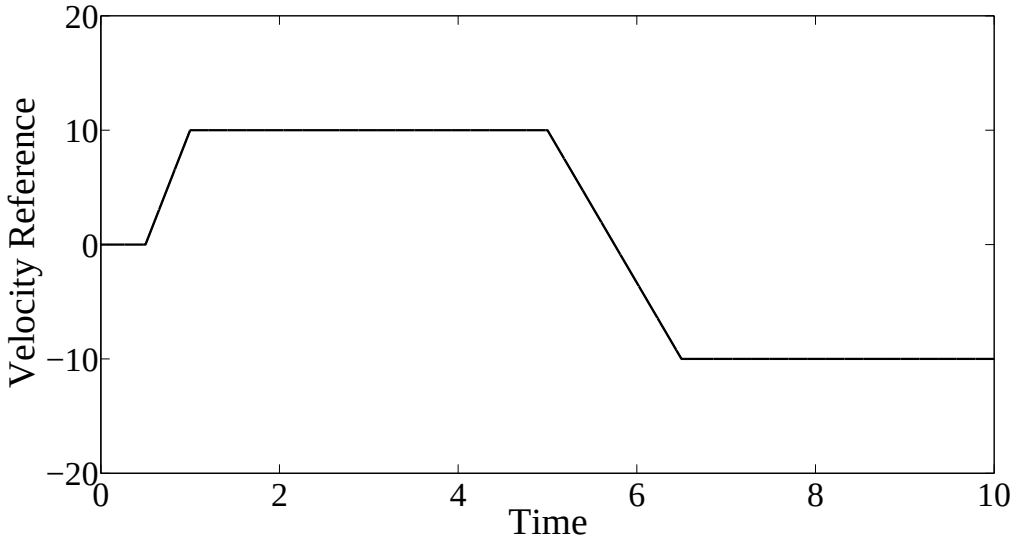

Figure 2.4: Velocity reference trajectory $(\dot{r})$.

The field current is used as the source of control and is denoted by u. The controller expression is $u = \dfrac{10}{w}(0.11\hat{x}_2 + \ddot{r} - 100(y - r) - 20(\hat{x}_2 - \dot{r}))$, where feedback linearization is applied. The functions above are defined as $\phi(x, z, u) = -0.1x_2 + 0.1zu$ and $\psi(x, z, u) = -2z - 0.2x_2u + 200$. The estimate $\hat{x}_2$

is saturated outside [-100, 100]. The saturation values are chosen such that the saturation is never active when the system is under state feedback control. The nominal value for ϕ used in the observer is $\phi_0(\hat{x}, w, u) = -0.11x_2 + 0.1wu$. The state z is measured and need not be estimated; thus, the observer is second-order. The gains for the observers are chosen as $\varepsilon_1 = 0.0005$ and $\varepsilon_2 = 0.01$. The remaining parameters are chosen as $\alpha_1 = 71$ and $\alpha_2 = 70$. The initial conditions are set at $x_1(0) = x_2(0) = \hat{x}_2(0) = 0$, $\hat{x}_1(0) = 0.02$ and $z(0) = 100$ to match values consistent with the physical system. The initial conditions are deliberately chosen to be unequal to ensure peaking in the transient response of the system, lending itself to a more realistic scenario. The measurement noise v is generated using the Simulink block "Uniform Random Number", where the magnitude is limited to [-0.0016, 0.0016] and the sampling time is set at 0.0008 seconds. The noise magnitude is based on a 1000 c/r encoder. The value of d for the switched observer threshold is 0.005. The parameters for the nonlinear-gain observers are $d = 0.0035$ and $d_2 = 0.05$. For the switched observer, it was shown in [2] that if the system switches before the transient response of the estimates of the higher order derivatives has subsided and entered a positively invariant set, the other transients can take $(y - \hat{x}_1)$ out of the strip and subsequently cause the value of ε to switch. If this occurs, the system could be susceptible to multiple switching until all of the trajectories recover from peaking. Thus, the switched observer requires the additional component of a switching timer, based on the peaking period, that prevents the observer from switching before the trajectories of the estimation error have reached a positively invariant set. The delay timer is set for 0.15 seconds; details on how to choose this value can be found in [2].

Figure 2.5 shows the transient response of the error $x_2 - \hat{x}_2$. As expected, the switched observer captures the behavior of the linear observer with the parameter ε_1. Unlike the switched observer, the two-piece nonlinear-gain observer does not wait for the transients to subside in both states before entering the strip. Therefore, the

system utilizing the nonlinear-gain observer does not perfectly mimic the transient response of the system with the linear observer shown in Figure 2.5(c). However, the nonlinear-gain observer is able to recover the performance of the observer in Figure 2.5(d) faster than the switched observer. As a result, the presence of noise is more noticeable in the estimation error generated with the switched observer than with the nonlinear-gain observer. Figure 2.6 compares the transient performance of the three-piece nonlinear-gain observer with the two-piece and two linear-gain high-gain observers. Clearly, there is no appreciable difference in the estimation error $x_2 - \hat{x}_2$ between the two nonlinear-gain observers.

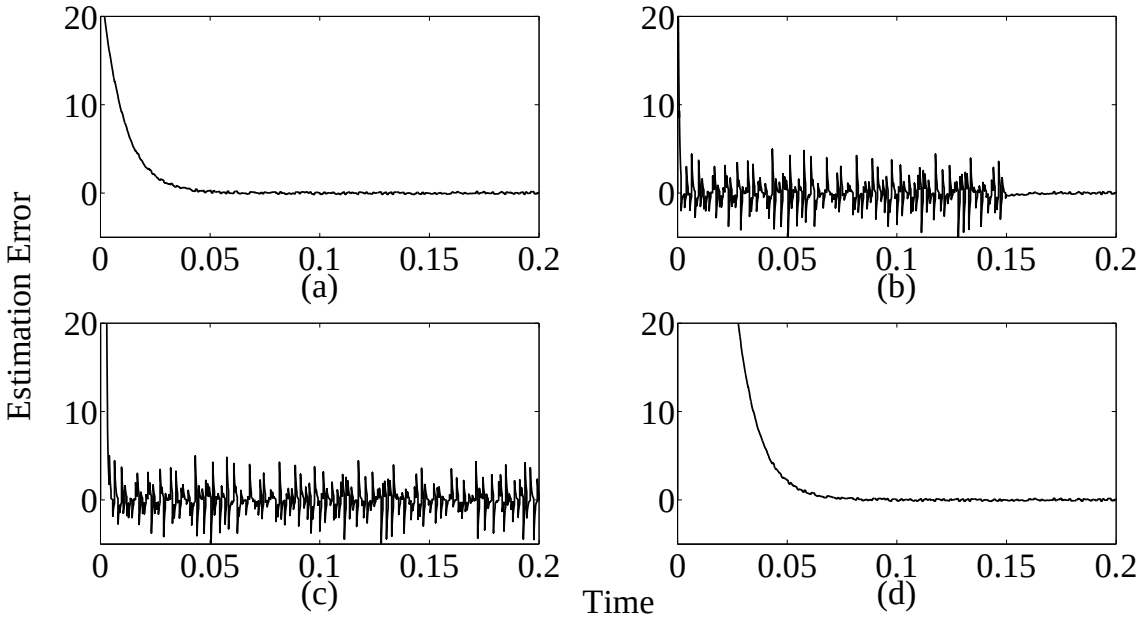

Figure 2.5: Transient response of the error $x_2 - \hat{x}_2$ vs. time for a (a) Two-Piece Nonlinear, (b) Switched, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

In Figures 2.7 and 2.8, the estimation error steady-state behavior is practically identical for all observers shown, with the exception of the linear ε_1 observer.

Figure 2.9 shows the tracking error $x_2 - \dot{r}$ during the transient response of the observer dynamics. The transient response resulting from the system using the nonlinear-gain observers is faster than both linear observers. As shown in Figures

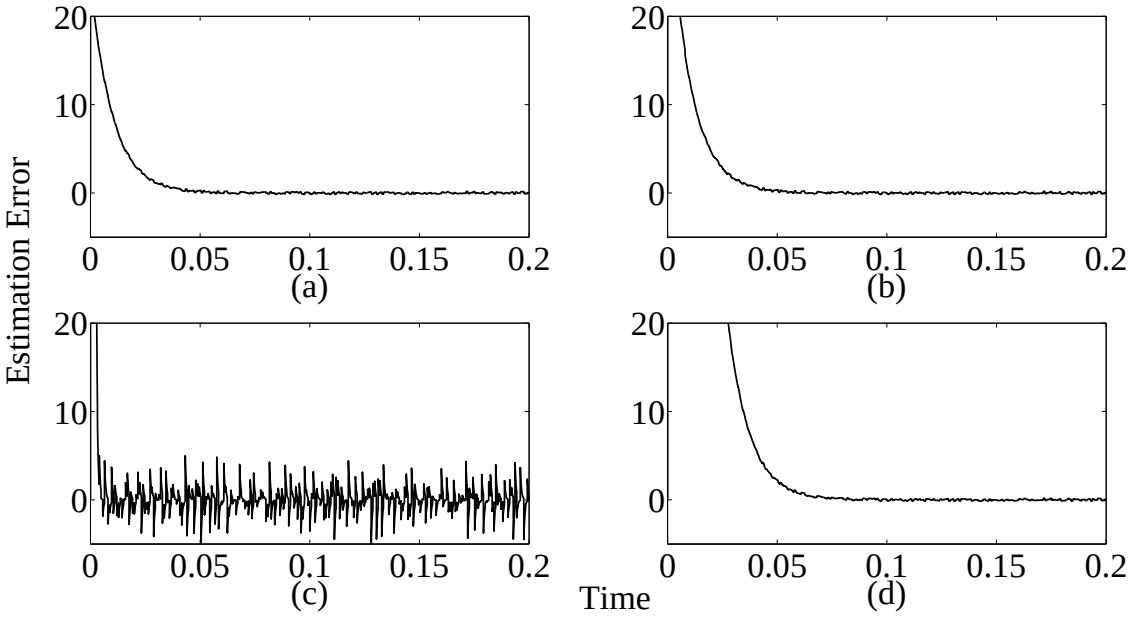

Figure 2.6: Transient response of the error $x_2 - \hat{x}_2$ vs. time for a (a) Two-Piece Nonlinear, (b) Three-Piece Nonlinear, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

2.11 and 2.12, the steady-state response of the tracking errors are nearly identical for four out of the five observers. The nonlinear gain observers exhibit similar system behavior, as reported in Figure 2.10 and Figure 2.12.

Overall, the nonlinear-gain observers are able to achieve better system performance than the two linear-gain observers, while bypassing the complications typically associated with switching. In general, the two-piece observer is able to perform just as well as an observer with three distinct gain regions.

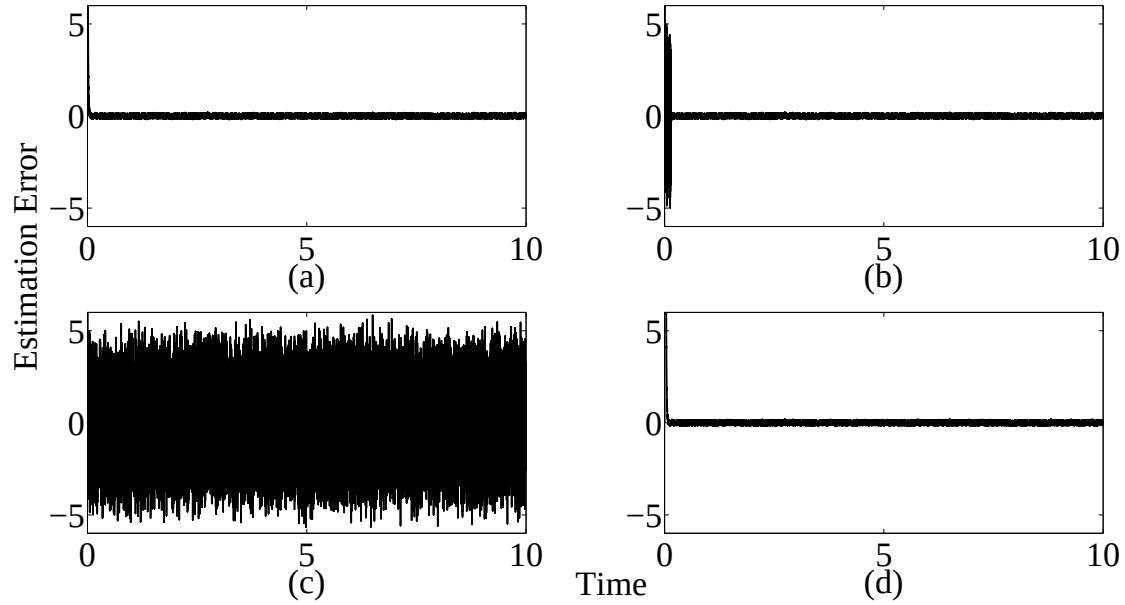

Figure 2.7: Steady-state response of the error $x_2 - \hat{x}_2$ vs. time for a (a) Two-Piece Nonlinear, (b) Switched, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

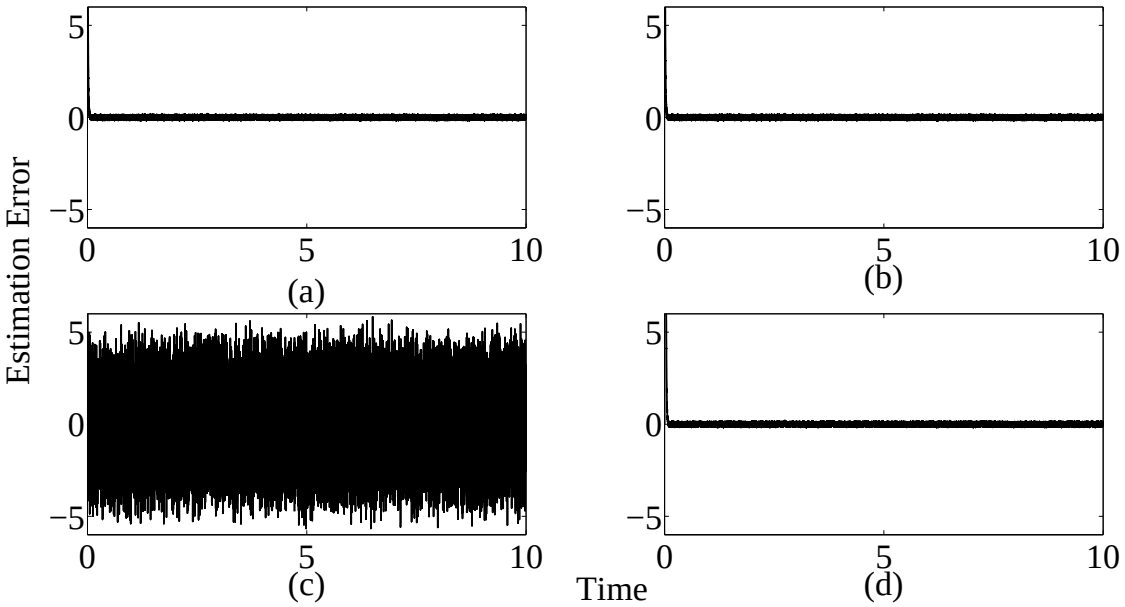

Figure 2.8: State-state response of the error $x_2 - \hat{x}_2$ vs. time for a (a) Two-Piece Nonlinear, (b) Three-Piece Nonlinear, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

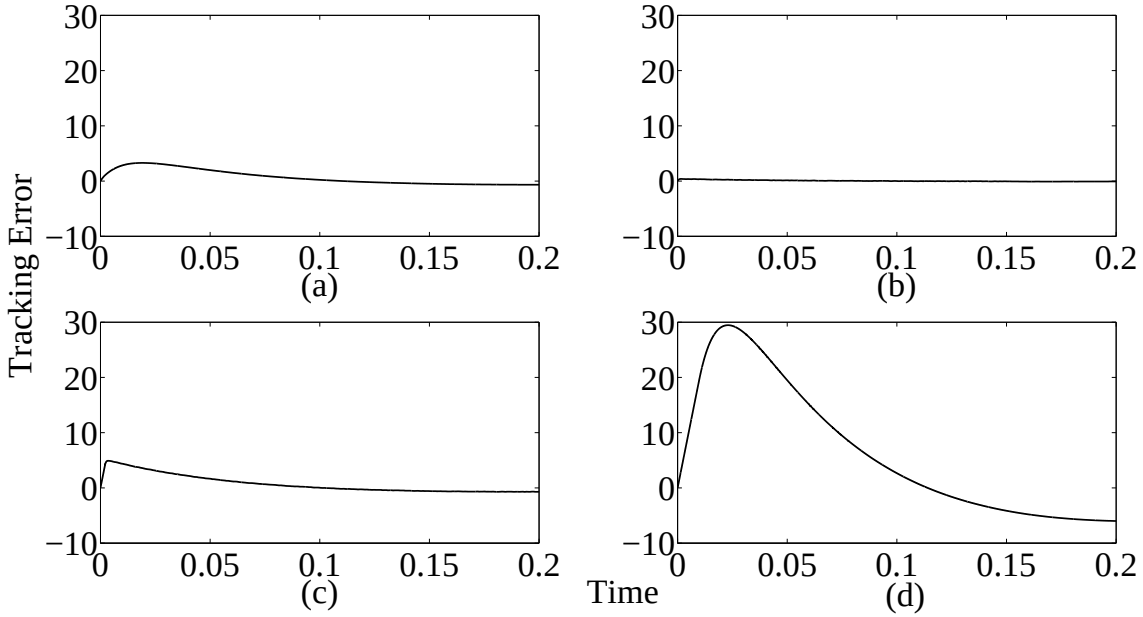

Figure 2.9: Transient response of the tracking error $x_2 - \dot{r}$ vs. time for a (a) Two-Piece Nonlinear, (b) Switched, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

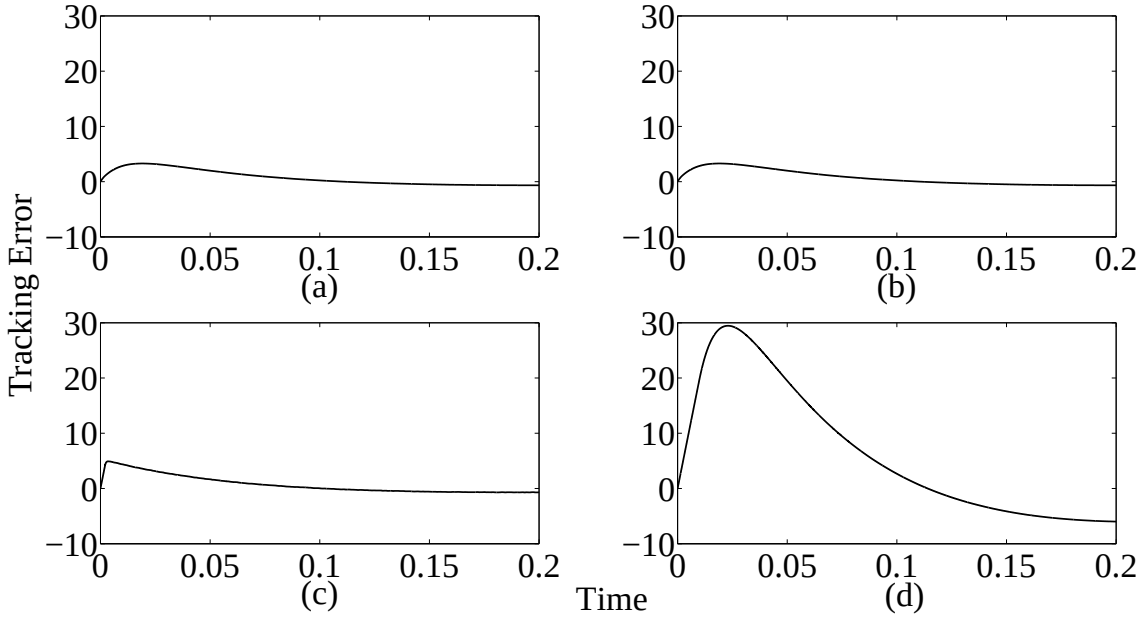

Figure 2.10: Transient response of the tracking error $x_2 - \dot{r}$ vs. time for a (a) Two-Piece Nonlinear, (b) Three-Piece Nonlinear, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

