AN INVESTIGATION OF TECHNIQUES FOR NONLINEAR STATE OBSERVATION

DEAN CHRISTIAAN TAIT MCBRIDE

A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science in Engineering.

Johannesburg, 2016

DECLARATION

I declare that this dissertation is my own unaided work. It is being submitted for the Degree of Master of Science to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other university.

…………………………………………………

(Signature of Candidate)

On the ………………….. day of ……………………………..……. year ………………….

ABSTRACT

An investigation and analysis of a collection of different techniques, for estimating the states of nonlinear systems, was undertaken. It was found that most of the existing literature on the topic could be organized into several groups of nonlinear observer design techniques, of which each group follows a specific concept and slight variations thereof.

From out of this investigation it was discovered that a variation of the adaptive observer could be successfully applied to numerous nonlinear systems, given only limited output information. This particular technique formed the foundation on which a design procedure was developed in order to asymptotically estimate the states of nonlinear systems of a certain form, using only partial state information available. Lyapunov stability theory was used to prove the validity of this technique, given that certain conditions and assumptions are satisfied. A heuristic procedure was then developed to get a linearized model of the error transient behaviour that could form the upper bounds of the transient times of the observer.

The technique above, characterized by a design algorithm, was then applied to three well-known nonlinear systems; namely the Lorenz attractor, the Rössler attractor, and the Van Der Pol oscillator. The results, illustrated through numerical simulation, clearly indicate that the technique developed is successful, provided all assumptions and conditions are satisfied.

ACKNOWLEDGMENTS

I would like to extend gratitude to my supervisors, Professor Brian Wigdorowitz and Professor Anton Van Wyk for the guidance and supervision during the duration of this research.

I wish to thank Yokogawa South Africa (YMA-SA), firstly for the financial support in my first year of research, and for the understanding and support throughout the duration of study.

I thank God for giving me this great opportunity, and the courage to accept it and embark on this journey. I also thank him for providing me with the ability to follow through with it and for its successful completion.

I would like to thank my parents, Brian and Frieda McBride for their encouragement, and for teaching me the value of perseverance and diligence.

Finally, I would like to sincerely thank my wife, Carmen, for the endless understanding, support, encouragement, and especially patience that she rewarded me with, during every hour of work.

TABLE OF CONTENTS

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

Table **Page**

GLOSSARY

Observability - a given system is said to be observable if the current states can be determined using only knowledge of the system outputs and the internal system parameters [1].

Detectability - a linear system is said to be detectable if all the modes that are unobservable, are stable. An implication of this is that an observable system, will also be detectable [2].

Hurwitz - a matrix is defined as Hurwitz (or Hurwitz stable) if it is square and all of its poles have negative real parts [3].

Diffeomorphism (diffeomorphic transformation) **–** a diffeomorphism is signified by an invertible (vector) function, that maps a continuous differentiable manifold to another under a change of coordinates such that the function (or matrix of functions) as well as its inverse are both continuous (smooth) [4].

Manifold **–** a manifold is signified by a topological structure that resembles a Euclidean space of the same dimension in a small neighborhood around each point situated on the manifold. A differentiable manifold comprises of charts that, in a neighborhood around a point on the chart, allows the rules of calculus to be applied (i.e. is differentiable) [5], [6]. For more on topology refer to [7].

1. INTRODUCTION

State observation, also known as state estimation, is a concept from control theory which entails estimating certain states of a system, given a mathematical model and certain output information of that system. This concept can be extremely useful, particularly in engineering industries since, given an arbitrary plant, there will always be certain state information that is either desired or needed, but cannot be measured due to certain practical constraints, such as inaccessibility of the sensor location, the cost of instrumentation, or due to the abstract nature of the state quantity to be measured. Linear observer theory (for deterministic [8], and stochastic [1], [9] systems) started coming into fruition around the 1960's and has attracted much research ever since. Consequently, the linear state observer is now well understood and has been successfully applied in numerous practical scenarios. It is however well known, that very few real systems are linear by nature, and a linear observer is usually inadequate when applied directly to the system. An approach to overcome this shortfall of the linear observer, is to linearize the system around its operating points and then only apply the linear observer to estimate the states at those points. The effectiveness of this approach is however limited, depending considerably on the degree and nature of the nonlinearity of the system, and hence it is desirable to change the approach such that an observer can be constructed for a class of observable, nonlinear systems, the objective of which is to be a successful estimator across numerous engineering applications. It is because of this requirement and the diverse nature of nonlinear systems that this research has been undertaken, and has led to the research question discussed in this dissertation.

Research Question

Assuming an arbitrary nonlinear system of differential equations of the form

$$\dot{x} = g(x) + Bu$$

$$y = h(x).$$

which nonlinear observer design techniques are appropriate in order to successfully estimate the states of a given system, assuming that the system is observable and that only partial state information is available?

This dissertation attempts to answer the research question, and is organized as follows:

Chapter 2: Literature Survey

Chapter 2 summarizes all the literature that has been reviewed and analyzed, in order to gain an understanding on what has been researched regarding the topic of both general linear and nonlinear estimation theory. The chapter starts off by briefly revising linear state estimation techniques developed by the well-known authors Luenberger [8], [10] and Kalman [1], [9]. A more extensive section then follows where literature on nonlinear state estimation is reviewed and analyzed. In this section, the reader discovers that there are in fact numerous techniques of estimating the state information of nonlinear systems, and is exposed to the most common and popular of these methods. It is also highlighted that the theory has indeed been developed over a few decades and even been applied in some practical applications; which then leads the reader to the section that follows. Here, a common application of nonlinear estimation theory is discussed; synchronization and secure communication. The literature in this section illustrates how numerous authors have applied or developed a diverse set of nonlinear estimation techniques in order to provide secure communication channels, usually masked using a chaotic master system.

Chapter 3: Methodology

In Chapter 3, an estimation technique is developed that is based on existing adaptive observer techniques reviewed in the literature survey, and more specifically, based on theory of [11] that shows the design of an adaptive controller to achieve chaotic synchronization. The proof makes use of Lyapunov stability theory and shows that the error between the state and observer is asymptotically stable. The application of this technique to any nonlinear system (assuming certain conditions are met) is made simple by a design algorithm that can be followed step by step in order to construct the observer. The chapter is then concluded by showing how an approximation of the error transient time can be made through certain 'linearizations' in order to achieve a stable linear error system. Since this theory provides approximations only, the resulting linear error transients can be seen as upper bounds of the actual nonlinear error transients.

Chapter 4: Application

The design methodology discussed in Chapter 3 is then applied to some nonlinear systems in Chapter 4 using numerical simulations. The systems considered are the Lorenz attractor, the Rössler attractor, and a 2-dimensional limit cycle. The numerical simulations show that by applying the technique of Chapter 3, an observer can be constructed with estimated states that asymptotically converge to the actual states. This section also shows how the approximated error transient times compare to the actual transient times.

Chapter 5: Conclusion

The dissertation is then concluded in Chapter 5 where the findings are summarized, the research question is reconsidered, and a decision is made concerning whether the research question has been answered or not. Implications of the findings are considered and recommended extended research ideas around the topic are then stated.

2. LITERATURE SURVEY

In order to gain a complete understanding of the topic, and consequently be in a position to address the research question, a survey of relevant literature was done. The survey covered literature published on linear state observation theory in the early 1960's, through to current day observer theory for nonlinear systems.

2.1. State Observation for Linear Systems

Luenberger Observer

The pioneering work in state observation was done by Luenberger [10], [8] in the early 1960's and 1970's. He developed the method to design a linear state observer in a deterministic setting. Also known as the Luenberger observer, it is designed for systems of the form

$$\dot{x} = Ax + Bu \tag{2.1}$$

$$y = Cx \tag{2.2}$$

where $x \in \mathcal{R}^n$ is the state vector, $u \in \mathcal{R}^m$ is the input vector, and $y \in \mathcal{R}^p$ is the output vector, and $A \in \mathcal{R}^{n\times n}$, $B \in \mathcal{R}^{n\times 1}$, $C \in \mathcal{R}^{1\times n}$ represents the system dynamics matrix, the input matrix and the output matrix respectively. The Luenberger observer assumes the form

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}) \tag{2.3}$$

$$\hat{y} = C\hat{x} \tag{2.4}$$

The objective is that the observed or estimated states $\hat{x}$, asymptotically tend toward the actual states *x*. Before the observer can be designed, the system above must be shown to be completely observable, and this is done by checking that the observability matrix

$$[\boldsymbol{C}' \;\; \boldsymbol{A}'\boldsymbol{C}' \;\; (\boldsymbol{A}')^{\boldsymbol{2}}\boldsymbol{C}' \ldots (\boldsymbol{A}')^{\boldsymbol{n-1}}\boldsymbol{C}']$$

has a rank of *n*, where *n* is the dimension of the system. Once observability is established, it is required to calculate the gain matrix *L*, such that the dynamics of the observer ensures that the estimated states converge to the actual states. According to Theorem 1 in [8], these dynamics are determined by the matrix $(A - LC)$ and the Lemma in [8] states that a matrix L can be appropriately chosen such that the poles of $(A - LC)$ can be equal to those of a specified matrix. Hence, for the observer, the poles are chosen to be negative (stable), thus ensuring the observer dynamics requirement, and must also be chosen such that the time constants are smaller than those of the plant. This will ensure that the observer converges to the actual states faster than any change due to the plant dynamics, ensuring accurate estimations. Numerous methods exist for calculating the L matrix, based upon the concept of pole placement. One of the most common of these methods is known as the Bass-Gura formula [12]. Unlike the Kalman filter (which is still to be discussed), this technique doesn't require there to be any Gaussian distribution models of the state and output, which makes this method considerably simpler and less computationally intensive. However, the drawback is that the model of the observer is required to have a high degree of accuracy, and any system modelling errors or measurement noise will immediately be reflected in the state estimations.

Kalman Filter

Around the same time, R.E. Kalman was developing a method [9], [1] for linear state estimation in a stochastic setting. This theory assumes that the state of a system at time t evolves from a previous state value at time $t - 1$, according to the equations

$$x_t = A_t x_{t-1} + B_t u_t + w_t \quad \textbf{(2.5)}$$

$$z_t = C_t x_t + v_t \quad \textbf{(2.6)}$$

where x_t is the state vector with A_t being the state transition model, and z_t is the output with C_t being the observation model. w_t represents the process noise vector for each state and is assumed to be drawn from a zero mean normal distribution with known covariance Q_t . v_t represents the measurement noise for each measurement and is similarly zero mean white Gaussian with known covariance R_t . The estimates of the state vector are determined by probability density functions, instead of absolute or discrete values, and hence it is required to know their variances and covariances. These values are stored in the state covariance matrix P, where the terms on the diagonal are the variances corresponding to each state, and the other terms are the covariances between states. The Kalman filter algorithm is a two-stage process made up of prediction steps and update steps.

The prediction steps are

$$\hat{x}_{t|t-1} = A_t \hat{x}_{t-1|t-1} + B_t u_t \quad \textbf{(2.7)}$$

$$P_{t|t-1} = A_t P_{t-1|t-1} A_t^T + Q_t \quad \textbf{(2.8)}$$

Here, a predicted value of the state $\hat{x}_t$ at time t is determined using the state transition matrix at t and the previous state value at $t-1$, added to the control input at t. P_t represents the predicted state covariance matrix at time t and evolves from the covariance matrix at $t-1$ and the transition matrix at t, added to the current process noise covariance at t.

The update steps are

$$K_t = P_{t|t-1} C_t^T (C_t P_{t|t-1} C_t^T + R_t)^{-1} \quad \textbf{(2.9)}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(z_t - C_t \hat{x}_{t|t-1}) \quad \textbf{(2.10)}$$

$$P_{t|t} = P_{t|t-1} - K_t C_t P_{t|t-1} \quad \textbf{(2.11)}$$

These equations illustrate how the state estimate $\hat{x}_{t|t}$ and the state covariance $P_{t|t}$ are updated at time t. This however can only be done once the Kalman gain K_t is computed (see references [9] and [1] for the derivations of K_t). The above steps are recursive for each time step over the specified time duration. The Kalman filter has successfully been applied in real-world applications such as

GPS navigation and tracking, and has been proven to be robust in noisy applications. However, the requirements for the state and output covariance to be Gaussian makes this technique unusable for nonlinear applications, since only a linear mapping of a Gaussian distribution can produce a Gaussian distribution.

The linear state observation theory mentioned above has been well developed over the past 50 years and has been useful in numerous real word applications, however, since most systems are nonlinear in nature, this theory has its limitations. Therefore an investigation into nonlinear state estimation techniques was undertaken.

2.2. State Observation for Nonlinear Systems

Extended Kalman Filter

The Extended Kalman Filter [13], [14] was designed for systems of the form

$$x_t = f(x_{t-1}, u_t) + w_t \quad \textbf{(2.12)}$$

$$z_t = h(x_t) + v_t \quad \textbf{(2.13)}$$

where $f(\cdot)$ is the nonlinear state transition model and $h(\cdot)$ is the observation model. Referring to the Kalman filter in Section 2.1 it was highlighted that a Gaussian covariance model applied to a nonlinear mapping will produce a distribution that no longer is Gaussian and hence the Kalman filter cannot be used in the nonlinear case. Consequently the extended Kalman filter requires that $f(\cdot)$ and $h(\cdot)$ be locally linearized around the operating points, by computing the Jacobian matrices. The remainder of the procedure is the same as with the Kalman filter and makes use of these linearized matrices. The prediction steps are

$$\hat{x}_{t|t-1} = f(\hat{x}_{t-1|t-1}, u_{t-1}) \quad \textbf{(2.14)}$$

$$P_{t|t-1} = F_{t-1} P_{t-1|t-1} F_{t-1}^{T} + Q_t \quad \textbf{(2.15)}$$

The update steps are

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R_t)^{-1} \quad \textbf{(2.16)}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t (z_t - h(\hat{x}_{t|t-1})) \quad \textbf{(2.17)}$$

$$P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1} \quad \textbf{(2.18)}$$

where $F_{t-1} = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{t-1|t-1}, u_{t-1}}$ and $H_t = \left.\frac{\partial h}{\partial x}\right|_{\hat{x}_{t|t-1}}$

The extended Kalman filter has proved to work reasonably well in practical applications where there are moderate nonlinearities, however, in systems that are highly nonlinear, this method starts to fail since it relies on local linearizations. When significant nonlinearities are present, the actual systems and the linearized systems start deviating too much and estimation accuracy is lost. Consequently, much research has been done to address the problem of estimating the state across all forms of nonlinear systems.

High Gain Observers

As the name suggests, the point behind high gain observation, is to design an observer that makes use of an adequately strong observer gain that can 'eliminate' the effects of the nonlinearity in the error dynamics. Initially developed by Thau [15] it soon found numerous applications in the nonlinear control field, [16], [17]. With reference to [18], and considering a system of dimension 2, the theory requires the form to be

$$\dot{x}_1 = x_2 \quad \textbf{(2.19)}$$

$$\dot{x}_2 = \varphi(x, u) \quad \textbf{(2.20)}$$

$$y = x_1 \quad \textbf{(2.21)}$$

The observer is designed as

$$\dot{\hat{x}}_1 = \hat{x}_2 + h_1(y - \hat{x}_1) \quad \textbf{(2.22)}$$

$$\dot{\hat{x}}_2 = \varphi_0(\hat{x}, u) + h_2(y - \hat{x}_1) \quad \textbf{(2.23)}$$

$$\hat{y} = \hat{x}_1 \quad \textbf{(2.24)}$$

where $\varphi_0(\hat{x}, u)$ is simply the observer model of the nonlinearity $\varphi(x, u)$. Then defining the error of each state as $\tilde{x}_1 = x_1 - \hat{x}_1$ and $\tilde{x}_2 = x_2 - \hat{x}_2$, the error dynamics becomes

$$\dot{\tilde{x}}_1 = -h_1\tilde{x}_1 + \tilde{x}_2 \quad \textbf{(2.25)}$$

$$\dot{\tilde{x}}_2 = -h_2\tilde{x}_1 + \delta(\tilde{x}, u) \quad \textbf{(2.26)}$$

where $\delta(\tilde{x}, u) = \varphi(x, u) - \varphi_0(\hat{x}, u)$. Here it is required to design $H = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$ in such a way that $A_0 = \begin{bmatrix} -h_1 & 1 \\ -h_2 & 0 \end{bmatrix}$ is Hurwitz (stable). However, this won't be sufficient since the effect of the nonlinearity will still be significant. Considering then the transfer function from $\delta(\tilde{x}, u)$ to $\tilde{x}$,

$$G_0(s) = \frac{1}{s^2 + h_1 s + h_2}\begin{bmatrix} 1 \\ s + h_1 \end{bmatrix} \quad \textbf{(2.27)}$$

the idea is to design H such that $\sup_{\omega \epsilon \mathcal{R}} \|G_0(j\omega)\|_\infty$ is minimized, where in this case $\|\cdot\|_\infty$ represents the L_∞- norm [19]. This can be done by letting $h_1 = \frac{\alpha_1}{\varepsilon}$ and $h_2 = \frac{\alpha_2}{\varepsilon^2}$ where $0 < \varepsilon \ll 1$. The transfer function can then be written as

$$G_0(s) = \frac{\varepsilon}{(\varepsilon s)^2 + \varepsilon\alpha_1 s + \alpha_2}\begin{bmatrix} \varepsilon \\ \varepsilon s + \alpha_1 \end{bmatrix} \quad \textbf{(2.28)}$$

The observer poles (eigenvalues) will be λ_1/ε and λ_2/ε where λ_1 and λ_2 are the chosen roots of the characteristic polynomial $\lambda^2 + \alpha_1\lambda + \alpha_2$. Consequently, the small magnitude of ε will cause $\sup_{\omega \epsilon \mathcal{R}} \|G_0(j\omega)\| = O(\varepsilon)$, and will therefore no longer have a significant effect on the error dynamics, resulting in an asymptotically stable observer. Due to the high gain, the main disadvantage of this technique is a peaking phenomenon (i.e. a large overshoot of the estimated states in comparison to the actual states) during the transient behaviour which can quickly cause any practical control effort to become saturated. This nonlinear state estimation technique has, however, been around for a number of years and much research has been put into attenuating the peaking phenomena as well as other issues [20], [21], [22] inter alia. One other drawback of this technique is the requirement for the system to be transformable into the form (2.19) - (2.21), and since not all systems can undergo this transformation, this technique will not always be a viable option. (The form (2.19) - (2.21) is known as the nonlinear observable canonical form and will be discussed in more detail further on in this dissertation).

A review of recent literature covering high gain observers will show that the theory is quite well developed and most of the research covers the application thereof [23], [24], [25]. Since the estimation method of high gain observers has become so common, the theory has found application in certain industries, of which some have started putting it into practice, as is discussed in [26], [27].

Nonlinear Observers through Output Injection

Krener and Isidori proposed in [28], that if it were needed to observe the state of a nonlinear system of the form

$$\dot{\xi} = f(\xi) \tag{2.29}$$

$$y = h(\xi) \tag{2.30}$$

then it is conceivable that there could be some sort of transformation that could yield a system of the form

$$\dot{x} = Ax + \varphi(y) \tag{2.31}$$

$$y = Cx \tag{2.32}$$

where the change of coordinates is done by a transformation $x = \Phi(\xi)$. Assuming this transformation is possible then designing an observer for this system would simply be designing a Luenberger observer for a similar linear system. An observer for the transformed system would then be

$$\dot{\hat{x}} = A\hat{x} + \varphi(y) + L(y - C\hat{x}) \quad \textbf{(2.33)}$$

with the error dynamics simply being

$$\dot{e} = (A - LC)e \quad \textbf{(2.34)}$$

where

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

The literature then proceeds to show and prove the method by which the transformation is constructed [28]. Notably, a limitation in this case, is that the transformation and its inverse must exist and not be singular, and must hence be a diffeomorphism. As can be seen, the transformed state is a function of the original state and output

$$x = \phi(\xi) = [h(\xi)\ L_f h(\xi) \ldots L_f^{n-1} h(\xi)]^T$$

where $L_f h(\xi)$ is the Lie derivative of $h(\xi)$ along the trajectories of $f(\xi)$. The observability assumption is that all terms of the transformation are linearly independent, i.e. $L_f^k(dh)(\xi), k = 0, \ldots, n-1$ are linearly independent. This method of transformation described above has formed the foundation of much research into nonlinear observer theory since then, resulting in the term 'nonlinear observable canonical form' which refers to the form of the system in the transformed coordinates. The reason for this being, after the original system has successfully been transformed, the resulting system will always have a structure where *A*, *C* is observable. The fact that the resulting transformed system (assuming a successful transformation is possible) is guaranteed to be in an observable form, makes this method very powerful, and hence it can be noted that the term 'nonlinear observable canonical form' is ubiquitous in the realms of nonlinear state estimation. Further literature concerning this transformation and observers of this form can be seen in [29], and [30].

Sliding Mode Observers

Although sliding mode estimation was developed around the 1990's by Drakunov [31], the notion of sliding modes had already been around for a few years [32] and had found application as a control method for nonlinear systems. The idea behind sliding mode control is to force the state trajectories onto a sliding surface (also referred to as a hypersurface) through a designed control input. This sliding surface is a stable manifold of the system and so all trajectories on this manifold will reach an equilibrium point. Once this manifold is reached, the controller continuously works to keep the trajectory on this manifold, and is usually characterized by high frequency switching. There are usually two parts to designing this control scheme: selection of the sliding surface or manifold, and designing the feedback in order to drive the trajectory onto this manifold. The sliding mode observer uses this concept, and it is designed in such a way that the error trajectories between the actual and estimated states are also forced onto a sliding surface via an equivalent control, and in turn reach a stable equilibrium point $e = 0$. The observer is designed for systems of the form

$$\dot{x} = f(x) \qquad \textbf{(2.35)}$$

$$y = h(x) \in \mathcal{R}^m \qquad \textbf{(2.36)}$$

where the observer takes the form

$$\dot{\hat{x}} = \left(\frac{\partial H(\hat{x})}{\partial x}\right)^{-1} M(\hat{x})\mathrm{sgn}\big(V(t) - H(\hat{x})\big) \qquad \textbf{(2.37)}$$

Here $H(x) = \mathrm{col}\{h_1(x), \dots, h_n(x)\}$ where

$$h_1(x) = h(x) \qquad \textbf{(2.38)}$$

$$h_i(x) = L_f^{i-1} h(x), \qquad i = 1, \dots, n \tag{2.39}$$

and $h_i(x)$ is the i[th] Lie derivative of $h(x)$ along the trajectories of $f(x)$.

$V(t) = \mathrm{col}\{v_1(t), \dots, v_n(t)\}$ where

$$v_1(t) = y(t) \tag{2.40}$$

$$v_{i+1}(t) = (m_i(\hat{x})\mathrm{sgn}(v_i(t) - h_i(\hat{x}(t)))_{eq} \tag{2.41}$$

and M is an $(n \times n)$ diagonal matrix, all entries (m_i) being positive. Following the procedure in [31], the error can be written as $e = H(x) - H(\hat{x})$ which produces the error dynamics

$$\dot{e} = \frac{dH(x)}{dt} - M(\hat{x})\mathrm{sgn}(V(t) - H(\hat{x}(t))) \tag{2.42}$$

and so

$$\begin{bmatrix} \dot{e}_1 \\ \dot{e}_2 \\ \vdots \\ \dot{e}_n \end{bmatrix} = \begin{bmatrix} \dot{h}_1(x) - m_1 \mathrm{sgn}\left(v_1(t) - h_1(\hat{x}(t))\right) \\ \dot{h}_2(x) - m_2 \mathrm{sgn}\left(v_2(t) - h_2(\hat{x}(t))\right) \\ \vdots \\ \dot{h}_n(x) - m_n \mathrm{sgn}\left(v_n(t) - h_n(\hat{x}(t))\right) \end{bmatrix} \tag{2.43}$$

$$= \begin{bmatrix} h_2(x) - m_1 \mathrm{sgn}\left(v_1(t) - h_1(\hat{x}(t))\right) \\ h_3(x) - m_2 \mathrm{sgn}\left(v_2(t) - h_2(\hat{x}(t))\right) \\ \vdots \\ L_f^n h(x) - \quad m_n \mathrm{sgn}\left(v_n(t) - h_n(\hat{x}(t))\right) \end{bmatrix} \tag{2.44}$$

Considering the first line of the error equation, note that $v_1(t) = y(t) = h_1(x)$. This causes the term $v_1(t) - h_1(\hat{x}(t)) = e_1$ and so the first equation becomes $\dot{e}_1 = h_2(x) - m_1 \mathrm{sgn}(e_1)$. As long as $m_1 \geq |h_2(x(t))|$, this will ensure that the first row of the error dynamics will enter the $e_1 = 0$ sliding mode. Considering the second equation of the error dynamics, on the $e_1 = 0$ surface, the $v_2(t) = \{m_1 \mathrm{sgn}(e_1)\}_{eq}$ equivalent control term will be equal to $h_2(x)$ and therefore $v_2(t) - h_2(\hat{x}) = h_2(x) - h_2(\hat{x}) = e_2$, and so similarly, as with the first equation, as long as

$m_2 \geq |h_3(x(t))|$, the second equation of the error dynamics will enter the sliding mode manifold for $e_2 = 0$. Once all error trajectories have entered onto the stable manifold, they will go to zero, and hence the observed states will converge to the actual states. For nonlinear systems where the operating points (or the state values) coincide with the equilibrium points, this method proves to be sufficient. However, further evaluation of this method will clearly highlight, that when it is required to observe the state of a system around an operating point that does not necessarily lie on a stable manifold, then this method will fail to produce accurate estimates of the states.

