MODELING, IDENTIFICATION AND CONTROL OF A BELT DRIVE SYSTEM

MODELING, SYSTEM IDENTIFICATION AND CONTROL OF A BELT DRIVE SYSTEM

BY

SHENJIN ZHU, B.S. (MECH. ENG.)

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Applied Science

McMaster University

MASTER OF APPLIED SCIENCE (2011) McMaster University

(Mechanical Engineering) Hamilton, Ontario

TITLE: Modeling, System Identification and Control of a Belt Drive System

AUTHOR: Shenjin Zhu, B.S. in Mechanical Engineering (Harbin University of Science and Technology)

SUPERVISOR: Professor Dr. Saeid Habibi

NUMBER OF PAGES: xiv, 205

Abstract

Belt drives have been serving the industry for a long period. Certain features of belt drives such as slippage, tension fluctuations, and sliding of the belt on the pulleys lead to highly nonlinear deformation, large rigid body motion, dynamical contact with sticking and slipping zones and cyclic tension. The performance of motion control for belt drives is important in many industrial fields and is affected by these factors. Advanced control can improve robustness of belt drive and result in a faster dynamic response and more accuracy. The Purpose of this project is to develop a mathematical model of an experimental belt drive system through physical modeling and system identification. This model is then used for the design of an advanced robust discrete-time controller. An extensive literature review is provided, covering modeling and control of belt drive system as well as sliding mode control (SMC) theory. Physical modeling is carried out for an experimental system followed by system identification. Both the physical and the identified models are used to analyze and investigate the characteristics of the system. Different control approaches such as discrete-time proportional integral derivative (DPID) and discrete-time sliding mode control (DSMC) are designed and implemented. The results are compared and conclusions are drawn from both control approaches.

To my wife, Jinbo Zhang.

Her encouragement and support

made this opportunity possible.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr. Saeid Habibi, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the control engineering subject. I am heartily thankful to Dr. Habibi for being patient and inspiring to keep me going during the years of this research, for reading my chapters over and over again with numerous major and minor corrections.

I would also like to thank Dr. Simon Haykin and Dr. Gary Bone, whose encouragement ignited my light of hope.

I thank my fellow graduate students for the support they provided during this research. I thank Mr. Kevin McCullogh, for his help and support during the experimental setup. I thank Mr. Andrew Gadsden and Mr. Mohammad El Sayed, for their advice and support during this research. I also thank Mr. Wanlin Zhang and Mr. Yu Song for their kind help on the experimental set up.

I would also like to thank Mr. Joe Verhaeghe and Mr. Ron Lodewyks for their help and constant support during the building of the experimental system.

I thank my examination committee members Dr. Stephen C. Veldhuis and Dr. Eu-Gene Ng for their advice.

I own thanks to my parents in-law, Deming Zhang, my wife, my son, Jason and my daughter, Cathy, who encouraged and helped me so much and were always supportive of me in pursuing my dreams.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Background

The sliding mode control (SMC) technique has been serving control engineering for over half a century since its origin in the former Soviet Union in the 1930s. Its development has undergone several stages. The SMC has been applied to different types of systems such as linear, nonlinear, discrete-time, stochastic, large-scale and infinite dimensional systems. It has been applied to numerous industrial problems including control of electric drives, power systems, robot manipulators, mobile drives, autonomous underwater vehicles, aircrafts, and space crafts. The essential feature of the SMC technique is the definition of a switching surface in terms of the state variables and associated desired closed-loop dynamics. A control law is designed in such a way that the state trajectories reach the surface in a finite time and remain on it thereafter assuring the desired dynamics specified by the switching hyperplane. The main advantages of the SMC are as follows (Perruquetti & Barbot, 2002): (1) the system becomes robust to external perturbations and internal parametric uncertainties, (2) the closed-loop system dynamics mimic a simplified reduced-order system, and (3) stability.

The primary motivation of this research is to take advantage of the robustness and the stability of the SMC by applying it to the control of a laboratory-scale belt drive system. The reason for choosing the belt drive system as the experimental device for this research is as follows.