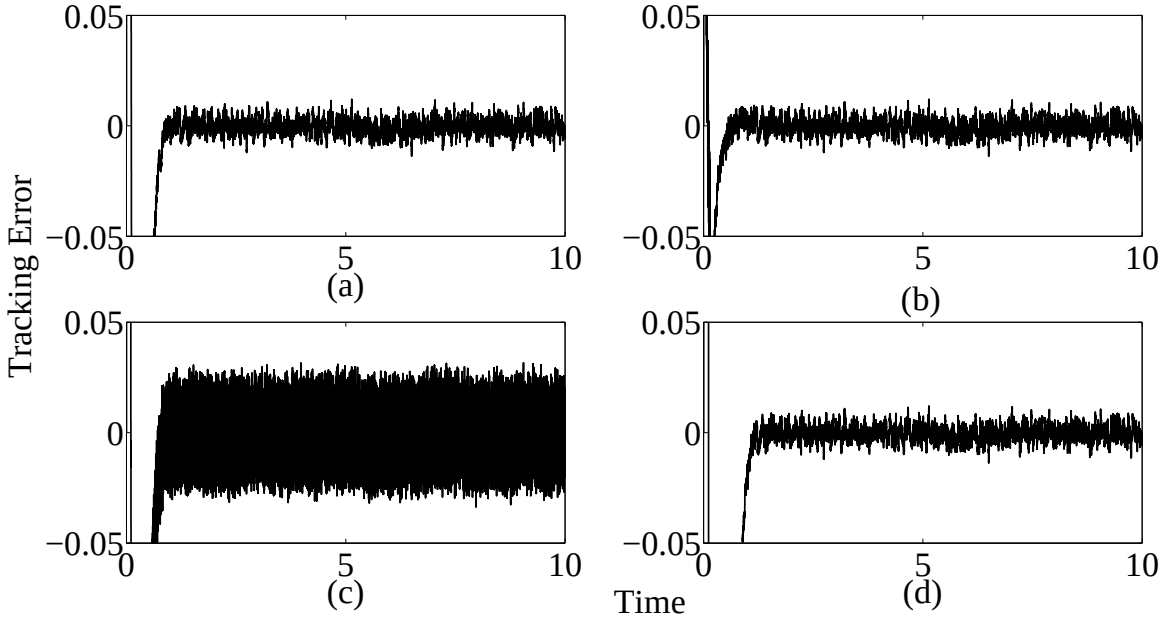

Figure 2.11: Steady-state response of the tracking error $x_2 - \dot{r}$ vs. time for a (a) Two-Piece Nonlinear, (b) Switched, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

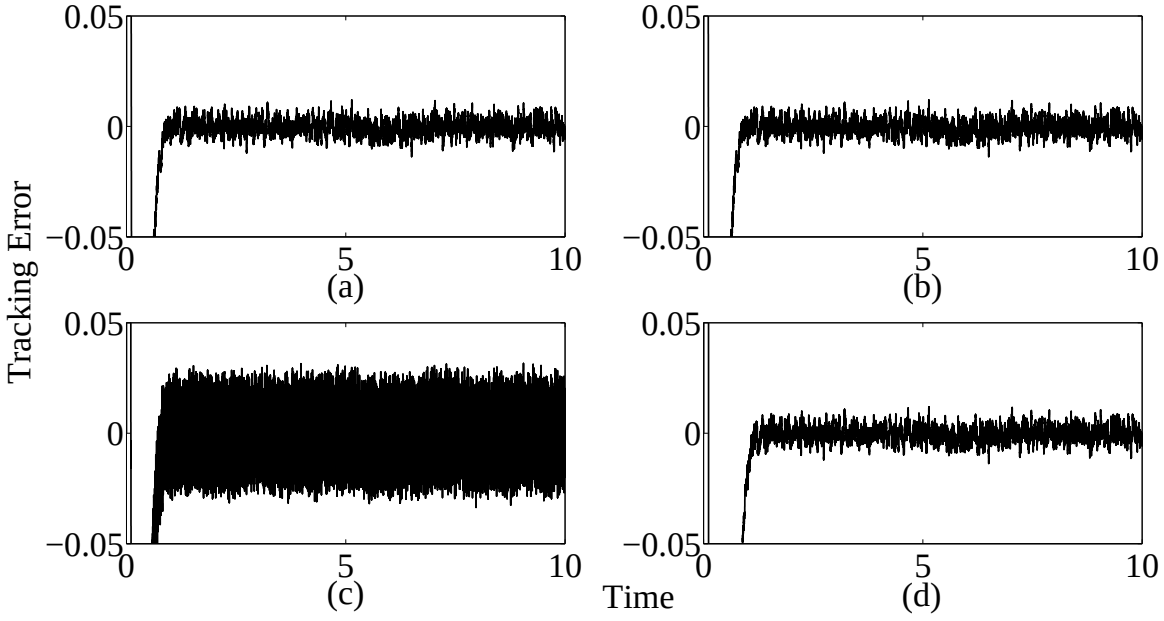

Figure 2.12: Steady-state response of the tracking error $x_2 - \dot{r}$ vs. time for a (a) Two-Piece Nonlinear, (b) Three-Piece Nonlinear, (c) Linear ε_1 and (d) Linear ε_2 gain high-gain observer.

2.6 Conclusions

When high-gain observers are utilized in the presence of measurement noise, there exists a tradeoff between fast state reconstruction and a reasonable state estimation error. The nonlinear-gain high-gain observers adequately captured the transient and steady-state performance seen in comparable linear-gain observers. Specifically, the nonlinearity was chosen to have a higher observer gain during the transient period and a lower gain afterwards, thus overcoming the tradeoff between fast state reconstruction and measurement noise attenuation. It appears as if altering the number of piecewise linear regions in the nonlinear gain function do not have an appreciable effect on the system performance, at least for the types of systems considered. Moreover, it was assumed that all assumptions hold globally, allowing the use of a converse Lyapunov theorem of [39] in the proof. By a slight modification of the proof, it is also possible to require the assumptions to hold only in a given region of the state space and invoke the converse Lyapunov theorem of [7], which is a regional version of the theorem of [39]. In particular, the first derivative of the exogenous signals, $\dot{\varsigma}$, is required to be bounded.

Chapter 3

High-Gain-Observer Tracking Performance in the Presence of Measurement Noise

3.1 Introduction

Prior to this chapter, a significant amount of emphasis has been placed on the performance of the estimation error in the presence of measurement noise. Likewise, the literature on high-gain observers subjected to measurement noise primarily focuses on qualifying the system performance and quantifying potential bounds on the estimation error. However, there are many practical problems that seek to achieve more than just stabilization; namely, path following and other goals that can be worked into a tracking structure. This is not to suggest that the estimation error becomes a moot point, however, given the observer estimates will be used in the feedback control. Yet, a greater insight into the challenges associated with measurement noise can be gained by investigating the coupling between the observer design choices and the effects on the tracking performance. For instance, simulation studies have suggested

that the effect of measurement noise on the tracking error is significantly less than the effect manifested in the estimation error. If this can be shown mathematically, even for just a special class of systems, designers may be able to achieve additional leeway in constructing the high-gain observer and enhanced performance.

In order to explicitly show the importance and prevalence of this topic, Section 3.2 begins the discussion with a nonlinear example. Section 3.3 poses the questions motivated by the previous section. Next, the class of systems investigated are defined in Section 3.4. Before delving into the complexity that arises when dealing with nonlinear systems, Section 3.5 investigates the effect of measurement noise on the tracking error from a linear systems perspective. The result obtained from the linear system analysis is extended to a class of nonlinear systems in Section 3.6, utilizing the framework provided by ordinary differential equations and the unique properties of multi-time scale systems. Ultimately, the tracking error is shown to be uniformly bounded in the observer parameter ε.

3.2 Motivation

Consider the example found in [30]

$$\begin{aligned} \dot{x}_1 &= x_2 && (3.1)\\ \dot{x}_2 &= x_2^3 + u && (3.2)\\ y &= x_1 + v\ , && (3.3) \end{aligned}$$

where the x_i's are the the system states, y the output, u the control and v the measurement noise. The control objective is to have the state x_1 track a sinusoid with an amplitude of 0.1 and a frequency of 0.3 rad/s. Using standard feedback linearization techniques, the state feedback control is chosen as $u = -x_2^3 - (x_1 -$

$r) - (x_2 - \dot{r}) + \ddot{r}$. However, only the first state is available for use in the controller. Therefore, the output feedback control is constructed by replacing the system state x_2 with the estimate obtained from the high-gain observer defined as

$$\dot{\hat{x}}_1 = \hat{x}_2 + \frac{2}{\varepsilon}(y - \hat{x}_1) \tag{3.4}$$

$$\dot{\hat{x}}_2 = \frac{1}{\varepsilon^2}(y - \hat{x}_1) \, . \tag{3.5}$$

Given the choice of output feedback, in order to prevent peaking in the plant during the transient period, the controller is saturated outside [-1, 1]. The bounds on the controller are chosen such that the saturation is not active under state feedback control. The initial conditions are set at $x_1(0) = 0.1$ and $x_2(0) = \hat{x}_1(0) = \hat{x}_2(0) = 0$. Note that $x_1(0)$ and $\hat{x}_1(0)$ are deliberately chosen to be unequal to depict a more realistic scenario. Additionally, when the initial conditions for the system and observer differ, peaking is induced and appears in the transient response. The measurement noise v is generated using the Simulink block "Uniform Random Number", where the magnitude is limited to [-0.00011, 0.00011] and the sampling time is set at 0.00005 seconds. In order to compare the effect the observer parameter ε has on the estimation and tracking error, two separate trials are run with $\varepsilon = 0.001$ and $\varepsilon = 0.0005$.

Figure 3.1 shows the steady-state response of the estimation error $x_2 - \hat{x}_2$ for the linear observers. In particular, as the value of the observer parameter ε is decreased, the magnitude of the error significantly increases. However, the error in tracking the reference signal, displayed in Figure 3.2, shows no appreciable change as ε is decreased. In fact, for values of $\varepsilon \in [0.0005, 0.01]$, the steady-state response of the tracking error is restricted to the range $[-0.000029, 0.00011]$. Hence, the tracking error is uniformly bounded in ε. The same phenomenon is exhibited in linear systems with measurement noise, and will be investigated in a subsequent section.

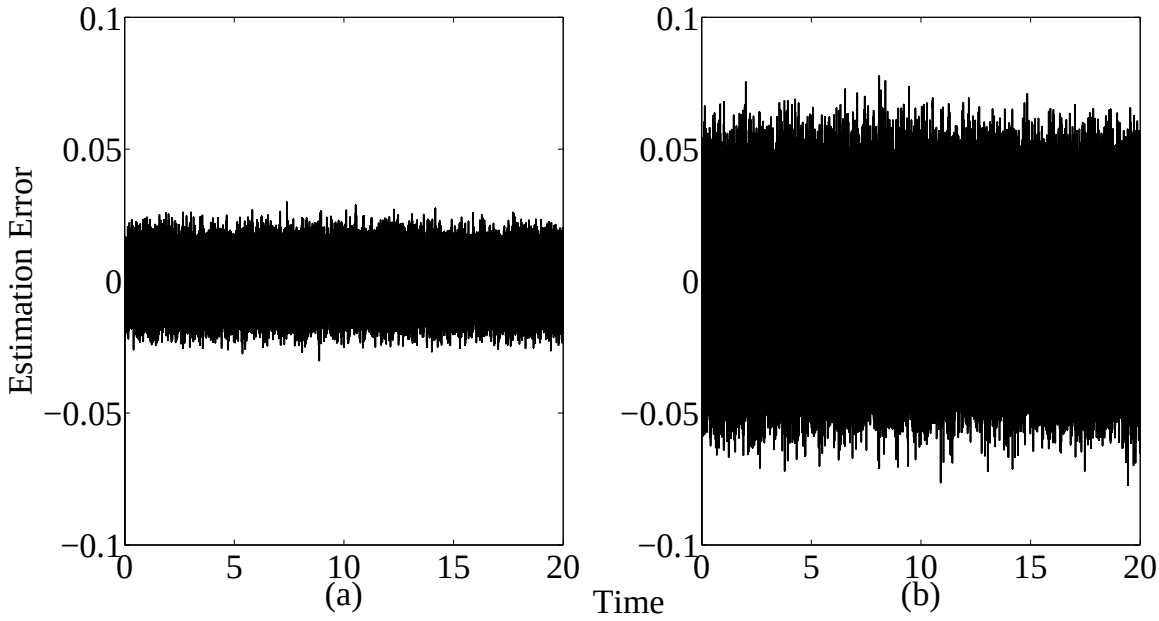

Figure 3.1: Steady-state response of the error $x_2 - \hat{x}_2$ vs. time for a high-gain observer with (a) $\varepsilon = 0.001$ and (b) $\varepsilon = 0.0005$.

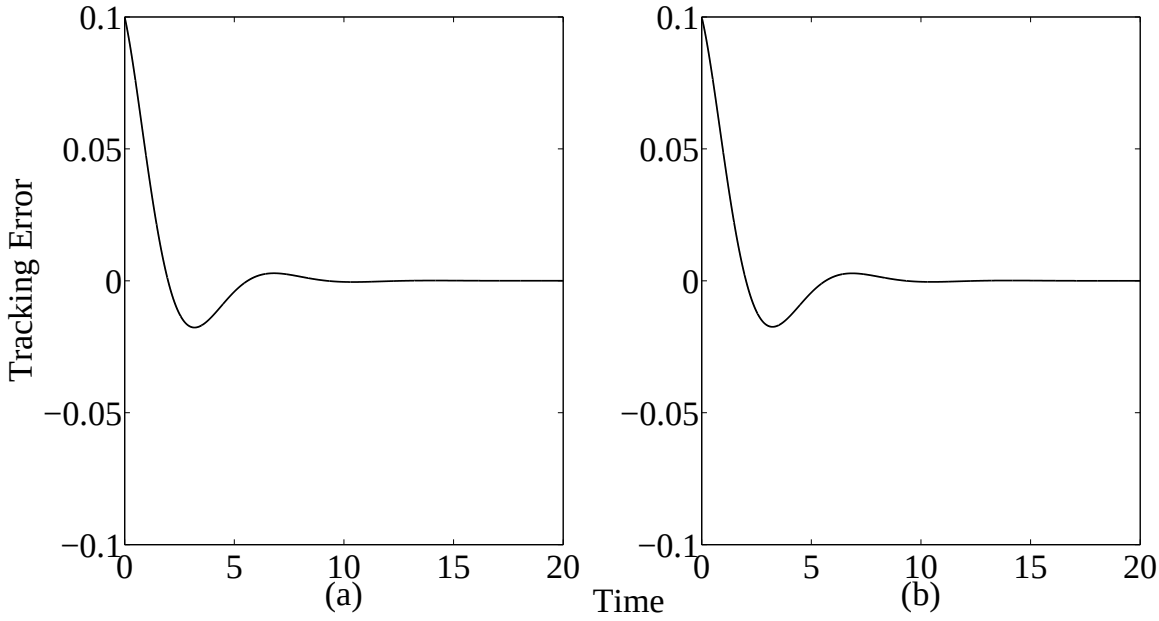

Figure 3.2: Steady-state response of the tracking error $x_2 - \hat{x}_2$ vs. time for a high-gain observer with (a) $\varepsilon = 0.001$ and (b) $\varepsilon = 0.0005$.

3.3 Tracking Performance

This section prepares the reader for the class of systems investigated, and poses the questions motivated by Section 3.2. Referring back to the example in Section 3.2, the simulation results suggest that the tracking error, $x_1 - r$, is uniformly bounded in ε. Let's investigate this claim further.

Unlike nonlinear representations, the transfer functions of a linear system can reveal how the measurement noise impacts the tracking error. Before delving into the details of a time domain analysis, consider the third-order system

$$\begin{aligned} \dot{x}_1 &= x_2 && (3.6)\\ \dot{x}_2 &= x_3 && (3.7)\\ \dot{x}_3 &= a_1 x_1 + a_2 x_2 + a_3 x_3 + bu && (3.8)\\ y &= x_1 + v\ , && (3.9) \end{aligned}$$

where the x_i's are the system states, y the output and u the control. The variable v is the measurement noise. Tracking can be achieved by the state feedback controller

$$u = \frac{1}{b}[-k_1(x_1 - r) - k_2(x_2 - r^{(1)}) - k_3(x_3 - r^{(2)}) + r^{(3)}]\ , \qquad (3.10)$$

where $r(t)$ and $r^{(j)}(t)$ are the tracking signal and j^{th} derivative of tracking signal, respectively. The coefficients k_1, k_2, and k_3 are chosen such that

$$s^3 + (k_3 - a_3)s^2 + (k_2 - a_2)s + (k_1 - a_1) \qquad (3.11)$$

is Hurwitz. The state estimates for the output feedback control are generated with

the linear high-gain observer

$$\dot{\hat{x}}_1 = \hat{x}_2 + \frac{\alpha_1}{\varepsilon}(y - \hat{x}_1) \tag{3.12}$$

$$\dot{\hat{x}}_2 = \hat{x}_3 + \frac{\alpha_2}{\varepsilon^2}(y - \hat{x}_1) \tag{3.13}$$

$$\dot{\hat{x}}_3 = \frac{\alpha_3}{\varepsilon^3}(y - \hat{x}_1)\ , \tag{3.14}$$

where the α_i's are designed such that

$$s^3 + \alpha_1 s^2 + \alpha_2 s + \alpha_3 \tag{3.15}$$

is Hurwitz. The output feedback control is constructed by substituting the state estimates, generated by (3.12)-(3.14), for the states x that appear in (3.10). In this discussion the control is not saturated, as typically employed to avoid peaking in the plant during the transient response. In the case of tracking performance, the analysis is concerned with the system behavior in steady-state, where the saturation is not active.

Define the change of variables

$$e_1 = x_1 - r \tag{3.16}$$

$$e_2 = x_2 - r^{(1)} \tag{3.17}$$

$$e_3 = x_3 - r^{(2)} \tag{3.18}$$

for the tracking error.

The transfer functions from the noise to the tracking errors are

$$\frac{E}{V} = \begin{bmatrix} \frac{E_1}{V} \\ \frac{E_2}{V} \\ \frac{E_3}{V} \end{bmatrix} = \begin{bmatrix} H_1 \\ H_2 \\ H_3 \end{bmatrix} , \tag{3.19}$$

where

$$\begin{aligned} H_1 &= -\frac{\Delta_2}{\Delta_1} \\ H_2 &= sH_1 \\ H_3 &= s^2 H_1 \end{aligned}$$

and

$$\begin{aligned} \Delta_1 = {} & \varepsilon^3 s^6 + (\varepsilon^2 \alpha_1 - \varepsilon^3 a_3) s^5 \\ & + (\varepsilon \alpha_2 - \varepsilon^2 a_3 \alpha_1 - \varepsilon^3 a_2) s^4 \\ & + (\alpha_3 - \varepsilon a_3 \alpha_2 - \varepsilon^2 a_2 \alpha_1 - a_1 \varepsilon^3) s^3 \\ & + (-a_1 \varepsilon^2 \alpha_1 - a_3 \alpha_3 - \varepsilon a_2 \alpha_2 + \alpha_1 \varepsilon^2 k_1 + \alpha_2 k_2 \varepsilon + \alpha_3 k_3) s^2 \\ & + (-a_2 \alpha_3 - a_1 \varepsilon \alpha_2 + \alpha_2 k_1 \varepsilon + \alpha_3 k_2) s + (\alpha_3 k_1 - a_1 \alpha_3) \end{aligned}$$

$$\Delta_2 = (\alpha_1 \varepsilon^2 k_1 + \alpha_2 k_2 \varepsilon + \alpha_3 k_3) s^2 + (\alpha_2 k_1 \varepsilon + \alpha_3 k_2) s + \alpha_3 k_1 .$$

The transfer functions H_1 and H_2 are two-frequency-scale transfer functions according to the definition of [40]. Therefore, their H_∞ norms are of order $\mathcal{O}(1)$, i.e. they are bounded uniformly in ε. The transfer function H_3 is not two-frequency-scale, because setting $\varepsilon = 0$ results in an improper transfer function. However, H_3 can be written as

$$H_3 = \frac{1}{\varepsilon} \bar{H}_3 , \tag{3.20}$$

where $\bar{H}_3$ is a two-frequency-scale transfer function. Hence, the H_∞ norm of $\bar{H}_3$ is $\mathcal{O}(1)$. This shows that the H_∞ norm of H_3 is $\mathcal{O}(1/\varepsilon)$. Furthermore, for any noise that is an $\mathcal{L}_2$ signal of order $\mathcal{O}(\mu)$, the output will also be an $\mathcal{L}_2$ signal of order $\mathcal{O}(\mu)$.

Can a similar result be shown without the aid of frequency domain tools for linear and nonlinear systems? What happens to the other states in the state vector e for a system of dimension n? The next two sections seek to answer these questions.

3.4 Problem Formulation

Consider the class of nonlinear systems that can be represented in the form

$$\dot{x} = Ax + B[b_x(x) + au] \tag{3.21}$$

$$y = Cx + v(t)\ , \tag{3.22}$$

where $x \in \mathbb{R}^n$ is the system states, $y \in \mathbb{R}$ the measured output, $u \in \mathbb{R}$ the control input and $v(t) \in \mathbb{R}$ the measurement noise. The matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^n$ and $C \in \mathbb{R}^{1\times n}$ are defined as

$$A = \begin{bmatrix} 0 & 1 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

and

$$C = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}\ ,$$

where it is assumed that $n \geq 2$, which implies that the relative degree of the system is greater than or equal to two. If this condition is not met, the construction of an observer is unnecessary; the measured output y is simply used. The function $b_x(x)$ may not be known, and $a > 0$. The measurement noise is assumed to be a bounded measurable function, $|v(t)| \leq \mu$. Given the purpose of this chapter is to investigate the effects of measurement noise on the high-gain observer manifested in the tracking error, the following change of variables is defined

$$e_i = \begin{cases} x_i - r(t) & \text{for } i = 1 \\ x_i - r^{(i-1)}(t) & \text{for } 1 < i \leq n \end{cases} , \tag{3.23}$$

where $r(t)$ and $r^{(j)}(t)$ are the tracking signal and its j^{th} derivative. Then, applying the change of variables in (3.23) to the system (3.21)-(3.22) results in the representation

$$\dot{e} = Ae + B[b(t,e) + au] \tag{3.24}$$

$$y = Ce + v(t) , \tag{3.25}$$

where $b(t,e)$ replaces $b_x(x)$.