Nonlinear State Observer by Output Vector

In more recent years, much of the development and research into estimation theory has been done on a technique that uses an output vector. Messaoud, Zanzouri and Ksouri [33] consider nonlinear systems of the form

$$\dot{x} = f(x) + Bu + Ed \quad \textbf{(2.45)}$$

$$y = Cx \quad \textbf{(2.46)}$$

where $f(x)$, Bu, y and Ed represent the system dynamics, the known input, the measured output and the unknown inputs respectively. B, E, C and the structure of $f(x)$ are known. It is supposed that $f(x)$ is continuously differentiable. The objective of the technique discussed in [33] is to design an observer that can asymptotically estimate the states of the nonlinear system (2.45) - (2.46) without any information on the input d. The observer takes the form

$$\hat{x} = z + Hy \quad \textbf{(2.47)}$$

where $\hat{x}$ is the state estimate and z is the output vector. z is determined by

$$\dot{z} = F_{\hat{x}}z + K_{\hat{x}}y + Tu - P(f_{\hat{x}} - D_x(f_{\hat{x}})\hat{x}) \quad \textbf{(2.48)}$$

and $F_{\hat{x}}$, $K_{\hat{x}}$, T, and H are

$$F_{\hat{x}} = -PD_x(f_{\hat{x}}) - K_{1\hat{x}}\, C \quad \textbf{(2.49)}$$

$$K_{\hat{x}} = K_{1\hat{x}} + F_{\hat{x}}H \quad \textbf{(2.50)}$$

$$PE = 0 \quad \textbf{(2.51)}$$

$$T = -PB \quad \textbf{(2.52)}$$

$$H = (I + P)C^{+} \quad \textbf{(2.53)}$$

Here, P and $K_{1\hat{x}}$ are to be chosen by the designers, C^{+} is the pseudoinverse of C (refer to [19]), and $D_x(f_{\hat{x}})$ represents the Jacobian of the nonlinearity f with respect to $\hat{x}$. The following necessary conditions are to be satisfied:

1. $\text{rank}(CE) = \text{rank}(E) = p$
2. $PD_x(f_{\hat{x}}) + K_{1\hat{x}}\, C$ is a positive definite matrix

where p is the dimension of the unknown input disturbance. The proof is based on Lyapunov stability theory and can be seen in more detail in [33].

Similar to the technique described above, Zhao, Shen, Ma and Gu have developed an output vector state observer in [34] for nonlinear Lipschitz uncertain systems. The systems are of the form

$$\dot{x} = Ax + Bu + \varphi(x, u, t) + Ed \quad \textbf{(2.54)}$$

$$y = Cx \quad \textbf{(2.55)}$$

where x, u, $\varphi(x, u, t)$, y and d represent the state, the known input, the system nonlinearity, the measured output and the unknown input. A, B, C and E are all known matrices. For the technique to successfully construct an observer, three assumptions are needed. Firstly, $\text{rank}(CE) = \text{rank}(E) = q$ (where q is the dimension of the unknown input d). Secondly, the nonlinearity $\varphi(x, u, t)$ must be Lipschitz continuous, i.e.

$$\|\varphi(x, u, t) - \varphi(\hat{x}, u, t)\| \le \gamma\, \|x - \hat{x}\|$$

where γ is the Lipschitz constant and $\|\cdot\|$ represents the L_2-norm of the vector $\varphi(x, u, t) - \varphi(\hat{x}, u, t)$. Thirdly, the unknown input d must be bounded, i.e.

$$\|d - \hat{d}\| \leq \delta$$

where δ is positive and real. *(Throughout the rest of the dissertation,* $\|\cdot\|$ *will always represent the* L_2*- norm* [19] *unless stated otherwise).* The authors make use of Lyapunov stability theory, as well as a lemma (refer to *Lemma 2.1*) to prove the observer converges (or, equivalently, that the error is stable).

Lemma 2.1

Consider systems that have been transformed into the form (2.54) - (2.55).

If there exist two matrices K, G and a positive symmetric matrix $P \in \mathcal{R}^{n\times n}$ *such that*

$$P(A - LC) + (A - LC)^T P < 0 \quad \textbf{(2.56)}$$

$$E^T P = GC \quad \textbf{(2.57)}$$

then an asymptotic state observer can be constructed, provided certain conditions on the unknown input are met (for the proof of Lemma 2.1 see [35]).

The observer, which asymptotically converges to the state, can be designed of the form

$$\hat{x} = z - Fy \quad \textbf{(2.58)}$$

where $\hat{x}$ is the state estimate and z is the output vector. z is determined by

$$\dot{z} = Nz + Ly + Du + M\varphi(\hat{x}, u, t) + R\hat{d} \quad \textbf{(2.59)}$$

N, L, D, M, R, and *F* are all matrices that need to be determined and are done so accordingly

$$Y = P^{-1}\overline{Y} \quad \textbf{(2.60)}$$

$$K = P^{-1}\overline{K} \quad \textbf{(2.61)}$$

$$U = -E(CE)^{+} \quad \textbf{(2.62)}$$

$$V = I_p - (CE)(CE)^{+} \quad \textbf{(2.63)}$$

$$F = U + YV \quad \textbf{(2.64)}$$

$$M = (I_n + FC) \quad \textbf{(2.65)}$$

$$R = (I_n + FC) \quad \textbf{(2.66)}$$

$$D = MB \quad \textbf{(2.67)}$$

$$L = K\left(I_p + CF\right) - MAF \quad \textbf{(2.68)}$$

$$N = MA - KC \quad \textbf{(2.69)}$$

$(CE)^+$ is known as the left inverse of (CE) and reference can be made to [19] for its calculation. In order to obtain P, $\bar{Y}$ and $\bar{K}$, it is required to find feasible solutions to a set of linear matrix inequalities (LMI) that have to be satisfied in order to guarantee error stability. This can be done numerically using an LMI computation toolbox, which makes use of convex optimization techniques (see [36], [37] for more information on solving LMI's).

A brief evaluation of the technique developed in [33] will make it clear to the reader that it is not constructive i.e. the matrices P and $K_{1\hat{x}}$ have to be chosen by the designers in order for the observer to work and cannot be constructed or calculated from given information. Methods such as these can prove to be frustrating to the estimator designer since they are not very reliable and the designer can spend considerable effort in attempting to obtain arbitrary matrices that work. The technique described in [34] however is different and it can be noted that all the required matrices are attainable through calculation, and finding plausible solutions for P, $\bar{Y}$ and $\bar{K}$ can be achieved through solving a set of LMI's. The drawback of this being that the designer must become familiar with the theory behind solving LMI's and the corresponding tools needed. Further literature concerning these methods, and the applications thereof can be seen in [38] and [39].

It can be noted, after extensively reviewing general observer literature, that the processing of Linear Matrix Inequalities has become ubiquitous in the field of estimation theory. This primarily stems from the use of Lyapunov theory, where it is required to prove stability of the error dynamics by proving that the derivative of the candidate function is negative definite. These LMI's are becoming more prominent in the literature, and much recent research in the field of nonlinear observers makes use of their solutions, see [40], [41], [42] and [43].

State-Dependent Riccati Equation (SDRE) Observer

SDRE methods for estimation started appearing in the research community around the mid 1990's [44], [45], and has been studied ever since [30], [46] . The research published by Mracek, Cloutier and D'Souza [44] shows that the design of the SDRE filter (as it was referred to) reveals a similar structure to that of the Kalman and extended Kalman filters, however the main difference being that where the extended Kalman filter requires there to be a linearization, the SDRE filter makes use of a state-dependent coefficient (SDC) matrix that parameterizes the system into a linear structure. Systems of the general form

$$\dot{x} = f(x) \tag{2.70}$$

$$y = h(x) \tag{2.71}$$

are assumed to be representable in the linear structure as

$$\dot{x} = A(x)x \tag{2.72}$$

$$y = C(x)x \tag{2.73}$$

where $A(x) \in \mathcal{R}^{n\times n}$ and $C(x) \in \mathcal{R}^{m\times m}$ are the state-dependent coefficient matrices. The SDRE filter is then designed as

$$\dot{\hat{x}} = A(\hat{x})\hat{x} + K(t)(y - C(\hat{x})\hat{x}) \tag{2.74}$$

with $K(t)$ as the gain matrix, determined by

$$K(t) = P(t)C^T(\hat{x})W^{-1} \tag{2.75}$$

Here, $P(t)$ is the solution to the state-dependent Riccati equation.

$$A(\hat{x})P(t) + P(t)A^T(\hat{x}) - P(t)C^T(\hat{x})W^{-1}(\hat{x})C(\hat{x})P(t) + V(\hat{x}) = 0 \tag{2.76}$$

where $W(x) \in \mathcal{R}^{m \times m}$ and $V(x) \in \mathcal{R}^{n \times n}$ are seen as tuning parameters and must satisfy the conditions $V(x) \geq 0,\ W(x) > 0,\ \forall\, x \in \mathcal{R}^n$. It is clear that the solution to the SDRE above is found using the SDC matrices, instead of the usual system and output matrices (A, C) with constant parameters. Consequently, the solution will continuously be changing due to the continuous changing of the parameters i.e. the observed state. This implies that the calculation of the solution to the SDRE can clearly become quite computationally intensive, since it must be determined at each time step of the estimation process, an unfortunate drawback of this technique.

Finite-time Observers

The discussion to this point has covered mainly observers that converge to the state asymptotically i.e. the error tends toward zero as time tends to infinity. However, certain applications require the estimate to become acceptably accurate within a certain period of time, and hence the finite-time observer has been developed [47], [48], [49], [50]. More recent literature uses existing finite-time observer theory but investigates semi-global finite-time observers on systems of a certain class [51], as well as those designed for more general systems, with less restrictions [52]. The design procedure in [49] considers systems that can be represented by the nonlinear observable canonical form

$$\dot{x} = Ax + f(y) \tag{2.77}$$

$$y = Cx \tag{2.78}$$

where $C = [1 \quad 0 \quad 0]$, and an observer for the system is designed as

$$\begin{bmatrix} \frac{d\hat{x}_1}{dt} \\ \vdots \\ \frac{d\hat{x}_n}{dt} \end{bmatrix} = A \begin{bmatrix} x_1 \\ \hat{x}_1 \\ \vdots \\ \hat{x}_n \end{bmatrix} + f(y) + \Phi_n(y - \hat{x}_1) \tag{2.79}$$

where

$$\Phi_n(y - \hat{x}_1) = \begin{bmatrix} k_1\lceil x_1 - \hat{x}_1 \rfloor^{\alpha} \\ k_2\lceil x_1 - \hat{x}_1 \rfloor^{2\alpha-1} \\ \vdots \\ k_n\lceil x_1 - \hat{x}_1 \rfloor^{n\alpha-(n-1)} \end{bmatrix} \quad \textbf{(2.80)}$$

and $\lceil x_1 - \hat{x}_1 \rfloor^{\alpha} = |x_1 - \hat{x}_1|^{\alpha}\text{sgn}(x_1 - \hat{x}_1)$. The error dynamics will consequently become

$$\begin{cases} \dot{e}_1 = e_2 - k_1\lceil e_1 \rfloor^{\alpha} \\ \dot{e}_2 = e_3 - k_2\lceil e_1 \rfloor^{2\alpha-1} \\ \vdots \\ \dot{e}_{n-1} = e_n - k_{n-1}\lceil e_1 \rfloor^{(n-1)\alpha-(n-2)} \\ \dot{e}_n = -k_n\lceil e_1 \rfloor^{n\alpha-(n-1)} \end{cases} \quad \textbf{(2.81)}$$

The observer design follows Theorem 2.1 that proves that the error dynamics tend to the origin in finite time.

Theorem 2.1

Consider the matrix A_0 that should be made stable by specifying the constant gains $[k_1, \ldots, k_n]$ accordingly.

$$A_0 = \begin{bmatrix} -k_1 & 1 & 0 & 0 & 0 \\ -k_2 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -k_{n-1} & 0 & 0 & 0 & 1 \\ -k_n & 0 & 0 & 0 & 0 \end{bmatrix} \quad \textbf{(2.82)}$$

Then there exists $\varepsilon \in \left[1 - \frac{1}{n-1}, 1\right)$ so that for all $\alpha \in (1 - \varepsilon, 1)$, the error system (2.81) becomes globally stable in finite time (see [49] for the proof of Theorem 2.1).

In summary Theorem 2.1 above implies that with appropriate choices of α and $(k_i)_{1\le i\le n}$, the error can converge to zero in finite time. As mentioned earlier, this method of estimating the state to zero error within a finite amount of time certainly has its advantages. For example, the designer will easily be able to tell if an estimate is accurate or not by simply looking at the time passed after a disturbance or change in the state value. This method (discussed in [49]) is however not constructive, and so a feasible solution depends on a manual iterative process to arrive at a working set of parameters.

It is recommended that the reader refer to [53], [54] for more recent literature on finite-time observers. It can be noted that the theory remains similar as discussed above, but is applied to specific types or classes of systems.

Adaptive Observer Design

The concept of an adaptive observer is to estimate certain states of a system as well as to simultaneously estimate any unknown parameters of that same system. Estimation techniques that make use of an adaptation law started with linear systems theory with Luders and Narendra [55], and Kreisselmeier [56]. The observer designed in [56] was seen as an equivalent observer to the Luenberger observer, but just with a different structure (which was called a parameterized observer), where it continuously updated the observer parameters through various adaptation schemes. The property of this observer to update parameters was clearly seen as useful since it was realized that the quality or accuracy of any designed observer for a given system could only be as good as the model of the system itself. Since, in real world applications, it is challenging to perfectly model a given system, observers that would otherwise successfully estimate the required states, were failing due to mismatched model parameters. Research into adaptive observers has then been a popular research topic in estimation theory. Early research on this topic for nonlinear systems includes that of Bastin and Gevers [57] where they propose a transformation of the system

$$\dot{z} = f(z, u, p) \tag{2.83}$$

$$y = z_1 \tag{2.84}$$

where z is the state vector, u is the known input, y is the measurable output, and p is a vector of unknown, possibly time-varying parameters. The continuous, smooth transformation

$$\begin{bmatrix} x \\ \theta \end{bmatrix} = T(z, p, c_2, \dots, c_n) \tag{2.85}$$

is to change the original system into what is referred to as the *canonical adaptive observer form*

$$\dot{x} = Rx + \Omega\big(\omega(t)\big)\theta(t) + g(t) \tag{2.86}$$

$$R = \begin{bmatrix} 0 & k \\ 0 & A_0 \end{bmatrix} \tag{2.87}$$

$$y = x_1 \tag{2.88}$$

Here, $x \in \mathcal{R}^n$, $\theta(t) \in \mathcal{R}^m$ is a vector of unknown time-varying parameters, $\Omega\big(\omega(t)\big)$ is an $n \times m$ matrix with entries of $\alpha_{ij}\big(\omega(t)\big)$ where α_{ij} is a known constant, $\omega(t) \in \mathcal{R}^m$ is a vector of known functions of the inputs and outputs, and $g(t)$ is a vector of known functions of t. R is an $n \times n$ matrix of constants where k is known and $k = [k_1, \dots, k_n]$ and $A_0(c_2, \dots, c_n)$ is an $(n-1) \times (n-1)$ matrix of arbitrary design constants that are to be chosen in such a way that A_0 is stable. The form of the system after transformation appears strange and designing the actual transformation $T(z, p, c)$ can be quite difficult, however, it does have advantages. Bastin and Gevers [57] point out that firstly, the apparent linear structure of the transformed system makes it easier to design an adaptive observer for it, and secondly, there are numerous time-varying linear, time-invariant bilinear, and nonlinear physical systems that can be transformed into this form. The adaptive observer is then designed as

$$\dot{\hat{x}} = R\hat{x} + \Omega\big(\omega(t)\big)\hat{\theta}(t) + g(t) + \begin{bmatrix} c_1\tilde{y} \\ V\hat{\theta} \end{bmatrix} \tag{2.89}$$

$$\hat{y} = \hat{x}_1 \ , \qquad \tilde{y} = y - \hat{y} \tag{2.90}$$

where c_1 is an arbitrary constant, and $(c_2, \dots, c_n)$ are arbitrarily chosen such that A_0 is stable. The parameter adaptation law is given by

$$\dot{\hat{\theta}} = \Gamma\varphi(t)\tilde{y} \qquad \textbf{(2.91)}$$

where $\Gamma = \text{diag}(\gamma_1, \dots, \gamma_m), \gamma_i > 0$ is chosen as a positive definite matrix. $V(t)$ and $\varphi(t)$ are found by the solution of an auxiliary filter

$$\dot{V} = FV + \bar{\Omega}\big(\omega(t)\big), \qquad V(0) = 0 \qquad \textbf{(2.92)}$$

$$\varphi(t) = V^T k + \Omega_1^T\big(\omega(t)\big) \qquad \textbf{(2.93)}$$

where Ω_1 and $\bar{\Omega}$ are the first rows and remaining rows of $\Omega\big(\omega(t)\big)$ respectively. The literature in [57] then continues to highlight sufficient conditions of the error system stability as well as useful reachability conditions for a successful transformation. Continued research into this topic has 'streamlined' and simplified the design process as is illustrated in [58], for example. In this literature, the system again undergoes a transformation into the form

$$\dot{x} = A_0 x + f(y) + g(x)u + \psi(t)\theta \qquad \textbf{(2.94)}$$

$$y = C_0 x \qquad \textbf{(2.95)}$$

$$A_0 = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \qquad C_0 = [1\ 0 \dots\ 0] \qquad \textbf{(2.96)}$$

where $f(\cdot)$ and $g(\cdot)$ are two nonlinear functions in triangular form

$$f(x) = \begin{bmatrix} f_1(x_1) \\ f_2(x_1, x_2) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) \end{bmatrix} \qquad \textbf{(2.97)}$$

$$g(x) = \begin{bmatrix} g_1(x_1) \\ g_2(x_1, x_2) \\ \vdots \\ g_n(x_1, x_2, \dots, x_n) \end{bmatrix} \qquad \textbf{(2.98)}$$

and $\psi(t) \in \mathcal{R}^{n\times n}$ is a known matrix and $\theta \in \mathcal{R}^n$ is a vector of unknown parameters. The conditions of the transformation are the same as those mentioned earlier for the Nonlinear Observer through Output Injection. The adaptive observer is then designed as

$$\dot{\hat{x}} = A_0\hat{x} + f(\hat{x}) + g(\hat{x})u + \psi(t)\hat{\theta}(t) + \rho\Lambda^{-1}K(y - C_0\hat{x}) \quad \textbf{(2.99)}$$

$$\dot{\hat{\theta}} = -2EG(C_0\hat{x} - y) \quad \textbf{(2.100)}$$

where

$$\Lambda = \begin{bmatrix} 1 & 0 & & \cdots & 0 \\ 0 & \rho^{-1} & & \cdots & 0 \\ \vdots & \vdots & & \cdots & \vdots \\ 0 & 0 & \cdots & & \rho^{-(n-1)} \end{bmatrix} \quad \textbf{(2.101)}$$

and ρ is any positive real number. K is given by

$$K = \frac{1}{2}S^{-1}C_0^T \quad \textbf{(2.102)}$$

where S is a positive definite matrix and is the solution to the equation

$$A_0^T S + SA_0 + S = C_0^T C_0 \quad \textbf{(2.103)}$$

In this instance, assumptions on the system are that $f(\cdot)$ and $g(\cdot)$ are globally Lipschitz nonlinearities and $\psi(t)$ is bounded, and that u and θ are also bounded. Other than the constraints on $f(\cdot)$ and $g(\cdot)$, and in certain circumstances the limitations of boundedness on u and θ, adaptive observers have proved to be very useful across a wide variety of systems. The simple concept of the unknown parameter vector of the model settling to the actual parameter value of the system as the error approaches zero, has allowed for greater estimator accuracy given mismatched model parameters. This concept of adaptation has also proved to be quite versatile in certain applications and it is this versatility that has provided the motivation for further investigation into this particular technique, which will be discussed in greater detail further on in this dissertation.

Research in this particular area of adaptive nonlinear observers continues and entails extensions of the theory and the application of the existing theory to different types of systems. The reader can refer to [59], [60], [61], [62], [63], [64], [65], [66], [67], [68] for more information and most recent advances on this topic.

2.3. A Common Application using Observer-based Methods

The theory researched and discussed in Section 2.2, has not only provided control systems engineers with useful tools to estimate the states of process variables in the absence of certain measured data, but has also provided a foundation for techniques in synchronization. Observer-based methods have proven useful in synchronizing a slave system with a given dynamical master system, by driving it with a scalar signal, or signal from only a single channel. An application of this being secure and encrypted communication, where a chaotic system is used to mask the message, and the slave system or receiver is designed to recover it. Early work in this area was done by Pecora and Carrol [69], Nijmeier and Mareels [29] as well as Cuomo, Oppenhiem and Strogatz [70]. The summaries that follow, indicate how the techniques in Section 2.2 have been successfully applied to this application. It also summaries some novel techniques, not mentioned in Section 2.2, that have been developed to address the problem of chaotic synchronization, but that are also based on general observer concepts.

Chaotic Synchronization using Output Injection and High Gain Observer Methods

The literature in [29] explains how the synchronization problem is addressed for linear systems, followed by systems with linearizable error dynamics, and finally for nonlinear systems. It is then illustrated how existing observer theory can be applied in certain conditions that will guarantee synchronization. For example, in the case for systems with linearizable error dynamics and with well-defined solutions, a coordinate transformation from systems of the form (2.29) – (2.30) to those of the form (2.31) – (2.32) is applied, as shown in the section of nonlinear observers through output injection. A simple high-gain observer is then designed as the slave system

$$\dot{\hat{x}} = A\hat{x} + \varphi(\hat{y}) + K_\theta(\hat{y} - y) \quad \text{OR} \quad \dot{\hat{x}} = F(\hat{x}) + K_\theta(\hat{y} - y) \tag{2.104}$$

$$\hat{y} = \hat{x}_1 \tag{2.105}$$

where

$$K_{\theta} = -S_{\theta}^{-1}C^{T} \tag{2.106}$$

and S_{θ} is a positive definite matrix and the solution to the algebraic Riccati equation

$$0 = \theta S_{\theta} + A^{T}S_{\theta} + S_{\theta}A - C^{T}C \tag{2.107}$$

The error is then proven to exponentially approach zero for $t \to \infty$.

If interested in the proofs, the reader is encouraged to analyze the literature in more detail. It will be noticed that most of the proofs in the above techniques as well as the literature to follow, rely on Lyapunov stability theory to prove that the error dynamics are stable. Also take note that a restriction on the nonlinearity to be Lipschitz continuous, will appear frequently in the literature for nonlinear observers and synchronization, since it is usually a vital assumption that is used when proving the error stability with Lyapunov theory.

Chaotic Synchronization using Finite-time Observation Methods after a Coordinate Transformation

The method of transformation, as described above, is also found to be ubiquitous amongst the theory developed for synchronization, as is demonstrated by Perruquetti, Floquet and Moulay in [49], where the slave system is synchronized to a transformed master system, through design of a finite-time observer. The synchronization technique via reduced observers explained by Zheng and Boutat in [71] also requires the master system to be in a type of nonlinear observer canonical form that is specifically conducive for the type of observer that is designed thereafter. This 'reduced-order canonical form' is slightly different to the form mentioned above, however the concept remains the same, and the procedure followed as well as the necessary and sufficient conditions to be satisfied for its existence is also similar.

A Novel Synchronization Technique Using an Injected Scalar Function

As useful and popular as it is, the method of first transformation and then synchronization, is not all that has been researched on this topic. In [72] a technique is shown that synchronizes one chaotic system to another. The synchronization problem is restated as a nonlinear observer design problem and requires the construction of a nonlinear function built into the slave system in order to achieve the synchronization. The two chaotic systems are given as

$$\dot{x} = f(x), \qquad x \in \mathcal{R}^n \tag{2.108}$$

$$\dot{y} = f(y), \qquad y \in \mathcal{R}^n \tag{2.109}$$

Synchronization is said to have been achieved when

$$e(t) = \big(y(t) - x(t)\big) \rightarrow 0 \quad \text{as} \quad t \rightarrow \infty \tag{2.110}$$

The observer is then designed as

$$\dot{y} = f(y) + g\big(z - s(y)\big) \tag{2.111}$$

where the output from the first system is represented as $z = s(x) \in \mathcal{R}^n$ and $g(\cdot)$ is any suitably chosen nonlinear function that ensures stability of the error dynamics. With application to secure communication, it is required that $s(x)$ be a scalar output i.e. $z = s(x) \in \mathcal{R}$. Then in order to design an appropriate nonlinear function $g(\cdot)$ it is first assumed that the master system can be written as

$$\dot{x} = Ax + bf(x) + c \tag{2.112}$$

where $A \in \mathcal{R}^{n\times n}$, $b \in \mathcal{R}^{n\times 1}$, $c \in \mathcal{R}^{n\times 1}$ and $f : \mathcal{R}^n \rightarrow \mathcal{R}$. The scalar output signal can be specified as

$$s(x) = f(x) + kx \tag{2.113}$$

with $k = (k_1, \dots, k_n)$ and let

$$g\big(s(x) - s(y)\big) = b\big(s(x) - s(y)\big) \tag{2.114}$$

Substituting in these two functions results in the linear error dynamics

$$\dot{e} = Ae - bke \quad \textbf{(2.115)}$$

where the term $-ke$ plays the role of a control feedback law. The values of $k = [k_1, \dots, k_n]$ must then be chosen in such a way that the error dynamics have stable poles. Examples are shown in [72] using well-known chaotic systems that illustrate the effectiveness of the method. However, the rather significant limitation of the dimension of the nonlinearity $f(\cdot)$ (as is quoted in the assumption of the form), makes this method impractical for many nonlinear systems.

Chaotic Synchronization Based on the Luenberger Observer Form and the Solution to the Riccati Inequality

Another technique that also doesn't require any transformations is designed in [73]. Here, the authors consider systems that can be represented in the form

$$\dot{x} = Ax + f(x, y) + B(\Phi(x, y)\theta + d) \quad \textbf{(2.116)}$$

$$y = Cx \quad \textbf{(2.117)}$$

where $x \in \mathcal{R}^n$, $y \in \mathcal{R}^m$, $\theta \in \mathcal{R}^q$, $d \in \mathcal{R}^p$ and $f : \mathcal{R}^n \to \mathcal{R}^n$. Assuming that the pair (*A*, *C*) is detectable, an observer can be designed as

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}, y) + B(\Phi(\hat{x}, y)\theta + d) + L(y - C\hat{x}) \quad \textbf{(2.118)}$$

$$\hat{y} = C\hat{x} \quad \textbf{(2.119)}$$

where *L* is the gain matrix that is to be determined to ensure synchronization. For the technique to work, two assumptions are made. Firstly, both $f(\hat{x}, y)$ and $\Phi(\hat{x}, y)$ must satisfy the Lipschitz continuity condition

$$\|f(x, y) - f(\hat{x}, y)\| \leq k_f \|x - \hat{x}\|$$

$$\|\Phi(x, y) - \Phi(\hat{x}, y)\| \leq k_\Phi \|x - \hat{x}\|$$

Secondly, the gain matrix L must exist, as well as a symmetric, positive definite matrix P such that the Riccati inequality is satisfied

$$(A - LC)^T P + P(A - LC) + \left(k_f + k_\Phi \|\theta\| \|B\|\right)^2 PP + I < 0 \qquad \textbf{(2.120)}$$

Feasible solutions to this inequality can be found using LMI numerical computation tools. To illustrate how a given system can be written in the form (2.118) an example using the Chua circuit is considered. With reference to [74], the Chua system is usually written as

$$\dot{x}_1 = \rho[x_2 - x_1 - f(x_1)]$$

$$\dot{x}_2 = x_1 - x_2 + x_3$$

$$\dot{x}_3 = -\sigma x_2 + x_3$$

where $f(x_1) = bx_1 + \frac{1}{2}(a - b)(|x_1 + 1| - |x_1 - 1|)$

with a, b, ρ, σ being constants with values of -1.28, -0.68, 10, 15.53 respectively. Re-writing it in the form (2.118)

$$\dot{x} = \begin{bmatrix} -10 & 10 & 0 \\ 1 & -1 & 1 \\ 0 & -14.28 & 0 \end{bmatrix} x + \begin{bmatrix} f(x_1) \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} (-1.25 x_2)$$

$$y = Cx = [0\ 0\ 1]x$$

where $\Phi(\hat{x}, y) = x_2$, $B = [0 \quad 0 \quad 1]^T$, $\theta = -1.25$ and $d = 0$. A synchronized slave system can then successfully be constructed according to the literature in [73] since both the assumptions are satisfied i.e. $f(\hat{x}, y)$ and $\Phi(\hat{x}, y)$ are Lipschitz continuous and (A, C) are detectable (observable).