The belt drives are currently a popular form of material transportation. They have been used in industry for over 200 years and have played an important role in the industrial revolution. The flat belts of leather and cotton ropes running in V-grooves transmitted power from steam engines or water wheels to various production machineries through a series of line shafts (Timing, 2009). Nowadays the belt drives are common in products such as home appliances, pumping equipment, heating, ventilation, and air conditioning systems. The synchronous belt drives, also called timing belts, are most commonly used technology in Europe for modern engine overhead camshaft drives (Abrate, 1992).

Some special features of the belt drive systems have made their modeling and control challenging. The belt drives have the following advantages compared to chain drives and gear trains: low price, cleanliness, simple installation, lubrication free, minimal and infrequent maintenance, ability to absorb sudden shocks or changes in loading, quietness and smoothness in operation, longevity, wide range of speed ratios, small power loss, relatively long distances between the drive and driven pulleys, and provision of visual warning of failure (Cepon & Boltezar, 2009). They also bear several drawbacks: (1) the belt usually can not be repaired after breaking; (2) slippage can occur, in particular, when the belt is not tensioned properly; (3) belts can be damaged in adverse service environments such as extreme temperature ranges, high moisture, oily or chemically filled atmospheres and result in severe slippage; (4) belt drives can demonstrate complex dynamical behaviours such as rotational and transverse vibrations, tension fluctuations, and sliding of the belt over the pulleys (Cepon & Boltezar, 2009); (5) even in a case of a

belt-drive transmitting a constant torque, the translating belt is subject to a cyclic tension variation (Dufva, Kerkkanen, Maqueda & Shabana, 2007). Such behaviours often lead to sliding wear, angular velocity loss between the drive and the driven pulleys, noisy operation and fatigue of the belt. In summary, belt drives possess peculiar features that make their dynamics nonlinear and their models highly uncertain (Leamy, 2002). This makes the control of a belt drive to be challenging.

The early belt drive studies were mainly focused on the belt drive mechanics to serve the belt drive design and manufacturing. The belt drive dynamic response studies appeared in 1980s due to a wide application of serpentine belt drives in the automotive industry. Nowadays, the study of the belt drive mechanics and dynamics have been merged together to generate a more accurate model for control design.

An alternative to obtaining a mathematical model of the belt drive system is the system identification method. In this method, a model is experimentally generated that does not relate directly to the detailed composition of the system. This model is generated by analyzing the responses of the system to certain input signals such as step signals, sinusoidal signals, and PRBS signals. In comparison to system identification, physical modeling is often unpredictable, costly, time consuming and often less accurate. System identification is fast, predictable, universal and can be very accurate (Habibi, 2009).

1.2 Objective of This Research

A motivation of this research is to use system identification techniques to investigate the linearity of belt drive systems and if possible to obtain a mathematical model of the

belt drive system for control designs. A further objective of this research is to determine the stability and the benefits of applying sliding mode control to belt drive systems. As such, a mathematical model of the belt drive system is physically derived and compared to an identified model obtained by using the system identification technique. A discrete-time sliding mode control (DSMC) and a proportional-integral-derivative (PID) controller are then designed using the identified model. The performance of the two controllers is then compared.

1.3 Outline of the Thesis

This thesis is organized as follows. A general classification of belt drive systems is provided and a laboratory-scale experimental set-up is introduced and modeled in Chapter 2. The physical model provides the *a-priori* information needed for the system identification. In Chapter 3, a parametric model of the belt drive system is obtained through system identification. An extensive literature review on sliding mode control is provided in Chapter 4. Control of the experimental belt drive system using the DSMC and the PID methods is presented in Chapter 5. A state observer and a Kalman filter are designed, for the DSMC, based on the identified model. These controllers are implemented and tested on the system without and with added modeling uncertainties. Conclusions on the performance of the controllers are provided in Chapter 6.

Chapter 2 Belt Drive Systems and Their Modeling

2.1 Belt Classification

There are a wide variety of belts classified according to their constructions. These include:

- Flat belt: it has been widely used in factories to drive machines in the past. Now flat belts are mainly employed in low power and high speed transmissions and belt conveying. The flat belt was generally made of leather. Nowadays it is mainly made of elastomer. An example of a flat belt drive is shown in Figure1 (Flat Belts, 2009).