The state feedback controller takes the following form

$$u = -Ke , \tag{3.26}$$

where the values of the vector K are chosen such that the origin of closed-loop system is asymptotically stable. However, this does not exclude the possibility of employing a nonlinear control technique. For instance, a continuously-implemented sliding mode control is one type of nonlinear control scheme that will fit the controller form in

(3.26). For instance, the control

$$u = -\beta_{\text{sat}}\left(\frac{k_1 e_1 + k_2 e_2 + \cdots + k_{n-1} e_{n-1} + e_n}{\rho}\right)$$

takes the form $u = -Ke$ inside the boundary layer. It is appropriate to consider the controller form inside the boundary layer, given the focus is on the steady-state behavior of the tracking error.

The high-gain observer used to estimate the derivatives of the tracking error is defined as

$$\dot{\hat{e}} = A\hat{e} + H(y - C\hat{e}) \tag{3.27}$$

and the gain is given by

$$H^T = \begin{bmatrix} \frac{\alpha_1}{\varepsilon} & \frac{\alpha_2}{\varepsilon^2} & \cdots & \frac{\alpha_n}{\varepsilon^n} \end{bmatrix} .$$

Consider the scaled estimation error

$$z_i = \frac{e_i - \hat{e}_i}{\varepsilon^{n-i}}, 1 \leq i \leq n \tag{3.28}$$

that when differentiated becomes

$$\varepsilon \dot{z} = A_0 z + \varepsilon B\delta - B_0 \frac{1}{\varepsilon^{n-1}} v \ , \tag{3.29}$$

where

$$A_0 = \begin{bmatrix} -\alpha_1 & 1 & \cdots & \cdots & 0 \\ -\alpha_2 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ -\alpha_{n-1} & \cdots & \cdots & 0 & 1 \\ -\alpha_n & \cdots & \cdots & \cdots & 0 \end{bmatrix}, B_0 = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ \alpha_n \end{bmatrix} \tag{3.30}$$

and $\delta = b(t,e) + au$.

The output feedback controller takes the form

$$u = -K\hat{e} = -Ke + KDz \ , \tag{3.31}$$

where $D = \text{diag}[\varepsilon^{n-1}, \varepsilon^{n-2}, \cdots, 1]$. Then, the closed-loop system under the output feedback control (3.31) can be written in the following form

$$\dot{e} = Ae + B[\gamma(t,e) + aKDz] \tag{3.32}$$

$$\varepsilon\dot{z} = A_0 z + \varepsilon B[\gamma(t,e) + aKDz] - \frac{1}{\varepsilon^{n-1}} B_0 v \ , \tag{3.33}$$

where $\gamma(t,e) = b(t,e) - aKe$.

To examine the dynamics of the "slow" system (3.32) separately from the dynamics of the "fast" system (3.33), a decoupled form is necessary. Given the noise enters $\dot{z}$ in (3.33), and z enters $\dot{e}$ in (3.32), z should be removed from (3.32). If z is not removed from $\dot{e}$, the noise entering $\dot{z}$ will lead to a more conservative bound on e. The next two sections address this concern as the bound on the tracking error is constructed.

3.5 Linear Systems Exploration

To adapt the general form given in (3.32)-(3.33) to a class of linear systems, we require that $b(t,e) = b_1(t) + \theta^T e$ be linear in e and $\gamma(t,e) = b_1(t) + \bar{\gamma}e$, where $\bar{\gamma} = \theta^T - aK$. In preparation for the change of variables that will decouple the differential equations, write (3.32)-(3.33) as

$$\dot{e} = A_{11}e + A_{12}z + Bb_1(t) \tag{3.34}$$

$$\varepsilon\dot{z} = A_{21}e + A_{22}z + \varepsilon Bb_1(t) - \frac{1}{\varepsilon^{n-1}} B_0 v \ , \tag{3.35}$$

where

$$A_{11} = A + B\bar{\gamma} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \\ \theta_1 - ak_1 & \theta_2 - ak_2 & \cdots & \cdots & \theta_n - ak_n \end{bmatrix},$$

$$A_{12} = BaKD = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 \\ \varepsilon^{n-1}ak_1 & \varepsilon^{n-2}ak_2 & \cdots & \cdots & ak_n \end{bmatrix},$$

$$A_{21} = \varepsilon B\bar{\gamma} = \varepsilon \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 \\ \theta_1 - ak_1 & \theta_2 - ak_2 & \cdots & \cdots & \theta_n - ak_n \end{bmatrix} \text{ and}$$

$$A_{22} = A_0 + \varepsilon BaKD = \begin{bmatrix} -\alpha_1 & 1 & 0 & \cdots & 0 \\ -\alpha_2 & 0 & 1 & \cdots & 0 \\ \vdots & & & \ddots & \vdots \\ -\alpha_{n-1} & \cdots & \cdots & 0 & 1 \\ -\alpha_n + \varepsilon^n ak_1 & \varepsilon^{n-1}ak_2 & \cdots & \cdots & -\varepsilon ak_n \end{bmatrix}.$$

In order to remove the "fast" dynamics from the "slow" system, and vis versa, the singularly perturbed system (3.34)-(3.35) is transformed into a block-diagonal form provided in [35]; see Appendix B for further details. The change of variables, that

will provide the desired setup is

$$\psi = e - \varepsilon M\phi \tag{3.36}$$

$$\varepsilon\phi = \varepsilon z + \varepsilon L(\varepsilon)e \tag{3.37}$$

and exist for matrices M and L that satisfy the linear algebraic equations

$$0 = \varepsilon(A_{11} - A_{12}L)M - M(A_{22} + \varepsilon LA_{12}) + A_{12} \tag{3.38}$$

$$0 = A_{21} - A_{22}L + \varepsilon L(A_{11} - A_{12}L), \tag{3.39}$$

respectively. Alternatively, (3.38)-(3.39) can be written as

$$\begin{aligned} 0 = \varepsilon(A + B\theta^T - BaK - BaKDL)M \\ - M(A_0 + \varepsilon BaKD + \varepsilon LBaKD) \\ + BaKD \end{aligned} \tag{3.40}$$

$$0 = -\varepsilon BaK - (A_0 + \varepsilon BaKD)L + \varepsilon L(A + B\theta^T - BaK - BaKDL), \tag{3.41}$$

and is the form that will be used for all subsequent derivations. Rewrite (3.40) as

$$\begin{aligned} MA_0 = BaKD + \varepsilon(A + B\theta^T - BaK - BaKDL)M \\ - \varepsilon M(I + L)BaKD \end{aligned}$$

and bring A_0 to the right-hand side

$$\begin{aligned} M = BaKDA_0^{-1} + \varepsilon(A + B\theta^T - BaK - BaKDL)MA_0^{-1} \\ - \varepsilon M(I + L)BaKDA_0^{-1}\ . \end{aligned} \tag{3.42}$$

Hence, for sufficiently small ε, (3.42) is a contraction mapping. Therefore, we can solve for M using successive approximations [30]. Define

$$\begin{aligned} M_{k+1} = BaKDA_0^{-1} + \varepsilon(A + B\theta^T - BaK - BaKDL)M_kA_0^{-1} \\ - \varepsilon M_k(I + L)BaKDA_0^{-1} \ . \end{aligned} \tag{3.43}$$

According to [35], after k iterations the exact solution M is approximated to within $\mathcal{O}(\varepsilon^k)$ error. Differentiating (3.36) yields

$$\begin{aligned} \dot{\psi} &= \dot{e} - M(\varepsilon\dot{\phi}) \\ &= (\cdots) - \frac{1}{\varepsilon^{n-1}} MB_0 v \ . \end{aligned}$$

Therefore, we need $MB_0 = \mathcal{O}(\varepsilon^{n-1})$ to eliminate the negative powers of ε from $\dot{\psi}$. Note the following properties

$$A_0^{-1}B_0 = \begin{bmatrix} 0 & 0 & \cdots & \cdots & -\frac{1}{\alpha_n} \\ 1 & 0 & \cdots & \cdots & -\frac{\alpha_1}{\alpha_n} \\ 0 & 1 & 0 & \cdots & -\frac{\alpha_2}{\alpha_n} \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & 1 & -\frac{\alpha_{n-1}}{\alpha_n} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ \alpha_n \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix},$$

$$A_0^{-2}B_0 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \text{ and}$$

$$DA_0^{-1}B_0 = \begin{bmatrix} \varepsilon^{n-1} & 0 & \cdots & \cdots & 0 \\ 0 & \varepsilon^{n-2} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & \cdots & \varepsilon & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} .$$

As the matrix A_0 is raised to progressively higher powers and multiplied by B_0, the only nonzero element in the product moves down the vector.

Now, show that $MB_0 = \mathcal{O}(\varepsilon^{n-1})$ with the following argument:

$$\begin{aligned} M_0 &= (\cdots)DA_0^{-1} \\ M_1 &= (\cdots)DA_0^{-1} + \varepsilon(\cdots)DA_0^{-2} \\ M_2 &= (\cdots)DA_0^{-1} + \varepsilon(\cdots)DA_0^{-2} + \varepsilon^2(\cdots)DA_0^{-3} \\ M_3 &= (\cdots)DA_0^{-1} + \varepsilon(\cdots)DA_0^{-2} + \varepsilon^2(\cdots)DA_0^{-3} + \varepsilon^3(\cdots)DA_0^{-4} \\ M_{n-2} &= (\cdots)DA_0^{-1} + \varepsilon(\cdots)DA_0^{-2} + \cdots + \varepsilon^{n-2}(\cdots)DA_0^{-(n-1)} \\ &+ \varepsilon^{n-1}(\cdots) . \end{aligned}$$

Then,

$$\begin{aligned} MB_0 &= (\cdots)DA_0^{-1}B_0 + \varepsilon(\cdots)DA_0^{-2}B_0 + \cdots \\ &\quad + \varepsilon^{n-2}(\cdots)DA_0^{-(n-1)}B_0 + \varepsilon^{n-1}(\cdots)B_0 \\ &= M_{n-2}B_0 + \mathcal{O}(\varepsilon^{n-1}) \end{aligned}$$

and

$$DA_0^{-i}B_0 = \mathcal{O}(\varepsilon^{n-i})$$

implying that

$$MB_0 = \mathcal{O}(\varepsilon^{n-1}) . \tag{3.44}$$

This shows that ψ is $\mathcal{O}(\mu)$.

Alternatively, we have

$$\begin{aligned} M_0 &= B(\cdots) \\ M_1 &= B(\cdots) + \varepsilon AB(\cdots) \\ M_2 &= B(\cdots) + \varepsilon AB(\cdots) + \varepsilon^2 A^2 B(\cdots) \\ M_{n-2} &= B(\cdots) + \varepsilon AB(\cdots) + \cdots + \varepsilon^{n-2} A^{n-2} B(\cdots) \\ &+ \varepsilon^{n-1}(\cdots) \, , \end{aligned}$$

where

$$B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} , \, AB = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix} , \, A^2 B = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } A^{n-2} B = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} .$$

Then,

$$M = M_{n-2} + \mathcal{O}(\varepsilon^{n-1})$$

and

$$M = \begin{bmatrix} \mathcal{O}(\varepsilon^{n-1}) \\ \mathcal{O}(\varepsilon^{n-2}) \\ \vdots \\ \mathcal{O}(\varepsilon) \\ \mathcal{O}(1) \end{bmatrix} . \tag{3.45}$$

As a consequence of linearity,

$$\phi = \mathcal{O}\left(\frac{\mu}{\varepsilon^{n-1}}\right) . \tag{3.46}$$

Then, combining (3.45) and (3.46) results in

$$\varepsilon M\phi = \begin{bmatrix} \mathcal{O}(\varepsilon\mu) \\ \mathcal{O}(\mu) \\ \mathcal{O}\left(\frac{\mu}{\varepsilon}\right) \\ \vdots \\ \mathcal{O}\left(\frac{\mu}{\varepsilon^{n-2}}\right) \end{bmatrix} . \tag{3.47}$$

Given ψ is $\mathcal{O}(\mu)$, we can see that both e_1 and e_2 are $\mathcal{O}(\mu)$, while the other components of e are of the order $\mathcal{O}(\mu/\varepsilon^i)$ with increasing powers of i.

Theorem 3.1: Consider the closed-loop system in (3.32)-(3.33), where $b(t, e)$ is assumed to be linear in e. Then, for a linear system of dimension n, the tracking error and subsequent derivatives satisfy

$$|e_i(t)| = \mathcal{O}(\mu), \forall i = 1, 2 \tag{3.48}$$

$$|e_i(t)| = \mathcal{O}\left(\frac{\mu}{\varepsilon^{i-2}}\right), \forall 2 < i \leq n \tag{3.49}$$

for $v \leq \mu$.

3.6 Nonlinear Systems Extension

To isolate the dynamics of the tracking error, the decomposition method proposed in [49] is utilized to eliminate the fast states, z, from the slow equation (3.32). In general, the decomposition is valid for singularly perturbed differential systems. The goal of the change of variables is to transform the system (3.32)-(3.33) into the following

form

$$\dot{\psi} = F(t, \psi, \varepsilon) \tag{3.50}$$

$$\varepsilon\dot{\phi} = G(t, \psi, \phi, \varepsilon) \; . \tag{3.51}$$

The change of variables that achieves the transformation (3.50)-(3.51) is found by setting $v = 0$ in (3.32)-(3.33) and following the method of [49]; see Appendix C. Note that the change of variables is actually applied to (3.32)-(3.33) for $v \neq 0$. Let

$$f(t, e, z, \varepsilon) = Ae + B[\gamma(t, e) + aKDz] \tag{3.52}$$

$$g(t, e, z, \varepsilon) = A_0 z + \varepsilon B[\gamma(t, e) + aKDz] \; . \tag{3.53}$$

The integral manifold is defined as $z = h(t, e, \varepsilon)$; see Appendix C for a definition of an integral manifold. A valid expression for $h(t, e, \varepsilon)$ is found by satisfying the equation

$$\varepsilon\frac{\partial h}{\partial t} + \varepsilon\frac{\partial h}{\partial e} f(t, e, h, \varepsilon) = g(t, e, h, \varepsilon)$$

or, equivalently,

$$\varepsilon\frac{\partial h}{\partial t} + \varepsilon\frac{\partial h}{\partial e}[Ae + B\gamma(t, e) + BaKDh] = A_0 h + \varepsilon B[\gamma(t, e) + aKDh] \; . \tag{3.54}$$

Furthermore, the function h can be represented as a Taylor series in ε, i.e. $h = h_0(t, e) + \varepsilon h_1(t, e) + \cdots$. The expressions for the h_i's can be found by substituting the expansion for h into the partial differential equation (3.54). Matching the coefficients of like powers of ε, it can be shown that the first set of coefficients are

$$h_0 = 0$$

$$h_1 = -A_0^{-1} B\gamma(t, e) \; ,$$

where E_i is defined as an $n \times n$ diagonal matrix, whose ith diagonal element is 1. Let $\phi = z - h(t, e, \varepsilon)$ and $w = e - \psi$. Then,

$$\begin{aligned} \dot{\psi} &= A\psi + B[\gamma(t,\psi) + aKDh(t,\psi,\varepsilon)] \\ &\triangleq f(t,\psi,h(t,\psi,\varepsilon),\varepsilon) \end{aligned} \tag{3.55}$$

$$\begin{aligned} \dot{w} &= f(t,\psi+w,\phi+h(t,\psi+w,\varepsilon),\varepsilon) - f(t,\psi,h(t,\psi,\varepsilon),\varepsilon) \\ &= Aw + B[\gamma(t,\psi+w) - \gamma(t,\psi) \\ &+ aKD[\phi + h(t,\psi+w,\varepsilon) - h(t,\psi,\varepsilon)]] \\ &\triangleq f_1(t,\psi,w,\phi,\varepsilon) \end{aligned} \tag{3.56}$$

$$\begin{aligned} \varepsilon\dot{\phi} &= g(t,\psi+w,\phi+h(t,\psi+w,\varepsilon),\varepsilon) \\ &- g(t,\psi+w,h(t,\psi+w,\varepsilon),\varepsilon) \\ &- \varepsilon\frac{\partial h(t,\psi+w,\varepsilon)}{\partial e}[f(t,\psi+w,\phi+h(t,\psi+w,\varepsilon),\varepsilon) \\ &- f(t,\psi+w,h(t,\psi+w,\varepsilon),\varepsilon)] \\ &= \left[A_0 + \varepsilon BaKD - \varepsilon\frac{\partial h}{\partial e}BaKD\right]\phi \\ &\triangleq Z(t,\psi,w,\phi,\varepsilon)\ . \end{aligned} \tag{3.57}$$

Define

$$\begin{aligned} F(t,\psi,\varepsilon) &= f(t,\psi,h(t,\psi,\varepsilon),\varepsilon) \\ &= A\psi + B\gamma(t,\psi) + BaKDh(t,\psi,\varepsilon) \end{aligned}$$

$$\begin{aligned} G(t,\psi,\phi,\varepsilon) &= Z(t,\psi,\varepsilon H(t,\psi,\phi,\varepsilon),\phi,\varepsilon) \\ &= \left[A_0 + \varepsilon BaKD + \varepsilon\frac{\partial h(t,\psi+\varepsilon H,\varepsilon)}{\partial e}BaKD\right]\phi \end{aligned}$$

introduced in (3.50)-(3.51). According to [49], $H(t,\psi,\phi,\varepsilon)$ satisfies the equation

$$\begin{aligned} 0 = \varepsilon\frac{\partial H}{\partial t} + \varepsilon\frac{\partial H}{\partial \psi}F(t,\psi,\varepsilon) + \frac{\partial H}{\partial \phi}Z(t,\psi,\varepsilon H,\phi,\varepsilon) \\ - f_1(t,\psi,\varepsilon H,\phi,\varepsilon) \end{aligned} \tag{3.58}$$

and has the asymptotic expansion $H = H_1 + \varepsilon H_2 + \varepsilon^2 H_3 + \cdots$, where $H_i = H_i(t,\psi,\phi)$. Expanding the function f_1 as a power series in ε results in

$$f_1(t,\psi,\varepsilon H,\phi,\varepsilon) = f_1^0(t,\psi,\phi) + \varepsilon f_1^1(t,\psi,\phi) + \cdots \ .$$

Solving for the coefficient f_1^0 yields the expression

$$f_1^0(t,\psi,\phi) = f_1(t,\psi,0,\phi,0) = aBKE_n\phi \ . \tag{3.59}$$

Solving for the coefficient f_1^1

$$\begin{aligned} f_1^1(t,\psi,\phi) &= \frac{\partial f_1(t,\psi,0,\phi,0)}{\partial w}H_1 + \frac{\partial f_1(t,\psi,0,\phi,0)}{\partial \varepsilon} \\ &= \left[A + B\frac{\partial \gamma(t,\psi)}{\partial e} + BaKE_n\frac{\partial h(t,\psi,0)}{\partial e}\right]H_1 \\ &+ BaKE_{n-1}\phi \ . \end{aligned}$$

However, consider that

$$\begin{aligned} h(t,e,\varepsilon) &= \varepsilon h_1(t,e) + \varepsilon^2 h_2(t,e) + \cdots \\ \frac{\partial h(t,e,\varepsilon)}{\partial e} &= \varepsilon\frac{\partial h_1}{\partial e} + \varepsilon^2\frac{\partial h_2}{\partial e} + \cdots \\ \frac{\partial h(t,e,0)}{\partial e} &= 0 \ . \end{aligned}$$

Further simplifying the right-hand side of f_1^1

$$f_1^1(t,\psi,\phi) = \left[A + B\frac{\partial\gamma(t,\psi)}{\partial e}\right] H_1 + BaKE_{n-1}\phi\,. \tag{3.60}$$

Next, the coefficient f_1^2 requires the following calculations

$$\begin{aligned} f_1^2(t,\psi,\phi) &= \frac{1}{2}\left[\frac{d^2 f_1(t,\psi,\varepsilon H(t,\psi,\phi,\varepsilon),\phi,\varepsilon)}{d\varepsilon^2}\right]\Bigg|_{\varepsilon=0} \\ &= AH_2 + BaKE_{n-2}\phi \\ &+ \frac{1}{2}B\frac{d^2}{d\varepsilon^2}\left[\gamma(t,\psi+\varepsilon H) + aKDh(t,\psi+\varepsilon H,\varepsilon)\right. \\ &\left.-aKDh(t,\psi,\varepsilon)\right] \end{aligned}$$

Realize that

$$\begin{aligned} h(t,\psi+\varepsilon H,\varepsilon) - h(t,\psi,\varepsilon) &= \varepsilon[h_1(t,\psi+\varepsilon H) - h_1(t,\psi)] \\ &+ \varepsilon^2[h_2(t,\psi+\varepsilon H) - h_2(t,\psi)] + \cdots \end{aligned}$$

is a valid expansion, where $h_0 = 0$. Then,

$$\begin{aligned} h_1(t,\psi+\varepsilon H) &= h_1^0(t,\psi+\varepsilon H) + \varepsilon h_1^1(t,\psi+\varepsilon H) + \cdots \\ &= h_1(t,\psi) + \frac{\partial h_1(t,\psi)}{\partial e}H_1 + \cdots \end{aligned}$$

leading to the conclusion that

$$h(t,\psi+\varepsilon H,\varepsilon) - h(t,\psi,\varepsilon) = \varepsilon^2\frac{\partial h_1(t,\psi)}{\partial e}H_1 + \cdots\,.$$

With a few more calculations

$$\begin{aligned}\frac{d\gamma(t,\psi+\varepsilon H)}{d\varepsilon} &= \frac{d\gamma(t,\psi+\varepsilon H_1+\varepsilon^2 H_2+\cdots)}{d\varepsilon}\\ &= \frac{\partial\gamma(t,\psi+\varepsilon H_1+\varepsilon^2 H_2+\cdots)}{\partial e}(H_1+2\varepsilon H_2+\cdots)\end{aligned}$$

and

$$\left.\frac{d^2\gamma}{d\varepsilon^2}\right|_{\varepsilon=0} = H_1^T\Gamma_h(t,\psi)H_1 + 2\frac{\partial\gamma(t,\psi)}{\partial e}H_2\ ,$$

where Γ_h is the Hessian matrix of γ with respect to e. Ultimately,

$$\begin{aligned}f_1^2(t,\psi,\phi) = AH_2 &+ BaKE_{n-2}\phi\\ &+ B\left[\frac{\partial\gamma(t,e)}{\partial e}H_2 + aKE_n\frac{\partial h_1(t,\psi)}{\partial e}H_1\right]\ .\end{aligned} \tag{3.61}$$

The remaining terms in the asymptotic expansion of f_1 can be found in a similar manner.