Chaotic Synchronization Based on the Luenberger Observer

In [75] the authors also use observer-based methods to synchronize a slave system to a chaotic master system, with exponential convergence. In [75], the form of the slave system resembles the form of the linear Luenberger observer, except with more conditions that are to be satisfied before the construction of the gain matrix L can take place. The gain matrix also contains a parameter that can be viewed as a tuning parameter that has the effect of changing the transient behavior of the observer. The systems considered are similar to those that are linear and are of the form

$$\dot{z} = Az + Bg(z) \quad \textbf{(2.121)}$$

$$y = C^T z \quad \textbf{(2.122)}$$

where $A \in \mathcal{R}^{n\times n}$, and $B \in \mathcal{R}^n$. Once again the Lipschitz continuity restriction is placed on the nonlinearity $g(z)$. As long as (*A*, *C*) is detectable, the observer can be designed as

$$\dot{\hat{z}} = A\hat{z} + Bg(\hat{z}) + L(y - \hat{y}) \quad \textbf{(2.123)}$$

where the gain matrix *L* must be determined to ensure that $(A - LC)$ is Hurwitz stable. The authors then continue to state a further proposition saying that the gain matrix *L* can be designed as

$$L(\theta) = O^{-1}L_o(\theta) \quad \textbf{(2.124)}$$

where $O = (C^T \quad C^T A \quad \cdots \quad C^T A^{n-1})^T$ is the observability matrix, and $L_o(\theta) = (\alpha_1\theta \quad \alpha_2\theta^2 \quad \cdots \quad \alpha_n\theta^n)^T$. The constant parameters $\alpha_i,\ i = 1, 2, \dots, n$ are chosen such that the characteristic polynomial, given by $p_n(s) = s^n + \alpha_1 s^{n-1} + \cdots + \alpha_n$, is stable. The parameter θ can be appropriately chosen and changed by the designer, and as long as $\theta_{min} > 1$, it is proved that the error $e \rightarrow 0$ exponentially. The proof involves a coordinate change through a transformation that is done by the invertible observability matrix O, i.e. $x = Oz$, $A_o = OAO^{-1}$, $B_o = OB$, and $C_o^T = C^T O^{-1}$. The convergence rate is shown to be directly linked to the pair (*A*, *C*) since when this pair is observable, the convergence rate can be changed by changing the value of the parameter θ, however, when the pair is only detectable (i.e. rank(*O*) < *n*) then the convergence rate is limited to that of the slowest mode or pole.

A Novel Synchronization Technique Based on the Sylvester Equation

The technique described in [46] proposes an observer based on the solution to the state-dependent Sylvester equation. The technique is also viewed as a nonlinear extension of the Luenberger observer and also highlights the advantages of analytically determining the solution of the state-dependent Sylvester equation as opposed to constantly determining the solution to the state-dependent Riccati equation online. Systems of the form (2.70) - (2.71) are expressed in a state-dependent linear representation as

$$\dot{x} = F(x)x \tag{2.125}$$

$$y = H(x)x \tag{2.126}$$

where $F(x) \in \mathcal{R}^{n\times n}$, $H(x) \in \mathcal{R}^{m\times n}$. The observer-based synchronized system can be constructed as

$$\dot{\hat{x}} = F(\hat{x})\hat{x} + X^{-1}(\hat{x})B(\hat{x})(y - H(\hat{x})\hat{x}) \tag{2.127}$$

where $X(x), B(x)$ are parameters of the Sylvester equation

$$X(x)F(x) = A(x)X(x) + B(x)H(x) \tag{2.128}$$

In the Sylvester equation $A(x) \in \mathcal{R}^{n\times n}$ and $B(x) \in \mathcal{R}^{n\times m}$ are free parameter matrices that are selected by the designer, and $X(x)$ is the calculated solution. The observer is proved to be asymptotically stable under three assumptions: the origin of the system given by (2.70) - (2.71) is Lyapunov stable, and for all $x \in \Omega$ the pair $\left(F(x), H(x)\right)$ is observable for $H(x) \in \mathcal{R}^{1\times n}$, and for all $x \in \Omega$, $F(x)$ and $H(x)$ are locally Lipschitz continuous. From the proof it is shown that the parameters of $A(x)$ must be chosen to make it a stable matrix in order to make $A(0)$ a constant stable matrix, which in turn guarantees asymptotic stability. The authors illustrate the method by applying it to the chaotic Lorenz system, and it's useful to show this example for clarity on setting up the Sylvester equation. It is shown that the parameter matrices of the Sylvester equation, $A(x)$ and $B(x)$, are chosen as

$$A(x) = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

and $B(x) = [1 \quad 1 \quad 1]^T$, and note that $A(x)$ is chosen as stable. The Sylvester equation is then solved for $X(x)$ (refer to Appendix A for details of the solution to the Sylvester equation) and the solution is then used to calculate the gain matrix $L(x) = X^{-1}(x)B(x)$. Here one is exposed to another technique that successfully estimates the states of the system. Although it has been shown to be an improvement of the method of observation by the state-dependent Riccati equation, it is still important to highlight the issues of this technique that could possibly be stumbling blocks in the design process. The first of these issues is the fact that the state-dependent matrices $\left(F(x), H(x)\right)$ have to contain only the states that will be available in the output, for use in the observer. Secondly, we once again encounter the situation of having to arbitrarily choose matrices $A(x)$ and $B(x)$ such that a valid solution to the Sylvester equation can be found, i.e. there is no constructive method highlighted in [46] that can be used to obtain these matrices deterministically. To add to these, there are also further assumptions that have to be considered to guarantee the validity of this method, such as the requirement for the equilibrium point to be stable, as well as the Lipschitz continuity and observability assumptions. All these requirements that have to be satisfied to enable the technique to work, unfortunately also mean that the technique is less versatile and will not be successful for a wide variety of applications.

Chaotic Synchronization Based on Adaptive Observer Methods

The adaptive observer design method has been used frequently in the field of synchronization as is illustrated in [11]. The method is based on the updating of an adaptive parameter, as described in Section 2.2, which is used as a changing gain that is dependent on the magnitude of the error term. The technique in [11] addresses nonlinear systems that are in the common form

$$\dot{x} = Ax + Bf(x) \quad \textbf{(2.129)}$$

$$y = Cx \quad \textbf{(2.130)}$$

This observer-based method then makes use of the usual correcting term as well as a control term to build the synchronizing slave system

$$\dot{\hat{x}} = A\hat{x} + Bf(\hat{x}) + Bu + \frac{1}{2}\hat{\theta}B(y - C\hat{x}) \tag{2.131}$$

where *u* is the control term and $\frac{1}{2}\hat{\theta}B(y - C\hat{x})$ is the corrector term. The adaptive parameter is given by $\hat{\theta}$ and is updated according to the ODE

$$\dot{\hat{\theta}} = l\|y - C\hat{x}\|^2 \tag{2.132}$$

The technique then states that it is not necessary for the nonlinear term $f(x)$ to be Lipschitz continuous as long as the control term *u* is chosen as

$$u = f(x) - f(\hat{x}) \tag{2.133}$$

This is shown in the proof in [11] where it can be noted that the nonlinearities cancel out when the error term is determined as $e = x - \hat{x}$. This places less restriction on the nature of the master system nonlinearity, but in turn places a larger restriction by the amount of state information that must be supplied by the master system output. For example, if the nonlinearity contains terms with all state information in it, then the control term *u* of the slave system must contain all that information, which results in *C* having to be a full matrix. This method is, however, simple in its design and is shown to provide fundamental concepts for extended techniques that prove to have diverse application in recent research into observer theory.

After much analysis and application of the observer and synchronization methods discussed throughout Chapter 2, it was found that the theory discussed in [11] seemed to be the most general and could be successfully applied to a variety of engineering systems. This generality proved to be the foundation on which a design methodology was developed, and the approach taken was based on the underlying concepts of adaptive observers and the theory analyzed in [11]. The objective, was to develop a method of state estimation, given only partial state information in the output, and to incorporate an adaptive law in the observer similar to that found in [11]. The details of this method will be discussed in Chapter 3 and the applications thereof in Chapter 4.

3. METHODOLOGY

The survey of literature in Chapter 2 has made it clear that design methods for nonlinear observers have been around for several decades. It has also highlighted the fact that the theory of the techniques themselves are diverse and can be applied to a wide variety of situations where it is needed to observe the states of a nonlinear system. This variety of techniques is due to the nature of nonlinear systems as well as the nature of practical engineering systems. Since there can be a diverse range of systems, as well as a diverse range of scenarios in which an observer is needed to estimate state information, each technique has been developed due to the inherent constraints on specific systems and/or scenarios. Hence, as has been discussed, each technique works in certain cases, but not in others. It is obvious to state that not only will it be ideal, but also practical to have a single technique available that can be applied to a diverse set of systems, instead of being forced to change the observer method depending on whether different conditions are met. Consequently, the methods in the literature were analyzed, and the techniques applied to a variety of nonlinear systems in order to gauge the degree of generality amongst them.

It was found that the most general theory of the studied literature was the adaptive observer theory, more specifically the theory from [11]. Referring to the literature of [11] we recall that the technique describes a way to construct a slave system that synchronizes to a master system by making use of the adaptive parameter as well as a control term. This control term is defined as $u = f(x) - f(\hat{x})$, which implies that the entire nonlinearity of the master system must be present in the slave system. If the nonlinearity contains terms in which all state information is present, then in order to construct the slave system, all state information must be available. This requirement clearly defeats the purpose of constructing an observer in the first place, since the nature of estimation theory is to construct all state information from only partial state measurements. Hence, a design methodology was developed and proved that eliminates the need for this control term and yet still manages to successfully estimate all unknown states.

3.1. Design of a Nonlinear Observer

Consider nonlinear systems

$$\dot{x} = g(x) + Bu \tag{3.1}$$

$$y = h(x) \tag{3.2}$$

that can either be re-written in the form

$$\dot{x} = Ax + f(x) + Bu \tag{3.3}$$

$$y = Cx \tag{3.4}$$

or that can undergo a coordinate change via a diffeomorphic transformation into the form

$$\dot{z} = Az + \varphi(y) + B_z u \tag{3.5}$$

$$y = Cz \tag{3.6}$$

where the nonlinearity is given by $f(x) = [f_1(x) \quad \cdots \quad f_n(x)]^T$ or $\varphi(y)$, and $x, z \in \mathcal{R}^n$, $A \in \mathcal{R}^{n\times n}$, $B, B_z \in \mathcal{R}^{n\times m}$ and $C \in \mathcal{R}^{p\times n}$. The following theorem can be stated

Theorem 3.1

Given then that the system is in the form (3.3) - (3.4) the observer for this system can be designed as

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}) + \frac{1}{2}\hat{\theta}D(y - C\hat{x}) + Bu \tag{3.7}$$

where the matrix D is to be determined to ensure that the error

$$e = x - \hat{x} \rightarrow 0 \; as \; t \rightarrow \infty \tag{3.8}$$

and where the adaptive parameter $\hat{\theta}$ is updated according to

$$\dot{\hat{\theta}} = l\|y - C\hat{x}\|^2 \quad \textbf{(3.9)}$$

Proof

In order for the error to be proved to be stable it is required to highlight some assumptions and conditions first (refer to [35]).

Assumption 1:

$$\text{rank}(CD) = \text{rank}(D)$$

Assumption 2:

$$\text{rank}\begin{bmatrix} A - \lambda I & D \\ C & 0 \end{bmatrix} = n + \text{rank}(D)$$

Condition 1:

Again referring to [35], the following condition can be stated in light of *Assumptions 1* and *2* above. There exists an $n \times p$ matrix L, and an $n \times n$, symmetric, positive definite matrix P such that

$$P(A - LC) + (A - LC)^T P = -Q \quad \textbf{(3.10)}$$

$$D^T P = C \quad \textbf{(3.11)}$$

Condition 1 holds, if and only if *Assumptions 1 & 2* hold.

Assumption 3:

The nonlinearity is Lipschitz continuous with Lipschitz constant γ

$$\|f(x) - f(\hat{x})\| \leq \gamma\|x - \hat{x}\| \quad \textbf{(3.12)}$$

where $\|\cdot\|$ represents the L_2- norm. Then let the error be given by

$$e = x - \hat{x} \quad \textbf{(3.13)}$$

and hence the error dynamics will be

$$\dot{e} = \dot{x} - \dot{\hat{x}} \quad \textbf{(3.14)}$$

$$\dot{e} = (Ax + f(x) + Bu) - \left(A\hat{x} + f(\hat{x}) + \frac{1}{2}\hat{\theta}D(y - C\hat{x}) + Bu\right) \quad \textbf{(3.15)}$$

$$\dot{e} = Ae + (\, f(x) - f(\hat{x})\,) - \frac{1}{2}\hat{\theta}DCe \tag{3.16}$$

Then, re-writing (3.16) we get

$$\dot{e} = (A - LC)e + LCe + (\, f(x) - f(\hat{x})\,) - \frac{1}{2}\hat{\theta}DCe \tag{3.17}$$

Consider the Lyapunov candidate function

$$V = e^T Pe + \frac{1}{2}l^{-1}\tilde{\theta}^2 \tag{3.18}$$

where

$$\tilde{\theta} = \theta - \hat{\theta} \tag{3.19}$$

Then, differentiating

$$\dot{V} = \dot{e}^T Pe + e^T P\dot{e} + l^{-1}\tilde{\theta}\dot{\tilde{\theta}} \tag{3.20}$$

Then, substituting (3.17) into (3.20)

$$\dot{V} = \left(e^T(A - LC)^T + e^T(LC)^T - \frac{1}{2}\hat{\theta}e^T(DC)^T + f^T(x) - f^T(\hat{x})\right)Pe$$

$$+ e^T P((A - LC)e + LCe - \frac{1}{2}\hat{\theta}DCe$$

$$+ f(x) - f(\hat{x})) + l^{-1}\tilde{\theta}\dot{\tilde{\theta}} \tag{3.21}$$

$$= e^T\big((A - LC)^T P + P(A - LC)\big)e + e^T((LC)^T P + PLC)e$$

$$-\left(\frac{1}{2}\hat{\theta}e^T((DC)^T P + PDC)e\right) + \big(f^T(x) - f^T(\hat{x})\big)Pe$$

$$+ e^T P\big(f(x) - f(\hat{x})\big) + l^{-1}\tilde{\theta}\dot{\tilde{\theta}} \tag{3.22}$$

Then, considering the term $e^T((LC)^T P + PLC)e$

$$e^T((LC)^T P + PLC)e = e^T(LC)^T Pe + e^T PLCe \tag{3.23}$$

However, $e^T(LC)^T Pe \in \mathcal{R}^{1\times 1}$ i.e. scalar, therefore

$$(e^T(LC)^T Pe)^T = e^T(LC)^T Pe \tag{3.24}$$

$$e^T P^T LCe = e^T (LC)^T Pe \tag{3.25}$$

And since $P = P^T$ from *Condition 1*,

$$e^T PLCe = e^T (LC)^T Pe \tag{3.26}$$

Therefore

$$PLC = (LC)^T P \tag{3.27}$$

Then, also considering the term $e^T((DC)^T P + PDC)e$, and applying the same procedure (with reference to (3.23) to (3.27)), a similar result is obtained

$$(DC)^T P = P^T DC = PDC$$

Using this result, as well as equation (3.10) from *Condition 1*, (3.22) then becomes

$$\dot{V} = -\, e^T Qe + \, 2e^T PLCe - \hat{\theta} e^T PDCe + \, l^{-1}\tilde{\theta}\dot{\tilde{\theta}}$$

$$+ \big(f^T(x) - \, f^T(\hat{x})\big)Pe \, + \, e^T P\big(f(x) - f(\hat{x})\big) \tag{3.28}$$

The, for simplification, let

$$\big(f^T(x) - \, f^T(\hat{x})\big)Pe \, + \, e^T P\big(f(x) - f(\hat{x})\big) = \, \alpha(x, \hat{x}) \tag{3.29}$$

Then, substituting in both (3.11) from *Condition 1* and (3.29) into (3.28) we get

$$\dot{V} = -\, e^T Qe + \, 2e^T PLD^T Pe - \hat{\theta} e^T PDD^T Pe + \, l^{-1}\tilde{\theta}\dot{\tilde{\theta}} + \, \alpha(x, \hat{x}) \tag{3.30}$$

$$\leq -\, e^T Qe + \, 2\|D^T Pe\|\|L^T Pe\| - \hat{\theta}\|D^T Pe\|^2 + \, l^{-1}\tilde{\theta}\dot{\tilde{\theta}} + \, \alpha(x, \hat{x}) \tag{3.31}$$

Applying the following result to the term $2\|D^T Pe\|\|L^T Pe\|$ in (3.31)

$$\text{if } (a - b)^2 \geq 0$$

$$\text{then } 2ab \leq a^2 + b^2 \tag{3.32}$$

we get

$$2\|D^T Pe\|\|L^T Pe\| \leq \|D^T Pe\|^2 + \|L^T Pe\|^2 \tag{3.33}$$

Also considering the following result from [19] and applying that to the term $\|L^T Pe\|^2$ in (3.33)

$$\|X\| = \sqrt{\lambda_{\max}(X^T X)}$$

$$\text{where } \lambda_{\max}(X^T X) \text{ is the maximum eigenvalue of } X^T X \qquad \textbf{(3.34)}$$

we get

$$\|L^T Pe\|^2 = \lambda_{\max}(PLL^T P). \|e\|^2 \qquad \textbf{(3.35)}$$

Substituting (3.33) and (3.35) into (3.31), we then get

$$\dot{V} \leq -e^T Qe + \|D^T Pe\|^2 + \lambda_{\max}(PLL^T P). \|e\|^2$$

$$-\hat{\theta}\|D^T Pe\|^2 + l^{-1}\tilde{\theta}\dot{\tilde{\theta}} + \alpha(x, \hat{x}) \qquad \textbf{(3.36)}$$

$$\leq e^T(-Q + \lambda_{\max}(PLL^T P)I)e + \|D^T Pe\|^2(1 - \hat{\theta})$$

$$+ l^{-1}\tilde{\theta}\dot{\tilde{\theta}} + \alpha(x, \hat{x}) \qquad \textbf{(3.37)}$$

Considering the adaptive parameter equation (3.19), and then differentiating, we get

$$\dot{\tilde{\theta}} = \dot{\theta} - \dot{\hat{\theta}} \qquad \textbf{(3.38)}$$

taking note that

$$\theta = \text{constant}$$

$$\text{therefore } \dot{\theta} = 0$$

Considering the term out of equation (3.37), and substituting (3.38) into it

$$l^{-1}\tilde{\theta}\dot{\tilde{\theta}} = l^{-1}(\theta - \hat{\theta})(\dot{\theta} - \dot{\hat{\theta}})$$

$$= l^{-1}(-\theta\dot{\hat{\theta}} + \hat{\theta}\dot{\hat{\theta}})$$

$$= l^{-1}\dot{\hat{\theta}}(-\theta + \hat{\theta}) \qquad \textbf{(3.39)}$$

Consequently, (3.37) becomes

$$\dot{V} \leq e^T(-Q + \lambda_{\max}(PLL^T P)I)e + \|D^T Pe\|^2(1 - \hat{\theta})$$

$$+ l^{-1}\dot{\hat{\theta}}(-\theta + \hat{\theta}) + \alpha(x, \hat{x}) \qquad \textbf{(3.40)}$$

Substituting in for the adaptive parameter (3.9)

$$\dot{V} \leq e^T(-Q + \lambda_{\max}(PLL^TP)I)e + \|D^TPe\|^2(1 - \hat{\theta})$$
$$+ l^{-1}l\|y - C\hat{x}\|^2(-\theta + \hat{\theta}) + \alpha(x, \hat{x}) \quad \textbf{(3.41)}$$

where

$$l\|y - C\hat{x}\|^2 = l\|C(x - \hat{x})\|^2 = l\|D^TPe\|^2 \quad \textbf{(3.42)}$$

This leads to

$$\dot{V} \leq e^T(-Q + \lambda_{\max}(PLL^TP)I)e + \|D^TPe\|^2(1 - \hat{\theta})$$
$$+ l^{-1}l\|D^TPe\|^2(-\theta + \hat{\theta}) + \alpha(x, \hat{x}) \quad \textbf{(3.43)}$$

$$\dot{V} \leq e^T(-Q + \lambda_{\max}(PLL^TP)I)e$$
$$+\|D^TPe\|^2(1 - \theta) + \alpha(x, \hat{x}) \quad \textbf{(3.44)}$$

Then, re-looking at the nonlinear terms (3.29)

$$\alpha(x, \hat{x}) = \left(f^T(x) - f^T(\hat{x})\right)Pe + e^TP\left(f(x) - f(\hat{x})\right)$$
$$= 2e^TP\left(f(x) - f(\hat{x})\right) \quad \textbf{(3.45)}$$

since $\alpha(x, \hat{x})$ is a scalar.

We also know that

$$2e^TP\left(f(x) - f(\hat{x})\right) \leq 2\left\|e^TP\left(f(x) - f(\hat{x})\right)\right\|$$
$$\leq 2\|Pe\|\left\|\left(f(x) - f(\hat{x})\right)\right\| \quad \textbf{(3.46)}$$

and since we are also under the assumption that $f(x)$ is Lipschitz continuous

$$2\|Pe\|\left\|\left(f(x) - f(\hat{x})\right)\right\| \leq 2\gamma\|Pe\|\|x - \hat{x}\| \quad \textbf{(3.47)}$$
$$\leq 2\gamma\|Pe\|\|e\|$$
$$\leq 2\gamma Pe^Te \quad \textbf{(3.48)}$$

Then substituting this into (3.44) we get

$$\dot{V} \leq e^T(-Q + \lambda_{\max}(PLL^TP)I)e + \|D^TPe\|^2(1-\theta) + 2\gamma Pe^Te \tag{3.49}$$

$$\leq e^T(-Q + \lambda_{\max}(PLL^TP)I)e + (1-\theta)\|D^TP\|^2e^Te + 2\gamma Pe^Te$$

$$\leq e^T\big(-Q + \lambda_{\max}(PLL^TP)I + \big((1-\theta)\|D^TP\|^2\big)I + 2\gamma P\big)e \tag{3.50}$$

Q and θ are to be chosen such that the following inequality is satisfied

$$\big(-Q + \lambda_{\max}(PLL^TP)I + \big((1-\theta)\|D^TP\|^2\big)I + 2\gamma P\big) < 0 \tag{3.51}$$

Then it is proven that

$$\dot{V} < 0 \tag{3.52}$$

This concludes the proof.

*Remark***:** It can be noted that the above proof is very conservative, and that asymptotic error stability can be obtained without such strict conditions on Q and θ. Therefore the method above can be seen as sufficient for error stability but not necessary.

3.2. Local Nonlinear Observability

Before the above observer can be designed, or any other observer for that matter, the system first has to be observable. In other words, given the internal dynamics of the system, there must be sufficient information supplied by the output of that system, in order to successfully reconstruct all states. It has already been highlighted in the literature survey, the observability conditions for a given linear system for the construction of a Luenberger observer. In this section, the conditions for nonlinear observability will be illustrated by Theorem 3.2, and it can be noted that linear observability is merely a special case with respect to these conditions. Reference can be made to [76] where systems of the general form below are considered

$$\dot{x} = g(x) \tag{3.53}$$

$$y = h(x) \tag{3.54}$$

with $x \in \mathcal{R}^n$ and $y \in \mathcal{R}^p$. The observability mapping or observation space $Q(x)$ for an autonomous system can be given by

$$\begin{bmatrix} y \\ y' \\ y'' \\ \vdots \\ y^{(n-1)} \end{bmatrix} = \begin{bmatrix} L_g^0 h(x) \\ L_g^1 h(x) \\ L_g^2 h(x) \\ \vdots \\ L_g^{n-1} h(x) \end{bmatrix} = Q(x) \tag{3.55}$$

$Q(x)$ is a set of nonlinear equations that illustrates how to determine the states *x* from $(y, y', y'', \ldots)$. It can be noted that for a given nonlinear system, the number of time derivatives are not fixed. In the case of linear systems, the observability matrix *O* can be determined similarly to $Q(x)$, however it is obvious to realize that the matrix entries will all be constants as they are all first derivatives of linear equations.

Theorem 3.2

With reference to [76]*, the definition of the observability rank condition then states that given the system (3.53) - (3.54) and the observation space (3.55), if* $\left.\frac{\partial Q(x)}{\partial x}\right|_{x_0}$ *contains n linearly independent row vectors, then the system is observable at the point* x_0*. Obviously if* $\partial Q/\partial x$ *remains linearly independent for all x, then the system is observable in its entire domain. Again, refer to* [76] *for the proof of this theorem.*

3.3. Design Algorithm

In this section, the observer design process for the adaptive observer discussed in Section 3.1 is summarized in a step-by-step algorithm. Take note that the form of the system should either be written in a linearity-nonlinearity form, or it should be transformed into the nonlinear observable canonical form. Most of the steps were done by simulation software, however the calculation of the Jacobian was done by hand, and hardcoded into the software.

Algorithm 3.3

Step 1

Referring to Section 3.2, check that the system is observable. If it is observable proceed to *Step 2*, but if not, then the observer cannot be constructed using the existing structures of $g(x)$ and $h(x)$. In the unobservable case, if possible, change the structure of $h(x)$ in an attempt to achieve observability (i.e. utilize a different state for the measured output or add more states to the measured output).