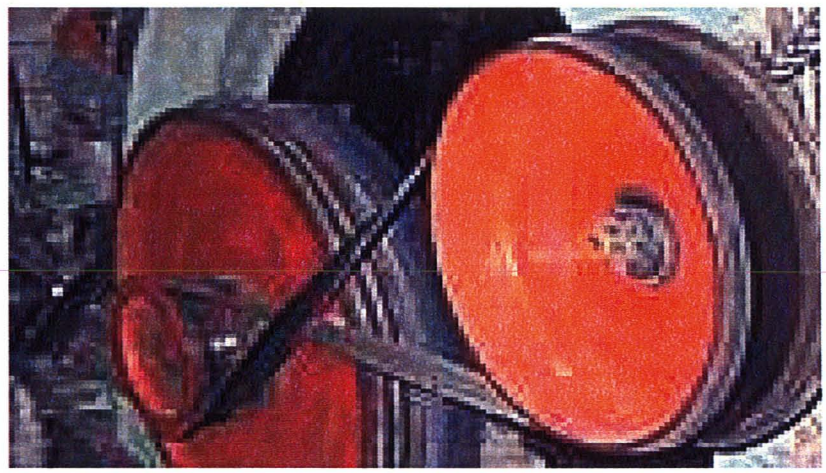

Figure 1 Flat Transmission Belt

- Round belt: it has a circular cross section and runs in pulleys with a circular groove. Round belts are supplied in various lengths or cut and jointed by stapling, gluing or welding. Round belts are used in low torque applications. An example is

shown in Figure 2 (Round Belt, 2009).

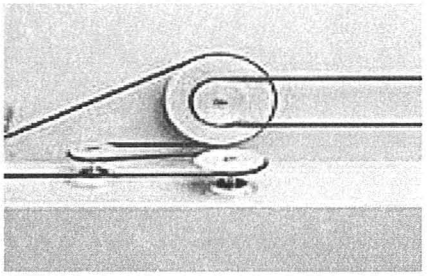

Figure 2 Round Transmission Belt

- Vee belt: it has a cross-section shaped like a trapezoid, and is also called the V-belt or the wedge-rope. The application of the Vee belt solved the slippage and the alignment problems, and is now commonly used for power transmission. It provides the best combination of traction, speed of movement, load bearing and long service life. An example of a Vee belt is shown in Figure 3 (Belt Drive, 2009).

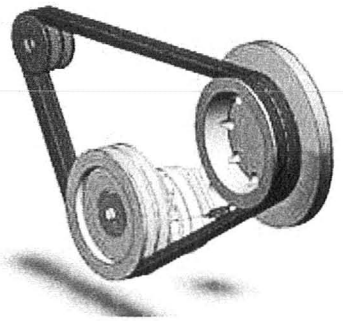

Figure 3 A Pair of Vee Belts

- Film belt: it is a new ultrasonically welded belt of film. A film belt provides uniformly high strength for instrumentation and light power transmission applications. It has excellent chemical resistance, so that it is applicable in areas

with contaminated solvents, acids, and oils. It is highly flexible and has length stability, making it ideal for disc memory drives, turntables, and miniature precision drives (Film Belts, 2009). Some film belts are shown in Figure 4 (Film Belts, 2009).

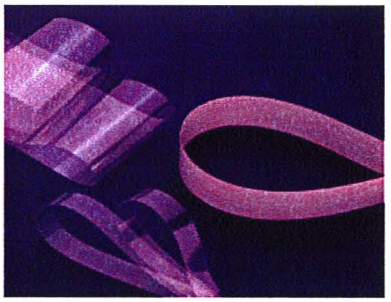

Figure 4 Samples of Film Belts

- Timing belt: it is a flat belt with integral teeth, used for high power transmission, especially in situations that require specific drive ratios. With proper tension, the timing belt provides a no-slip operation. A timing belt is shown in Figure 5 (Belt Mechanical, 2009).

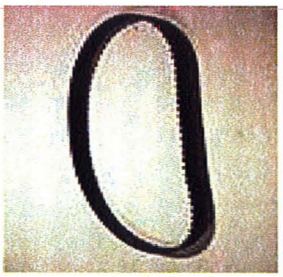

Figure 5 A Timing Belt

- Wire Belt: It is normally made of stainless steel. Wire belts usually have exceptional thermal and chemical resistances and allow lateral and longitudinal

flow of air or liquid. They are ideal for applications involving coating, drying, cooling, heating, washing, flushing and draining in food processing, bakery, ceramic and glass industry (Belt and Chain Drives, 2009). A wire belt conveyor is shown in Figure 6 (Belt and Chain Drives, 2009).