Return to (3.58), substitute in the expansion for $H(t,\psi,\phi,\varepsilon)$ and solve for the H_i's by matching the coefficients of like powers of ε. To obtain the expression for H_1, gather the coefficients of (3.58) that do not contain ε and set that equation to zero. For the remaining H_i's, gather the coefficients that contain ε^{i-1} and set the resulting equation to zero to solve for H_i. For H_1

$$0 = \frac{\partial H_1}{\partial\phi}A_0\phi - BaKE_n\phi\ ,$$

where the coefficients without ε are included, and the equation set to zero. Solving for H_1

$$H_1 = BaKE_nA_0^{-1}\phi\ . \tag{3.62}$$

For H_2

$$\begin{aligned} 0 = BaKE_nA_0^{-1}BaKE_n\phi + \frac{\partial H_2}{\partial \phi}A_0\phi - ABaKE_nA_0^{-1}\phi \\ - B\frac{\partial \gamma(t,\psi)}{\partial e}BaKE_nA_0^{-1}\phi \\ - BaKE_{n-1}\phi \ , \end{aligned}$$

where the coefficients of ε are included and the equation set to zero. Solving for H_2

$$\begin{aligned} H_2 = [ABaKE_nA_0^{-1} + B\frac{\partial \gamma(t,\psi)}{\partial e}BaKE_nA_0^{-1} \\ + BaKE_{n-1} \\ - BaKE_nA_0^{-1}BaKE_n]A_0^{-1}\phi \ . \end{aligned} \tag{3.63}$$

The coefficients of ε^2 set to zero are

$$\begin{aligned} 0 = \frac{\partial H_2}{\partial t} + \frac{\partial H_2}{\partial \psi}[A\psi + B\gamma(t,\psi)] \\ + \frac{\partial H_3}{\partial \phi}A_0\phi + \frac{\partial H_2}{\partial \phi}BaKE_n\phi \\ + \frac{\partial H_1}{\partial \phi}\left[BaKE_{n-1} + \frac{\partial h_1(t,\psi)}{\partial e}BaKE_n\right]\phi \\ - AH_2 - BAKE_{n-2}\phi \\ - B\left[\frac{1}{2}H_1^T\Gamma_h(t,\psi)H1 + \frac{\partial \gamma(t,\psi)}{\partial e}H_2 \right. \\ \left. + aKE_n\frac{\partial h_1(t,\psi)}{\partial e}H_1\right] \ , \end{aligned} \tag{3.64}$$

where

$$\frac{\partial H_2}{\partial \psi} = \frac{\partial}{\partial \psi}\left[B\frac{\partial \gamma(t,\psi)}{\partial e}BaKE_nA_0^{-2}\phi\right] ,$$

$$\frac{\partial \gamma}{\partial e} B = \frac{\partial \gamma}{\partial e_n},$$

$$\frac{\partial H_2}{\partial t} = B \frac{\partial}{\partial t} \left[\frac{\partial \gamma(t, \psi)}{\partial e_n} \right] a K E_n A_0^{-2} \phi$$

and $aKE_nA_0^{-2}\phi$ is a scalar. Substituting in the values above into (3.64) reveals that all terms, except one, have the vector ϕ only on the very right most side of the expression. The term that is the exception is

$$-\frac{1}{2} B \phi^T A_0^{-T} E_n K^T a B^T \Gamma_h(t, \psi) B a K E_n A_0^{-1} \phi \, , \tag{3.65}$$

where the vector ϕ appears in two locations. Aside from (3.65), the vector ϕ always appears paired with $aKE_nA_0^{-2}$ and becomes a scalar quantity; dividing both sides of the equation by the scalar quantity eliminates ϕ from all terms except (3.65). To solve for H_3, the integral, with respect to ϕ, is taken of (3.64). However, the solution supplied by the term in (3.65) is unclear. To eliminate (3.65) from (3.64), notice that

$$B^T \Gamma_h(t, \psi) B = \frac{\partial^2 \gamma(t, \psi)}{\partial e_n^2}$$

and require that the function b be linear in e_n. This requirement implies that

$$\frac{\partial^2 b}{\partial e_n^2} = 0 \, .$$

Then, the term (3.65) is zero. For higher-order systems, more complicated terms like (3.65) appear. However, there is no obvious way in which to eliminate their presence, and solve for the remaining H_i's. Therefore, the remainder of this discussion will focus on the case for $n = 3$, where H_3 will not appear in the expansion of H and no additional restrictions are placed on the system structure.

According to [49],

$$e = \psi + \varepsilon H(t, \psi, \phi, \varepsilon) \tag{3.66}$$

$$z = \phi + \varepsilon \bar{h}(t, \psi + \varepsilon H(t, \psi, \phi, \varepsilon), \varepsilon) \tag{3.67}$$

exists for choices of H and h that satisfy the partial differential equations (3.58) and (3.54), respectively. Moreover, for (3.32)-(3.33), the right-hand side of (3.67) can be written with $\varepsilon\bar{h}$ instead of h, because $h_0 = 0$. Hence, the inverse transformation must also take the form of an $\mathcal{O}(\varepsilon)$ perturbation, namely

$$\psi = e + \varepsilon Q(t, e, z, \varepsilon) \tag{3.68}$$

$$\phi = z + \varepsilon q(t, e, z, \varepsilon) \ . \tag{3.69}$$

Taking the derivative of (3.68) and (3.69) yields

$$\begin{aligned} \dot{\psi} &= \dot{e} + \varepsilon \frac{\partial Q}{\partial t} + \varepsilon \frac{\partial Q}{\partial e}\dot{e} + \frac{\partial Q}{\partial z}\dot{z} \\ &= F(t, \psi, \varepsilon) \end{aligned} \tag{3.70}$$

$$\begin{aligned} \varepsilon\dot{\phi} &= \varepsilon\dot{z} + \varepsilon^2 \frac{\partial q}{\partial t} + \varepsilon^2 \frac{\partial q}{\partial e}\dot{e} + \varepsilon \frac{\partial q}{\partial z}\dot{z} \\ &= G(t, \psi, \phi, \varepsilon) \ , \end{aligned} \tag{3.71}$$

where $v = 0$ in the expressions for $\dot{e}$ and $\dot{z}$ in (3.32) and (3.33), respectively. For $v \neq 0$

$$\dot{\psi} = F(t, \psi, \varepsilon) - \frac{1}{\varepsilon^{n-1}} \frac{\partial Q}{\partial z} B_0 v \tag{3.72}$$

$$\varepsilon\dot{\phi} = G(t, \psi, \phi, \varepsilon) - \frac{1}{\varepsilon^{n-1}} \left[I + \frac{\partial q}{\partial z} \right] B_0 v \tag{3.73}$$

which reveals that the fast states are eliminated from the slow equation, and the

presence of measurement noise adds an additional term multiplied by the noise.

For a system with dimension $n = 3$,

$$H = H_1 + \varepsilon H_2 + \varepsilon^2 H_R \,, \tag{3.74}$$

where $H_R = H_R(t, \psi, \phi, \varepsilon)$ and is $\mathcal{O}(1)$. The change of variables in (3.66)-(3.67) can be written as

$$\begin{aligned} e &= \psi + \varepsilon H_1 + \varepsilon^2 H_2 + \varepsilon^3 H_R \\ &= \psi + \varepsilon BaKE_3 A_0^{-1}\phi \\ &+ \varepsilon^2 \left[ABaKE_3 A_0^{-1} + B\frac{\partial \gamma(t,\psi)}{\partial e} BaKE_3 A_0^{-1} \right. \\ &\left. + BaKE_2 - BaKE_3 A_0^{-1} BaKE_3 \right] A_0^{-1}\phi \\ &+ \varepsilon^3(\cdots) \end{aligned} \tag{3.75}$$

$$\begin{aligned} z &= \phi + \varepsilon \bar{h}(t, \psi + \varepsilon H(t,\psi,\phi,\varepsilon), \varepsilon) \\ &= \phi + \varepsilon h_1(t, \psi + \varepsilon H) + \varepsilon^2 h_2(t, \psi + \varepsilon H) + \varepsilon^3(\cdots) \\ &= \phi + \varepsilon h_1(t,\psi) + \varepsilon^2 \frac{\partial h_1(t,\psi)}{\partial e} H_1 + \varepsilon^2 h_2(t,\psi) + \varepsilon^3(\cdots) \,. \end{aligned} \tag{3.76}$$

Applying (3.76) to (3.33), for $v \neq 0$, results in

$$\begin{aligned} \varepsilon \dot{z} &= A_0 z + \varepsilon B[\gamma(t,e) + aKDz] - \frac{1}{\varepsilon^2} B_0 v \\ &= \varepsilon \dot{\phi} + \varepsilon^2 \left[\frac{\partial h_1}{\partial t} + \frac{\partial h_1}{\partial \psi} \dot{\psi} \right] \\ &+ \varepsilon^2 \frac{\partial h_1}{\partial e} BaKE_3 A_0^{-1}(\varepsilon \dot{\phi}) + \varepsilon^3(\cdots) \end{aligned} \tag{3.77}$$

and solving for $\varepsilon\phi$ yields

$$\varepsilon\dot{\phi} = -\left[I + \varepsilon^2 \frac{\partial h_1(t,\psi)}{\partial e} BaKE_3A_0^{-1}\right]^{-1} P \ , \tag{3.78}$$

where

$$P = \varepsilon^2\left[\frac{\partial h_1}{\partial t} + \frac{\partial h_1}{\partial \psi}\dot{\psi}\right] - A_0 z - \varepsilon B\left[\gamma(t,e) + aKDz\right] + \frac{1}{\varepsilon^2}B_0 v + \varepsilon^3(\cdots) \ .$$

The term of interest contains the measurement noise, and can be represented as

$$\left[-I + \varepsilon^2 \frac{\partial h_1(t,\psi)}{\partial e} BaKE_3A_0^{-1} + \varepsilon^3(\cdots)\right]\frac{1}{\varepsilon^2}B_0 v \ , \tag{3.79}$$

where the matrix identity $(I+L)^{-1} = I - L + L^2 - L^3 + \cdots$ is used. Notice that $E_3A_0^{-1}B_0 = [0\ 0\ \cdots\ 0]^T$. Hence, (3.79) reduces to

$$-\frac{1}{\varepsilon^2}B_0 v + \varepsilon(\cdots)v$$

showing that

$$\phi = \mathcal{O}\left(\frac{\mu}{\varepsilon^{n-1}}\right) = \mathcal{O}\left(\frac{\mu}{\varepsilon^2}\right) \ . \tag{3.80}$$

Move $\dot{\psi}$ to the left-hand side and take the derivative of (3.75). Then, according to (3.72),

$$\begin{aligned}\dot{\psi} = F(t,\psi,\varepsilon) &+ \frac{1}{\varepsilon^2}BaKE_3A_0^{-1}B_0 v \\ &+ \frac{1}{\varepsilon}\left[ABaKE_3A_0^{-1} + B\frac{\partial\gamma(t,\psi)}{\partial e}BaKE_3A_0^{-1}\right. \\ &\left.+BaKE_2 - BaKE_3A_0^{-1}BaKE_3\right]A_0^{-1}B_0 v \\ &+ \varepsilon^2(\cdots + \frac{1}{\varepsilon^2}B_0 v + \cdots) \ .\end{aligned} \tag{3.81}$$

The first two terms in (3.81) with $\varepsilon\dot{\phi}$ are eliminated, due to the property $E_3 A_0^{-1} B_0 = E_3 A_0^{-2} B_0 = E_2 A_0^{-1} B_0 = 0$. However, the ε^2 term does not possess this property and remains present in (3.81). Given the origin of the closed-loop system in (3.32)-(3.33) is designed to be exponentially stable and the last term in (3.81) is $\mathcal{O}(\mu)$, $\psi = \mathcal{O}(\mu)$. Repeating the argument from Section 3.5,

$$B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, AB = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ and } A^2 B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

showing that raising the A matrix to a power determines which terms of (3.75) enter each component of the e vector.

Proposition 3.1: For a closed-loop nonlinear system of the form (3.32)-(3.33), with bounded measurement noise $v \leq \mu$, the tracking error, first derivative of the tracking error and second derivative of the tracking error are of the following orders of magnitude

$$e_1 = \psi_1 + \varepsilon^3(\cdots) = \mathcal{O}(\mu) \tag{3.82}$$

$$e_2 = \psi_2 + \varepsilon^2(\cdots)\phi + \varepsilon^3(\cdots) = \mathcal{O}(\mu) \tag{3.83}$$

$$e_3 = \psi_3 + \varepsilon(\cdots)\phi + \varepsilon^3(\cdots) = \mathcal{O}\left(\frac{\mu}{\varepsilon}\right), \tag{3.84}$$

where the expression in (3.75) is used.

Moreover, the relationship shown in (3.48)-(3.49) for the linear system explored in Section 3.5 holds for a nonlinear system of dimension $n = 3$.

3.7 Conclusions

When high-gain observers are employed in the presence of measurement noise, there exists a tradeoff between fast state reconstruction and a reasonable state estimation error. However, this sort of compromise does not exist when the primary interest is in the system tracking error. It was argued by constructing the system transfer functions from the noise to the tracking error and its derivatives, that the error and its first derivative are bounded uniformly in ε. Using singular perturbation analysis, the results were extended to a class of linear systems of dimension n. After the tracking error and its first derivative, all remaining derivatives of the tracking error are inversely proportional to increasing powers of ε. Subsequently, a similar result was derived for a class of nonlinear systems using the special features of singularly perturbed systems, further generalizing the results reported for linear systems. Due to the form of the nonlinearity, the result for nonlinear systems is restricted to a third-order system.

Although it has been shown that the tracking error is more immune to the effects of measurement noise than the estimation error for both the linear and nonlinear forms considered, this does not mean ε can be made arbitrarily small. It is important to keep in mind that the estimates of the states will still be used in the controller. However, the control may be able to tolerate a larger amount of error than the state estimates, providing some additional flexibility in choosing the value of ε when tracking is the focus.

Chapter 4

Enhancing High-Gain Observer Performance with Wavelet Denoising

Using wavelets to extract various signal components for compression, feature detection and denoising is a common approach in offline signal processing; see [17, 18, 36, 42] and the references therein. Recently, the techniques for signal denoising offline have been extended to produce acceptable results for online systems, [10, 34].

The goal of this chapter is to explore the role of wavelet denoising in improving the performance of high-gain observers subjected to measurement noise. In order to provide a basis for understanding the complexity of wavelets, the chapter begins with a brief introduction in Section 4.1. Section 4.2 is intended to familiarize the reader with the issues commonly addressed in designing a denoising algorithm with wavelets. The discussion in Section 4.2 is extended to include the intricacies associated with a real-time wavelet filter in Section 4.3. A simulation investigating the design parameters, their varying success in denoising the measurement and a comparison with the traditional lowpass filter is provided in Section 4.4. The last section addresses

directions for future work.

4.1 A Wavelet Introduction

Often times, we ignore the existence of noise in our systems to live in an idealized world where control algorithms are simplified, proofs are elegant and assumptions are abundant. However, a system with no noise is hardly a realistic scenario. When noise is accounted for, it is generally assumed to have some special characteristics that allow us to apply a particular type of filter (i.e. Kalman, lowpass, bandpass, etc.). Yet, what if the noise is not so easily compartmentalized? The noise may not be localized within a particular bandwidth, or may occupy the same frequency space as the signal. When this occurs, traditional denoising techniques only utilizing the temporal or frequency data cannot successfully remove the presence of noise without causing attenuation of the desired signal. Moreover, if an algorithm could localize in time and frequency where the noise occurs, a more thorough removal of the undesirable signal components can be accomplished. Frequently, the noise is not obviously periodic, and a non-stationary approach is necessary. This leads to the introduction of wavelets.

Wavelets, like Fourier analysis, provide another domain in which to analyze signals. In Fourier analysis, a Fourier transform is used to decompose the signal into its frequency components. Similarly, a wavelet can also be used to transform a signal into the frequency domain. The primary difference is that a wavelet transform provides the concentration of frequencies at each time instant. To follow the development of the wavelet transform from the Fourier perspective, see [42]. Another mathematically formal history of wavelets can be found in [14]. Some concerns that should be addressed when approaching the task of denoising with wavelets include how the noise enters the system (additive, multiplicative), the noise profile (smooth, erratic, impul-

sive, etc.) and the signal to noise ratio (SNR), see [54] and [9]. Moreover, it is not necessarily trivial to find the optimal wavelet and denoising technique pair; however, it can be done for certain classes of problems, for instance [36].

4.1.1 The Anatomy of a Wavelet Transform: Continuous-Time

A wavelet system is composed of two components: the wavelet function, known as the mother wavelet and the scaling function. The scaling function is used to shift, compress and stretch the mother wavelet. In particular, the wavelet expansion is used to transform the signal into a domain where both time and frequency are localized. Moreover, the representation of the signal is compressed into a few wavelet (expansion) coefficients. Then, these altered versions of the mother wavelet paired with the transform coefficients can be used to reconstruct the original signal. This idea is analogous to the the Fourier series using sinusoids to build a desired signal. Unlike the expansion set (sinusoids) for the Fourier series, there is no definite form for the mother wavelet function. In fact, there is an infinite number of choices available.

This chapter will focus on the most commonly used wavelets to cleanly capture the signal multi-scale behavior; Haar, Daubechies, Coiflets and Symlets. The Haar wavelet is part of the well-known and commonly used Daubechies family and is equivalent to the Daubechies order 1 transform. Some properties beyond simplicity, in the case of the Haar wavelet, that make these wavelets prime candidates for signal denoising, is that all are compactly supported, orthogonal and have discrete implementations. Moreover, the transforms not only conserve the signal energy, but compress it into a small number of coefficients; this is an important characteristic in the application of denoising. Furthermore, the least amount of asymmetry appears in the Symlet wavelets. The discrete wavelet transform contains the mother wavelet

$$\Psi_{j,k}(t) = 2^{j/2}\Psi(2^j t - k) \tag{4.1}$$

and the scaling function

$$\Phi_{j,k}(t) = 2^{j/2}\Phi(2^j t - k)\ , \tag{4.2}$$

where $j \in \mathbb{Z}$ and $k \in \mathbb{Z}$. The variables j and k are used to manipulate the mother wavelet. From a practical standpoint, the value of j controls the resolution of the signal (as interpreted by the scaling function). For $j > 0$ the scaling function becomes narrower; this translates to capturing smaller details of the signal under the transform. Conversely, for $j < 0$ the information provided by the wavelet transform is coarser. The notion of multiple scales is at the heart of what makes wavelets so attractive for denoising and compression applications. The fact that the signal components can be separated, and represented on different time scales, allows us to single out the various frequencies occurring at different times. The variable k simply shifts the functions to cover the entire signal space. Specifically, the Haar mother wavelet and the scaling functions are, respectively,

$$\Psi(t) = \begin{cases} 1,\, 0 \leq t < \dfrac{1}{2} \\ -1,\, \dfrac{1}{2} \leq t < 1 \\ 0,\, \text{else} \end{cases} \tag{4.3}$$

and

$$\Phi(t) = \begin{cases} 1,\, 0 \leq t < 1 \\ 0,\, \text{else} \end{cases}\ . \tag{4.4}$$

A closed-form expression is not available for all wavelet families, or necessarily as

easy to express in a clear fashion. For illustration and ease of understanding, the expression for the Haar wavelet is provided.

Hence, the above functions can be used to form an orthonormal and compact support basis for the signal of interest. The original signal $f(t)$ can be represented as

$$f(t) = \sum_{k=-\infty}^{\infty} a_{j_0}(k)\Phi_{j_0,k}(t) + \sum_{k=-\infty}^{\infty}\sum_{j=j_0}^{\infty} d_j(k)\Psi_{j,k}(t) \ , \qquad (4.5)$$

where the integer j_0 dictates the coarsest scale whose space is spanned by the scaling function; in general, the choice of j_0 depends on the signal itself and the desired resolution. In the context of this work, $f(t)$ always refers to the output signal y corrupted by measurement noise. The approximation or average (low frequency) coefficients are

$$a_{j_0}(k) = \int_{-\infty}^{\infty} f(t)\Phi_{j_0,k}(t)dt \qquad (4.6)$$

and the fluctuation or detail (high frequency) coefficients are

$$d_j(k) = \int_{-\infty}^{\infty} f(t)\Psi_{j,k}(t)dt \ . \qquad (4.7)$$

It should be noted that the above notation is intended for a discrete wavelet, even though the function itself is piecewise continuous. In general, discrete wavelets are not actually discrete in terms of the time variable. Rather, the dilation and translation effects are discrete. The analysis and synthesis equations listed above assume that the signal is infinite, and that all calculations are performed offline. Naturally, a finite and truly discrete model is necessary to implement this in real-time and digitally.