Step 2

Perform a check to determine whether the nonlinearity $f(x)$ is Lipschitz continuous. If it is found that the nonlinearity is indeed Lipschitz continuous, then asymptotic convergence of the observer can be guaranteed. If it is found that the nonlinearity is NOT Lipschitz continuous, it might still be possible to construct a successful observer, however convergence of the observer cannot be guaranteed. (Refer to Appendix B for more details on Lipschitz continuity and Lipschitz constants).

Step 3

Using matrices *A* and *C*, determine an *L* matrix using linear pole placement methods and arbitrarily chosen stable poles for $(A - LC)$. Take note that the choice of poles will have an effect on the transient behaviour of the observer.

If it is found that matrices *A* and *C* are structured in such a way that matrix *L* cannot be determined as specified in *Step 3*, then proceed to *Step 3(a)*. Otherwise, proceed to *Step 4*.

Step 3(a)

Find the equilibrium points of the system $\dot{x} = Ax + f(x)$ by setting $\dot{x} = 0$

Step 3(b)

Perform a linearization on $f(x)$ by determining its Jacobian $F(x)$ at the system equilibrium points:

$$F(x) = \begin{bmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_n/\partial x_1 & \cdots & \partial f_n/\partial x_n \end{bmatrix}$$

Step 3(c)

Determine an L matrix using linear pole placement methods by using $\left(A + F(x)\right)$ at an arbitrary equilibrium point, the output matrix C, and arbitrarily chosen poles for $\left(\left(A + F(x)\right) - LC\right)$ that are stable. Once again take note that the choice of poles will have an effect on the transient behaviour of the observer.

If the L matrix cannot be determined, re-attempt by changing the poles of $\left(\left(A + F(x)\right) - LC\right)$ and/or choosing a different equilibrium point, and repeat *Step 3(c)*. Once the L matrix has been determined, proceed to *Step 4*.

Step 4

Using the L matrix that has been determined, calculate the 'error' matrix as

$$A_e = A - LC \tag{3.56}$$

Step 5

Choose a matrix Q such that it is symmetric and positive definite.

Step 6

Solve the Lyapunov equation to determine a symmetric, positive definite matrix P, using A_e and Q

$$A_e P + P A_e^T + Q = 0 \tag{3.57}$$

Step 7

Using matrices P and C, and equation (3.11), determine matrix D as

$$D = (CP^{-1})^T \tag{3.58}$$

Step 8

Referring to Section 3.1, and using the Lipschitz constant $\lambda_{\max}$ obtained in *Step 2*, ensure that a value for θ can be chosen as to satisfy equation (3.51), which will guarantee asymptotic convergence of the observer.

This concludes the algorithm.

3.3.1 Illustrative Example 1

Consider the system

$$\left.\begin{aligned} \dot{x}_1 &= x_2 + x_2 x_1^2 \\ \dot{x}_2 &= -x_1 - x_1^3 \end{aligned}\right\} \; g(x)$$

$$y = x_1 \;\; \} \;\; h(x)$$

which can also be written in the form

$$\dot{x} = Ax + f(x) + Bu$$

$$y = Cx$$

as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} x_2 x_1^2 \\ -x_1^3 \end{bmatrix}$$

$$y = [1 \quad 0]x$$

Step 1

Determine observability (referring to Section 3.2).

$$Q(x) = \begin{bmatrix} L_g^0 h(x) \\ L_g^1 h(x) \\ L_g^2 h(x) \\ \vdots \\ L_g^{n-1} h(x) \end{bmatrix}$$

$$L_g^0 h(x) = h(x) = x_1$$

$$L_g^1 h(x) = \frac{\partial h}{\partial x} g = [1 \quad 0] \begin{bmatrix} x_2 + x_2 x_1^2 \\ -x_1 - x_1^3 \end{bmatrix} = x_2 + x_2 x_1^2$$

$$Q(x) = \begin{bmatrix} x_1 \\ x_2 + x_2 x_1^2 \end{bmatrix}$$

$$\frac{\partial Q}{\partial x} = \begin{bmatrix} 1 & 0 \\ 2x_1 x_2 & 1 + x_1^2 \end{bmatrix}$$

Noting that n = 2, and since $\frac{\partial Q}{\partial x}$ contains 2 linear independent row vectors, this system is observable.

Step 2

Check Lipschitz continuity (referring to Appendix B

Consider the nonlinearity $f(x)$, of $g(x)$

$$f(x) = \begin{bmatrix} x_2 x_1^2 \\ -x_1^3 \end{bmatrix}$$

where $-1.5 \leq x_1 \leq 1.5$ and $-1.5 \leq x_2 \leq 1.5$

$$\left|\frac{\partial f_1}{\partial x_1}\right| = |2x_1 x_2| = 4.5 = \gamma_{f_1 x_1}$$

$$\left|\frac{\partial f_1}{\partial x_2}\right| = |x_1^2| = 2.25 = \gamma_{f_1 x_2}$$

$$\left|\frac{\partial f_2}{\partial x_1}\right| = |-3x_1^2| = 6.75 = \gamma_{f_2 x_1}$$

$$\left|\frac{\partial f_2}{\partial x_2}\right| = |0| = 0 = \gamma_{f_2 x_2}$$

We can already conclude that the nonlinearity $f(x)$ is Lipschitz continuous since all the derivatives exist. Then, since $\gamma_{f_1 x_1} > \gamma_{f_1 x_2}$ and $\gamma_{f_2 x_1} > \gamma_{f_2 x_2}$, we can determine the Lipschitz constant as

$$\gamma = \sqrt{\left(\gamma_{f_1}\right)^2 + \left(\gamma_{f_2}\right)^2} = \sqrt{(4.5)^2 + (6.75)^2} = 8.1$$

Step 3

Determine the L matrix.

Consider matrices $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and the arbitrarily chosen poles of $A - L^T C$ of $\begin{bmatrix} -4 & -1 \end{bmatrix}$. These poles are chosen to achieve certain desired transient behaviour of the error term. An L matrix can then be determined through pole placement as $L = \begin{bmatrix} 5 & 3 \end{bmatrix}$.

Step 4

Calculate the 'error' matrix.

Using the L matrix that has been determined, the 'error' matrix is determined as

$$A_e = A - L^T C$$

$$A_e = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -5 & 1 \\ -4 & 0 \end{bmatrix}$$

Step 5

Choose a matrix Q such that it is symmetric and positive definite.

$$Q = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

Step 6

Solve the Lyapunov equation to determine a symmetric, positive definite matrix *P.*

$$A_e P + P A_e^T + Q = 0$$

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 4.5 \end{bmatrix}$$

Noting that

$$P = P^T$$

$$P > 0$$

Step 7

Determine matrix D using equation (3.11).

$$D = (CP^{-1})^T$$

$$D = \left(\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 4.5 \end{bmatrix}^{-1} \right)^T = \begin{bmatrix} 2.25 \\ -0.25 \end{bmatrix}$$

Step 8

Ensure that a value for θ exists as to satisfy the inequality

$$\left(-Q + \lambda_{\max}(PLL^TP)I + \left((1-\theta)\|D^TP\|^2\right)I + 2\gamma P\right) < 0$$

Choosing $\theta = 350$ will result in

$$\left(-Q + \lambda_{\max}(PLL^TP)I + \left((1-\theta)\|D^TP\|^2\right)I + 2\gamma P\right) = \begin{bmatrix} -72.9 & 8.1 \\ 8.1 & -8.1 \end{bmatrix}$$

Note that this resulting matrix is symmetric (due to Q being symmetric) and hence negative eigenvalues of this matrix will mean that is it indeed negative definite. The eigenvalues of $\begin{bmatrix} -72.9 & 8.1 \\ 8.1 & -8.1 \end{bmatrix}$ are $[-73.8 \quad -7.1] < 0$, clearly showing that the inequality is satisfied and that observer convergence can be guaranteed.

3.3.2 Illustrative Example 2

With reference to [77], consider the Lorenz system

$$\left.\begin{aligned}\dot{x}_1 &= -\sigma x_1 + \sigma x_2\\ \dot{x}_2 &= \rho x_1 - x_2 - x_1 x_3\\ \dot{x}_3 &= -\beta x_3 + x_1 x_2\end{aligned}\right\}\ g(x)$$

$$y = x_1 \ \}\ h(x)$$

Note that for this example the constant parameters were chosen as $\sigma = 10, \rho = 28, \beta = \frac{8}{3}$.

Step 1

Determine observability (referring to Section 3.2).

$$Q(x) = \begin{bmatrix} L_g^0 h(x) \\ L_g^1 h(x) \\ L_g^2 h(x) \\ \vdots \\ L_g^{n-1} h(x) \end{bmatrix}$$

$$L_g^0 h(x) = h(x) = x_1$$

$$L_g^1 h(x) = \frac{\partial (L_g^0 h)}{\partial x} g = [1 \quad 0 \quad 0] \begin{bmatrix} -\sigma x_1 + \sigma x_2 \\ \rho x_1 - x_2 - x_1 x_3 \\ -\beta x_3 + x_1 x_2 \end{bmatrix} = -\sigma x_1 + \sigma x_2$$

$$\begin{aligned} L_g^2 h(x) = \frac{\partial (L_g^1 h)}{\partial x} g &= [-\sigma \quad \sigma \quad 0] \begin{bmatrix} -\sigma x_1 + \sigma x_2 \\ \rho x_1 - x_2 - x_1 x_3 \\ -\beta x_3 + x_1 x_2 \end{bmatrix} \\ &= -\sigma f_1(x) + \sigma f_2(x) \end{aligned}$$

$$Q(x) = \begin{bmatrix} x_1 \\ -\sigma x_1 + \sigma x_2 \\ -\sigma(-\sigma x_1 + \sigma x_2) + \sigma(\rho x_1 - x_2 - x_1 x_3) \end{bmatrix}$$

$$\frac{\partial Q}{\partial x} = \begin{bmatrix} 1 & 0 & 0 \\ -\sigma & \sigma & 0 \\ \sigma(\sigma + \rho - x_3) & \sigma(-\sigma - 1) & -\sigma x_1 \end{bmatrix}$$

Noting that n = 3, and since $\frac{\partial Q}{\partial x}$ contains 3 linear independent row vectors, this system is observable.

Step 2

Check Lipschitz continuity (referring to Appendix B

Consider the nonlinearity $f(x)$, of $g(x)$

$$f(x) = \begin{bmatrix} 0 \\ -x_1 x_3 \\ x_1 x_2 \end{bmatrix}$$

$$\begin{gathered} -17 \leq x_1 \leq 20 \\ -23 \leq x_2 \leq 24 \\ 6 \leq x_3 \leq 45 \end{gathered}$$

$$\left|\frac{\partial f_1}{\partial x_1}\right| = |0| = 0 = \gamma_{f_1 x_1}$$

$$\left|\frac{\partial f_1}{\partial x_2}\right| = |0| = 0 = \gamma_{f_1 x_2}$$

$$\left|\frac{\partial f_1}{\partial x_3}\right| = |0| = 0 = \gamma_{f_1 x_3}$$

$$\left|\frac{\partial f_2}{\partial x_1}\right| = |-x_3| = 45 = \gamma_{f_2 x_1}$$

$$\left|\frac{\partial f_2}{\partial x_2}\right| = |0| = 0 = \gamma_{f_2 x_2}$$

$$\left|\frac{\partial f_2}{\partial x_3}\right| = |-x_1| = 20 = \gamma_{f_2 x_3}$$

$$\left|\frac{\partial f_3}{\partial x_1}\right| = |x_2| = 24 = \gamma_{f_3 x_1}$$

$$\left|\frac{\partial f_3}{\partial x_2}\right| = |x_1| = 20 = \gamma_{f_3 x_2}$$

$$\left|\frac{\partial f_3}{\partial x_3}\right| = |0| = 0 = \gamma_{f_3 x_3}$$

We can already conclude that the nonlinearity $f(x)$ is Lipschitz continuous since all the derivatives exist. Then, taking $\max(\gamma_{f_1})$, $\max(\gamma_{f_2})$, $\max(\gamma_{f_3})$, we can determine the Lipschitz constant as

$$\gamma = \sqrt{(\gamma_{f_1})^2 + (\gamma_{f_2})^2 + (\gamma_{f_3})^2} = \sqrt{(0)^2 + (45)^2 + (24)^2} = 51$$

Step 3

Determine the L matrix.

Considering the structure of the matrices $A = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho & -1 & 0 \\ 0 & 0 & -\beta \end{bmatrix}$ and $C = [1 \quad 0 \quad 0]$, we can conclude that an L matrix cannot be determined through pole placement at this point. Consequently, proceed to *Step 3(a)*.

Step 3(a)

Find the equilibrium points of the system.

Set $\dot{x} = 0$

$$\begin{aligned} 0 &= -\sigma x_1 + \sigma x_2 \\ 0 &= \rho x_1 - x_2 - x_1 x_3 \\ 0 &= -\beta x_3 + x_1 x_2 \end{aligned}$$

Then solving for x_1, x_2, x_3 we get equilibrium points of $(0, 0, 0)$, $(\sqrt{72}, \sqrt{72}, 27)$ and $(-\sqrt{72}, -\sqrt{72}, 27)$.

Step 3(b)

Perform a linearization on the nonlinearity $f(x)$.

$$\mathrm{F}(x) = \begin{bmatrix} \partial f_1/dx_1 & \partial f_1/dx_2 & \partial f_1/dx_3 \\ \partial f_2/dx_1 & \partial f_2/dx_2 & \partial f_2/dx_3 \\ \partial f_3/dx_1 & \partial f_3/dx_2 & \partial f_3/dx_3 \end{bmatrix}$$

$$\mathrm{F}(x) = \begin{bmatrix} 0 & 0 & 0 \\ -x_3 & 0 & -x_1 \\ x_2 & x_1 & 0 \end{bmatrix}$$

Step 3(c)

Determine an L matrix through pole placement by using the matrix $\left(\left(A + \mathrm{F}(x)\right) - LC\right)$ and arbitrary poles that are stable.

Since the equilibrium point $(0, 0, 0)$ will result in a matrix structure of $\left(A + \mathrm{F}(x)\right)$ where an L matrix cannot be determined through pole placement, the equilibrium point $\left(-\sqrt{72}, -\sqrt{72}, 27\right)$ is chosen. The poles are then chosen as $(-40, -10, -8)$, and consequently

$$L = [44.3 \quad 53.6 \quad -35.1]^T$$

is determined.

Step 4

Calculate the ‘error’ matrix.

Using the L matrix that has been determined, the ‘error’ matrix is determined as

$$A_e = A - LC$$

$$A_e = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho & -1 & 0 \\ 0 & 0 & -\beta \end{bmatrix} - \begin{bmatrix} 44.3 \\ 53.6 \\ -35.1 \end{bmatrix} [1 \quad 0 \quad 0]$$

$$= \begin{bmatrix} -54.3 & 10 & 0 \\ -25.6 & -1 & 0 \\ 35.1 & 0 & -8/3 \end{bmatrix}$$

Step 5

Choose a matrix Q such that it is symmetric and positive definite.

$$\mathrm{Q} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

Step 6

Solve the Lyapunov equation to determine a symmetric, positive definite matrix *P.*

$$A_e P + P A_e^T + Q = 0$$

$$P = \begin{bmatrix} 0.048 & 0.060 & 0.059 \\ 0.060 & 0.456 & 0.167 \\ 0.059 & 0.167 & 1.524 \end{bmatrix}$$

Noting that

$$P = P^T$$

$$P > 0$$

Step 7

Determine matrix D using equation (3.11).

$$D = (CP^{-1})^T$$

$$D = \left(\begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0.048 & 0.060 & 0.059 \\ 0.060 & 0.456 & 0.167 \\ 0.059 & 0.167 & 1.524 \end{bmatrix}^{-1} \right)^T = \begin{bmatrix} 25.63 \\ -3.15 \\ -0.64 \end{bmatrix}$$

Step 8

Ensure that a value for θ exists as to satisfy the inequality

$$\left(-Q + \lambda_{\max}(PLL^T P)I + \left((1-\theta)\|D^T P\|^2\right)I + 2\gamma P\right) < 0$$

Choosing $\theta = 2450$ will result in

$$\left(-Q + \lambda_{\max}(PLL^T P)I + \left((1-\theta)\|D^T P\|^2\right)I + 2\gamma P\right) = \begin{bmatrix} -223 & 6 & 6 \\ 6 & -181 & 17 \\ 6 & 17 & -72 \end{bmatrix}$$

where the eigenvalues are $[-223 \quad -183 \quad -69] < 0$, clearly showing that the inequality is satisfied and that observer convergence can be guaranteed. Note that the value of θ has to be quite large in order to *guarantee* stability in this case, however smaller values can be used where it will be noticed that the observer still converges. This result can be expected since conservative measures have been used throughout the proof in order to prove stability of the error terms.

Note that the simulation results for this particular example can be seen in Chapter 4.

3.4. Approximation of Transient Behaviour

An Introduction to the Error Transient Behaviour

In many practical instances, transient behaviour can have a significant effect on how the system as a whole performs. For example, if the transient behaviour of a reconstructed state has an amplitude too large, then this could cause controllers that are using these values, to become saturated too quickly and cause undesired control actions. Also, in high speed control systems, if the observer error does not reach zero fast enough and the actual states are constantly changing, then it becomes meaningless to estimate such states. Consequently the transient behaviour of the above method was analyzed and a simple process proposed in order to gauge an approximation of the upper bound (or largest value in seconds) of the transient time.

An Introduction to the Error Transient Approximation Algorithm

The idea behind Algorithm 3.4 below, was to find a linear ‘version’ of the error equation such that the poles of this new linearized system would provide the designer with an indication of the transient time of the observer. *It must be emphasized that this Algorithm is purely **heuristic**, and no mathematical basis has been formulated to substantiate the steps of the algorithm*. It is to be applied to gain a rough idea of a possible “worst-case” for the transient time, however, it is not definitive, and there is no guarantee that this procedure will indeed provide the worst possible case (as is demonstrated further on).

If, after working through the algorithm, it is found that the transient time is unacceptably large for this error approximation, one can repeat step 3 of the design algorithm (Algorithm 3.3), and by choosing faster poles, the result will be a new matrix *D*, directly influencing the transient behaviour.

Algorithm 3.4

Consider the error dynamics of the adaptive observer given by equation (3.16)

$$\dot{e} = Ae + \left(f(x) - f(\hat{x}) \right) - \frac{1}{2}\hat{\theta}DCe$$

noting that this equation is nonlinear.

Step 1

Replace the matrix entries of $f(x)$ with terms that include only the estimated states $\hat{x}_1, \ldots, \hat{x}_n$, and the error terms $e_1, \ldots, e_n$, by substituting in the equations

$$x_1 = \hat{x}_1 + e_1$$
$$\vdots$$
$$x_n = \hat{x}_n + e_n$$

thus resulting in

$$\left(f(x) - f(\hat{x}) \right) \approx \left(f(\hat{x}, e) \right) \quad \textbf{(3.59)}$$

Step 2

Proceed to construct a new matrix $\bar{f}(\hat{x}, e) \in \mathcal{R}^{n \times n}$ such that

$$\left(\bar{f}(\hat{x}, e) \right)(e) \approx \left(f(\hat{x}, e) \right) \quad \textbf{(3.60)}$$

where $e = [e_1 \quad \cdots \quad e_n]^T$. Assume that any error terms of order 2 and higher (i.e. $e_1 e_2$, e_1^3 etc.) are close to zero and can be ignored.

Step 3

Re-write (3.16) using $\bar{f}(\hat{x}, e)$

$$\dot{e} = \left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)(e) \quad \textbf{(3.61)}$$

Step 4

With the continued objective of ensuring that $\left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)$ is a stable matrix, an iterative process should be taken for the unknown variables that will guarantee stability and produce upper and lower bounds of the transient times of $\left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)$. This iterative process is started in *Step 4* and continues through *Step 5*.

The process is started by first substituting the initial conditions of $\hat{x}$ and $\hat{\theta}$ into (3.61), since these values are known *a priori*, however leaving the known (measured) components of $\hat{x}$ as a free variable (i.e. if x_1 is part of the measured set of states, then don't substitute the initial condition for $\hat{x}_1$). Then estimates for e can be substituted (also using the initial condition values).

Step 5

Then, considering the bounded domain of the free variable, substitute the lower bound value into the matrix $\left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)$, noting whether the resulting matrix is stable or not. Increment the value of the free variable and again substitute into the matrix, noting again whether it is stable or not. Proceed until the upper bound of the free variable is reached. The designer can expect numerous feasible values for the free variables that will lead to a stable matrix, or none at all that result in a stable matrix (it must be stressed to the reader that this is a heuristic approach and there is no mathematical proof or guarantee that such a matrix will exist, and hence that an approximation using these steps will be possible).

Step 6

It is assumed that *Step 5* has resulted in numerous viable values of the free variable for which the matrix $\left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)$ is stable. Use those values of the variable that correspond to the largest transient times (upper bound) which will provide an approximation of the "worst-case" for the error transient response.

This concludes the algorithm.

3.4.1 Illustrative Example 1

Consider the system

$$\begin{aligned}\dot{x}_1 &= x_2 + x_2 x_1^2 \\ \dot{x}_2 &= -x_1 - x_1^3\end{aligned}$$

$$y = x_1$$

Designing an observer of the form (3.7), the resulting error system is

$$\dot{e} = Ae + \left(f(x) - f(\hat{x})\right) - \frac{1}{2}\hat{\theta}DCe$$

$$\dot{e} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} e + \left[\begin{bmatrix} x_2 x_1^2 \\ -x_1^3 \end{bmatrix} - \begin{bmatrix} \hat{x}_2 \hat{x}_1^2 \\ -\hat{x}_1^3 \end{bmatrix}\right] - 0.5\hat{\theta}\begin{bmatrix} 2.25 \\ -0.25 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} e$$

Step 1

Replace the matrix entries of $f(x)$ with terms

$$\begin{aligned}x_1 &= \hat{x}_1 + e_1 \\ &\vdots \\ x_n &= \hat{x}_n + e_n\end{aligned}$$

which results in

$$\left(f(x) - f(\hat{x})\right) = \left(f(\hat{x}, e)\right) = \left[\begin{bmatrix} (\hat{x}_2 + e_2)(\hat{x}_1 + e_1)^2 \\ -(\hat{x}_1 + e_1)^3 \end{bmatrix} - \begin{bmatrix} \hat{x}_2 \hat{x}_1^2 \\ -\hat{x}_1^3 \end{bmatrix}\right]$$

Step 2

Simplify $\left(f(\hat{x},e)\right)$ and proceed to construct a new matrix $\bar{f}(\hat{x},e) \in \mathcal{R}^{n\times n}$ such that

$$\left(\bar{f}(\hat{x},e)\right)(e) \equiv \left(f(\hat{x},e)\right)$$
$$\left(\bar{f}(\hat{x},e)\right)(e) = \begin{bmatrix} 2\hat{x}_1\hat{x}_2 + \hat{x}_2e_1 + 2\hat{x}_1e_2 & \hat{x}_1^2 \\ -3\hat{x}_1^2 - 3\hat{x}_1e_1 & 0 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

Step 3

Re-write the error dynamics equation using $\bar{f}(\hat{x},e)$.

$$\dot{e} = \left(A + \bar{f}(\hat{x},e) - \frac{1}{2}\hat{\theta}DC\right)e$$
$$\dot{e} = \left[\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 2\hat{x}_1\hat{x}_2 + \hat{x}_2e_1 + 2\hat{x}_1e_2 & \hat{x}_1^2 \\ -3\hat{x}_1^2 - 3\hat{x}_1e_1 & 0 \end{bmatrix} - 0.5\hat{\theta}\begin{bmatrix} 2.25 \\ -0.25 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix}\right]e$$

Step 4

Substitute the initial conditions of $\hat{x}$ and $\hat{\theta}$, and estimates for e.

$$\hat{x}_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, e_0 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \ \hat{\theta}_0 = 0$$

and since the output $y = x_1$, replace $\hat{x}_1$ with the free variable $\bar{x}_1$. The error dynamics then becomes

$$\dot{e} = \left[\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 2\bar{x}_1(-1) + (-1)(2) + 2\bar{x}_1(2) & \bar{x}_1^2 \\ -3\bar{x}_1^2 - 3\bar{x}_1(2) & 0 \end{bmatrix} - 0.5(0)\begin{bmatrix} 2.25 \\ -0.25 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix}\right]e$$

Step 5

Substitute values for $\hat{x}_1$ and iterate across the entire range of $\hat{x}_1$. Note and extract all values of $\hat{x}_1$ that produce a stable linear system. In this case the range of $\hat{x}_1$ is

$$-1.5 \leq \hat{x}_1 \leq 1.5$$

and the values for $\hat{x}_1$ and corresponding stable eigenvalues

$\bar{x}_1$	$Re(\lambda_1)$	$Re(\lambda_2)$
0.1	-0.9	-0.9
0.2	-0.8	-0.8
0.3	-0.7	-0.7
0.4	-0.6	-0.6
0.5	-0.5	-0.5
0.6	-0.4	-0.4
0.7	-0.3	-0.3
0.8	-0.2	-0.2
0.9	-0.1	-0.1

Table 3. 1: Stable Modes of the Approximate Error Dynamics

Note that in this case the increment between iterations is 0.1. This value is chosen at the discretion of the designer.

Step 6

The values of $\bar{x}_1$ that will provide the upper and lower bounds of the approximated error transient times are chosen as $\bar{x}_1 = 0.1$ and $\bar{x}_1 = 0.9$. The approximated error dynamics systems (upper and lower bounds) then respectively become

$$\dot{e} = \left[\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 2(0.9)(-1) + (-1)(2) + 2(0.9)(2) & (0.9)^2 \\ -3(0.9)^2 - 3(0.9)(2) & 0 \end{bmatrix}\right] e$$

$$\dot{e} = \begin{bmatrix} -0.2 & 1.81 \\ -8.8 & 0 \end{bmatrix} e$$

and

$$\dot{e} = \left[\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 2(0.1)(-1) + (-1)(2) + 2(0.1)(2) & (0.1)^2 \\ -3(0.1)^2 - 3(0.1)(2) & 0 \end{bmatrix}\right] e$$

$$\dot{e} = \begin{bmatrix} -1.8 & 1.0 \\ -1.6 & 0 \end{bmatrix} e$$

with eigenvalues given in 3. 1, corresponding to those values of $\bar{\boldsymbol{x}}_1 = \mathbf{0.1}$ and $\bar{\boldsymbol{x}}_1 = \mathbf{0.9}$.

METHODOLOGY CONCLUDED

This chapter has revealed and proved that, based on an existing technique for chaotic synchronization, a method can be formulated that successfully constructs a nonlinear observer, given that the system can be written in the form (3.3) - (3.4).