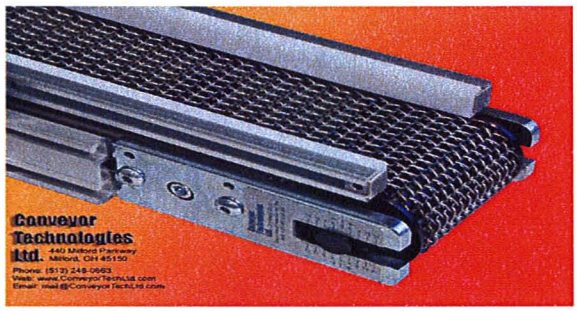

Figure 6 A Wire Belt Conveyor

2.2 Modeling of the Belt Drive System

Extensive research has been done on the modeling of belt drives over the past few decades. According to Leamy & Wasfy (2002), belt drive research can be categorized into belt-drive mechanics studies (Bechtel, Vohra, Jacob & Carlson, 2000) and dynamic-response studies (Leamy & Perkins 1998).

Representative papers of belt drive mechanics have been reviewed up to 1992 in (Abrate, 1992). These studies cover the topics: (1) fractional mechanics of the belt drives under steady operating conditions (Fawcett, 1981), (2) the classical creep theory of belt drive operation (Leamy & Wasfy, 2002a), (3) the mechanics of the belt drive with belt shear effects, seating/unseating effects and radial compliance effects (Firbank, 1970), (4) derivation of the power loss expression and the efficiency limit of a belt drive assuming

the validity of the classical creep theory (Bechtel, Vohra, Jacob & Carlson, 2000).

Dynamic Response studies of the belt drives initiated in the 1990s due to the wide application of automotive serpentine belt drives to crankshaft excitation. These studies have concentrated on the rotational response of the pulleys and/or the transverse response of the axially moving belt with the simplified linear stretching and viscous damping models of the belt-pulley contacts. Barker, Oliver and Brieg (1991) studied the belt drive tensions resulting from engine acceleration. Hwang, Perkins, Ulsoy and Mechstroth (1994) studied the periodic response of the serpentine belt drive systems. Beikmann, Perkins and Ulsoy (1996) studied the coupled rotational and transverse response of a prototypical serpentine belt drive system. Leamy and Perkins (1998) studied the periodic response of serpentine belt drives to harmonic excitation from the crankshaft. Kraver, Fan and Shah (1994) studied the effects of damping on the rotational response of a multiple pulley/flat belt systems.

Two groups of studies on the belt drive mechanics and the belt drive dynamics have little connection to each other due to: (1) the omission of dynamic excitation in the belt drive mechanics studies and (2) the lack of the true fractional belt pulley modeling in the serpentine drive dynamic response studies (Leamy & Wasfy, 2002). Some efforts have been made to bridge the gap by studying simplified dynamic models for small and large rotational speeds. Leamy, Barber and Perkins (1998) investigated the dynamic response of a belt drive subject to a train of harmonic tension waves by using a simplified model for the belt pulley contact. Leamy and Wasfy (2002) developed a dynamic finite element model to determine the transient and steady state response of a prototypical belt drive

system. In the dynamic finite element modeling, the belt is modeled using truss elements, the pulleys are modeled as rotating circular constraints and the fractional contact between the pulleys and the belt is modeled using a penalty function together with a Coulomb-like tri-linear creep-rate dependent friction law. Kerkkanen, Vallejo and Mikkola (2006) used an approximate Coulomb friction law to model the belt-pulley contact forces. This approximation avoided the numerical problems caused by the discontinuity of the Coulomb's friction law and removed the need for different sets of equations to model the sticking contacts. Stewart and Trinkle (1996) introduced a time stepping method to address unilateral contact problems. This method was applied by Jean (1999) to model contacts between deformable bodies, and has been widely employed in recent years (Heemels, Camhbel & Schumacher, 2000; Chakraborty, Berard, Akella & Trinkle, 2009).

This research is focused on physical modeling, system identification, and control of an experimental two-pulley laboratory-scale belt drive system. This system consists of an AC drive, a Permanent Magnet Synchronous Motor (PMSM), a two-pulley belt drive and two digital encoders as shown in Figure 7.