4.1.2 The Anatomy of a Wavelet Transform: Discrete-Time

We will now proceed to discuss the discrete-time wavelet implementation, focusing on the Haar wavelet. However, the process for Daubechies order 4 wavelet is similar. As the above discussion implies, the information contained in the signal is encapsulated by the wavelet transform coefficients (4.6) and (4.7), and the basis provided by the functions shown in (4.1) and (4.2). The discrete-time representations for the Haar wavelet approximation and detail coefficients are, respectively,

$$a(n) = \frac{\hat{f}(2n-1) + \hat{f}(2n)}{\sqrt{2}} \tag{4.8}$$

and

$$d(n) = \frac{\hat{f}(2n-1) - \hat{f}(2n)}{\sqrt{2}} \ , \tag{4.9}$$

where $\hat{f}$ is the sampled signal with a sampling period of T, and $n \in [1, \frac{N}{2}]$ where N is the total number of samples. The inverse mapping is

$$\hat{f}(n) = \begin{cases} \dfrac{a(n) + d(n)}{\sqrt{2}}, \ \forall n \text{ odd} \\ \dfrac{a(n) - d(n)}{\sqrt{2}}, \ \forall n \text{ even} \end{cases} \ . \tag{4.10}$$

At this point, we could easily take (4.8) through (4.10) and construct the necessary matrices to transform the desired signal into the time-frequency ("wavelet") domain, and back into the time-domain. However, the expressions for the approximation and detail coefficients given in (4.8) and (4.9) are generally not that simplistic nor informative. We seek to divide the coefficients into separate frequency scales, where the signal is separated into its frequency subbands via a filter bank. The perspective gained from multiresolution analysis, via filter banks, will subsequently provide

invaluable insight when choosing a denoising scheme.

Typically, filter banks are chosen to implement discrete wavelet transforms, because they provide a relatively fast computation time ($\mathcal{O}(N)$ complexity), reduced data storage and convenient signal processing interpretation. With filter banks in mind, the average coefficients are now represented as a convolution

$$a(n) = \hat{f} * \bar{l}[2n] \tag{4.11}$$

between the signal of interest and the lowpass filter $\bar{l}$. The detail coefficients

$$d(n) = \hat{f} * \bar{h}[2n] \tag{4.12}$$

are found via a convolution between the signal and the highpass filter $\bar{h}$. Moreover, both sets of coefficients are down-sampled by 2, given the process of filtering with both a lowpass and highpass filter creates twice the information necessary to reconstruct the original signal. These analysis filters are wavelet transform dependent, and are constructed using the mother wavelet (4.1) and scaling (4.2) functions defined above. Specifically, the lowpass filter in (4.11) is calculated from

$$\Phi(t) = \sum_m l(m)\sqrt{2}\Phi(2t - m) \ , \tag{4.13}$$

where $\bar{l}[n] = l[-n]$, and the highpass filter in (4.12) is calculated from

$$\Psi(t) = \sum_m \bar{h}(m)\sqrt{2}\Phi(2t - m) \ , \tag{4.14}$$

where $\bar{h}[n] = h[n]$. For a Haar filter $l = [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$ and $h = [-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$. Thus, the implementation of the discrete wavelet transform can be viewed as nothing more than FIR filter design, generally referred to as conjugate mirror filters. The

reconstructed signal is given by

$$\hat{f}(n) = \check{a} * \tilde{l}[n] + \check{d} * \tilde{h}[n] \; . \tag{4.15}$$

The coefficients (4.11) and (4.12) are up-sampled by 2, resulting in $\check{a}$ and $\check{d}$, respectively. The details concerning the construction of lowpass and highpass synthesis filters $\tilde{l}$ and $\tilde{h}$, along with the necessary and sufficient conditions on the analysis and synthesis filter pairs, is available in [42]. For this discussion, the synthesis filters for a Haar transform are $\tilde{l} = [\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}]$ and $\tilde{h} = [\frac{\sqrt{2}}{2}, \; -\frac{\sqrt{2}}{2}]$. These filter pairs allow for a perfect reconstruction of the original signal, barring any intentional manipulation of the wavelet coefficients in between applying the analysis and synthesis filters.

Lifting is another widely used method. Lifting seeks to improve the wavelet properties, in the context of perfect reconstruction filters (the methodology used above). It can be used to both design the form of the wavelets, in addition to implementing the discrete wavelet transform. Lifting tends to generate filter implementations that result in a faster runtime, when compared to other approaches. In theory, every perfect reconstructible filter bank can be expressed in terms of lifting, [27, 42]. In the context of this work, we did not employ lifting techniques, given perfect reconstruction filters yielded satisfactory results.

4.2 Denoising: Offline

The material covered thus far is a small portion of the entire picture. Unfortunately, the above ideas in isolation do not guarantee a successful implementation of a real-time discrete-time wavelet transform. Thus, before tackling the construction of a real-time filter, we first address the components associated with wavelet denoising that are necessary for both offline and real-time denoising applications. In this section,

we will temporarily neglect the issues that arise when implementing the filter in real-time. Details not covered here on wavelet denoising, and wavelet packet denoising, are discussed in [15].

4.2.1 Wavelet Type

The first step in developing a wavelet denoising algorithm, is to settle on a wavelet type. There are a plethora of mother wavelets and scaling functions available; in some cases, you may even want to create your own. The more common wavelet families include Daubechies, Symlets, Coiflets, Meyer and other biorthogonal functions, to name a few. For a more lengthy discussion on the various wavelet families available and their numerous properties, see [9, 14, 42]. Although we have focused on the first three transform families in this work, that does not preclude the inclusion of another type of wavelet. In fact, a different choice may achieve a smaller bound on the error in steady-state than what can be achieved with the wavelets presented. However, there will most likely be a tradeoff in the complexity. The authors of [11, 41, 51] suggest a number of tools that can help determine the best wavelet for a given system. However, only [41] and [11] also consider whether or not the choice will result in effective denoising.

The type of wavelet function is heavily dependent on the type of noise in the system, and potentially the signal itself. If the goal is to remove short and sporadic noise, a wavelet function that captures this type of behavior is more appropriate. For noise that is relatively smooth and omnipresent, a continuous and sufficiently smooth wavelet is better suited, [9] and [54].

4.2.2 Wavelet Transform Levels

The number of levels is the number of times the signal is concurrently subjected to the wavelet transform. For instance, recall the filter bank implementation of the discrete wavelet transform. Initially, the output is simultaneously passed through a lowpass (L) and highpass (H) filter, then down-sampled. This generates the level-1 coefficients. If the approximation coefficients (from the lowpass filter) are subsequently passed through another set of low and highpass filters, followed by down-sampling, the resultant would be the level-2 coefficients. We could repeat these steps for a N-level wavelet transform, where N is the number of cascaded filter banks. This technique is displayed in Figure 4.1.

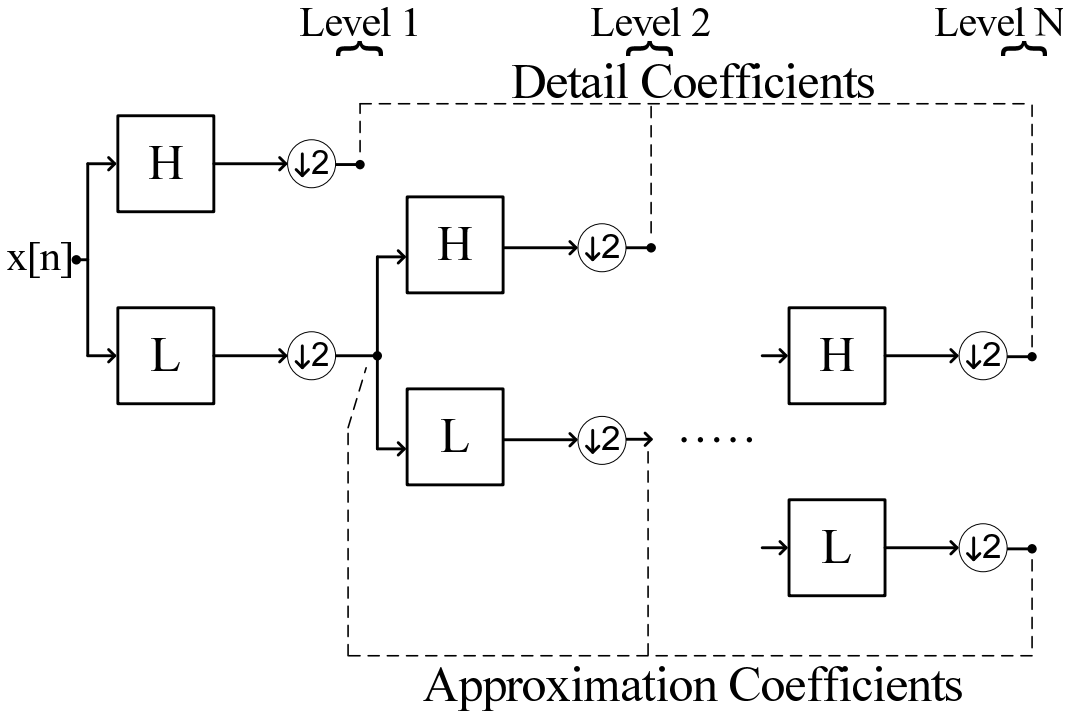

Figure 4.1: Diagram of a discrete wavelet transform implementation.

Each subsequent level results in a more detailed scale of the signal behavior. The first level is always the coarsest approximation of the signal. The coefficients from the remaining levels are intended to add finer detail to the signal reconstruction. It is worth noting, that different signal characteristics will appear under different scales

of the transform. Thus, the idea of multiresolution is to capture the majority of the noise within a finite number of scales. Hence, more scales does not always mean better results. In fact, the fewer number of coefficients necessary to capture the majority of the signal energy, the better the chance of removing the noise without eliminating significant features of the desired signal. The goal is to isolate the signal energy in a limited number of (relatively) large coefficients, while eliminating the remaining coefficients that are likely to contain the noise.

4.2.3 Thresholding Scheme

The basic premise behind thresholding, involves eliminating wavelet coefficients containing noise; analogously, keeping coefficients with a high percentage of the signal energy. There are numerous ways in which to accomplish this task; some complex, others remarkably simple. The first attempt at formalizing the wavelet coefficient thresholding for removal of additive noise from deterministic signals was recorded in [18].

Two of the most common (and simple) thresholding schemes are referred to as soft and hard thresholding; incidentally, terms that tend to have loose definitions. For a precise definition and interpretation of soft thresholding, see [16]. Both soft and hard thresholding eliminate coefficients that are below a chosen threshold. In hard thresholding, all values above the threshold are kept. Figure 4.2 is an example of a hard thresholding function, where x is the coefficient input and y the adjusted coefficient after thresholding is applied. In soft thresholding, the coefficients at or above the threshold value are altered (typically decreased) in some predefined fashion (i.e. dead-zone nonlinearity). Figure 4.3 is an example of a soft thresholding function, where x is the coefficient input and y the adjusted coefficient after thresholding is applied. Hard thresholding is a poor choice when the signal and noise coefficients are close in magnitude. Hard thresholding is not continuous, so it exaggerates even

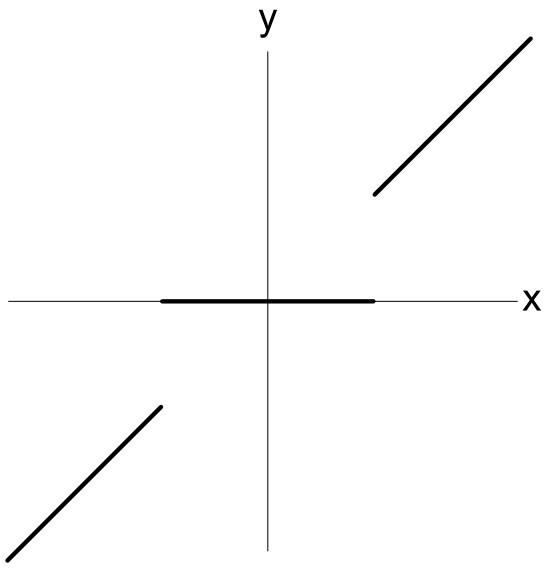

Figure 4.2: Potential hard thresholding function.

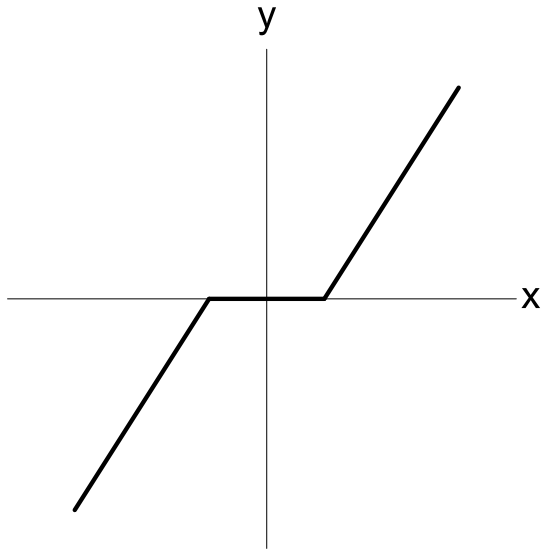

Figure 4.3: Potential soft thresholding function.

the slightest difference in coefficients near the threshold value. This can cause undesired discontinuities/artifacts in the reconstructed signal. For a smoother transition between the various signal transform components, soft thresholding is a valid choice. This form of thresholding is not sensitive to coefficient values near the thresholding value. However, soft thresholding decreases coefficients that contain a large percent of the signal energy, regardless the proximity to the threshold. Meaning that no coefficient remains unchanged after leaving the thresholding stage.

Another important detail of thresholding is the cutoff value. The cutoff value determines which coefficients are left untouched, and the others that will be diminished or removed. Common ways to determine an appropriate thresholding cutoff include Stein's Unbiased Risk, the universal threshold and minimax thresholding. In this work, comparable values were generated with all three methods. Hence, the remaining discussions (unless otherwise specified) will focus on the minimax principle, given this is an approach commonly used in statistics when designing optimal estimators, in the mean square error sense. Essentially, the minimax principle is used to predict the likelihood that a coefficient contains mostly noise or signal information; a minimum bound is found, and the threshold is chosen to respect that bound.

4.3 Denoising: Real-time

This section builds on the knowledge presented for offline denoising by adapting the algorithms for real-time implementation.

4.3.1 Delay

One of the key obstacles in real-time implementation is time delay. In [12], the authors investigate the use of wavelet denoising to attenuate the rotor vibration caused by step changes in sinusoidal forcing for a flexible rotor-magnetic bearing system. They

observed that the system states experience lag that originates between the signal measurement and the generation of the wavelet coefficients (prior to entering the controller). One suggestion to reduce the delay is to choose wavelets that are more asymmetric (relative to the Daubechies family). Given Daubechies wavelets tend to be localized around the center of their time duration, the denoising performance is reduced. A wavelet transform that is concentrated more towards the current point in time will produce a more immediate response in the coefficients, and hence allow a higher bandwidth of disturbance attenuation.

Another method to reduce delay is restricting the wavelet transform to the negative axis [10]. This allows the wavelet transform to only operate on past values of the signal. In utilizing an average-interpolation method, the authors discovered that taking estimates of the signal near the current time, and not precisely at the time of interest, resulted in a more fruitful noise removal. However, an additional amount of delay is introduced into the system that is proportional to the distance of the measurement taken from the current estimate. The authors were also able to increase the signal to noise ratio by using a redundant transform algorithm that rendered the wavelet transitionally invariant; the coefficients remain the same regardless of the point in time the calculation is made. This is accomplished though cycle spinning [17], which averages the coefficient results found at various time shifts of the signal.

In [1], the authors used wavelet filtering in an adaptive controller for a structural system, where the denoising does not take place in the feedback loop; thus, delay is not a primary concern. In [19], wavelets are used to remove noise from process data resulting from slowly-varying systems. In particular, they use wavelets to denoise signals from a pilot-scale distillation column, where the process signals are updated every 5 seconds. Although this algorithm is in real-time, the time scale is fairly slow.

Not surprisingly, increasing the wavelet levels can lead to an increase in the system

latency [34]. This is particularly the case when using filter banks that operate on the same time-scale as the system itself, and/or when the discrete wavelet transform is viewed as its own dynamic system [48].

4.3.2 Thresholding Scheme

If the bound on the measurement noise is unknown, the technique of coefficient thresholding, or wavelet shrinkage, can still be successful in eliminating noise. The central idea is to avoid setting a constant threshold value a priori, and allow the evolution of the system to define the threshold value. In this case, a larger window size (without introducing a detrimental amount of delay) would probably yield better results. We have not explored such methods in this work, given the level of complexity for those algorithms and the potential of the excessive time-delay destabilizing the closed-loop system.

4.3.3 Windowing

Given the denoising scheme is implemented online, we do not have access to the entire signal at once. Thus, we must choose to view and manipulate only a portion of that signal at anytime. This notion will be referred to as windowing. The size of the window is limited by the design choices made in the previous sections. The number of samples chosen dictates how high of a resolution the signal coefficients can be (level), and what kind of thresholding scheme is appropriate. Moreover, if data size is an issue (which it typically is in hardware implementation), the number of samples necessary to complete a satisfactory denoising should be minimized.

One windowing approach taken in [46], suggests that the window should increase as the number of data points increases; the interpretation is that by the time the last data point is reached, the entire data set should be encapsulated in the window.

4.4 Example

This section investigates the performance of high-gain observers utilizing wavelet denoising techniques.

4.4.1 Simulation

Consider the following simple pendulum stabilization problem

$$\dot{x}_1 = x_2 \tag{4.16}$$

$$\dot{x}_2 = -\sin(x_1 + \theta_r) + u \ , \tag{4.17}$$

where the pendulum arm is to be regulated at a constant reference signal of θ_r, by the control u. The state variable x_1 is defined as the difference between the actual angular position, θ, and the desired angular position, θ_r; x_2 is the velocity of the pendulum arm. The state feedback controller

$$u = \sin(x_1 + \theta_r) - 29x_1 - 10x_2 \tag{4.18}$$

is used to linearize the system and assign the closed-loop eigenvalues at $-5 \pm 2j$. To obtain a globally bounded control, we saturate x_1 at ± 1.2 and x_2 at ± 1.7. The bounds on the controller are chosen such that the saturation is never active when the system is under state feedback control. Assuming the only measurement is x_1, the linear high-gain observer will take the following form

$$\dot{x}_1 = \hat{x}_2 + \frac{2}{\varepsilon}(y - \hat{x}_1) \tag{4.19}$$

$$\dot{x}_2 = \frac{1}{\varepsilon^2}(y - \hat{x}_1) \, , \tag{4.20}$$

where the measurement y is corrupted by a bounded noise v, i.e. $y = x_1 + v$; the observer eigenvalues are assigned to $-\frac{1}{\varepsilon}$ and $-\frac{1}{\varepsilon}$. In the interest of exploring a more realistic approach, this setup will be studied in the context of a sampled-data output feedback controller. The actual implementation of the control algorithm is in discrete-time, modeled after the problem explored in [32]. The control is sampled using a zero-order-hold, where the control in (4.18) is held constant in-between the uniformly spaced sampling points. The continuous-time observer (4.19)-(4.20) is discretized using the Bilinear-Transformation method. The sampling period is chosen according to the guidelines presented in [13]. Namely, the sampling period T is designed as $T = \alpha\varepsilon$, where $0 < r_1 < \alpha < r_2 < \infty$ for some positive constants r_1 and r_2, independent of ε. For this example, $\alpha = 1$ and $T = \varepsilon$.

The discrete-time observer is taken as

$$q(k+1) = A_d q(k) + B_d y(k) \tag{4.21}$$

$$\hat{x}(k) = D^{-1}[C_d q(k) + D_d y(k)] \, , \tag{4.22}$$

where

$$A_d = \frac{1}{9}\begin{bmatrix} -1 & 4 \\ -4 & 7 \end{bmatrix}, B_d = \frac{2}{9}\begin{bmatrix} 5 \\ 2 \end{bmatrix}, C_d = \frac{2}{9}\begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix} \text{ and } D_d = \frac{1}{9}\begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

The sampled-data output feedback control is given by

$$u(k) = \sin\left(\hat{x}_1(k) + \theta_r\right) - 29\ \mathrm{sat}_{1.2}(\hat{x}_1(k)) - 10\ \mathrm{sat}_{1.7}(\hat{x}_2(k))\ ,$$

where $\mathrm{sat}_k(z) = \min\{|z|,k\}\mathrm{sign}(z)$. The initial states are taken as $x_1(0) = -1$, $x_2(0) = 0$, $q_1(0) = 0$ and $q_2(0) = 0$.

The noise is generated using the "Uniform Random Number" block in Simulink, where the bound is selected as ± 0.01 with a sample time of 0.001 seconds. The high-gain observer parameter, ε, is set to 0.02. Figure 4.4 shows the system states for both the continuous and sampled-data output feedback controllers.