Although Chapter 2 has clearly revealed that there are many applicable techniques that can successfully estimate the unknown states of a nonlinear system, the technique developed in Chapter 3 was considered to be most appropriate in nonlinear systems applications. As previously stated, the technique was developed based on the existing chaotic synchronization work of [11], which proved to be applicable to numerous and different nonlinear (chaotic) systems. This method was then further developed in Chapter 3 in order to design an observer where only **partial** state information is available. It was also then shown to address estimation problems of nonlinear systems that can be written in the general form (3.3) - (3.4), which covers a large class of nonlinear systems that can be separated into their respective linear and nonlinear parts. These aspects make this observer technique highly applicable to a wide range of nonlinear systems, provided that the conditions and assumptions, as listed in Section 3.1, are considered.

A heuristic algorithm was then developed that approximates the transient times of the error response, based on a linearization of the error dynamics, which can then give the designer an idea upfront of how quickly the observer can potentially respond.

4. APPLICATION

The method discussed in Chapter 3 was successfully applied to numerous nonlinear systems, and the results of these numerical simulations are given below. The simulations were executed using MATLAB/Simulink software and the code can be viewed in Appendix C.

4.1. Lorenz Attractor

A state observer, as described in Chapter 3, was designed to estimate the unknown states of the Lorenz attractor, with limited state information given in the output. The system dynamics of the Lorenz attractor (with reference to [77]) are given by

$$\dot{x}_1 = -\sigma x_1 + \sigma x_2 \tag{4.1}$$

$$\dot{x}_2 = \rho x_1 - x_2 - x_1 x_3 \tag{4.2}$$

$$\dot{x}_3 = -\beta x_3 + x_1 x_2 \tag{4.3}$$

$$y = x_1 \tag{4.4}$$

where the system behaves chaotically for parameter values of $\sigma = 10$, $\rho = 28$, and $\beta = \frac{8}{3}$.

Designing the Observer

The observer for the Lorenz system was designed according to the algorithm discussed in Section 3.3, and the numerical simulation results are illustrated and discussed below. For the step-by-step details of this particular application, the reader is referred to Illustrative Example 2 (Section 3.3.2). By applying the design algorithm of Chapter 3, the state observer was designed as

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}) + \frac{1}{2}\hat{\theta}D(y - C\hat{x})$$

where

$$A = \begin{bmatrix} 10 & 10 & 0 \\ 28 & -1 & 0 \\ 0 & 0 & \frac{-8}{3} \end{bmatrix},\ f(\hat{x}) = [0 \quad -\hat{x}_1\hat{x}_3 \quad \hat{x}_1\hat{x}_2]^T,\ C = [1 \quad 0 \quad 0]$$

and recall that the gain matrix D was determined as $D = [25.63 \quad -3.15 \quad -0.64]^T$, with the adaptive parameter $\hat{\theta}$ given by equation (3.9).

Results

These particular simulations were done with initial conditions of $x_0 = [1 \quad 1 \quad 1]^T$ and $\hat{x}_0 = [5 \quad 5 \quad 5]^T$. The figures below clearly confirm the expected outcome. After certain transient behaviour, the estimated states converge to the actual state values and the error becomes stable (or tends to zero). In this case stability is reached around the 2 second mark for both estimated states. It can also be noted that for this example there is no noticeable, radical transient behaviour before convergence occurs

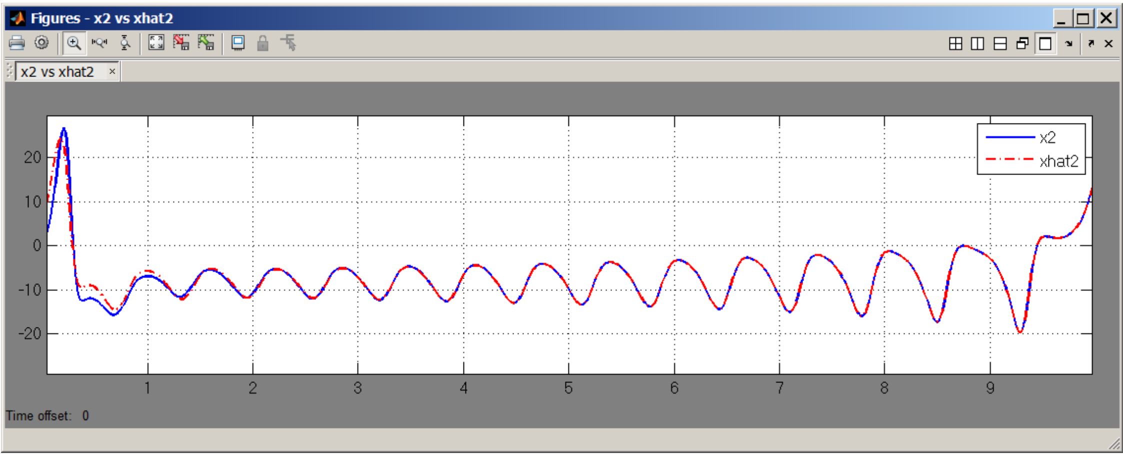

Figure 4.1: Estimated state $\hat{x}_2$ vs actual state x_2

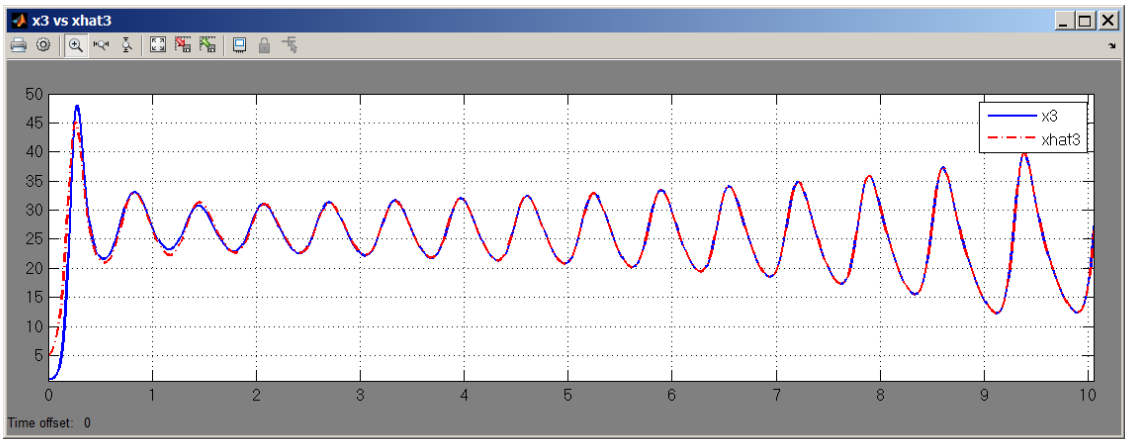

Figure 4.2: Estimated state $\hat{x}_3$ vs actual state x_3

Quantifying the Error

In order to gain a more detailed understanding of how accurate the estimations were, it was required to do more than just inspect the error response visually. As a result the mean squared error (MSE) was determined as per

$$\mathrm{MSE}_{x_1} = \frac{1}{N}\sum_{i=1}^{N}\left(x_{1_i} - \hat{x}_{1_i}\right)^2$$

Quantifying the Error with Varying Disturbance

Table 4. 1 below shows how the MSE for each state changes when the initial conditions (or magnitude of perturbations) change. The adaptive gain l from the equation $\dot{\hat{\theta}} = l\|y - C\hat{x}\|^2$ was kept constant with a value of 1 for all the simulations below.

$e_{1_0} = (x_1 - \hat{x}_1)$	MSE_{x_1}
-4 = (1 – 5)	0.07
-14 = (1 – 15)	0.12
-49 = (1 – 50)	0.27

$e_{2_0} = (x_2 - \hat{x}_2)$	MSE_{x_2}
-4 = (1 – 5)	1.02
-14 = (1 – 15)	1.79
-49 = (1 – 50)	7.78

$e_{3_0} = (x_3 - \hat{x}_3)$	MSE_{x_3}
-4 = (1 – 5)	1.62
-14 = (1 – 15)	3.81
-49 = (1 – 50)	22.3

Table 4. 1: Mean Squared Error with Varying Initial Conditions (Lorenz)

Quantifying the Error with Varying Adaptive Gain

Table 4. 2 below shows how the MSE for each state changes when the parameter of the adaptive equation changes i.e. the adaptive gain l in the equation $\dot{\hat{\theta}} = l\|y - C\hat{x}\|^2$. All other parameters remained constant, and the same initial conditions of $e_0 = (x_0 - \hat{x}_0) = [-4 \quad -4 \quad -4]^T$ were used throughout.

l	$\mathbf{MSE}_{x_1}$	$\mathbf{MSE}_{x_2}$	$\mathbf{MSE}_{x_3}$
5	0.02	0.58	0.76
10	0.01	0.43	0.54
30	0.006	0.28	0.33

Table 4. 2: Mean Squared Error with Varying Adaptive Gain (Lorenz)

Quantifying the Error with Varying Q Matrices

Table 4. 3 below shows how the MSE for each state changes when the constant parameters of the Q matrix are changed. Recall from the design procedure that this parameter must be freely chosen by the designer for a stability guarantee, however, the choice of this matrix indeed has an effect on the performance of the observer.

Q	$\mathbf{MSE}_{x_1}$	$\mathbf{MSE}_{x_2}$	$\mathbf{MSE}_{x_3}$
$0.1I$	0.0047	0.2488	0.3002
I	0.0253	0.6336	0.8462
$2I$	0.0422	0.8225	1.188
$4I$	0.069	1.0256	1.6271

Table 4. 3: Mean Squared Error with Varying Q matrix (Lorenz)

Calculating the MSE has provided interesting results concerning the performance of the designed observer and some important conclusions can be highlighted:

- The overall error does not remain the same irrespective of the magnitude of the initial conditions or perturbations, and hence the overall performance of the observer will depend on initial estimates of the unknown states
- The adaptive parameter gain *l* has a significant effect on the overall error, and consequently provides the designer with a useful parameter that can be tweaked in order to optimize the observer performance
- From the data it is clear that the choice of Q matrix not only determines whether stability can be guaranteed, but also influences the performance of the observer, however its effect is not as significant as the other parameters. In this case, the performance decreases as the parameters in the Q matrix are increased in magnitude. This will be something that the designer must consider when constructing this observer, with the goal of keeping the Q matrix as small as is possible.

Approximating the Error Transients

In order to model the transient behaviour of the error, a linear approximation was done based upon the algorithm described in Section 3.4. The details of each step have been omitted in this case, however, the final dynamical equation of the linear approximation was determined as

$$\dot{e} = \left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)e$$

$$\dot{e} = \left[\begin{bmatrix} 10 & 10 & 0 \\ 28 & -1 & 0 \\ 0 & 0 & \frac{-8}{3} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ \hat{x}_3 & 0 & x_1 \\ \hat{x}_2 & x_1 & 0 \end{bmatrix} - \frac{1}{2}\hat{\theta}\begin{bmatrix} 25.6 \\ -3.1 \\ -0.6 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}\right]e$$

where the initial condition values for $\hat{x}_2$ and $\hat{x}_3$ were used (recalling that $\hat{x}_0 = [5 \quad 5 \quad 5]^T$), the adaptive parameter was chosen as $\hat{\theta} = 0$, and the value for x_1 was iterated across its range in order to get the upper and lower bounds of the error transients. Figure 4.3 and Figure 4.4, respectively show the responses of the approximated and actual error signals for the estimated states, $\hat{x}_2$ and $\hat{x}_3$.

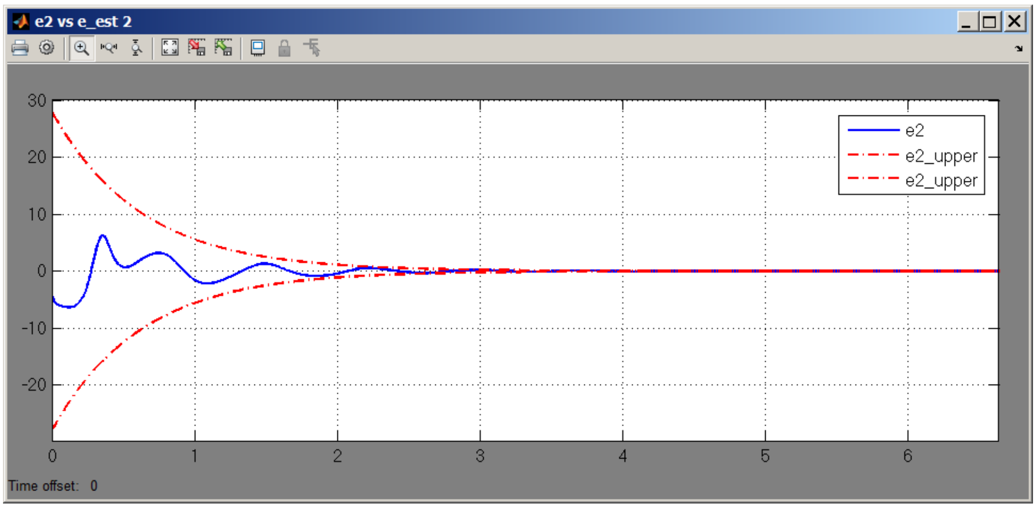

Figure 4.3: Actual and approximated error, e_2

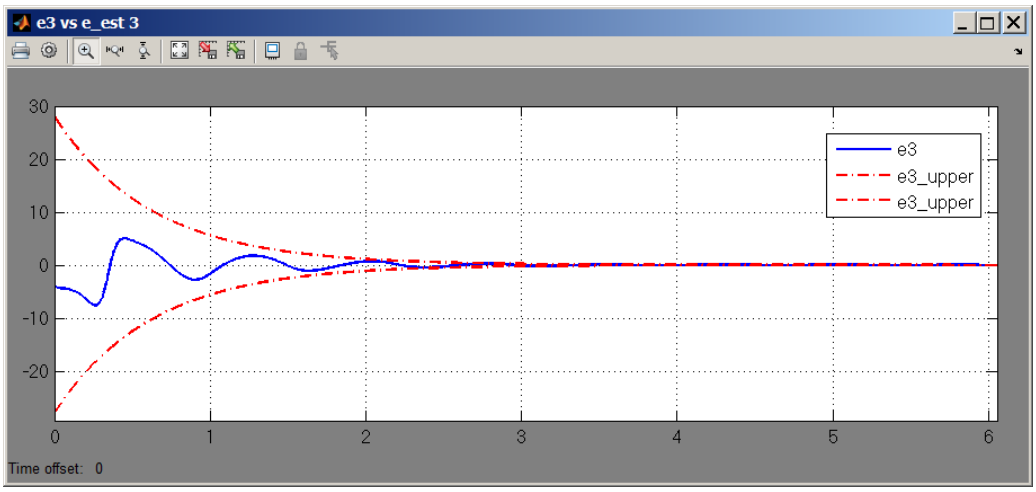

Figure 4.4: Actual and approximated error, e_3

With reference to Figure 4.3, the error bounds of e2_upper represent the **approximated "worst case"** transient response of the error for x_2. Similarly with reference to Figure 4.4, the error bounds of e3_upper represent the **approximated "worst case"** transient responses of the error for x_3.

The settling transient time of the error approximation was then compared to the settling transient time of the actual error in order to determine the degree of accuracy of the approximation. This was done by determining the time taken for the error terms to reach e^{-1} of the peak error, where $e^{-1} = 0.367$. These results can be seen in Table 4. 4 where the actual error decay and the approximate error decay have been compared. The time duration was measured, in seconds, from the peak amplitude of the disturbance up to the point where it decays to below 36.7% of the peak value.

The results in Table 4. 4 indicate that the error approximation, in this particular case, was indeed an upper bound of the actual error for both estimated states, since the settling time is longer for the approximation.

Time Taken to reach e^{-1} of the Peak Error			
Error Approximation for e_2 (sec)	0.62 (0.96-0.34)	**Actual Error e_2 (sec)**	0.45 (0.79-0.34)
Error Approximation for e_3 (sec)	0.65 (0.8-0.15)	**Actual Error e_3 (sec)**	0.08 (0.33-0.25)

Table 4. 4: Time to achieve e-1 of Actual Error and Error Approximation, e_2 and e_3

4.2. Rössler Attractor

A state observer was also designed to estimate the unknown states of the Rössler attractor, once again with limited state information provided in the output. The system dynamics of the Rössler attractor (with reference to [78]) are given by

$$\dot{x}_1 = -x_2 - x_3 \tag{4.5}$$

$$\dot{x}_2 = x_1 + a x_2 \tag{4.6}$$

$$\dot{x}_3 = b + x_3(x_1 - c) \tag{4.7}$$

$$y = x_2 \tag{4.8}$$

where the system behaves chaotically for parameter values of $a = 0.2$, $b = 0.2$, and $c = 5$.

Designing the Observer

The observer for the Rössler system was designed according to the algorithm discussed in Section 3.3, and the numerical simulation results are illustrated and discussed below.

Following Step 1 of the algorithm, the observability of the system was first checked

$$Q(x) = \begin{bmatrix} L_g^0 h(x) \\ L_g^1 h(x) \\ L_g^2 h(x) \\ \vdots \\ L_g^{n-1} h(x) \end{bmatrix}$$

$$L_g^0 h(x) = h(x) = x_2$$

$$L_g^1 h(x) = \frac{\partial (L_g^0 h)}{\partial x} g = [0 \quad 1 \quad 0][g] = x_1 + a x_2$$

$$L_g^2 h(x) = \frac{\partial (L_g^1 h)}{\partial x} g = [1 \quad a \quad 0] g = (-x_2 - x_3) + a(x_1 + a x_2)$$

$$Q(x) = \begin{bmatrix} x_2 \\ x_1 + a x_2 \\ (-x_2 - x_3) + a(x_1 + a x_2) \end{bmatrix}$$

$$\frac{\partial Q}{\partial x} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & a & 0 \\ a & -1 + a^2 & -1 \end{bmatrix}$$

Noting that $n = 3$, and since $\frac{\partial Q}{\partial x}$ contains 3 linear independent row vectors, this system is observable.

The Lipschitz continuity was then checked as per Step 2. Considering the nonlinearity of the Rössler attractor $f(x) = [0 \quad 0 \quad 0.2 + x_1 x_3]^T$ and the domain of x, the theory in Appendix B

$$f(x) = \begin{bmatrix} 0 \\ 0 \\ 0.2 + x_1 x_3 \end{bmatrix}$$

$$-7.7 \le x_1 \le 9.8$$
$$-9.2 \le x_2 \le 6.7$$
$$0 \le x_3 \le 15.5$$

$$\left|\frac{\partial f_1}{\partial x_1}\right| = |0| = 0 = \gamma_{f_1 x_1}$$

$$\left|\frac{\partial f_1}{\partial x_2}\right| = |0| = 0 = \gamma_{f_1 x_2}$$

$$\left|\frac{\partial f_1}{\partial x_3}\right| = |0| = 0 = \gamma_{f_1 x_3}$$

$$\left|\frac{\partial f_2}{\partial x_1}\right| = |0| = 0 = \gamma_{f_2 x_1}$$

$$\left|\frac{\partial f_2}{\partial x_2}\right| = |0| = 0 = \gamma_{f_2 x_2}$$

$$\left|\frac{\partial f_2}{\partial x_3}\right| = |0| = 0 = \gamma_{f_2 x_3}$$

$$\left|\frac{\partial f_3}{\partial x_1}\right| = |x_3| = 15.5 = \gamma_{f_3 x_1}$$

$$\left|\frac{\partial f_3}{\partial x_2}\right| = |0| = 0 = \gamma_{f_3 x_2}$$

$$\left|\frac{\partial f_3}{\partial x_3}\right| = |x_1| = 9.8 = \gamma_{f_3 x_3}$$

It can clearly be conclude that the nonlinearity $f(x)$ is Lipschitz continuous since all the derivatives exist. Then, taking $\max(\gamma_{f_1})$, $\max(\gamma_{f_2})$, $\max(\gamma_{f_3})$, the Lipschitz constant is determined as

$$\gamma = \sqrt{(\gamma_{f_1})^2 + (\gamma_{f_2})^2 + (\gamma_{f_3})^2} = \sqrt{(0)^2 + (15.5)^2 + (9.8)^2} = 18.3$$

Although not shown in the dissertation, Steps 3 to 8 of the algorithm were followed in the sequence as discussed in Section 3.3, and the observer was consequently designed accordingly

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}) + \frac{1}{2}\hat{\theta}D(y - C\hat{x})$$

where $A = \begin{bmatrix} 0 & -1 & -1 \\ 1 & 0.2 & 0 \\ 0 & 0 & -5 \end{bmatrix}$, $f(\hat{x}) = [0 \quad 0 \quad 0.2 + \hat{x}_1\hat{x}_3]^T$, $C = [0 \quad 1 \quad 0]$, and D calculated as $D = [-0.0002 \quad 0.008 \quad -0.0003]^T$ .The adaptive parameter $\hat{\theta}$ is given by (3.9).

Results

This particular simulation was done using initial conditions of $x_0 = [1 \quad 1 \quad 1]^T$ and $\hat{x}_0 = [5 \quad 5 \quad 5]^T$. The figures below show how the error becomes stable and the estimated states converge to the actual state values, just as expected. From inspection is can be noted that there also is no erratic transient behaviour before stability is reached. The time taken until both estimated states reach the actual state values is about 3 seconds in both cases.

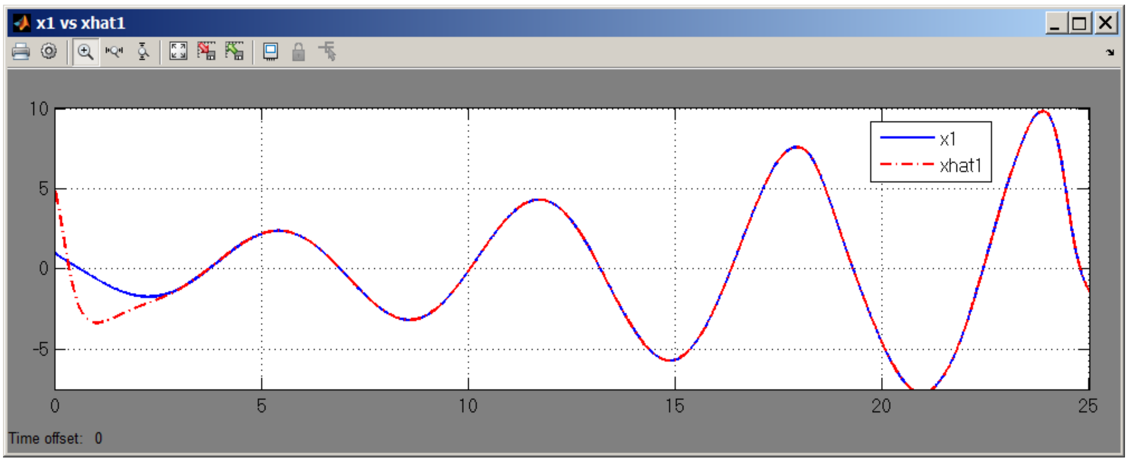

Figure 4.5: Estimated state $\hat{x}_1$ vs actual state x_1

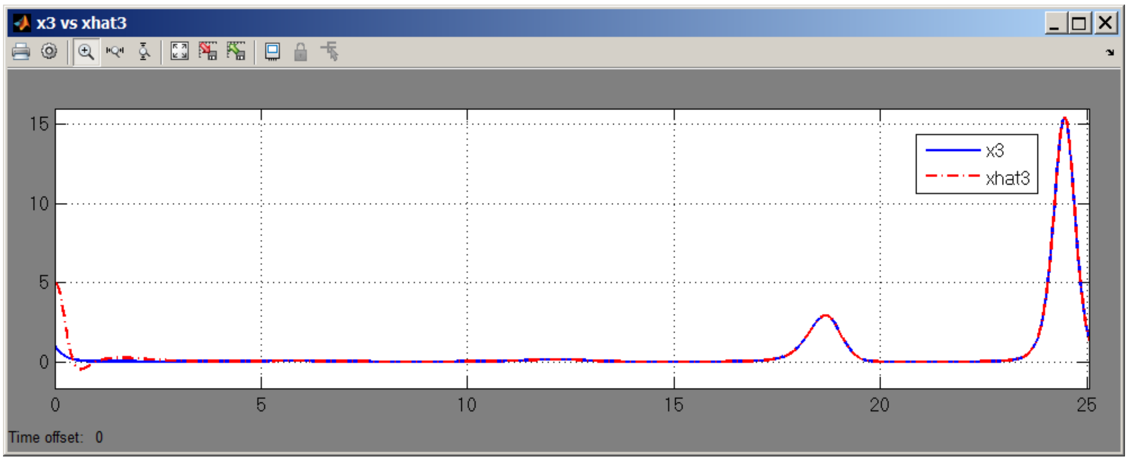

Figure 4.6: Estimated state $\hat{x}_3$ vs actual state x_3

Quantifying the Error

With the same objectives as stated in Section 4.1, the mean squared error (MSE) was determined in order to gain a more detailed understanding of how accurate the estimations were.

Quantifying the Error with Varying Disturbance

Table 4. 5 below shows how the MSE for each state changes when the initial conditions (or magnitude of perturbations) change. The adaptive gain l from the equation $\dot{\hat{\theta}} = l||y - C\hat{x}||^2$ was kept constant with a value of 1 for all the simulations below.

$e_{1_0} = (x_1 - \hat{x}_1)$	MSE_{x_1}
-4 = (1 - 5)	0.18
-9 = (1 - 10)	0.4909
-29 = (1 - 30)	2.4742

$e_{2_0} = (x_2 - \hat{x}_2)$	MSE_{x_2}
-4 = (1 - 5)	0.14
-9 = (1 - 10)	0.2966
-29 = (1 - 30)	0.9461

$e_{3_0} = (x_3 - \hat{x}_3)$	MSE_{x_3}
-4 = (1 - 5)	0.071
-9 = (1 - 10)	0.4648
-29 = (1 - 30)	8.4174

Table 4. 5: Mean Squared Error with Varying Initial Conditions (Rössler)

Quantifying the Error with Varying Adaptive Gain

Table 4. 6 shows how the MSE for each state changes when the parameter of the adaptive equation changes i.e. the adaptive gain l in the equation $\dot{\hat{\theta}} = l||y - C\hat{x}||^2$. All other parameters remained constant, and the same initial conditions of $e_0 = (x_0 - \hat{x}_0) = [-4 \quad -4 \quad -4]^T$ were used throughout.

l	MSE_{x_1}	MSE_{x_2}	MSE_{x_3}
5	0.0736	0.0591	0.0432
10	0.0497	0.0411	0.0332
30	0.0271	0.0236	0.0211

Table 4. 6: Mean Squared Error with Varying Adaptive Gain (Rössler)

Quantifying the Error by Varying the Q Matrix

Table 4. 7 below shows how the MSE for each state changes when the constant parameters of the Q matrix are changed. Note that the magnitude of Q is generally bigger than that used in the Lorenz simulations which was necessary to guarantee stability of the observer.