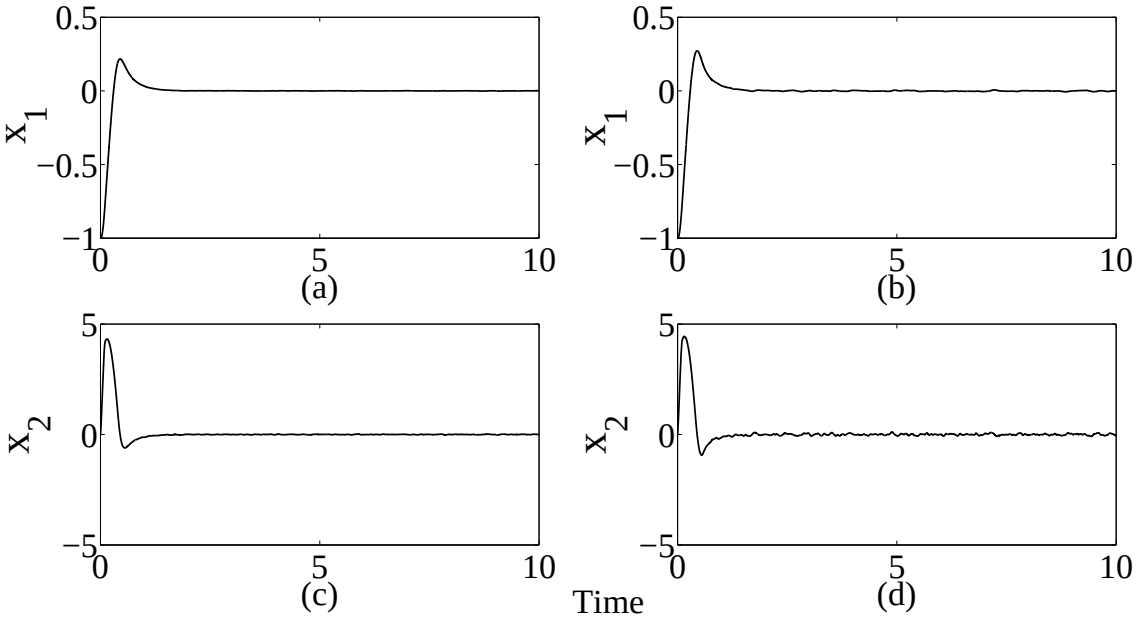

Figure 4.4: Comparison of the trajectories under (a) continuous-time, (b) sampled-data, (c) continuous-time and (d) sampled-data output feedback.

The difference in the signal magnitudes at steady-state is due to aliasing introduced via the discretization. Hence, an anti-aliasing filter (i.e. Butterworth order 8 and cutoff frequency of 25Hz) can be used to remove the aliasing introduced by sampling the control. As shown in [32], the pre-filter will reduced some of the noise.

The goal is to further reduce the noise level via wavelet denoising.

Consider the addition of a wavelet filter after the zero-order-hold sampling of y, but before y is injected into the high-gain observer. The filter is constructed from a Haar wavelet level 2 transform and performed on 4 samples (window length) at a time. A soft threshold with a dead-zone nonlinearity is used. The threshold value is chosen as the maximum value of the noise wavelet coefficients, and is the same for all levels of the transform. In Figure 4.5, it is evident that the filter is indirectly improving the performance of the state variables by eliminating a significant amount of the remaining noise from the signal injected into the high-gain observer.

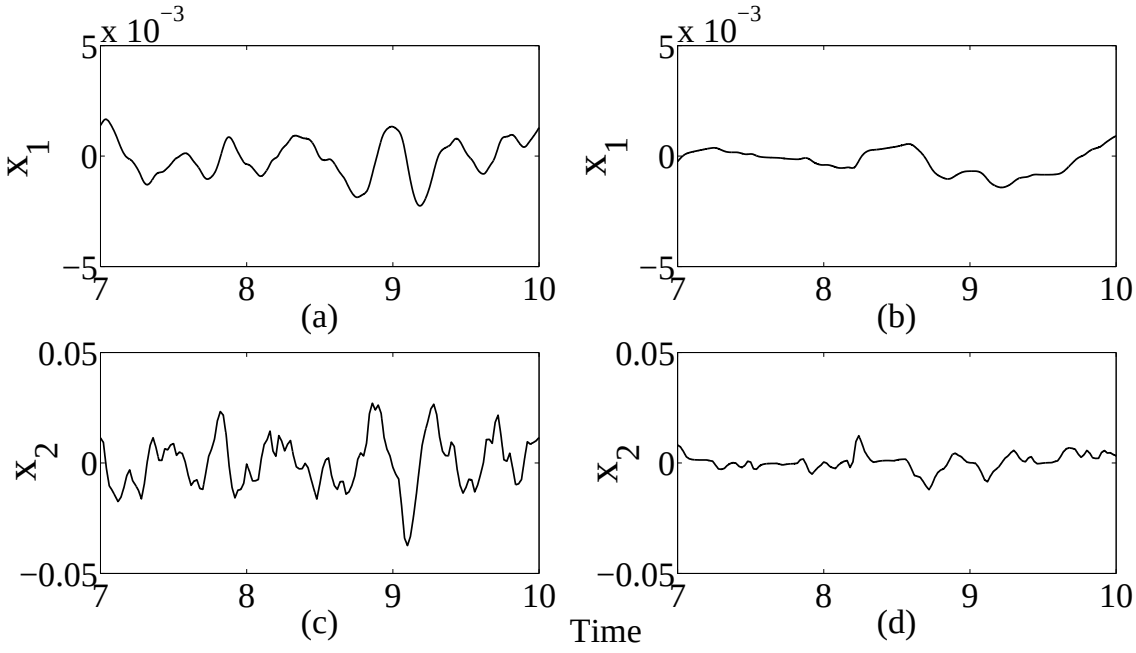

Figure 4.5: Comparison of the trajectories in steady-state (a) without wavelet denoising, (b) with wavelet denoising, (c) without wavelet denoising and (d) with wavelet denoising.

The nonlinear-gain observer developed in Chapter 2 was designed to attenuate the measurement noise in steady-state, while maintaining fast state reconstruction in the transient period. The notion of using a wavelet pre-filter is intended to remove some of the noise from the output before entering the high-gain observer. However, these

two techniques are not mutually exclusive. The wavelet scheme can be used to filter out additional noise from the measurement y before the signal reaches the high-gain observer with the nonlinear-gain. In fact, using these methods in a cascade fashion can result in the ability to choose smaller values of ε, consequently increasing the speed of the estimator. Consider the nonlinear-gain high-gain observer introduced in Chapter 2, where $\varepsilon_1 = 0.001$, $\varepsilon_2 = 0.02$ and $d = 0.2$. Figure 4.6 shows the transient performance of the state x_2, where the waveforms for the system with and without the wavelet pre-filter look identical. The steady-state value of x_2 for the system utilizing the nonlinear-gain observer with the wavelet pre-filter is significantly reduced when compared to the system without the benefit of the pre-filter; see Figure 4.7. Clearly, the denoising algorithm further reduces the noise in the steady-state trajectories.

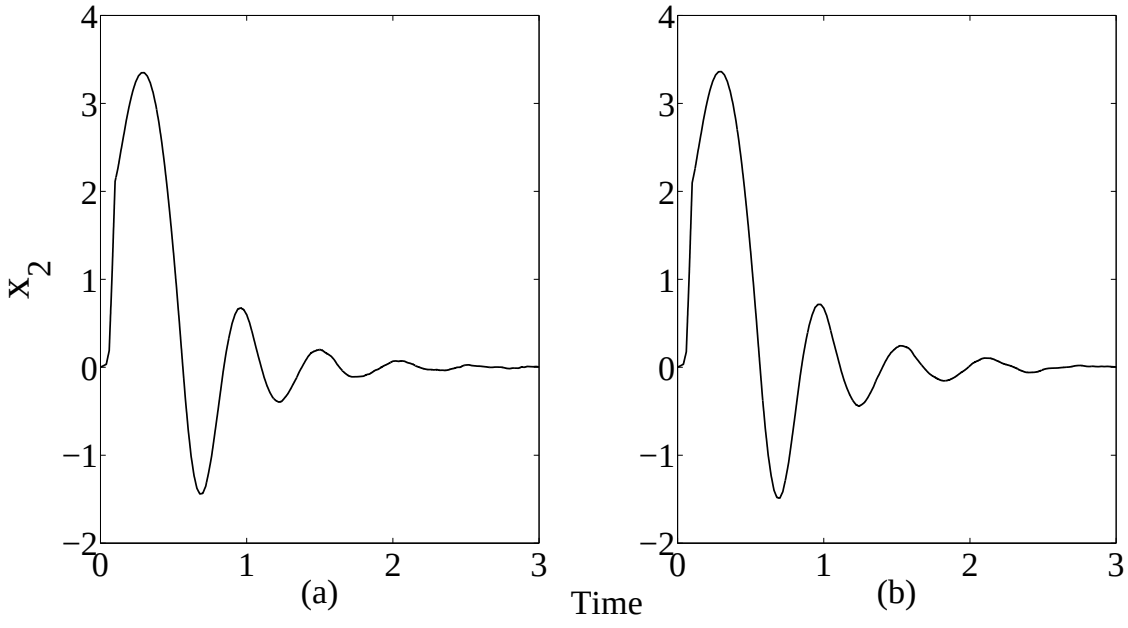

Figure 4.6: Transient performance comparison of the nonlinear-gain high-gain observer (a) without a wavelet pre-filter and (b) with a wavelet pre-filter.

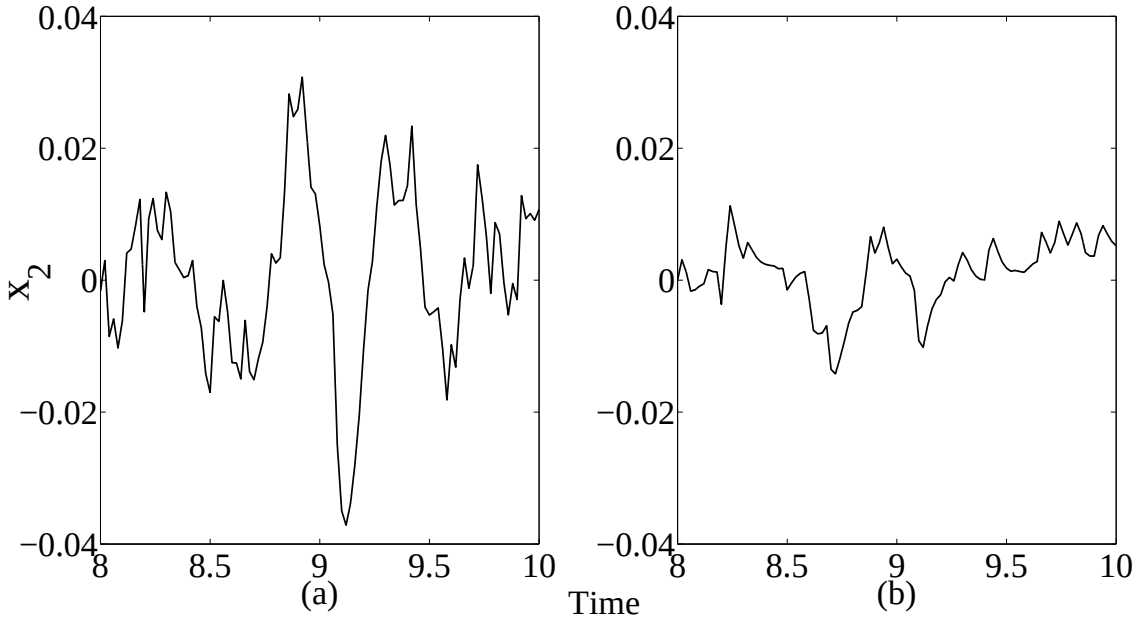

Figure 4.7: Steady-state performance comparison of the nonlinear-gain high-gain observer (a) without a wavelet pre-filter and (b) with a wavelet pre-filter.

4.4.2 Altering the Wavelet

Suppose we chose a different type of wavelet; perhaps a smoother set of functions, while leaving the level at 2, window size at 4 and thresholding soft. Figure 4.8, where the wavelet filter is active, shows that the steady-state values of the state x_2 contain more noise as the order of the Daubechies wavelet increases. For this example, the smoother the wavelet the poorer the noise approximation and subsequent removal. Most likely, the decreased performance is due to the increasing length of support as the order of the wavelet increases. In other words, more data (larger window) may be required to properly utilize the higher-order wavelets. We have limited this comparison to the Daubechies family, given the Haar transform is the lowest order transform available in the Daubechies subset of functions.

Consider that the signal at steady-state is presumably constant (or almost constant) for a stabilization problem. An interpretation for the successful denoising with this algorithm is that the shape of the Haar function can be used to approximate

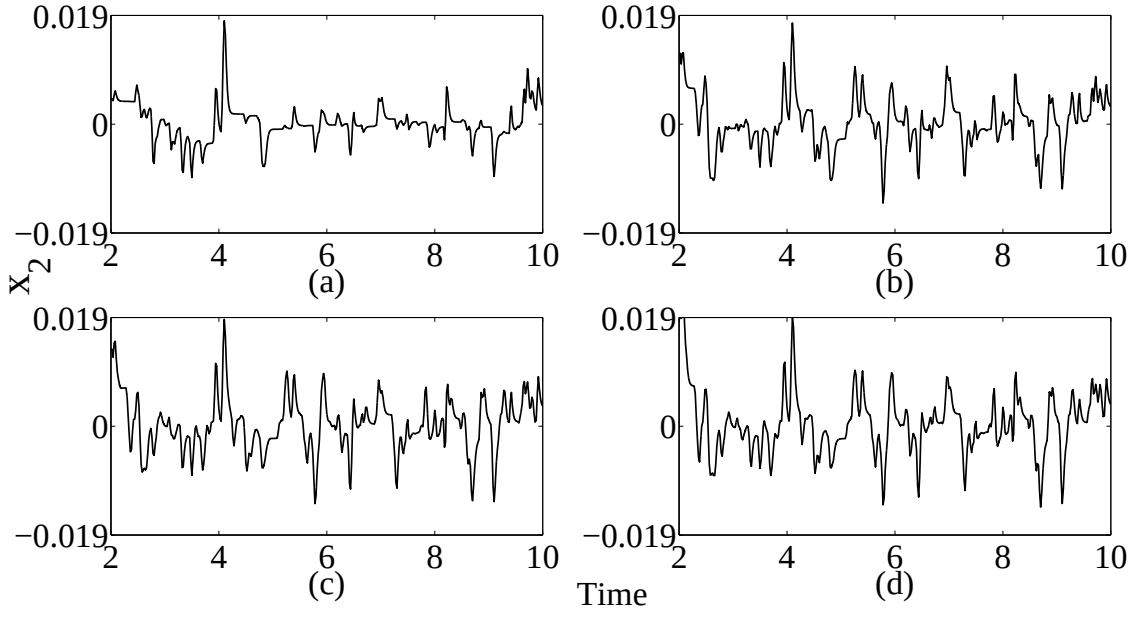

Figure 4.8: Denoising performance in steady-state with the (a) Daubechies 1, (b) Daubechies 4, (c) Daubechies 10 and (d) Daubechies 20 wavelets.

the signal on multiple scales, and separate it from the varying (non-constant) noise signal. In this case, the noise profile greatly differs from that of the signal, and the wavelet is able to distinguish that difference. Hence, the signal characteristics will still be clustered into a few large value wavelet coefficients, whereas the noise will occupy a larger number of small detail coefficients. This discovery is welcomed, given the Haar wavelet is one of the simplest to manipulate and implement, as previously noted.

However, we would be remiss to not compare the denoising performance of wavelets outside of the Daubechies family. Figure 4.9 provides a comparison of the steady-state denoising performance for order 2 Daubechies, Coiflets and Symlets wavelet. The level is set at 1, window size 10 and thresholding soft; the Matlab "thselect" function is used with the "minimax" option to determine the cutoff value. Clearly, all three families produce similar denoising results. Comparing Figure 4.8 with Figure 4.9, suggests that the ad-hoc method of determining the cutoff threshold may be more

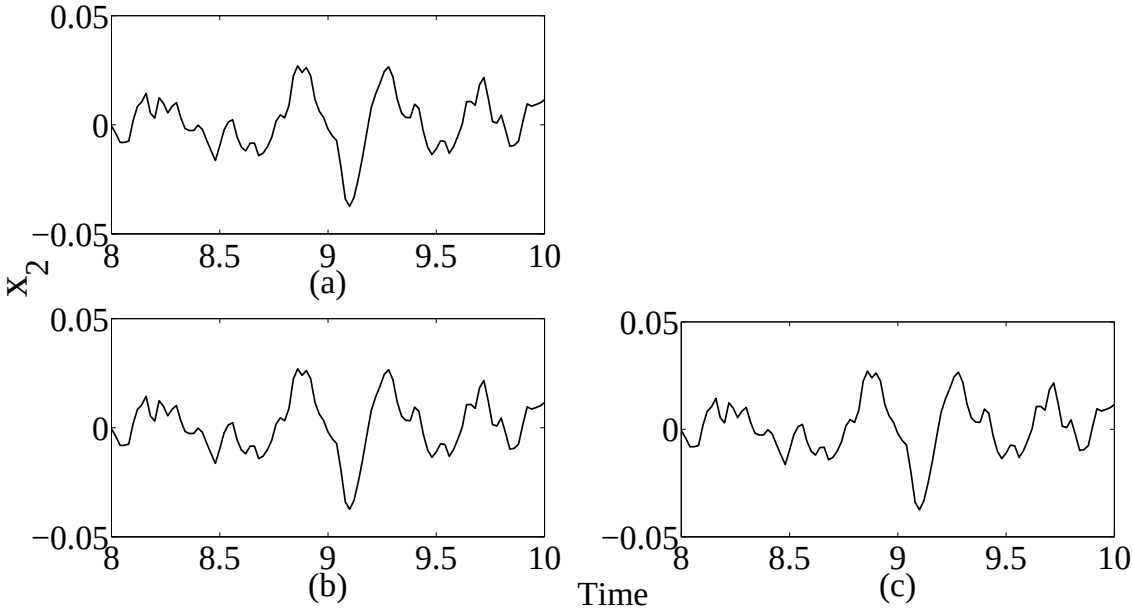

Figure 4.9: Denoising performance in steady-state with the (a) Daubechies 2, (b) Coiflets 2 and (c) Symlets 2 wavelets.

effective than the minimax method for smaller windows.

4.4.3 Levels

Although level 1 and level 2 transforms were chosen, we could have chosen a setup that would provide us with a larger number of detail coefficients. Yet, in the case of this feedback control stabilization problem, more detail is not more accuracy. In fact, Figure 4.10 shows that as additional levels are added, the level of noise in the state x_2 increases to more than twofold the amount seen with a level 1 Haar transform; the window size is 16 and the thresholding type is a dead-zone nonlinearity for the simulation considered. This result is partially a by-product from the way the thresholding value was chosen. The bound on the measurement noise was known a priori, implying that a bound on the wavelet coefficients is also known. The bound on the coefficients was used as the threshold in this example, meaning that all wavelet coefficients at or above that threshold will be eliminated before reconstructing the

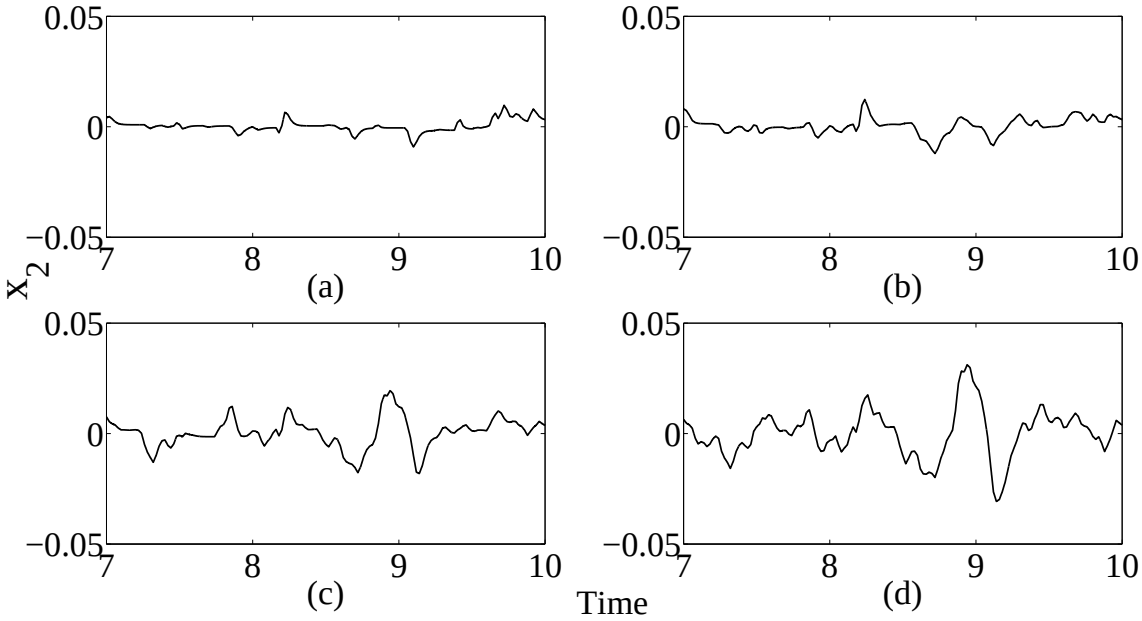

Figure 4.10: Denoising performance in steady-state with Haar (a) level 1, (b) level 2, (c) level 3 and (d) level 4 wavelet transforms.

signal. If this approach is taken with a low level wavelet transform (i.e. 2 or lower), the measurement signal can be recovered with minimal noise. However, even using an online algorithm to determine the cutoff value still results in a decreased performance as the level increases, although not as significantly as in the fixed threshold case; in this work the minimax method is used to determine a cutoff value online.

We did not observe any obvious delay effects irrespective of the level. Unlike the papers previously mentioned, the wavelet transforms in the denoising algorithms discussed in this work were implemented as algebraic calculations.

4.4.4 Thresholding Logic

For stabilization problems with bounded measurement noise, soft thresholding reduces the ultimate bound on the estimation error more so than hard thresholding; this is in the context of choosing the threshold value a priori from the bound on the measurement error. Figure 4.11 provides a comparison between a hard and soft

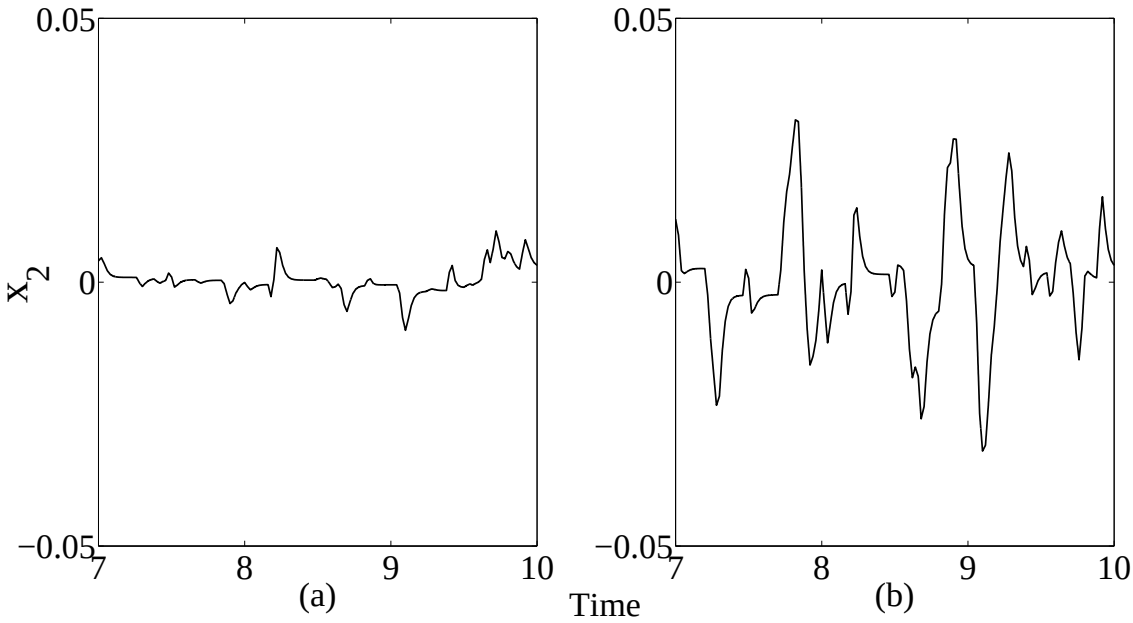

Figure 4.11: Denoising performance in steady-state with (a) soft thresholding and (b) hard thresholding.

thresholding approach for a Haar, window size 4, 1 level denoising scheme. Clearly, the bound on the error in x_2 is larger for the hard thresholding scheme. Thus, soft thresholding does a better job of preserving the signal, and eliminating the noise.

4.4.5 Windowing

As an example, the minimum number of samples (per window) for a n-level Haar transform is 2^n. If more sophisticated (online) methods are used to determine the threshold value, a larger window size can be required to achieve an accurate estimate of the noise mean and variance; otherwise large portions of the signal could be removed in error. This is not something previously addressed, given the measurement noise is a bounded uniformly distributed value, where the bound is known a priori.

Given that past values are used to construct the sample set, the window must utilize some set of initial conditions. These values could be significantly different from the true values, and will remain in the window until pushed out. After $(k-1)$

samples, where k is the sample size, the window will be populated with the true system values. If too large of a window is chosen, latency may be introduced into the system. For the system discussed here, there is no appreciable difference in system performance for a larger versus a smaller window size (in the range of 2 samples to 16 samples). This appears to be the case, given all algorithms are implemented algebraically.

4.5 Lowpass Filters

Lowpass filters are a simple and common way to eliminate high-frequency noise. However, if the frequency band of the signal overlaps that of the noise, eliminating the noise can also remove a significant portion of the signal we wish to preserve. Moreover, to produce a smoother denoised signal, we may want to increase the order of the filter. However, the higher the order the greater the phase lag introduced into the feedback loop. Thus, wavelets should have the potential to surpass lowpass filters in noise removable; otherwise, the additional complexity cannot be justified. However, this is not to suggest that all wavelet denoising schemes will outperform the classic lowpass filter. In fact, some schemes shown in this chapter do not remove more noise than a first-order lowpass filter. Thus, the many factors discussed in this chapter must be carefully considered when constructing an appropriate wavelet denoising scheme.

Consider an order eight digital Butterworth lowpass filter with a normalized cutoff frequency of 0.9. This filter is realized by entering the order and normalized cutoff frequency in the Matlab function "butter". Next, a wavelet denoising scheme is constructed for a level 2, window size 4, soft thresholding setup with an ad-hoc cutoff value. For the sake of argument the simplest waveform, the Haar wavelet, is chosen. Figure 4.12 shows the denoising performance for the Butterworth filter and the Haar

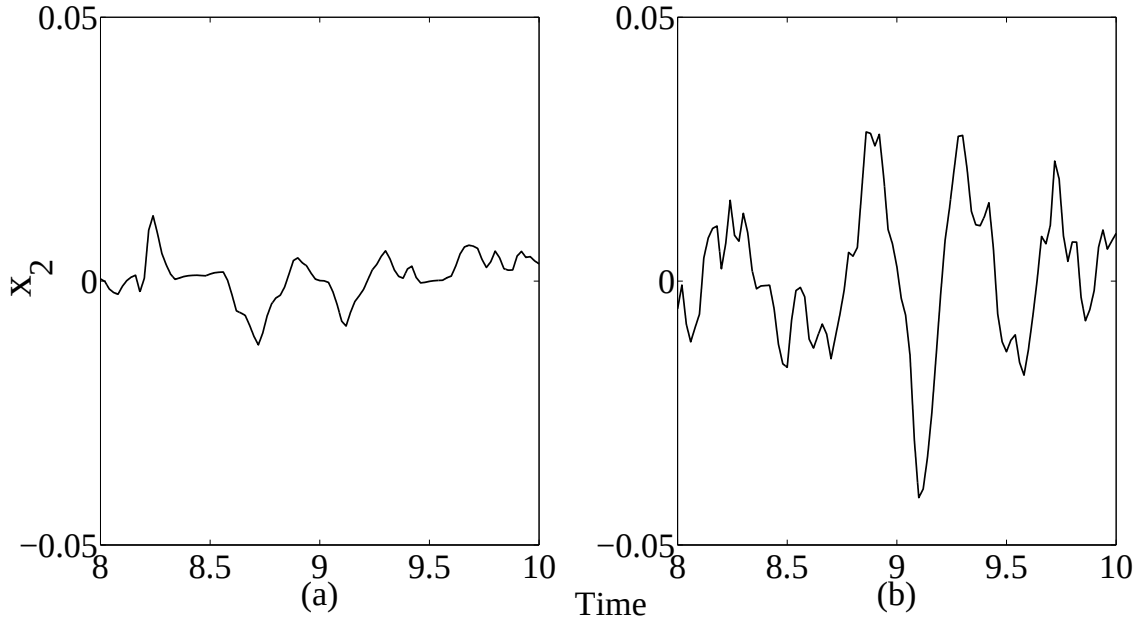

Figure 4.12: Steady-state performance comparison of a (a) Haar wavelet denoising scheme and (b) Butterworth filter.

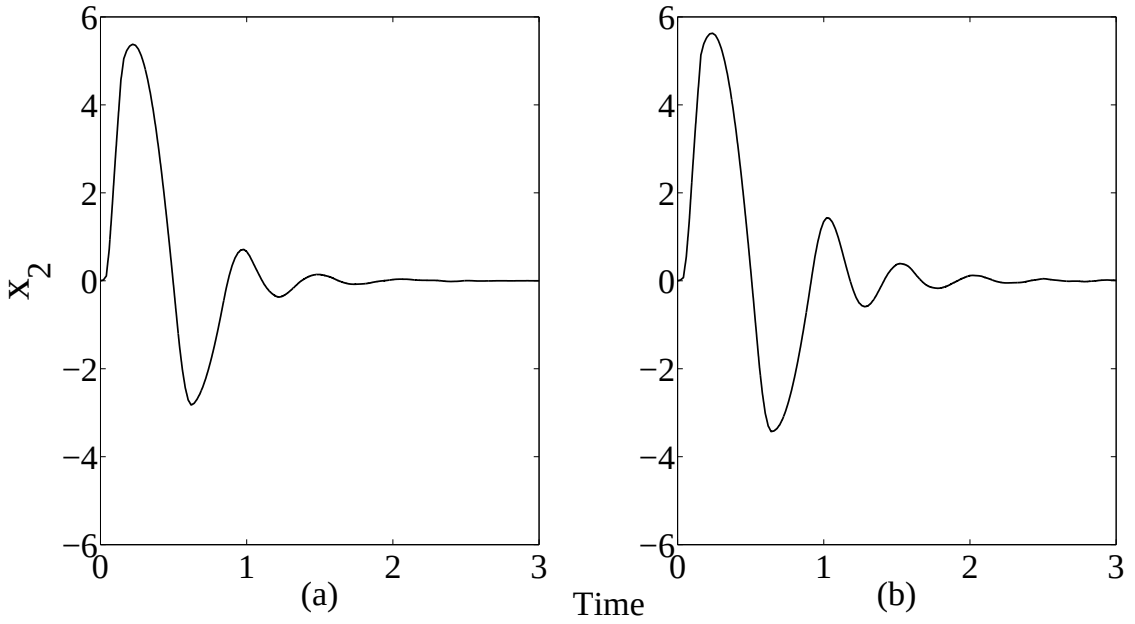

Figure 4.13: Transient performance comparison of a (a) Haar wavelet denoising scheme and (b) Butterworth filter.

wavelet denoising scheme for the steady-state signal of x_2. Clearly, the Haar wavelet is able to remove more of the noise from the system state than the Butterworth filter. Furthermore, Figure 4.13 shows that the high-gain observer utilizing the wavelet denoising scheme results in the system state x_2 having a slightly faster settling time than the observer using the Butterworth filter.

4.6 Conclusions

This chapter provided a simulation-based feasibility study on wavelet denoising to reduce the amount of measurement noise entering the high-gain observer. Logically, the wavelet design can greatly affect the system performance. We found that for denoising schemes with a relatively small window (small data set), an increase in the wavelet order leads to an increase in the amount of noise in the system states. To alleviate this deficiency, while maintaining the same wavelet function, a larger window size can be chosen. However, if there is a limit on the amount of data that can be stored, a different family of wavelets with a smaller support length is a viable option. Moreover, if data storage is a primary concern, extra attention will have to be taken in choosing the proper wavelet. It was shown that fewer wavelet coefficients are necessary to reconstruct the output signal for a wavelet that quickly captures the behavior of the noise. Moreover, excessive levels can lead to a decrease in performance, given the signal energy is being distributed across a larger number of coefficients that will likely be effected by the thresholding scheme. For the type of system and bounded measurement noise investigated in this work, the way in which we arrived at the cutoff value for the thresholding scheme appeared to have little to no effect on the denoising outcome. However, the soft thresholding function resulted in smoother and less noisy signals than the hard thresholding approach.

The ample degrees of freedom in wavelet design allow this denoising approach to be

extremely versatile. However, it is this flexibility that makes the problem incredibly challenging. Overall, the simplistic Haar transform was shown to provide superior steady-state and transient performance when compared to a lowpass Butterworth filter, for additive bounded measurement noise.

Chapter 5

Conclusions

It is a well-known fact that high-gain observers are susceptible to measurement noise. In particular, we discussed the tradeoff between fast state reconstruction, minimizing the bound on the steady-state estimation error and rejecting the model uncertainty. Hence, the focus of this dissertation has been to address issues concerned with observer design and analysis in the presence of measurement noise through three major thrust areas: observer structure, tracking performance and filtering.

Initially, a nonlinear-gain high-gain observer was constructed to capture the transient and steady-state performance present in comparable linear-gain observers. Specifically, the nonlinearity was chosen to have a higher observer gain during the transient period and a lower gain afterwards. Throughout the course of the investigation, we considered altering the number of piecewise linear regions in the nonlinear-gain function. However, we concluded that a two-piece function produced satisfactory results, suggesting that altering the number of piecewise linear regions is generally unnecessary. This approach allowed us to reduce the tradeoff between fast state reconstruction and measurement noise attenuation. The stability of the closed-loop system was proven for the case where all assumptions hold globally. Although, it was noted in Chapter 2 that a slight modification would specialize the proof for a

regional version of the theorem provided. Through the simulation results, we showed that the new gain structure successfully addresses the criticism encountered when implementing high-gain observers in systems with measurement noise.

The compromise between the transient and steady-state performance obvious in the estimation error is not apparent in the system tracking error. Motivating this notion with a nonlinear example, it became apparent that the tracking error is uniformly bounded in ε for systems without zero dynamics. Before attempting to confirm this phenomenon rigorously, it was argued by constructing the system transfer functions from the noise to the tracking error and its derivatives, that the error and its first derivative are bounded uniformly in ε for a third-order nonlinear system. Using singular perturbation analysis, the results were extended to a class of linear systems of dimension n. In particular, aside from the tracking error and its first derivative, all remaining derivatives of the tracking error are inversely proportional to increasing powers of ε. Subsequently, a similar result was derived for a class of nonlinear systems using the special features of singularly perturbed systems, further generalizing the results reported for linear systems. Due to the form of the nonlinearity, the result for nonlinear systems was restricted to a third-order system. Although it has been shown that the tracking error is more immune to the effects of measurement noise than the estimation error for both the linear and nonlinear forms considered, this does not mean that ε can be made arbitrarily small. It is important to keep in mind that the estimates of the states will still be used in the controller. However, the control may not be sensitive to a slight increase in the state estimation error, meaning we acquire some additional flexibility in choosing the value of ε when the tracking performance is the primary focus. This work can be extended to included nonlinear systems with dimension greater than three, additional control structures and the inclusion of zero dynamics.

In order to maximize the performance possible with the newly developed nonlinear-

gain high-gain observer, we performed a simulation-based feasibility study on wavelet denoising. The idea was to reduce the amount of measurement noise entering the high-gain observer from the on-set. We found that for denoising schemes with a relatively small window (small data set), an increase in the wavelet order leads to an increase in the amount of noise in the system states. To alleviate this deficiency, while maintaining the same wavelet function, a larger window size can be chosen. However, if there is a limit on the amount of data that can be stored, a different family of wavelets with a smaller support length was a viable option. Moreover, if data storage is a primary concern, extra attention will have to be taken in choosing the proper wavelet. It was shown that for a wavelet that quickly captures the behavior of the noisy signal, fewer wavelet coefficients are necessary to reconstruct the system output. Interestingly, iterating the wavelet transform excessively lead to a decrease in performance, because the signal energy was distributed across a larger number of coefficients that were reduced in the thresholding phase. Overall, the simplistic Haar transform was shown to provide superior steady-state and transient performance when compared to a lowpass Butterworth filter, for additive bounded measurement noise. The ample degrees of freedom in wavelet design allow this denoising approach to be extremely versatile. However, it is this flexibility that makes the problem incredibly challenging. There are many possibilities for future work. Namely, generalizing the results in Chapter 4 to a broader class of nonlinear systems, finding the ideal construction for an online wavelet denoising scheme and extending the work to include the noise statistics in the filter design.

APPENDICES

Appendix A

Nonnegative Impulse Response

Consider the transfer function

$$G(s) = \frac{\beta_1 \left(1 - \frac{\varepsilon_1}{\varepsilon_2}\right) s^{n-2} + \cdots + \beta_{n-1} \left(1 - \left(\frac{\varepsilon_1}{\varepsilon_2}\right)^{n-1}\right)}{s^{n-1} + \beta_1 s^{n-2} + \cdots + \beta_{n-1}} . \tag{A.1}$$

The poles and zeros of $G(s)$ can always be chosen real and distinct such that

$$G(s) = K \left(\prod_{i=1}^{\bar{m}} \frac{s + \bar{z}_i}{s + \bar{p}_i} \right) \left(\prod_{j=\bar{m}+1}^{\bar{n}} \frac{1}{s + \bar{p}_i} \right) \tag{A.2}$$

for an appropriate choice of $\beta_1, \cdots, \beta_{n-1}$ and $\varepsilon_1/\varepsilon_2$, where $\bar{m} \leq \bar{n}$, $\bar{z}_i > \bar{p}_i$ for $i = 1, ...\bar{m}$.

Example 1: System with relative degree 3 (i.e. $n = 3$)

For $n = 3$, (A.1) becomes

$$G(s) = \frac{\beta_1 \left(1 - \frac{\varepsilon_1}{\varepsilon_2}\right) + \beta_2 \left(1 - \left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right)}{s^2 + \beta_1 s + \beta_2} \,. \tag{A.3}$$

By matching like terms of (A.3) to (A.2) we see that

$$(s + p_1)(s + p_2) = s^2 + \beta_1 s + \beta_2 \tag{A.4}$$

$$p_2 > p_1 > 0 \tag{A.5}$$

meaning

$$\beta_1 = p_1 + p_2 \tag{A.6}$$

$$\beta_2 = p_1 p_2 \,. \tag{A.7}$$

Then,

$$\begin{aligned} G(s) &= \frac{\beta_1 \left(1 - \frac{\varepsilon_1}{\varepsilon_2}\right) \left[s + \frac{\beta_2}{\beta_1} \left(1 + \frac{\varepsilon_1}{\varepsilon_2}\right)\right]}{(s + p_1)(s + p_2)} \\ &= \frac{\beta_1 \left(1 - \frac{\varepsilon_1}{\varepsilon_2}\right) (s + z)}{(s + p_1)(s + p_2)} \,, \end{aligned} \tag{A.8}$$

where

$$\begin{aligned} z &= \frac{\beta_2}{\beta_1} \left(1 + \frac{\varepsilon_1}{\varepsilon_2}\right) \\ &= \frac{p_1 p_2}{p_1 + p_2} \left(1 + \frac{\varepsilon_1}{\varepsilon_2}\right) . \end{aligned}$$

Suppose $\varepsilon_1/\varepsilon_2 \geq p_1/p_2$. Then,

$$z \geq \frac{p_1 p_2}{p_1 + p_2}\left(1 + \frac{p_1}{p_2}\right) = p_1 \ . \tag{A.9}$$

Therefore, choose $p_1 = \mu$, $p_2 = 1$, $\varepsilon_1/\varepsilon_2 = \mu < 1$ and $z = \dfrac{\mu}{1+\mu}(1+\mu)$. With these choices, $G(s)$ can be written as

$$\begin{aligned} G(s) &= \frac{(1+\mu)(1-\mu)}{s+1} \\ &= \frac{1-\mu^2}{s+1} \ , \end{aligned} \tag{A.10}$$

where the condition in (A.2) is satisfied, and the poles and zeros of $G(s)$ are real and distinct.

Example 2: System with relative degree 4 (i.e. $n = 4$)

For $n = 4$, (A.1) becomes

$$G(s) = \frac{\beta_1\left(1 - \dfrac{\varepsilon_1}{\varepsilon_2}\right)s^2 + \beta_2\left(1 - \left(\dfrac{\varepsilon_1}{\varepsilon_2}\right)^2\right)s + \beta_3\left(1 - \left(\dfrac{\varepsilon_1}{\varepsilon_2}\right)^3\right)}{s^3 + \beta_1 s^2 + \beta_2 s + \beta_3} \ . \tag{A.11}$$

Choose the poles of (A.11) as $-\mu^2 p$, $-\mu p$ and $-p$. The denominator of (A.11) can be written as

$$\begin{aligned} (s+\mu^2 p)(s+\mu p)(s+p) = s^3 &+ (1+\mu+\mu^2)ps^2 \\ &+ (1+\mu+\mu^2)\mu p^2 s + \mu^3 p^3 \ , \end{aligned} \tag{A.12}$$

where $\mu < 1$. From (A.12), we can see that $\beta_1 = (1+\mu+\mu^2)p$, $\beta_2 = (1+\mu+\mu^2)\mu p^2$ and $\beta_3 = \mu^3 p^3$. Focusing on the numerator,

$$\begin{aligned}\beta_1\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right)s^2+\beta_2\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right)s+\beta_3\left(1-\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^3\right) = \\ = \beta_1\left(1-\frac{\varepsilon_1}{\varepsilon_2}\right)\left[s^2+\frac{\beta_2}{\beta_1}\left(1+\frac{\varepsilon_1}{\varepsilon_2}\right)s \right. \\ \left. +\frac{\beta_3}{\beta_1}\left(1+\frac{\varepsilon_1}{\varepsilon_2}+\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right)\right] ,\end{aligned} \tag{A.13}$$

where

$$\begin{aligned}\frac{\beta_2}{\beta_1}\left(1+\frac{\varepsilon_1}{\varepsilon_2}\right) &= \mu p\left(1+\frac{\varepsilon_1}{\varepsilon_2}\right) \\ \frac{\beta_3}{\beta_1}\left(1+\frac{\varepsilon_1}{\varepsilon_2}+\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right) &= \frac{\mu^3 p^3\left(1+\frac{\varepsilon_1}{\varepsilon_2}+\left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right)}{(1+\mu+\mu^2)p} .\end{aligned}$$

Take $p = 1$ and $\varepsilon_1/\varepsilon_2 = \mu$. Then, (A.13) becomes

$$(1+\mu+\mu^2)(1-\mu)(s+\mu)(s+\mu^2) . \tag{A.14}$$

Hence, $G(s)$ can be written as

$$G(s) = \frac{(1+\mu+\mu^2)(1-\mu)(s+\mu)(s+\mu^2)}{(s+1)(s+\mu)(s+\mu^2)} = \frac{1-\mu^3}{s+1} , \tag{A.15}$$

where the condition in (A.2) is satisfied, and the poles and zeros of $G(s)$ are real and distinct.