Q	$\mathbf{MSE}_{x_1}$	$\mathbf{MSE}_{x_2}$	$\mathbf{MSE}_{x_3}$
$100I$	0.3093	0.429	0.1079
$125I$	0.4356	0.5616	0.1081
$150I$	0.5729	0.7023	0.109
$175I$	0.7191	0.8496	0.1116

Table 4. 7: Mean Squared Error with Varying Q matrix (Rössler)

As expected, calculating the MSE has again provided interesting results concerning the performance of the designed observer. The conclusions drawn are similar to those discussed in Section 4.1.

Approximating the Error Transients

As was done in Section 4.1, a linear approximation of the error transient behaviour was modelled, by applying the algorithm in Section 3.4. The details of each step have been omitted in this case, however, the final dynamical equation of the linear approximation was determined as

$$\dot{e} = \left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)e$$

$$\dot{e} = \left[\begin{bmatrix} 0 & -1 & -1 \\ 1 & 0.2 & 0 \\ 0 & 0 & 5 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \hat{x}_3 + e_3 & 0 & \hat{x}_1 \end{bmatrix} - \frac{1}{2}\hat{\theta}\begin{bmatrix} -1.1 \\ 25.8 \\ -4.6 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}\right]e$$

Note in this case that the measured variable x_2 occurs nowhere in the nonlinearity. This implies that, according to the error approximation algorithm, there can be no step involving the iterations for the measured variable. Due to this finding, the step was skipped and the resulting matrix was analyzed after simply substituting in the initial conditions for $\hat{x}_3$, e_3 and $\hat{x}_1$. However, this procedure led to an unstable mode. It was then decided to substitute only the initial condition for $\hat{x}_3$, where $\hat{x}_0 = [5 \quad 5 \quad 5]^T$, and iterate through the (expected) range of $\hat{x}_1$, which resulted in numerous stable matrices for certain values of $\hat{x}_1$. Consequently, approximations for the upper and lower bounds could then be extracted. The value used for the adaptive parameter was $\hat{\theta} = 0$. Figure 4.7 and Figure 4.8 below show the responses of the approximated and actual error signals for the estimated states of $\hat{x}_1$ and $\hat{x}_3$ respectively.

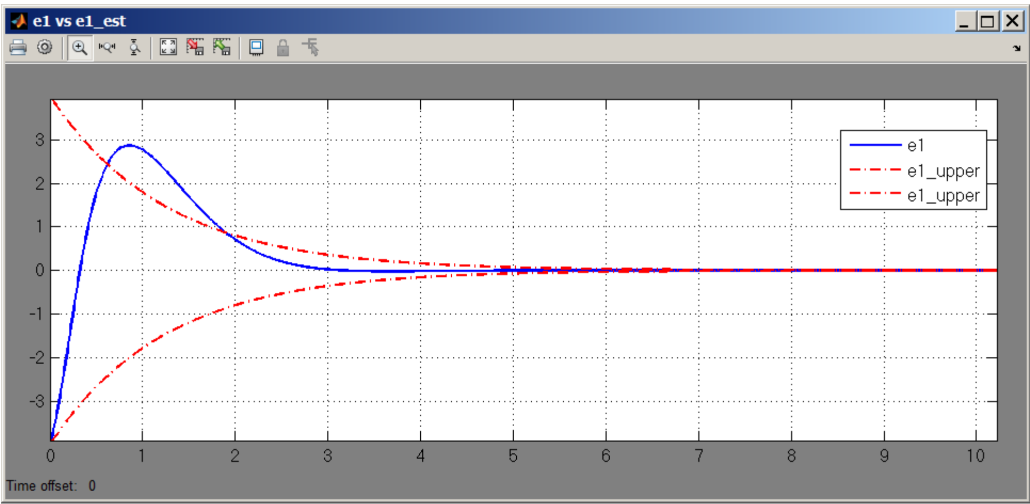

Figure 4.7: Actual and approximated error, e₁

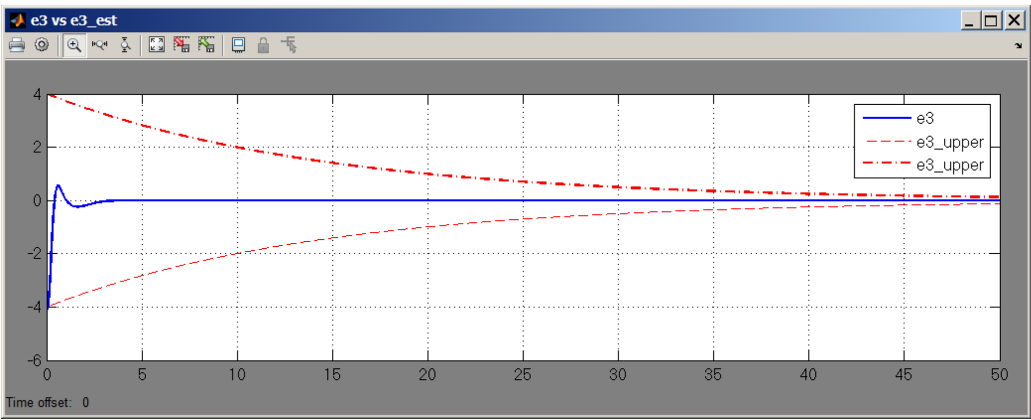

Figure 4.8: Actual and approximated error, e₃

With reference to Figure 4.7, the error bounds of e1_upper represent the slower transient response of the error for x_1. Similarly with reference to Figure 4.8, the error bounds of e3_upper represent the slower transient response of the error for x_3.

The settling time of the error approximation was then compared to the settling time of the actual error in order to determine the degree of accuracy of the approximation. This was done by determining the time taken for the error terms to reach e^{-1} of the initial error, where $e^{-1} = 0.367$. These results can be seen in Table 4. 8 where the actual error decay and the approximate error decay have been compared. The time duration was measured, in seconds, from the peak amplitude of the disturbance up to the point where it decays to below 36.7% of the peak value.

Although Table 4. 8 clearly shows that the e^{-1} settling time for the error approximation of e_1 is not an upper bound, by considering Figure 4.7 it can be noted that the approximation is indeed an upper bound, since it eventually approaches zero after the actual error does. This indicates that the actual error approaches zero at a steeper gradient as it does during the latter part of the transient time period (between 1 and 3 seconds), as opposed to the approximation, that approaches at a slower rate during that same time period. The transient time approximation for e_3 is clearly an upper bound.

Time Taken to reach e^{-1} of the Initial Error			
Error Approximation for e_1 (sec)	1.24	**Actual Error e_1 (sec)**	1.58
Error Approximation for e_3 (sec)	14.3	**Actual Error e_3 (sec)**	0.293

Table 4. 8: Time to achieve e-1 of Actual Error and Error Approximation, e_1 and e_3

4.3. Van Der Pol Oscillator

This section shows the results of the design method of Chapter 3 being applied to the Van Der Pol oscillator. An observer was once again designed using only partial state information. The system dynamics of this limit cycle (with reference to [79]) are given by

$$\dot{x}_1 = x_2 \tag{4.9}$$

$$\dot{x}_2 = \mu(1 - x_1^2)x_2 - x_1 \tag{4.10}$$

$$y = x_1 \tag{4.11}$$

where the parameter μ has been chosen as $\mu = 3$ for this particular application.

Designing the Observer

The observer for the Van Der Pol oscillator was again designed according to the algorithm discussed in Section 3.3, and the numerical simulation results are illustrated and discussed below. Following Step 1 of the algorithm, the observability of the system was first checked

$$Q(x) = \begin{bmatrix} L_g^0 h(x) \\ L_g^1 h(x) \\ L_g^2 h(x) \\ \vdots \\ L_g^{n-1} h(x) \end{bmatrix}$$

$$L_g^0 h(x) = h(x) = x_1$$

$$L_g^1 h(x) = \frac{\partial h}{\partial x} g = [1 \quad 0] \begin{bmatrix} x_2 \\ \mu(1 - x_1^2)x_2 - x_1 \end{bmatrix} = x_2$$

$$Q(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\frac{\partial Q}{\partial x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Noting that $n = 2$, and since $\frac{\partial Q}{\partial x}$ contains 2 linear independent row vectors, this system is observable.

The Lipschitz continuity was then checked as per Step 2. Considering the nonlinearity of the Van Der Pol oscillator $f(x) = [0 \quad -\mu x_1^2 x_2]^T$ and the domain of *x*, the theory in Appendix B

$$f(x) = \begin{bmatrix} 0 \\ -\mu x_1^2 x_2 \end{bmatrix}$$

$$-0.8 \le x_1 \le 0.8$$
$$-1.55 \le x_2 \le 1.55$$

$$\left|\frac{\partial f_1}{\partial x_1}\right| = |0| = 0 = \gamma_{f_1 x_1}$$

$$\left|\frac{\partial f_1}{\partial x_2}\right| = |0| = 0 = \gamma_{f_1 x_2}$$

$$\left|\frac{\partial f_2}{\partial x_1}\right| = |-2\mu x_2 x_1| = 7.44 = \gamma_{f_2 x_1}$$

$$\left|\frac{\partial f_2}{\partial x_2}\right| = |-\mu x_1^2| = 1.92 = \gamma_{f_2 x_2}$$

It can clearly be conclude that the nonlinearity $f(x)$ is Lipschitz continuous since all the derivatives exist. Then, taking $\max(\gamma_{f_1}), \max(\gamma_{f_2})$, the Lipschitz constant is determined as

$$\gamma = \sqrt{(\gamma_{f_1})^2 + (\gamma_{f_2})^2} = \sqrt{(7.44)^2 + (1.92)^2} = 7.68$$

Although not shown in the dissertation, Steps 3 to 8 of the algorithm were followed in the sequence as discussed in Section 3.3, and the observer was consequently designed accordingly

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}) + \tfrac{1}{2}\hat{\theta}D(y - C\hat{x})$$

where $A = \begin{bmatrix} 0 & 1 \\ -1 & \mu \end{bmatrix}$, $f(\hat{x}) = [0 \quad -\mu x_1^2 x_2]^T$, $C = [1 \quad 0]$ and the gain matrix *D* was calculated to be $D = [2.46 \quad -0.27]^T$. The adaptive parameter $\hat{\theta}$ is given by (3.9).

Results

The simulation was done for initial conditions $x_0 = [0.05 \quad 0.05]^T$ and $\hat{x}_0 = [-1 \quad -1]^T$. The figure below shows how the error becomes stable and the estimated state converges to the actual state value, as was expected out of the result of the design algorithm methodology. Error stability is achieved after about 5 seconds, and from inspection it can be noted that there is no erratic transient dynamics after the disturbance occurs (or after the initial conditions offset).

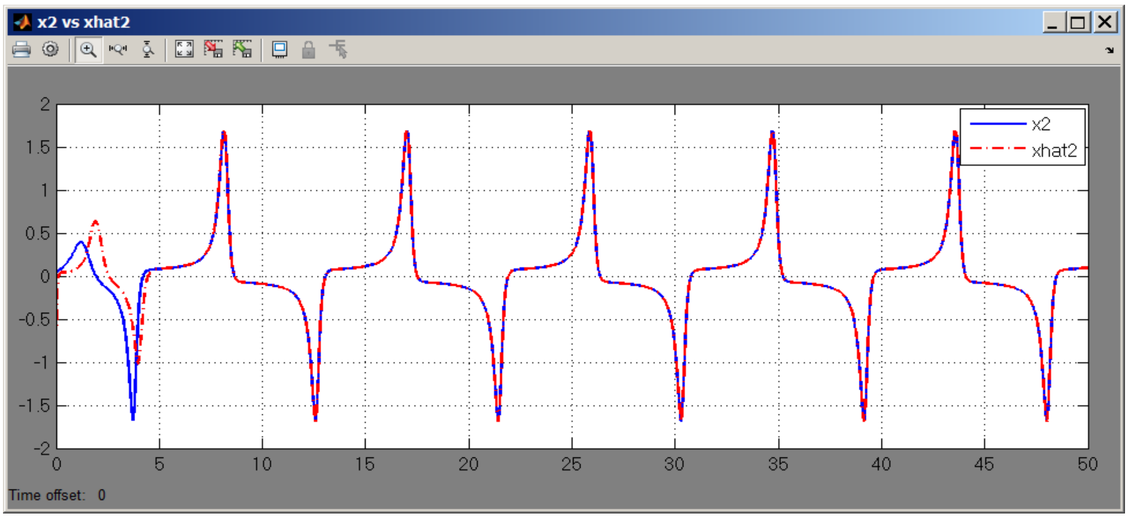

Figure 4.9: Estimated state $\hat{x}_2$ vs actual state x_2

Quantifying the Error

With the same objectives as stated in Section 4.1, the mean squared error (MSE) was determined in order to gain a more detailed understanding of how accurate the state estimations were.

Quantifying the Error by Varying Disturbance

Table 4. 9 below shows how the MSE for each state changed when the initial conditions (or magnitude of perturbations) changed. The adaptive gain l from the equation $\dot{\hat{\theta}} = l\|y - C\hat{x}\|^2$ was kept constant with a value of 1 for all the simulations below.

$e_{1_0} = (x_1 - \hat{x}_1)$	$\mathbf{MSE}_{x_1}$
1.05 = (0.05 + 1)	0.023
-4.95 = (0.05 - 5)	0.0985

$e_{2_0} = (x_2 - \hat{x}_2)$	$\mathbf{MSE}_{x_2}$
1.05 = (0.05 + 1)	0.015
-4.95 = (0.05 - 5)	0.3300

-9.95 = (0.05 - 10)	0.1801		-9.95 = (0.05 - 10)	0.0326

Table 4. 9: Mean Squared Error with Varying Initial Conditions (Van Der Pol)

Quantifying the Error by Varying Adaptive Gain

Table 4. 10 shows how the MSE for each state changed when the parameter of the adaptive equation changed i.e. the adaptive gain l in the equation $\dot{\hat{\theta}} = l||y - C\hat{x}||^2$. All other parameters remained constant, and the same initial conditions of $e_0 = (x_0 - \hat{x}_0) = [1.05 \quad 1.05]^T$ were used throughout.

l	MSE_{x_1}	MSE_{x_2}
5	0.0091	0.0010
10	0.0063	0.0006
15	0.0052	0.0013

Table 4. 10: Mean Squared Error with Varying Adaptive Gain (Van Der Pol)

Quantifying the Error by Varying the Q Matrix

Table 4. 11 below shows how the MSE for each state changes when the constant parameters of the Q matrix are changed. All other parameter remained constant.

Q	MSE_{x_1}	MSE_{x_2}
$4I$	0.0057	0.0017
$5I$	0.0063	0.0014
$6I$	0.0069	0.0013
$7I$	0.0074	0.0013

Table 4. 11: Mean Squared Error with Varying Q matrix (Van Der Pol)

As expected, calculating the MSE has again provided interesting results concerning the performance of the designed observer.

Considering Table 4. 9, an unexpected change in the magnitude of MSE_{x_2} occurred when the initial conditions increased from a magnitude of 4.95 to 9.95. Unlike the other systems investigated, the MSE did not increase with an increase in initial error but unexpectedly decreased from 0.33 to 0.03, by a rather significant factor of 10. It did however increase from the first initial condition of 1.05 to the second of 4.95. A similar occurrence took place when the adaptive parameter was varied (refer to Table 4. 10). Once again it was the magnitude of MSE_{x_2} that unexpectedly changed. In this instance it increased from 0.0006 to 0.0013 when the adaptive parameter was increased from 10 to 15, the opposite effect to what occurred when the parameter was initially changed from 5 to 10. Additionally (with reference to Table 4. 11) the MSE seems to decrease for the estimated state, as the magnitude of the Q matrix increases. This is opposite to what has been observed with the Lorenz and Rössler attractors. This questionable behaviour was not investigated and will not be covered in this dissertation, however, this provides an interesting point that can be addressed in further research, especially with respect to the Van Der Pol oscillator.

Approximating the Error Transients

In order to get a linear approximation of the error transient behaviour, the algorithm in Section 3.4 was applied. Although the details of each step have been omitted, the final dynamical equation of the linear approximation was determined as

$$\dot{e} = \left(A + \bar{f}(\hat{x}, e) - \frac{1}{2}\hat{\theta}DC\right)e$$

$$\dot{e} = \left[\begin{bmatrix} 0 & 1 \\ -1 & \mu \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -2\mu x_1 \hat{x}_2 & -\mu x_1^2 \end{bmatrix} - \frac{1}{2}\hat{\theta}\begin{bmatrix} 2.46 \\ -0.27 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix}\right]e$$

where the initial condition value for $\hat{x}_2$ was used (recalling that $\hat{x}_0 = [-1 \quad -1]^T$), the adaptive parameter was chosen as $\hat{\theta} = 2$, and the value for x_1 was iterated across its range in order to get the upper and lower bounds of the error transients. Figure 4.10 below shows the response of the approximated and actual error signals for the estimated state $\hat{x}_2$.

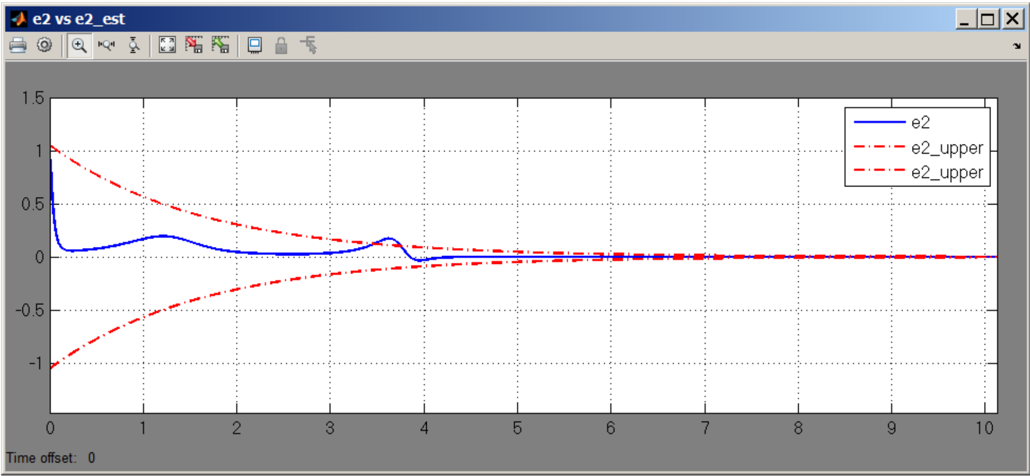

Figure 4.10: Actual and approximated error, e_2

With reference to Figure 4.10 above, the error bounds of e2_upper represent the slower transient responses of the error for x_1.

The settling time of the error approximation was then compared to the settling time of the actual error in order to determine the degree of accuracy of the approximation. This was done by determining the time taken for the error terms to reach e^{-1} of the initial error, where $e^{-1} = 0.367$. These results can be seen in Table 4. 12 where the actual error decay and the approximate error decay have been compared. The time duration was measured, in seconds, from the peak amplitude of the disturbance up to the point where it decays to below 36.7% of the peak value.

Considering Figure 4.10 and Table 4. 12, it is clear that in this particular case, the approximated transient time forms an upper bound of the actual error transient time.

Time Taken to reach e^{-1} of the Initial Error	
Error Approximation for e_2 (sec)	**Actual Error e_2 (sec)**
1.605	0.042

Table 4. 12: Time to achieve e-1 of Actual Error and Error Approximation

APPLICATION CONCLUDED

This chapter has illustrated, through numerical simulation, that an observer designed according to the methodology in Chapter 3, can successfully estimate the unknown states of the system, given only partially measured state information. All 3 systems, namely the Lorenz attractor, the Rössler attractor, and the Van Der Pol oscillator, display highly nonlinear, and in some cases, chaotic behaviour, and were selected in order to show the efficacy of the technique. Further studies were completed to analyze the accuracy of the estimated states, for varying design parameters and perturbations.

5. CONCLUSION

An extensive investigation into the different techniques of nonlinear state estimation has been discussed in this dissertation. From the literature survey we can conclude that many years of research have been done and much effort aimed at designing and developing techniques for estimating the states of nonlinear systems. Each of the papers mentioned in the literature review, describes how an observer is successfully constructed for a given system, proving that there are indeed numerous nonlinear estimation techniques. In fact, since the theory has been researched for the last 40 years or so, some of these techniques have been developed to such an extent that they have found successful application in industry. From the literature survey, we find that most of the techniques can be classified into different types of nonlinear observers, namely, the extended Kalman filter, high gain observers, observers through output injection (or observers through transformations into nonlinear observer canonical forms), sliding mode observers, observers using an output vector, the state-dependent Riccati equation observer, finite time observers, and adaptive observers.

After analysis of many of the papers, numerous techniques were applied to a diverse set of systems, and it was found that the adaptive observers provided a solid framework that seemed to be applicable to most of these example systems. An approach was then taken, using the theory from [11] as the foundation, to design a nonlinear observer based on adaptive observer techniques that can be applied to systems that can be written in the form

$$\dot{x} = Ax + f(x) + Bu$$

$$y = Cx$$

or that can be applied to systems of the form

$$\dot{x} = g(x) + Bu$$

$$y = h(x)$$

assuming they can be transformed into the nonlinear observer canonical form

$$\dot{z} = Az + \varphi(y) + B_z u$$

$$y = Cz$$

The methodology considers the designed observer as

$$\dot{\hat{x}} = A\hat{x} + f(\hat{x}) + \frac{1}{2}\hat{\theta}D(y - C\hat{x}) + Bu$$

where the adaptive parameter $\hat{\theta}$ is calculated by the equation

$$\dot{\hat{\theta}} = l\|y - C\hat{x}\|^2$$

This observer was proved to estimate the state of the system with asymptotic convergence, provided that certain assumptions on the master system were satisfied.

The degree of accuracy of the estimate was then determined by calculating the mean squared error of the actual and estimated states. At first, the initial conditions were varied (with all other parameters constant), and expectedly the mean squared error increased with an increased magnitude of the difference in initial conditions. This result is axiomatic, since a larger perturbation will obviously produce a larger difference between estimated and actual states during the transients. Thereafter, while maintaining constant initial conditions, the mean squared error was analyzed for varying values of the adaptive parameter gain l. It was found, in most cases that the mean squared error decreased with an increase in the gain l. The Van Der Pol oscillator, however, proved to be different and showed that the mean squared error in fact started to increase again after a certain value of l was exceeded. Consequently, the conclusions drawn from these results are two-fold. Firstly, the accuracy of the estimated states (during the transients) will decrease with an increase in magnitude of a perturbation. Secondly, it can be deduced that the accuracy of the estimated states will increase with a limited increase of the adaptive parameter gain l.

Finally, the error transients were analyzed. A basic procedure was developed with the objective of finding an upper bound of the error transient time, by using a linear approximation of the error dynamics. The results of this approximation procedure (quantified by the time taken to reach 36.7% of the peak error), proved to be diverse across the three different applications. When applied to the Lorenz attractor, the transient time of the approximated error was not far from the transient time of the actual error. When applied to the Rössler attractor, and considering e_1, the upper bound transient time was in fact less than the actual transient time. Then for e_2, although the upper bound transient time was greater than the actual transient time, it was very conservative. In the case of the Van Der Pol oscillator, it was also found that the approximated upper bound transient was very conservative when compared to the actual transient. Take note again that the objective of approximating these transient times was to simply obtain a 'worst case scenario' for the actual error transient time, such that this approximated information can be made available *a priori*, giving the designer a benchmark on what might be the worst transients that can be expected from designing the observer.

Implications and Further Research

By analyzing the proof and the assumptions that are to be satisfied in order to guarantee asymptotic stability of the error, it is clear that this technique does have limitations. One important example is that the nonlinearity of the master system has to be Lipschitz continuous. This is an area where further research can be directed in order to remove this constraint on the master system, thus making the application of this technique even more general.

Another recommendation is borne out of the fact that the transient times of the error vector are merely asymptotic. It would be practically desirable for the designer to know upfront, with certainty, that the transient time will always be less than a certain value. Hence, another recommendation is to either extend the theory in Chapter 3, such that convergence takes place in finite time, or otherwise to apply finite time estimation theory to a similar variety of diverse systems.

Further research can also be conducted in order to improve the procedure for the linear approximations of the error. The first objective would be to ensure that all upper bounds of the transient times are indeed greater than the actual transient times themselves. Secondly, less conservative (‘tighter’) upper bounds on the approximated error transient times would ensure a more accurate approximation, providing the designer with higher quality information on the expected error transient behaviour.

A final recommendation addresses the proof of Chapter 3. Much of the theory used in the proof provides a very conservative means of showing that the derivative of the Lyapunov candidate function is in fact negative definite. This implies that there could be numerous systems containing a given set of parameters for which an observer can be successfully designed, but without any guarantee of error stability using the theory in the proof as it is in this dissertation. It is therefore recommended that further analysis and development of the proof of the convergence be undertaken in such a manner as to decrease the conservativeness thereof, allowing the method to be applied to a wider variety of systems, with the guarantee of stability.

REFERENCES

[1] R. Kalman, "On the General Theory of Control Systems," *IRE Transactions on Automatic Control,*, vol. 4, no. 3, p. 110, 1959.

[2] D. Seborg, "Advanced Process Control," *Aiche Journal,* vol. 30, pp. 172–172, 1984.

[3] S. H. Lehnigk, "On the Hurwitz matrix," *Zeitschrift für Angewandte Mathematik und Physik (ZAMP),* 1970.

[4] M. Hazewinkel, "Diffeomorphism," in *Encyclopedia of Mathematics,* Springer, 2001.

[5] J. Lee, Introduction to Topological Manifolds, Springer-Verlag, 2000.

[6] H. Whitney, "Differentiable Manifolds," *Annals of Mathematics,* vol. 37, no. 3, pp. 645–680, 1936.

[7] R. Engelking, General Topology, Heldermann Verlag, 1989.

[8] D. Luenberger, "An Introduction to Observers," *IEEE Transactions on Automatic Control,* vol. 16, no. 6, pp. 596–602, 1971.

[9] R. Kalman, "New Results in Linear Filtering and Prediction Problems," *Journal of Basic Engineering, Transactions ASME,* vol. 82, pp. 35–45, 1960.

[10] D. Luenberger, "Observing the State of a Linear System," *IEEE Transactions on Military Electronics,* vol. 8, no. 2, pp. 74–80, 1964.

[11] L. Huang, Q. Yin, G. Sun, L. Wang and Y. Fu, "An Adaptive Observer-Based Nonlinear Control for Chaos Synchronization," in *2nd International Symposium on Systems and Control in Aerospace and Astronautics (ISSCAA),* Shenzhen , 2008.

[12] R. Bass and I. Gura, "High-Order System Design via State-Space Considerations," in *Proceedings of the Joint Automatic Control Conference,* Troy, NY, 1965.

[13] S. Julier and J. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE,* vol. 92, no. 3, pp. 401–422, 2005.

[14] G. Einicke and L. White, "Robust Extended Kalman Filtering," *IEEE Transactions on Signal Processing,* vol. 47, no. 9, pp. 2596–2599, 1999.

[15] F. Thau, "Observing the State of Nonlinear Dynamical Systems," *International Journal of Control* , vol. 17, no. 3, pp. 471–479, 1973.

[16] S. Nicosia and A. Tornambe, "High Gain Observer in the State and Parameter Estimation of Robots having Elastic Joints," *Systems & Control Letters,* vol. 13, no. 4, pp. 331–337, 1989.

[17] A. Tornambe, "Use of High-Gain Observers in the Inverse Kinematic Problem," *Applied Mathematics Letters,* vol. 3, no. 1, pp. 97–100, 1990.

[18] H. Khalil, Nonlinear Systems, Prentice Hall, 2002.

[19] G. Golub and C. V. Loan, Matrix Computations, John Hopkins Press, 1989.

[20] E. ElYaagoubi, A. E. Assoudi and H. Hammouri, "High Gain Observer: Attenuation of the Peak Phenomena," in *Proceedings of the American Control Conference,* Boston, MA, 2004.

[21] T. Oliveira, A. Peixoto and L. Hsu, "Dwell-time and Monitoring Schemes for Peaking Avoidance in High-Gain Observer based Output-Feedback Control," in *Proceedings of the*

48th IEEE Conference on Decision and Control, and the 28th Chinese Control Conference, Shanghai , 2009.

[22] K. Busawon and J. De Leon-Morales, "An Improved High Gain Observer for Single-Output Uniformly Observable Systems," in *European Control Conference (ECC)*, Karlsruhe, 1999.

[23] D. Astolfi and L. Marconi, "A High-Gain Nonlinear Observer With Limited Gain Power," *IEEE Transactions on Automatic Control,* vol. 60, no. 11, pp. 3059–3064, 2015.

[24] L. Menini and A. Tornambè, "High-gain Observers for Nonlinear Systems with Trajectories Close to Unobservability," *European Journal of Control,* vol. 20, no. 3, pp. 118–131, 2014.

[25] S. Lagmiri, M. Amghar and N. Sbiti, "Synchronization between a new Chaotic System and Rössler System by a High Gain Observer," in *14th Mediterranean Microwave Symposium (MMS)*, Marrakech, 2014 .

[26] S. Velardia, H. Hammourib and A. Barresia, "Development of a High Gain Observer for In-Line Monitoring of Sublimation in Vial Freeze Drying," *Drying Technology: An International Journal,* vol. 28, no. 2, pp. 256–268, 2010.

[27] F. Lafont, E. Busvelle and J. Gauthier, "An Adaptive High-Gain Observer for Water Treatment Systems," *Journal of Process Control,* vol. 21, no. 6, pp. 893–900, 2011.

[28] A. Krener and A. Isidori, "Linearization by Output Injection and Nonlinear Observers," *Systems & Control Letters,* vol. 3, no. 1, pp. 47–52, 1983.

[29] H. Nijmeijer and I. Mareels, "An Observer Looks at Synchronization," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications,* vol. 44, no. 10, pp. 882–890, 1997.

[30] J. Amirazodi, E. Yaz, A. Azemi and Y. Yaz, "Nonlinear Observer Performance in Chaotic Synchronization with Application to Secure Communication," in *Proceedings of the 2002 International Conference on Control Applications*, 2002.

[31] S. Drakunov, "Sliding-mode Observers based on Equivalent Control Method," in *Proceedings of the 31st IEEE Conference on Decision and Control*, Tucson, AZ, 1992.

[32] V. Utkin, "Sliding Modes and their Application in Variable Structure Systems," MIR Publications, Moscow, 1978.

[33] R. Messaoud, N. Zanzouri and M. Ksouri, "Robust Local State Observers for Nonlinear Systems," in *8th International Multi-Conference on Systems, Signals and Devices (SSD)*, Sousse, 2011.

[34] J. Zhao, Z. Shen, B. Ma and X. Gu, "Design of a Robust Nonlinear Observer for Nonlinear Lipschitz Uncertain Systems," in *Chinese Control and Decision Conference*, Yantai, Shandong , 2008.

[35] M. Corless and J. Tu, "State and Input Estimation for a Class of Uncertain Systems," *Automatica,* vol. 34, no. 6, pp. 757–764, 1998.

[36] A. Nemirovskii and P. Gahinet, "The Projective Method for Solving for Solving Linear Matrix Inequalities," in *Proceedings of the American Control Conference*, Baltimore, 1994.

[37] P. Gahinet, A. Nemirovskii, A. Laub and M. Chilali, "The LMI Control Toolbox," in *Proceedings of the 33rd IEEE Conference on Decision and Control*, Lake Buena Vista, FL, 1994.

[38] T. Sedighi, A. Koshkouei and K. Burnham, "Nonlinear Unknown Input Observer Design for Nonlinear Systems," in *UKACC International Conference on Control* , Coventry , 2010.

[39] M. Boutayeb, M. Darouach and H. Rafaralahy, "Generalized State Space Observers for Chaotic Synchronization and Secure Communication," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications,* vol. 49, no. 3, pp. 345–349, 2002.

[40] V. Estrada-Manzo, T. Guerra and Z. Lendek, "An LMI Approach for Observer Design for Takagi-Sugeno Descriptor Models," in *IEEE International Conference on Automation, Quality and Testing, Robotics,* Cluj-Napoca, 2014 .

[41] T. Dezuo and A. Trofino, "LMI Conditions for Designing Rational Nonlinear Observers," in *American Control Conference (ACC)*, Portland, 2014 .

[42] A. Zemouche and M. Boutayeb, "On LMI Conditions to Design Observers for Lipschitz Nonlinear Systems," *Automatica,* vol. 49, no. 2, pp. 585–591, 2013.

[43] H. Guannan, J. Hongyu, J. Jing and Y. Wensheng, "Robust State Estimation for a Class of Nonlinear Systems: Fuzzy-model-based LMI Approach," in *33rd Chinese Control Conference (CCC)*, Nanjing, 2014.

[44] C. Mracek, J. Cloutier and C. D'Souza, "A New Technique for Nonlinear Estimation," in *Proceedings of the IEEE International Conference on Control Applications,* Dearborn, MI , 1996.

[45] D. Erler and S. Vadali, "Examination of the Optimal Nonlinear Regulator Problem," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference,* Minneapolis, MN, 1988.

[46] K. Ueno, N. Sakamoto, M. Suzuki and T. Oguchi, "New Nonlinear Observer by State-Dependent Sylvester Equation Approach," in *Proceedings of the SICE Annual Conference,* Nagoya, 2013.

[47] R. Engel and G. Kreisselmeier, "A Continuous-time Observer which Converges in Finite Time," *IEEE Transactions on Automatic Control,* vol. 47, no. 7, pp. 1202–1204, 2002.

[48] F. Sauvage, M. Guay and D. Dochain, "Design of a Nonlinear Finite-time Converging Observer for a Class of Nonlinear Systems," *Journal of Control Science and Engineering,* vol. 2007, pp. 1-9, 2007.

[49] W. Perruquetti, T. Floquet and E. Moulay, "Finite-Time Observers: Application to Secure Communication," *IEEE Transactions on Automatic Control,* vol. 53, no. 1, pp. 356–360, 2008.

[50] S. Kou, T. Tarn and D. Elliott, "Finite-time Observer for Nonlinear Dynamic Systems," in *IEEE Conference on Decision and Control including the 12th Symposium on Adaptive Processes,* San Diego, CA, 1973.

[51] W. Tian, H. Du and C. Qian, "A Semi-global Finite-time Convergent Observer for a Class of Nonlinear Systems with Bounded Trajectories," *Nonlinear Analysis: Real World Applications,* vol. 13, no. 4, pp. 1827–1836, 2012.

[52] Y. Shen and X. Xia, "Semi-global Finite-time Observers for Nonlinear Systems," *Automatica,* vol. 44, no. 12, pp. 3152–3156, 2008.

[53] J. Bohn and A. Sanyal, "Almost Global Finite-time Stable Observer for Rigid Body Attitude Dynamics," in *American Control Conference (ACC),* Portland, 2014 .

[54] C. Ben Njima, W. Ben Mabrouk and H. Messaoud, "Finite Time Observers for a Class of Nonlinear System," in *11th International Multi-Conference on Systems, Signals & Devices (SSD)*, Barcelona, 2014 .

[55] G. Luders and K. Narendra, "An Adaptive Observer and Identifier for a Linear System," *IEEE Transactions on Automatic Control,* vol. 18, no. 5, pp. 496–499, 1973.

[56] G. Kreisselmeier, "Adaptive Observers with Exponential Rate of Convergence," *IEEE Transactions on Automatic Control,* vol. 22, no. 1, pp. 2–8, 1997.

[57] G. Bastin and M. Gevers, "Stable Adaptive Observers for Nonlinear Time-varying Systems," *IEEE Transactions on Automatic Control,* vol. 33, no. 7, pp. 650–658, 1988.

[58] Y. Dong and S. Mei, "Adaptive Observer for a Class of Nonlinear Systems," *Acta Automatica Sinica,* vol. 33, no. 10, pp. 1081–1084, 2007.

[59] M. Farza, I. Bouraoui, T. Menard, R. Ben Abdennour and M. M'Saad, "On Adaptive Observers Design for a Class of Uniformly Observable Systems," in *15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, Hammamet, Tunisia, 2014 .

[60] M. Ekramian, F. Sheikholeslam, S. Hosseinnia and M. Yazdanpanah, "Adaptive State Observer for Lipschitz Nonlinear Systems," *Systems & Control Letters,* vol. 62, no. 4, pp. 319–323, 2013.

[61] T. Menard, A. Maouche, B. Targui, I. Bouraoui, M. Farza and M. M'Saad, "Adaptive High Gain Observer for Uniformly Observable Systems with Nonlinear Parametrization," in *European Control Conference (ECC)*, Strasbourg, France, 2014.

[62] F. Xu, Y. Wang and X. Luo, "An Adaptive Soft Sensor Based on Nonlinear Differential-Algebraic Observer for Chemical Processes," *Sensors Journal, IEEE,* vol. 15, no. 6, pp. 3249–3257, 2015.

[63] Y. Shen, Q. Jiang and P. Miao, "Continuous but Non-smooth Adaptive Observer of a Class of Nonlinear Systems," in *2nd International Conference on Systems and Informatics (ICSAI)*, Shanghai, 2014.

[64] A. Boulkroune, M. Tadjine, M. M’Saad and M. Farza, "Design of a Unified Adaptive Fuzzy Observer for Uncertain Nonlinear Systems," *Information Sciences,* vol. 265, pp. 139–153, 2014.

[65] Y. Yingjuan and X. Pengzhang, "Design of a Nonlinear Adaptive Observer for a Class of Lipschitz Systems," in *33rd Chinese Control Conference (CCC)*, Nanjing, 2014 .

[66] I. N'Doye, H. Voos, T.-M. Laleg-Kirati and M. Darouach, "H-infinity Adaptive Observer Design and Parameter Identification for a Class of Nonlinear Fractional-order Systems," in *53rd Annual Conference on Decision and Control (CDC),* Los Angeles, 2014.

[67] X. Wu, W. Li, J. Zhang and Z. Tan, "Robust Adaptive Observer for a Class of Nonlinear Systems," in *33rd Chinese Control Conference (CCC)* , Nanjing, 2014.

[68] S. Ibrir, "Simultaneous State and Parameter Estimation of a Class of MIMO Nonlinear Systems," in *8th GCC Conference and Exhibition (GCCCE)*, Muscat, 2015 .

[69] L. Pecora and T. Carroll, "Synchronzation in Chaotic Systems," *Physical Review Letters,* vol. 64, no. 8, pp. 821–824, 1990.

[70] K. Cuomo, A. Oppenheim and S. Strogatz, "Synchronization of Lorenz-Based Chaotic Circuits with Applications to Communications," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing,* vol. 40, no. 10, pp. 626–633, 1993.

[71] G. Zheng and D. Boutat, "Synchronisation of Chaotic Systems via Reduced Observers," *Institute of Engineering and Technology - Control Theory & Applications,* vol. 5, no. 2, pp. 308–314, 2010.

[72] G. Grassi and S. Mascolo, "Nonlinear Observer Design to Synchronize Hyperchaotic Systems via a Scalar Signal," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications,* vol. 44, no. 10, pp. 1011–1014, 1997.

[73] L. Zhao, X. Li and G. Zhao, "Chaotic Synchronization Based on Riccati Inequality to Secure Communication," in *5th World Congress on Intelligent Control and Automation*, 2004.

[74] L. Chua, T. Matsumoto and M. Komuro, "The Double Scroll," *IEEE Transactions on Circuits and Systems,* vol. 32, no. 8, pp. 798–818, 1985.

[75] J. Alvarez-Ramirez, H. Puebla and I. Cervantes, "Convergence Rate of Observer-Based Approach for Chaotic Synchronization," *Physics Letters A,* vol. 289, no. 4-5, pp. 193–198, 2001.

[76] R. Hermann and A. Krener, "Nonlinear Controllability and Observability," *IEEE Transactions on Automatic Control,* vol. 22, no. 5, pp. 728–740, 1977.

[77] E. Lorenz, "Deterministic Nonperiodic Flow," *Journal of Atmospheric Sciences,* vol. 20, no. 2, pp. 130–148, 1963.

[78] O. Rössler, "An Equation for Continuous Chaos," *Physics Letters A,* vol. 57, no. 5, pp. 397–398, 1976.

[79] B. V. D. Pol, "On Relaxation-oscillations," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science: Series 7,* vol. 2, no. 11, pp. 978–992, 1926.

[80] A. Wu, G. Duan and B. Zhou, "Solution to Generalized Sylvester Matrix Equations," *IEEE Transactions on Automatic Control,* vol. 53, no. 3, pp. 811–815, 2008.

[81] K. Eriksson, D. Estep and C. Johnson, Applied Mathematics: Body and Soul: Calculus in Several Dimensions, Springer Science & Business Media, 2013.

APPENDICES

Appendix A – Analytical Solution to the Sylvester Equation

With reference to [80] and appendix A in [46], the solution to the generalized Sylvester equation

$$EXF = AX + BH \tag{6. 1}$$

is shown below, where in this case it is assumed $E = I$. For the matrix $A \in \mathcal{R}^{n\times n}$

$$\det(sI - A) = q(s) = q_n s^n + \cdots + q_1 s + q_0 \tag{6. 2}$$

$$\mathrm{adj}(sI - A) = R_{n-1} s^{n-1} + \cdots + R_1 s + R_0 \tag{6. 3}$$

where $q_i, i = 0, \ldots, n$ are scalar constant coefficients and $R_i, i = 0, \ldots, n-1$ are constant coefficient matrices. According to the Leverrier-Faddeev algorithm, R_k and q_k are determined as

$$R_k = \begin{cases} I & k = n-1 \\ AR_{k+1} + q_{k+1} I & k = n-1, \ldots, 1, 0 \end{cases} \tag{6. 4}$$

$$q_k = \begin{cases} 1 & k = n \\ -\dfrac{1}{k} tr(AR_k) & k = n-1, \ldots, 1, 0 \end{cases} \tag{6. 5}$$

The theorem for the solution then states that when $A \in \mathcal{R}^{n\times n}$, $B \in \mathcal{R}^{n\times r}$, and $F \in \mathcal{R}^{p\times p}$ are given, then the solution to the matrix Sylvester equation (6. 1) is

$$X = \sum_{k=0}^{n-1} R_k BZF^k \tag{6. 6}$$

$$H = \sum_{k=0}^{n} q_k ZF^k = Zq(F) \tag{6. 7}$$

where the matrix Z is a free parameter matrix. Iff $\sigma(A) \cap \sigma(F) = \emptyset$, then $q^{-1}(F)$ exists. When H is given and $\sigma(A) \cap \sigma(F) = \emptyset$, then Z is fixed as $Z = Hq^{-1}(F)$ and consequently X can be calculated (as a unique solution).

Appendix B – Lipschitz Continuity and Lipschitz Constants

With reference to [81] it is said that *f* is Lipschitz continuous if there is a constant γ such that

$$\|f(x) - f(\hat{x})\| \leq \gamma \|x - \hat{x}\|$$

where $\|\cdot\|$ represents the L_2- norm. To check if the function *f* is Lipschitz continuous on a given subset $D \rightarrow \mathcal{R}^n$, it suffices to check that all the component functions $f_i \in \mathcal{R}$ are Lipschitz continuous. This is the case since

$$\|f_i(x) - f_i(\hat{x})\| \leq \gamma_i \|x - \hat{x}\| \text{ for } i = 1, \dots, n$$

implies

$$\|f(x) - f(\hat{x})\|^2 = \sum_{i=1}^{n} |f_i(x) - f_i(\hat{x})|^2 \leq \sum_{i=1}^{n} \gamma_i^2 \|x - \hat{x}\|^2$$

which results in

$$\|f(x) - f(\hat{x})\| \leq \gamma \|x - \hat{x}\| \text{ with } \gamma = \left(\sum_i \gamma_i^2\right)^{\frac{1}{2}}$$

A useful theorem that is a sufficient condition for Lipschitz continuity is also stated below.

Theorem B.1

Suppose that $f(x_1, x_2)$ is defined on a set $D \rightarrow \mathcal{R}^2$. If a constant $\gamma > 0$ exists with

$$\left|\frac{\partial f}{\partial x_1}(x_1, x_2)\right| \leq \gamma \textit{ for all } (x_1, x_2) \textit{ in } D$$

then f is Lipschitz continuous on D with respect to x_1 and with a Lipschitz constant of γ.

Appendix C – Numerical Simulation Code

C.1. Lorenz Attractor

Gain Matrix Calculation

```
clear
clc

%% SYSTEM MATRICES

A = [ -10.0000   10.0000     0;
       28.0000   -1.0000     0 ;
        0           0       -2.6667 ];

C = [1 0 0];

F2 = [         0         0         0;
        -27.0000         0   -8.4853;
          8.4853    8.4853         0];

F3 = [         0         0         0;
        -27.0000         0    8.4853;
         -8.4853   -8.4853         0];

x0  = [1;1;1];
xhat0 = [5;5;5];
theta_hat = 0 % 7.1 ;

%% -----------------------------------------------------------------------

gamma = 51;                           % Lipschitz  constant
poles = [-40 -10 -8] ;
Q = 4*eye(3);

%% GAIN MATRIX

A_lin = A + F3;                       % Linearized system matrix created so that 'place' can be
used on an observable system (A,C)

L = place(A_lin',C',poles);

Ae = A - L'*C;                        % Error matrix ( A-LC )

P = lyap(Ae,Q);                       % P matrix (solution to Lyapunov equation)
P_eig = eig(P)                        % Confirming that P is positive definite

D = (C*inv(P))';                      % Observer Gain Matrix
```

Error Transient Approximation

```
%%

x3hat_sample = 16;
x2hat_sample = -7.7;

for i = 1:20 - (-15)


        x1test = -15 + i-1;

        A_error = (A + [0  0  0; -(x3hat_sample)  0  -(x1test) ; (x2hat_sample)  (x1test)  0 ] -
0.5*theta_hat*D*C);  % at t = about 0s

        eig_A_error = eig(A_error)';
        negativedefinite = all(eig(A_error) < 0);

   if negativedefinite == 1

        all_eig(i,:) =  [x1test eig_A_error];
        neg_magnitude(i,:) = [x1test norm(eig_A_error)];

   else
       % do nothing
   end

end

x1_sample1 = 11;
x1_sample2 = 19;

x1test_eig1_eig2_eig3 = real(all_eig)
neg_magnitude
```

Mean Squared Error Calculation

```
%% - MEAN SQUARED ERROR - Lorenz Attractor

Err_vec = lorenz_states - lorenz_est_states;

% - STATE1
Err_vec1 = Err_vec(:,1);
Sum_Err_Sqrd1 = 0;

MSE1 =  mse(Err_vec1)

% - STATE2
Err_vec2 = Err_vec(:,2);
Sum_Err_Sqrd2 = 0;

MSE2 =  mse(Err_vec2)

% - STATE3
Err_vec3 = Err_vec(:,3);
Sum_Err_Sqrd3 = 0;

MSE3 =  mse(Err_vec3)
```

Observer Model

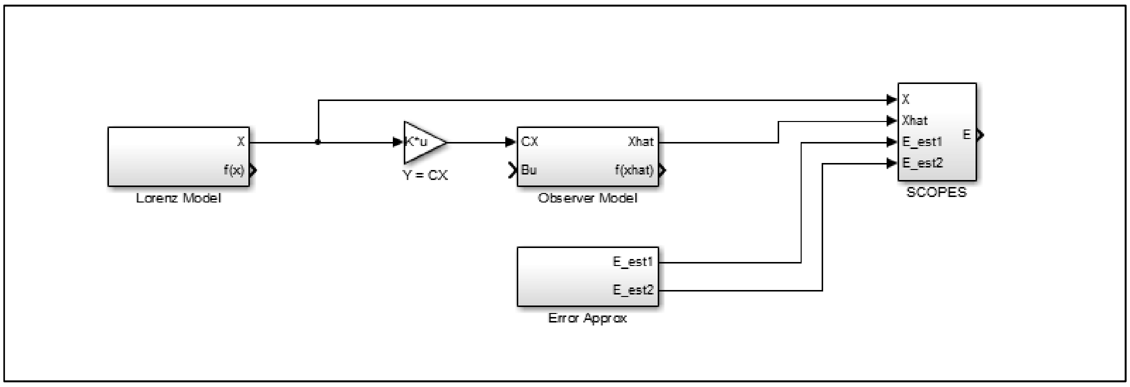

Figure C. 1. 1: Lorenz Attractor Observer - Simulation Overview

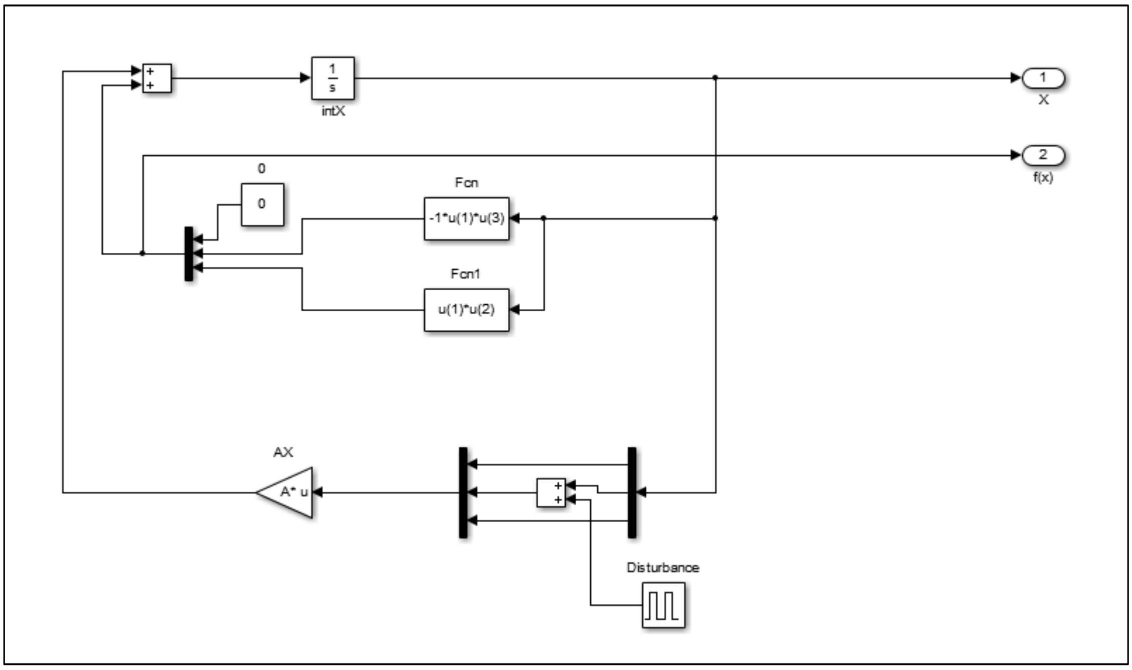

Figure C. 1. 2: Lorenz Attractor Observer - Lorenz Model

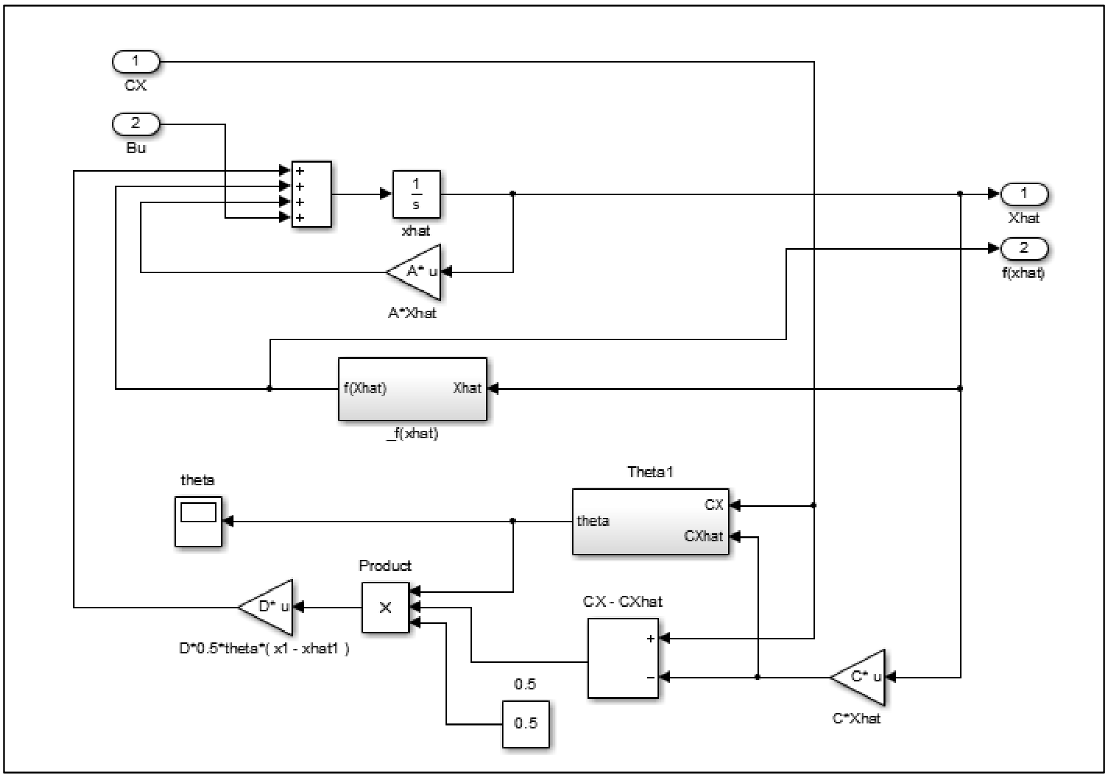

Figure C. 1. 3: Lorenz Attractor Observer - Observer Model

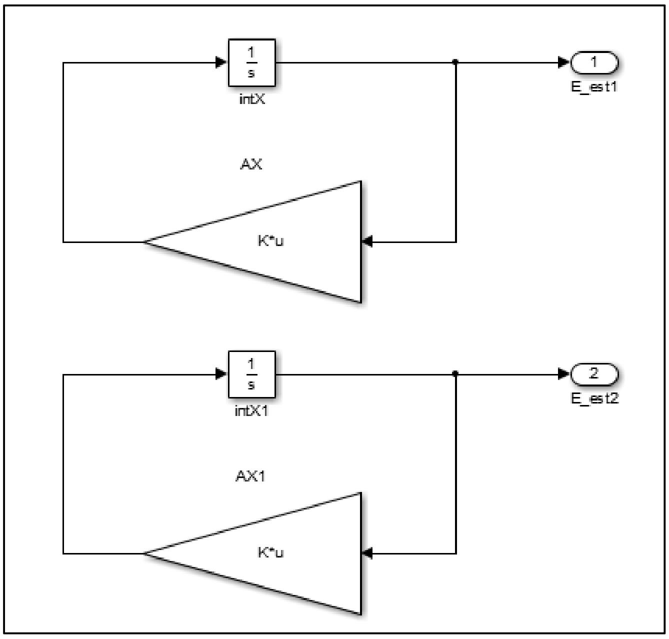

Figure C. 1. 4: Lorenz Attractor Observer - Error Approx

C.2. Rössler Attractor

Gain Matrix Calculation

```
clear
clc

%% SYSTEM MATRICES

a=0.2;
b=0.2;
c=5;

A =   [  0   -1   -1;
         1    a    0;
         0    0   -c ] ;

C = [0 1 0];

F2 = [   0   0   0;
         0   0   0;
         25  0   5];

F3 = [   0     0   0;
         0     0   0;
         0.04  0   0.008];

x0  = [1;1;1];
xhat0 = [5;5;5];
theta_hat = 0;

%% -----------------------------------------------------------------------

gamma = 18.3;                               % Lipschitz  constant
poles = [-0.1 -0.2 -0.4] ;
Q = 175*eye(3);

%% GAIN MATRIX

A_lin = A + F2;                             % Linearized system matrix

L = place(A_lin',C',poles);                 % Gain matrix

Ae = A_lin - L'*C;                              % Error matrix ( A-LC )

P = lyap(Ae,Q);                             % P matrix (solution to Lyapunov equation)
P_eig = eig(P)

D = (C*inv(P))';
```