Example 3: System with relative degree 5 (i.e. $n = 5$)

For $n = 5$, (A.1) becomes

$$G(s) = \frac{\beta_1 \left(1 - \frac{\varepsilon_1}{\varepsilon_2}\right) s^3 + \beta_2 \left(1 - \left(\frac{\varepsilon_1}{\varepsilon_2}\right)^2\right) s^2}{s^4 + \beta_1 s^3 + \beta_2 s^2 + \beta_3 s + \beta_4} + \frac{\beta_3 \left(1 - \left(\frac{\varepsilon_1}{\varepsilon_2}\right)^3\right) s + \beta_4 \left(1 - \left(\frac{\varepsilon_1}{\varepsilon_2}\right)^4\right)}{s^4 + \beta_1 s^3 + \beta_2 s^2 + \beta_3 s + \beta_4} . \tag{A.16}$$

Choose the poles of (A.16) as $-\mu^3$, $-\mu^2$, $-\mu$, -1 and set $\varepsilon_1/\varepsilon_2 = \mu$. The denominator of (A.16) can be written as

$$\begin{aligned} (s+\mu^3)(s+\mu^2)(s+\mu)(s+1) & \\ &= s^4 + (1+\mu)(1+\mu^2)s^3 \\ &+ (\mu^5 + \mu^2(1+\mu)^2 + \mu)s^2 \\ &+ \mu^3(1+\mu)(1+\mu^2)s + \mu^6 . \end{aligned} \tag{A.17}$$

From (A.17), we can see that $\beta_1 = (1+\mu)(1+\mu^2)$, $\beta_2 = (\mu^5+\mu^2(1+\mu)^2+\mu)$, $\beta_3 = \mu^3(1+\mu)(1+\mu^2)$ and $\beta_4 = \mu^6$. The numerator in (A.16) can be represented as

$$\begin{aligned} \beta_1 (1-\mu) \Big[s^3 + \frac{\beta_2}{\beta_1}(1+\mu)s^2 + \frac{\beta_3}{\beta_1}(1+\mu+\mu^2)s & \\ + \frac{\beta_4}{\beta_1}(\mu^3+\mu^2+\mu+1) \Big] , & \end{aligned} \tag{A.18}$$

where substituting in the values for the β_i's in the bracketed term results in

$$s^3 + \mu(\mu^2+\mu+1)s^2 + \mu^3(\mu^2+\mu+1)s + \mu^6 . \tag{A.19}$$

Thus, $G(s)$ is

$$G(s) = \frac{\beta_1(1-\mu)(s+\mu)(s+\mu^2)(s+\mu^3)}{(s+1)(s+\mu)(s+\mu^2)(s+\mu^3)} = \frac{1-\mu^4}{s+1}\,, \tag{A.20}$$

where the condition in (A.2) is satisfied, and the poles and zeros of $G(s)$ are real and distinct.

Appendix B

A Block-Diagonal Form for Linear Systems

The ideas in this section are borrowed from [35]. Consider the following linear two-time-scale system

$$\dot{x} = A_{11}x + A_{12}z \tag{B.1}$$

$$\varepsilon\dot{z} = A_{21}x + A_{22}z \tag{B.2}$$

for $0 < \varepsilon < 1$. Before transforming (B.1)-(B.1) into a block-diagonal form, we first seek to bring the system into a block-triangular form. In particular, the change of variables

$$\eta(t) = z(t) + L(\varepsilon)x(t) \tag{B.3}$$

will bring the system into what is known as actuator form. Essentially, actuator form means that the states of the "slow" equation (B.1) are removed from the dynamics of the "fast" equation (B.2). This similarity transform will bring the system (B.1)-(B.2)

into the form

$$\begin{bmatrix} \dot{x}(t) \\ \varepsilon\dot{\eta}(t) \end{bmatrix} = \begin{bmatrix} A_{11} - A_{12}L & A_{12} \\ R(L,\varepsilon) & A_{22} + \varepsilon L A_{12} \end{bmatrix} \begin{bmatrix} x(t) \\ \eta(t) \end{bmatrix} , \tag{B.4}$$

where the matrix $R(L,\varepsilon)$ must be zero for the state x to be removed from the $\dot{\eta}$ equation. In order for $R(L,\varepsilon)$ to be zero, the matrix $L(\varepsilon)$ should satisfy the algebraic equation

$$R(L,\varepsilon) = A_{21} - A_{22}L + \varepsilon L A_{11} - \varepsilon L A_{12} L = 0 . \tag{B.5}$$

The system (B.1)-(B.2) is partially decoupled, providing a separate fast subsystem

$$\varepsilon\dot{\eta}(t) = (A_{22} + \varepsilon L A_{12})\eta(t) , \tag{B.6}$$

where x does not appear.

However, another change of variables is necessary to achieve a complete separation of the fast and slow states of the system (B.1)-(B.2), leading to a block-diagonal form. Applying the change of variables

$$\zeta(t) = x(t) - \varepsilon M \eta(t) \tag{B.7}$$

results in

$$\begin{bmatrix} \dot{\zeta}(t) \\ \varepsilon\dot{\eta}(t) \end{bmatrix} = \begin{bmatrix} A_{11} - A_{12}L & S(M,\varepsilon) \\ 0 & A_{22} + \varepsilon L A_{12} \end{bmatrix} \begin{bmatrix} \zeta(t) \\ \eta(t) \end{bmatrix} , \tag{B.8}$$

where the matrix M is required to satisfy the linear algebraic equation

$$S(M,\varepsilon) = \varepsilon(A_{11} - A_{12}L)M - M(A_{22} + \varepsilon L A_{12}) + A_{12} = 0 . \tag{B.9}$$

The exact slow system is given by

$$\dot{\zeta} = (A_{11} - A_{12}L)\zeta(t) \ . \tag{B.10}$$

Therefore, the system (B.1)-(B.2) assumes the block-diagonal form

$$\begin{bmatrix} \dot{\zeta}(t) \\ \varepsilon\dot{\eta}(t) \end{bmatrix} = \begin{bmatrix} A_{11} - A_{12}L & 0 \\ 0 & A_{22} + \varepsilon L A_{12} \end{bmatrix} \begin{bmatrix} \zeta(t) \\ \eta(t) \end{bmatrix} \tag{B.11}$$

and is completely decoupled. Moreover, (B.11) has a unique solution for sufficiently small ε.

Appendix C

Decomposition of Nonlinear Singularly Perturbed Systems

The block-diagonal form for linear systems presented in Appendix B completely decouples the system dynamics of the "fast" and "slow" states. An analogous block-diagonal form does not exist for nonlinear singularly perturbed systems. However, there is a nonstandard change of variables that can accomplish partial decoupling of the states. The decomposition method presented in [49] begins by removing the slow input from the fast equation, resulting in an upper triangular form. When an additional change of variables is applied to remove the fast input from the slow equation, some of the slow input is reintroduced into the fast equation; this results in a lower triangular form. Ultimately, the complete transformation is able to eliminate the fast input from the slow equation, but not the slow input from the fast equation. This section contains additional details concerning the decomposition process.

For convenience, when referring to the original work, the naming conventions

in [49] are adopted here. Consider the following singularly perturbed system

$$\dot{x} = f(t, x, y, \varepsilon) \tag{C.1}$$

$$\varepsilon\dot{y} = g(t, x, y, \varepsilon) \tag{C.2}$$

for $0 < \varepsilon < 1$. The goal of this transform is to bring the system (C.1)-(C.2) into a block-triangular form. In particular, the transform should reduce the system to

$$\dot{u} = F(t, u, \varepsilon) \tag{C.3}$$

$$\varepsilon\dot{v} = G(t, u, v, \varepsilon)\ , \tag{C.4}$$

where the "fast" input v is eliminated from the "slow" equation (C.3). The change of variables that will take the system into the desired form is

$$x = u + \varepsilon H(t, u, v, \varepsilon) \tag{C.5}$$

$$y = v + h(t, x, \varepsilon) = v + h(t, u + \varepsilon H(t, u, v, \varepsilon), \varepsilon)\ . \tag{C.6}$$

Under the assumptions listed in [49], the system (C.1)-(C.2) has an integral manifold defined as $y = h(t, x, \varepsilon)$. Those assumptions for $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$ are:

- The function $g(t, x, y, \varepsilon)$ in (C.2) evaluated at $\varepsilon = 0$ is zero, and has the isolated solution $y = h_0(t, x)$;
- The functions f, g and h_0 are twice continuously differentiable for $|y - h_0(t, x)| \leq \rho$ and $0 \leq \varepsilon \leq \varepsilon_0$;
- The eigenvalues of
$$\frac{\partial g}{\partial y}(t, x, h_0(t, x), 0)$$
are negative.

In Chapter 3, the system (3.32)-(3.33) satisfies all of the above conditions.

Definition: Integral Manifold

Consider the differential equation

$$\dot{x} = X(t, x)$$

where $x, X \in \mathbb{R}^n$. The set $\mathcal{S} \subset \mathbb{R} \times \mathbb{R}^n$ is said to be an integral manifold if for $(t_0, x_0) \in \mathcal{S}$, the solution $(t, x(t))$, $x(t_0) = x_0$ is in $\mathcal{S}$ for $t \in \mathbb{R}$.

Referring back to the system of interest (C.1)-(C.2), the dynamics on the manifold can be described by the differential equation

$$\dot{x} = f(t, x, h(t, x, \varepsilon), \varepsilon) \ , \tag{C.7}$$

where the state y is replaced by the continuously differentiable function h. For convenience when solving for h, the asymptotic expansion

$$h = h_0(t, x) + \varepsilon h_1(t, x) + \varepsilon^2 h_2(t, x) + \cdots \tag{C.8}$$

is defined, where $h(t, x, 0) = h_0$. In order for the transform in (C.5)-(C.6) to exist, h must satisfy the partial differential equation

$$\varepsilon \frac{\partial h}{\partial t} + \varepsilon \frac{\partial h}{\partial x} f(t, x, h, \varepsilon) = g(t, x, h, \varepsilon) \ . \tag{C.9}$$

The coefficients of the expansion of h in (C.8) can be found by matching like powers of ε in (C.9). Define the variables $z = y - h(t, x, \varepsilon)$ and $w = x - u$. Then,

consider the functions

$$f_1 = f(t, u + w, z + h(t, u + w, \varepsilon), \varepsilon) - f(t, u, h(t, u, \varepsilon), \varepsilon) \tag{C.10}$$

$$\begin{aligned} Z(t, u, w, z, \varepsilon) = {} & g(t, u + w, z + h(t, u + w, \varepsilon), \varepsilon) \\ & - g(t, u + w, h(t, u + w, \varepsilon), \varepsilon) \\ & - \varepsilon \frac{\partial h}{\partial x}(t, u + w, \varepsilon)[f(t, u + w, z + h(t, u + w, \varepsilon), \varepsilon) \\ & - f(t, u + w, h(t, u + w, \varepsilon), \varepsilon)] \end{aligned} \tag{C.11}$$

for the auxiliary differential system

$$\dot{u} = f(t, u, h(t, u, \varepsilon), \varepsilon) \tag{C.12}$$

$$\dot{w} = f_1(t, u, w, z, \varepsilon) \tag{C.13}$$

$$\varepsilon \dot{z} = Z(t, u, w, z, \varepsilon) \, . \tag{C.14}$$

The system (C.12)-(C.14) has the integral manifold $w = \varepsilon H(t, u, z, \varepsilon)$, where the function H satisfies the partial differential equation

$$\begin{aligned} \varepsilon \frac{\partial H}{\partial t} + \varepsilon \frac{\partial H}{\partial u} F(t, u, \varepsilon) + \frac{\partial H}{\partial v} Z(t, u, \varepsilon H, v, \varepsilon) \\ = f_1(t, u, \varepsilon H, v, \varepsilon) \, , \end{aligned} \tag{C.15}$$

where

$$F(t, u, \varepsilon) = f(t, u, h(t, u, \varepsilon), \varepsilon)$$

and

$$G(t, u, v, \varepsilon) = Z(t, u, \varepsilon H(t, u, v, \varepsilon), v, \varepsilon) \, .$$

Often times, the function H can be found as an asymptotic expansion of

$$\varepsilon H = \varepsilon H_1(t,u,v) + \varepsilon^2 H_2(t,u,v) + \varepsilon^3 H_3(t,u,v) + \cdots \tag{C.16}$$

from the expression in (C.15) by matching like coefficients in ε. Therefore, if both h and H exist and satisfy the partial differential equations in (C.9) and (C.15), respectively, (C.1)-(C.2) can be transformed into the system representation in (C.3)-(C.4) using the change of variables in (C.5)-(C.6).

REFERENCES

REFERENCES

[1] H. Adeli and H. Kim. Wavelet-hybrid feedback-least mean square algorithm for robust control of structures. *Journal of Structural Engineering*, pages 128–137, 2004.

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

[3] J.H. Ahrens and H.K. Khalil. Multirate sampled-data output feedback control with application to smart material actuated systems. *IEEE Trans. Automat. Contr.*, 54:2518–2529, 2009.

[4] A.N. Atassi. *A separation principle for the control of a class of nonlinear systems.* PhD thesis, Michigan State University, East Lansing, 1999.

[5] A.N. Atassi and H.K. Khalil. A separation principle for the control of a class of nonlinear systems. In *Proc. IEEE Conf. on Decision and Control*, pages 855–860, Tampa, FL, December 1998.

[6] A.N. Atassi and H.K. Khalil. A separation principle for the stabilization of a class of nonlinear systems. *IEEE Trans. Automat. Contr.*, 44:1672–1687, 1999.

[7] A.N. Atassi and H.K. Khalil. A separation principle for the control of a class of nonlinear systems. *IEEE Trans. Automat. Contr.*, 46, 2001.

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

[9] C.S. Burrus, R.A. Gopinath, and H. Guo. *Introduction to Wavelets and Wavelet Transforms A Primer.* Prentice Hall, New Jersey, 1998.

[10] F. Chaplais, P. Tsiotras, and D. Jung. Redundant wavelet processing on the half-axis with applications to signal denoising with small delays: Theory and experiments. *International Journal of Adaptive Control and Signal Processing*, 20:447–474, 2006.

[11] R. Coifman and M. Wickerhauser. Entropy-based algorithms for the best basis selection. *IEEE Trans. on Information Theory*, 38:713–718, 1992.

[12] M.O.T. Cole, P.S. Keogh, C.R. Burrows, and M.N Sahinkaya. Wavelet domain control of rotor vibration. *J. Mechanical Engineering Science*, 220:177–184, 2005.

[13] A.M. Dabroom and H.K. Khalil. Output feedback sampled-data control of nonlinear systems using high-gain observers. *IEEE Trans. Automat. Contr.*, 46:1712–1725, 2001.

[14] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA, 1992.

[15] D.L. Donoho. Nonlinear wavelet methods for recovering signal, images, and densities from indirect and noisy data. In *'American Mathematical Society Proc. of Symposia in Applied Mathematics*, San Antonio, Texas, 1993.

[16] D.L. Donoho. De-noising by soft thresholding. *IEEE Trans. on Information Theory*, 41:613–627, 1995.

[17] D.L. Donoho and R Coifman. Translation-invariant de-noising. In *In Wavelets and Statistics, Lecture Notes in Statistics*. Springer-Verlag, 1995.

[18] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[19] F. Doymaz, A. Bakhtazad, J.A. Romagnoli, and A. Palazoglu. Wavetlet-based robust filtering of process data. *Computers and Chemical Engineering*, 25:1549–1559, 2001.

[20] F. Esfandiari and H.K. Khalil. Observer-based design of uncertain systems: recovering state feedback robustness under matching conditions. In *Proc. Allerton Conf.*, pages 97–106, Monticello, IL, September 1987.

[21] F. Esfandiari and H.K. Khalil. Output feedback stabilization of fully linearizable systems. *Int. J. Contr.*, 56:1007–1037, 1992.

[22] J.P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems application to bioreactors. *IEEE Trans. Automat. Contr.*, 37(6):875–880, 1992.

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

[24] T. Hu, T. Thibodeau, and A.R. Teel. Analysis of oscillation and stability for systems with piecewise linear components via saturation functions. In *Proc. American Control Conf.*, pages 1911–1916, St. Louis, MO, June 2009.

[25] A. Isidori. A remark on the problem of semiglobal nonlinear output regulation. *IEEE Trans. Automat. Contr.*, 42(12):1734–1738, 1997.

[26] S. Jayasuriya and M.A. Franchek. A class of transfer functions with non-negative impulse response. *J. Dyn. Sys. Measurement and Control*, 113:313–315, 1991.

[27] A. Jensen and A.la Cour-Harbo. *Ripples in Mathematics The Discrete Wavelet Transform*. Springer, New York, 2001.

[28] Z.P. Jiang, D.J. Hill, and Y. Guo. Semi-global output feedback stabilization for the nonlinear benchmark example. In *Proc. European Control Conf.*, Brussels, Belgium, July 1997. FR-A-K-8.

[29] H.K. Khalil. Robust servomechanism output feedback controllers for a class of feedback linearizable systems. *Automatica*, 30(10):1587–1599, 1994.

[30] H.K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, New Jersey, 3rd edition, 2002.

[31] H.K. Khalil. High-gain observers in nonlinear feedback control. In *International Conf. on Contr., Automat. and Syst.*, Seoul, Korea, 2008.

[32] H.K. Khalil. Analysis of sampled-data high-gain observers in the presence of measurement noise. *European Journal of Control*, 15:166–176, 2009.

[33] H.K. Khalil and E.G. Strangas. Robust speed control of induction motors using position and current measurement. *IEEE Trans. Automat. Contr.*, 41:1216–1220, 1996.

[34] S. Khorbotly. *Design and Implementation of Low Cost De-noising Systems for Real-time control applications*. PhD thesis, University of Akron, Akron, 2007.

[35] P. Kokotovic, H.K. Khalil, and J. O'Reilly. *Singular Perturbation Methods in Control Analysis and Design*. SIAM, Philadelphi, PA, 1999.

[36] Hamid Krim, Dewey Tucker, Stephane Mallat, and David Donoho. On denoising and best signal representation. *IEEE Trans. on Information Theory*, 45:2225–2238, 1999.

[37] P. Krishnamurthy and F. Khorrami. High-gain output-feedback control for nonlinear systems based on multiple time scaling. *Systems and Control Letters*, 56:7–15, July 2007.

[38] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems*. Wiley-Interscience, New York, 1972.

[39] Y. Lin, E. Sontag, and Y. Wang. A smooth converse lyapunov theorem for robust stability. *SIAM J. Contr. Optim.*, 34:124–160, 1996.

[40] D. W. Luse and H. K. Khalil. Frequency domain results for systems with slow and fast dynamics. *IEEE Trans. Automat. Contr.*, 30(12):1171–1179, December 1985.

[41] X. Ma, C. Zhou, and I. Kemp. Automated wavelet selection and thresholding for pd detection. *IEEE Electrical Insulation Magazine*, 18:37–45, 2002.

[42] S. Mallat. *A Wavelet Tour of Signal Processing*. Elsevier, San Diego, CA, second edition edition, 2001.

[43] D.Q. Mayne, R.W. Grainger, and C.G. Goodwin. Nonlinear filters for linear signal models. *IEE Proc. Control Theory Appl.*, 144:281–286, 1997.

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

[45] T. Raff. *Impulsive observers for continuous-time systems and global output feedback control*. PhD thesis, Institute for Systems Theory and Automatic Control of the University of Stuttgart, Düsseldorf, 2010.

[46] X. Rui, M. Ke, Q. Feng, and W. Zhen-Lei. Online wavelet denoising via a moving window. *Acta Automatica Sinica*, 33:897–901, 2007.

[47] R.G. Sanfelice and L. Praly. A technical result for the study of high-gain observers with sign-indefinite gain adaptation. In *8th IFAC Symposium on Nonlinear Control Systems*, pages 284–289, University of Bologna, Italy, 2010.

[48] M.G.E. Schneiders. Wavelets in control engineering. Master's thesis, Eindhoven University of Technology, Eindhoven, Netherlands, 2001.

[49] V. A. Sobolev. Integral manifolds and decomposition of singularly perturbed systems. *Syst. Contr. Lett.*, 5(3):169–179, 1984.

[50] A. Teel and L. Praly. Global stabilizability and observability imply semi-global stabilizability by output feedback. *Syst. Contr. Lett.*, 22:313–325, 1994.

[51] A. Tewfik, D. Sinha, and P. Jorgensen. On the optimal choice of a wavelet for signal representation. *IEEE Trans. on Information Theory*, 38:747–765, 1992.

[52] A. Tilli and M. Montanari. A low-noise estimator of angular speed and acceleration from shaft encoder measurements. *Journal Automatika*, 42:169–176, 2001.

[53] L.K. Vasiljevic and H.K. Khalil. Error bounds in differentiation of noisy signals by high-gain observers. *Syst. Contr. Lett.*, 57:856–862, 2008.

[54] J.S. Walker. *Primer on Wavelets and their Scientific Applications*. CRC Press LLC, Boca Raton, FL, 1999.