Error Transient Approximation

```
for i = 1:10 - (-10)

        x1test = -10 + i-1;

        A_error = (A + [0 0 0 ;0 0 0 ;5 0 x1test] - 0.5*theta_hat*D*C); % at t = 0s
        eig_A_error = eig(A_error)';
        negativedefinite = all(eig(A_error) < 0);

   if negativedefinite == 1

        all_eig(i,:) =  [x1test eig_A_error];
```

```
        neg_magnitude(i,:) = [x1test norm(eig_A_error)];
    else

        % do nothing

    end

end


x1test_eig1_eig2_eig3 = real(all_eig)

neg_magnitude
```

Mean Squared Error Calculation

```
%% - MEAN SQUARED ERROR - Rossler Attractor

Err_vec = rossler_states - rossler_est_states;

% - STATE1
Err_vec1 = Err_vec(:,1);
Sum_Err_Sqrd1 = 0;

MSE1 =  mse(Err_vec1)

% - STATE2
Err_vec2 = Err_vec(:,2);
Sum_Err_Sqrd2 = 0;

MSE2 =  mse(Err_vec2)

% - STATE3
Err_vec3 = Err_vec(:,3);
Sum_Err_Sqrd3 = 0;

MSE3 =  mse(Err_vec3)
```

Observer Model

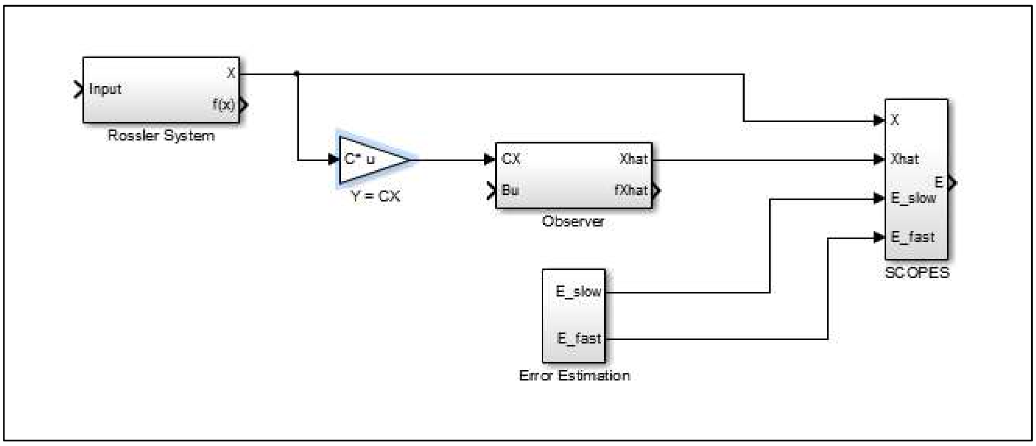

Figure C. 2. 1: Rössler Attractor Observer - Simulation Overview

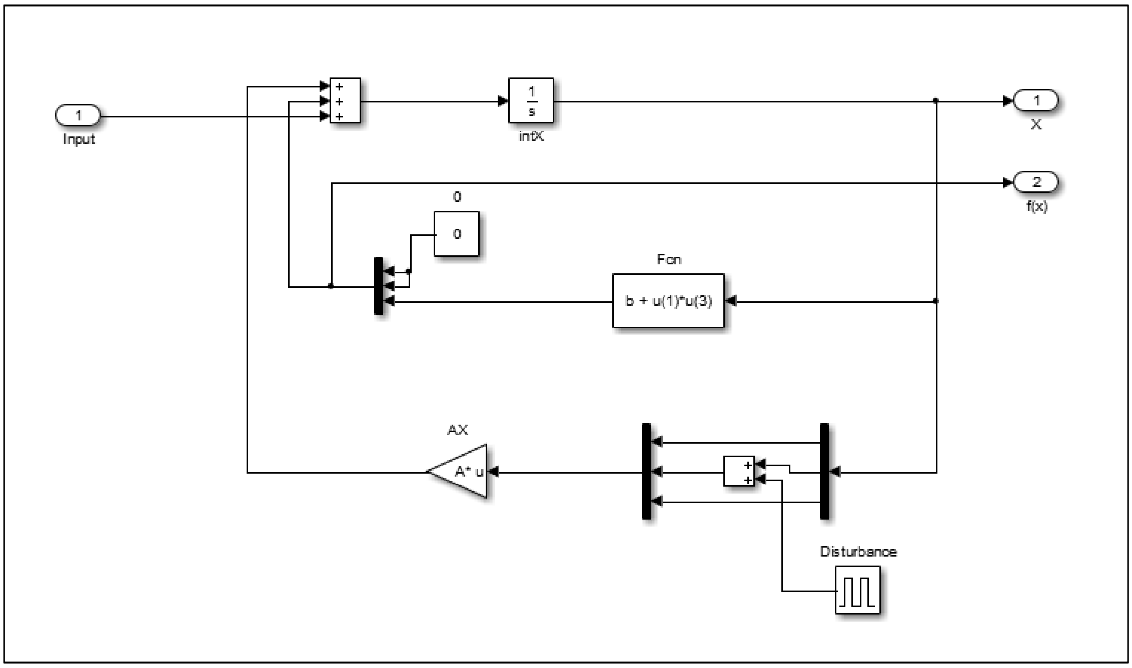

Figure C. 2. 2: Rössler Attractor Observer - Rössler Model

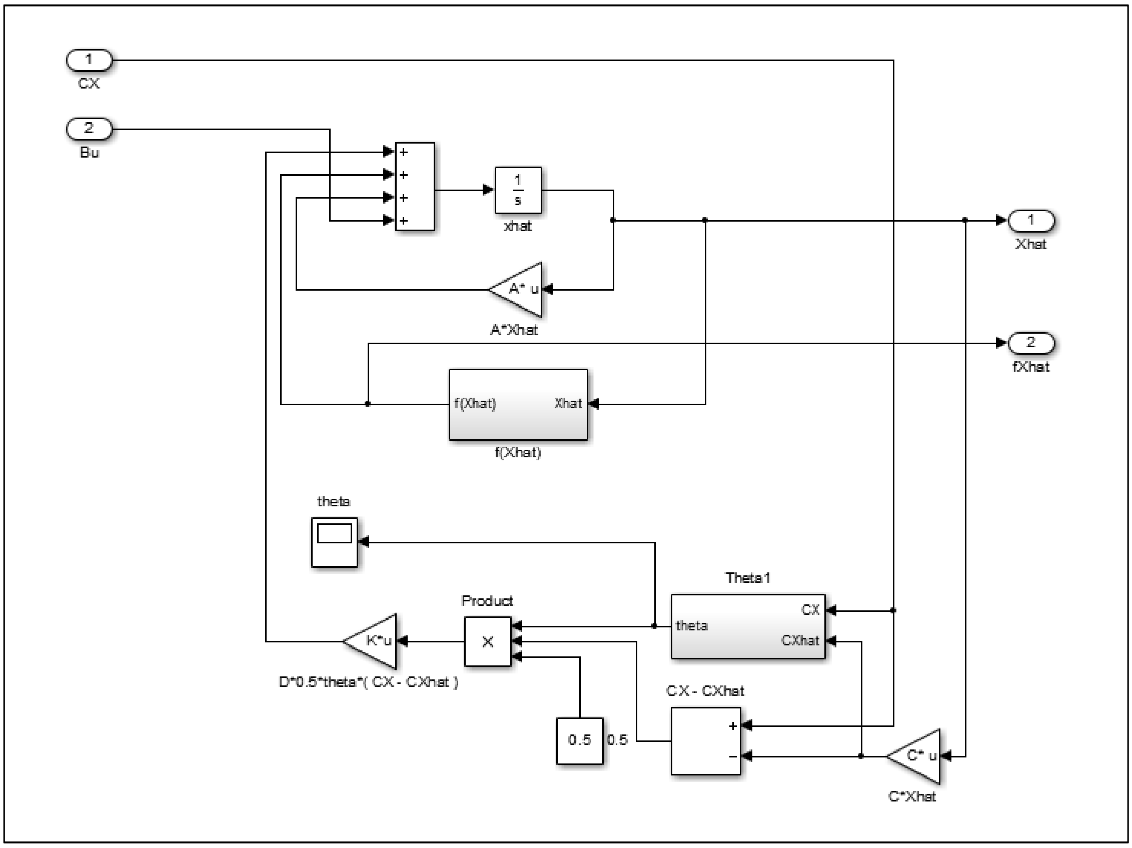

Figure C. 2. 3: Rössler Attractor Observer - Observer Model

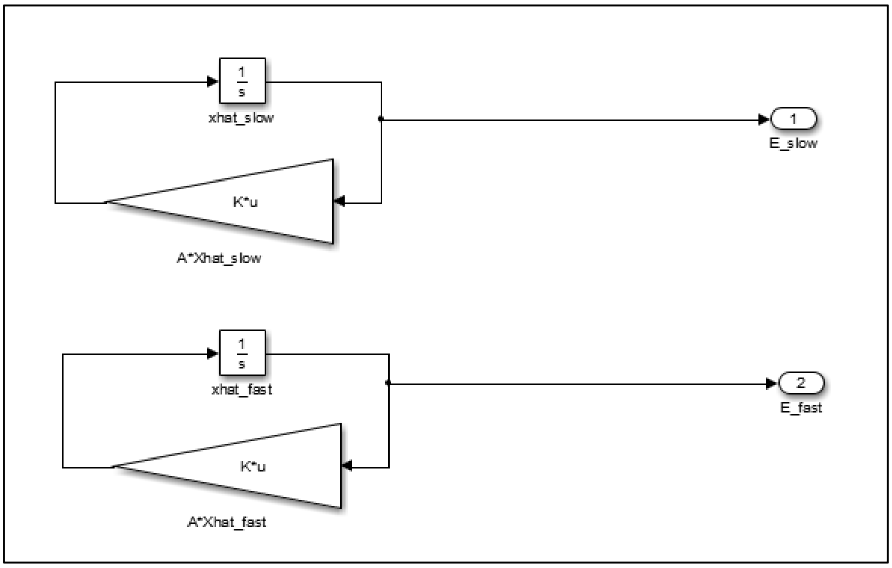

Figure C. 2. 4: Rössler Attractor Observer - Error Approx

C.3. Van Der Pol Oscillator

Gain Matrix Calculation

```
clear
clc

mu = 3;

A = [ 0  1 ; -1  mu ];

C = [ 1 0 ];

x0  = [0.05;0.05];
xhat0 = [-1;-1];


%% ---------------------------------------------------------------------

gamma = 7.68;                                % Lipschitz  constant
poles = [-4 -1 ] ;
Q = 4*eye(2)

%% GAIN MATRIX


A_lin = A + [0 0;0 0];                       % Linearized system matrix

L = place(A_lin',C',poles);                  % Gain matrix

Ae = A - L'*C;                               % Error matrix ( A-LC )

P = lyap(Ae,Q);                              % P matrix (solution to Lyapunov equation)
P_eig = eig(P)

D = (C*inv(P))'
```

Error Transient Approximation

```
theta_hat = 2  ;

x1max = 0.8;
x1min = -0.8;

xhat1_sample = xhat0(1,1);
xhat2_sample = xhat0(2,1);
e1_sample = x0(1,1) - xhat1_sample;
e2_sample = x0(2,1) - xhat2_sample;
x2_sample = 1;

res_fact = 0.01; % resolution factor

for i = 1:(x1max-x1min)/res_fact + 1

        x1test = x1min + i*res_fact - res_fact;

        A_error = (A + [ 0  0 ; -2*mu*x1test*xhat2_sample  -mu*x1test^2 ] - 0.5*theta_hat*D*C);

        eig_A_error = eig(A_error)';

        negativedefinite = all(eig(A_error) < 0);

   if negativedefinite == 1

        all_eig(i,:) =  [x1test eig_A_error];

        neg_magnitude(i,:) = [x1test norm(eig_A_error)];

   else
       % do nothing
   end

end


x1test_eig1_eig2_eig3 = real(all_eig)

neg_magnitude
```

Mean Squared Error Calculation

```
%% - MEAN SQUARED ERROR - Van Der Pol Oscillator

Err_vec = VDP_states - VDP_est_states;

% - STATE1
Err_vec1 = Err_vec(:,1);
Sum_Err_Sqrd1 = 0;

MSE1 =  mse(Err_vec1)

% - STATE2
Err_vec2 = Err_vec(:,2);
Sum_Err_Sqrd2 = 0;

MSE2 =  mse(Err_vec2)
```

Observer Model

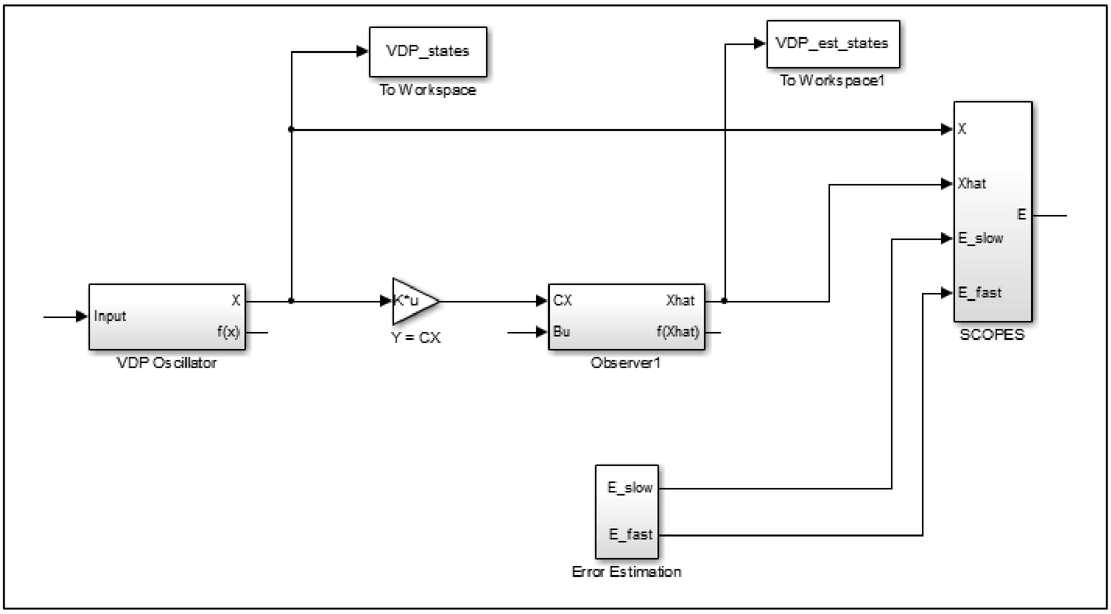

Figure C. 3. 1: Van Der Pol Oscillator Observer - Simulation Overview

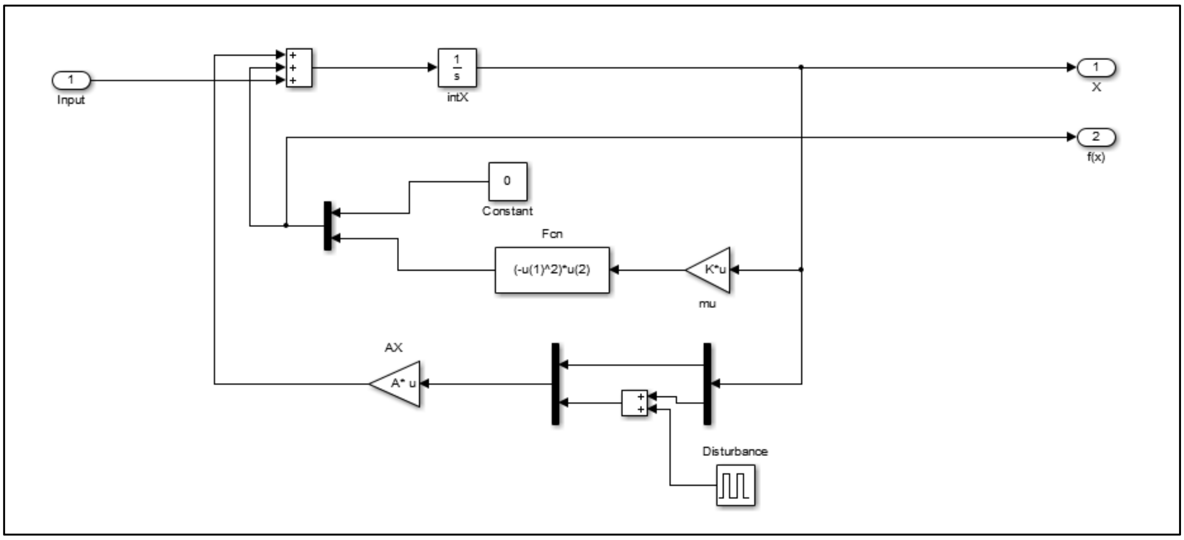

Figure C. 3. 2: Van Der Pol Oscillator Observer - Limit Cycle Model

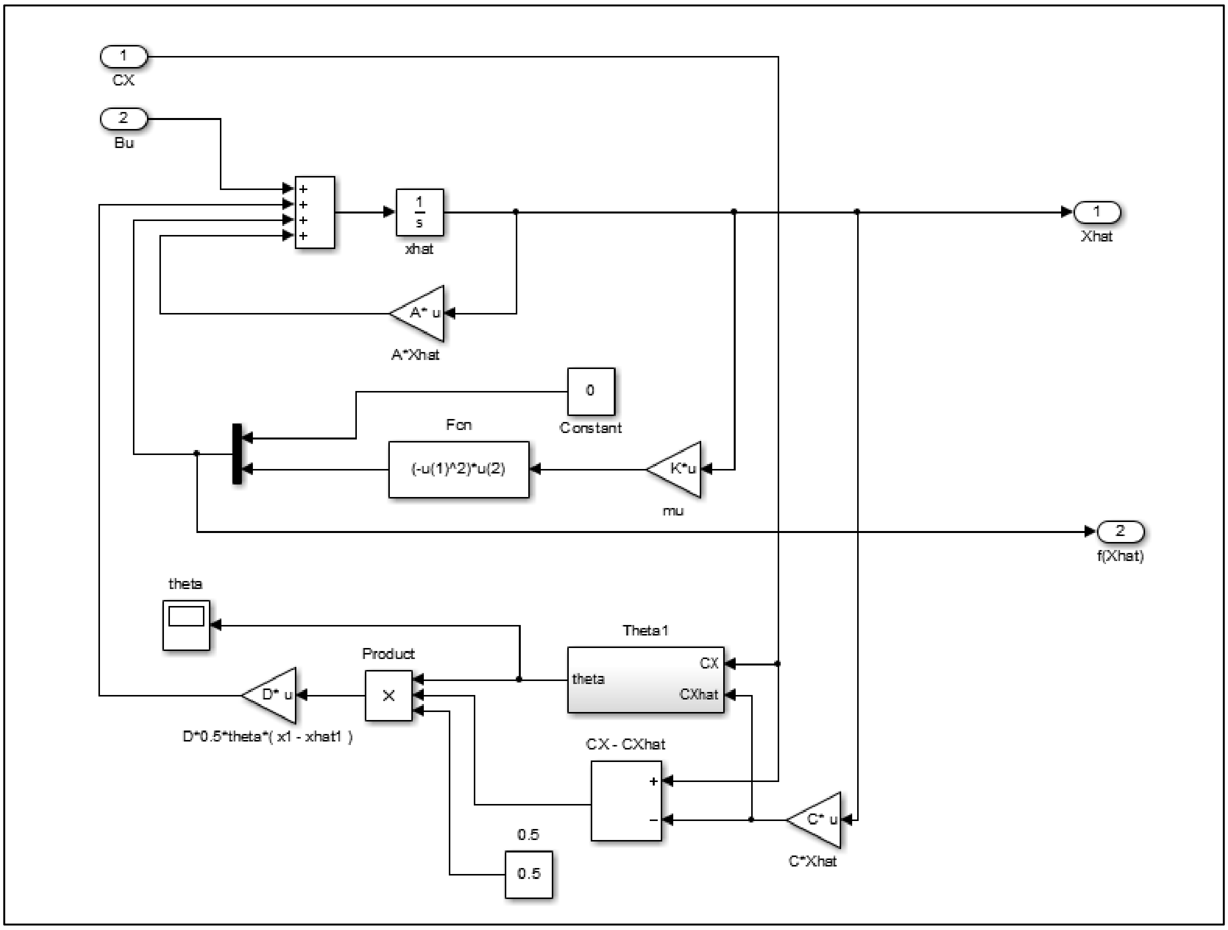

Figure C. 3. 3: Van Der Pol Oscillator Observer - Observer Model

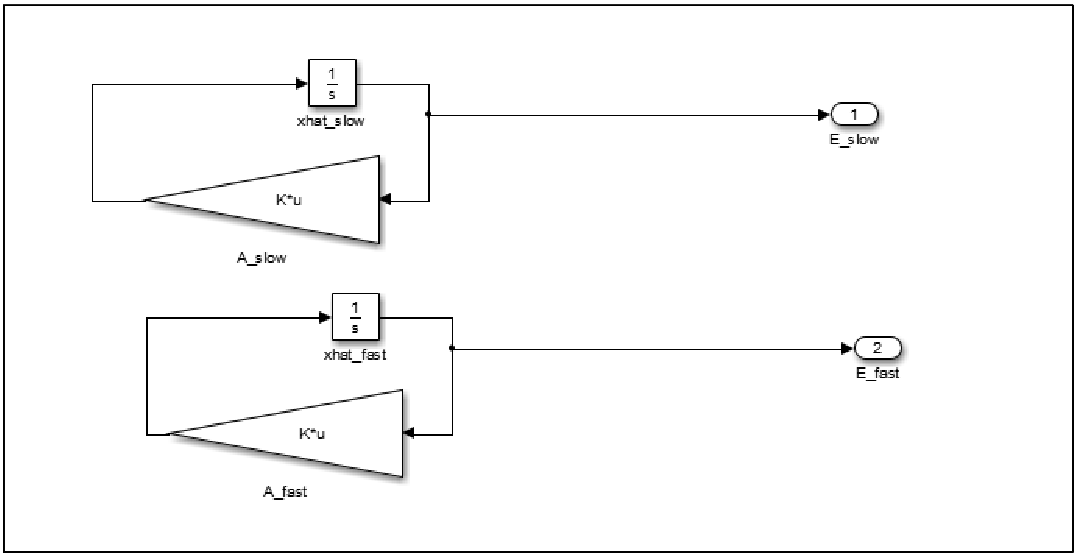

Figure C. 3. 4: Van Der Pol Oscillator Observer - Error Approx

BIBLIOGRAPHY

Full name

Dean Christiaan Tait McBride

Candidate for the Degree of

Master of Science

Dissertation

An Investigation of Techniques for Nonlinear State Observation

Major Field

Nonlinear Systems and Control Theory

Prior Education

BSc Electro-Mechanical Engineering

Personal Data

The author was born on the 25th of January, 1983 in Johannesburg, South Africa. In 2005 he graduated with a degree in engineering from the University of Cape Town. After working in various industries over a period of 8 years, he enrolled at the University of the Witwatersrand to study for an MSc in Nonlinear Systems and Control theory. This degree was undertaken part-time, as the author was concurrently employed full time as a control systems engineer at an industrial automation and control engineering company.

Adviser's Approval

Professor Brian Wigdorowitz

Professor Anton Van Wyk