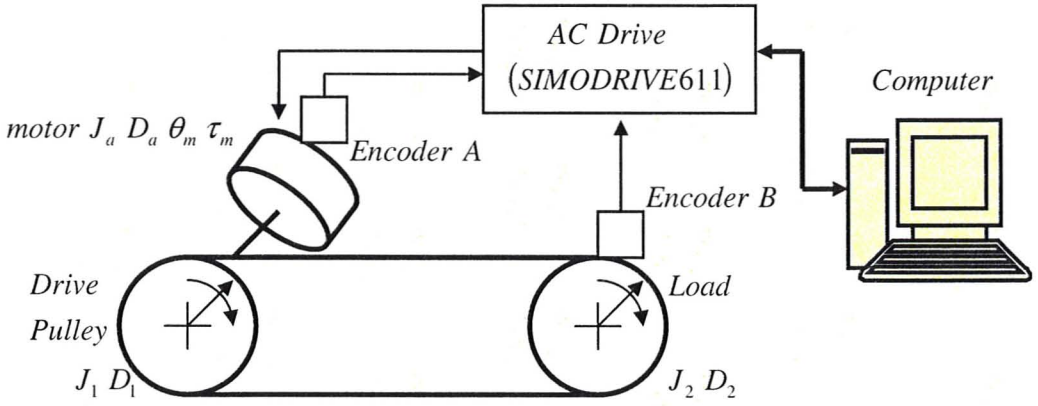

Figure 7 Structure of a Belt Drive System

The belt drive consists of a Vee belt and two equal-radius pulleys. The Vee belt is pre-tensioned by adjusting the centerline-distance between the drive and the driven pulleys. The drive pulley is rigidly coupled to the output shaft of the PMSM, and grips the belt through a dry frictional interface. A similar interfacing happens between the belt and the driven pulley. The driven pulley is made of stainless steel which is treated as the load of the system.

The belt drive system is mathematically modeled to provide the *a priori* information needed for system identification. The AC drive, the electric motor, and the belt drive are modeled separately, then jointed together to form an overall white-box model. Later in Section 3, a black-box model of the belt drive system is obtained by using system identification.

The drive pulley, powered by the PMSM grips the belt through static friction to move the load. Since the belt drive system operates in one direction only, modeling of backlash is not considered in this research. The AC drive and the PMSM are modeled as

an electrical sub-system as described later in Section 2.2.1. The belt drive is modeled as a mechanical sub-system with a torque input and a velocity output as described later in Section 2.2.2. All modeled subsystems are combined to obtain an overall system transfer function model in Section 2.2.3.

2.2.1 Modeling of the Electrical Subsystem

A synchronous motor operates at a constant speed uniquely related to the AC supply frequency. A Permanent Magnet Synchronous Motor (PMSM) consists of a stator containing three phase windings, ether in wye or delta form, and a rotor including a permanent magnet. The AC motor considered in this research has a wye winding configuration. The PMSMs are widely used in industry, particularly at low power range up to 100 kW.

A three-phase motor stator can be represented by an equivalent two-phase representation with d and q corresponding to the direct and quadrature axes (Bose, 2001) as shown in Figure 8.

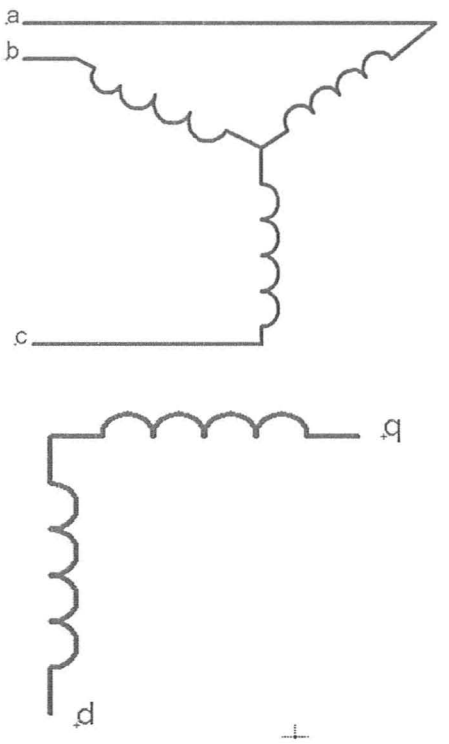

Figure 8 Three-Phase Balanced Stator Windings and Two-Phase Equivalent

The variables of the three-phase windings in the stationary frame *a-b-c* can be transformed into the variables in the two-phase rotatory frame *d-q* by:

$$\begin{bmatrix} v_a \\ v_b \\ v_c \end{bmatrix} = \begin{bmatrix} \cos(\omega_s t) & \sin(\omega_s t) & 1 \\ \cos(\omega_s t - 120^o) & \sin(\omega_s t - 120^o) & 1 \\ \cos(\omega_s t + 120^o) & \sin(\omega_s t + 120^o) & 1 \end{bmatrix} \begin{bmatrix} v_q \\ v_d \\ v_o \end{bmatrix} \qquad (2.1)$$

where v_a, v_b, and v_c are the phase voltages in the stator windings, and v_d, v_q are the voltages in the *d-q* frame, and v_o is the zero sequence[1] component required to yield a

[1] The zero sequence is a set of components with equal magnitudes and no angular displacement between the phases. The zero sequence components are always balanced (Fehr, 2010), so it produces no net flux linkage to rotor permanent magnet, and it may or may not be present.

unique transformation of the three stator phase quantities; ω_s is the inverter frequency. Inversely, the phase variables in *d-q* frame can be resolved from Equation (2.1) by:

$$\begin{bmatrix} v_q \\ v_d \\ v_o \end{bmatrix} = \frac{2}{3}\begin{bmatrix} \cos(\omega_s t) & \cos(\omega_s t - 120^o) & \cos(\omega_t + 120^o) \\ \sin(\omega_s t) & \sin(\omega_s t - 120^o) & \sin(\omega_s t + 120^o) \\ 0.5 & 0.5 & 0.5 \end{bmatrix}\begin{bmatrix} v_a \\ v_b \\ v_c \end{bmatrix} \qquad (2.2)$$

The permanent magnets used in the PMSM are of high resistivity, the current induced in the rotor is negligible. It is assumed that: there is no inductance leakage in the rotor; the permeability of the magnetic material is considered unity; the air gap inductances in the direct axis and the quadrature axis are the same; the magnetic flux saturation is negligible; the induced electromagnetic force (EMF) is sinusoidal; there are no field current dynamics; and there is no cage on the rotor. A *d-q* equivalent circuit of the PMSM (Shahat & Shewy, 2010) is shown in Figure 9.

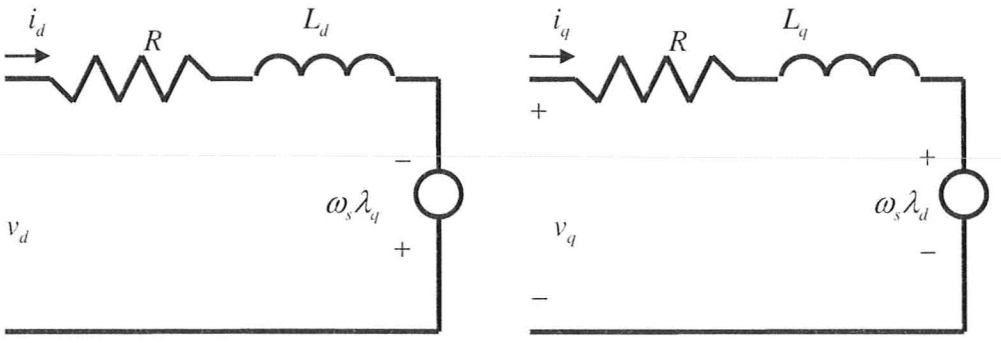

Figure 9 Equivalent d-q Circuits of a PMSM Stator

Application of the Kirchhoff's voltage law to the stator *d-q* frame gives:

$$v_d = Ri_d + \frac{d\lambda_d}{dt} - \omega_s\lambda_q \qquad (2.3)$$

$$v_q = Ri_q + \frac{d\lambda_q}{dt} + \omega_s \lambda_d \qquad (2.4)$$

$$\lambda_d = L_d i_d + \lambda_f \qquad (2.5)$$

$$\lambda_q = L_q i_q \qquad (2.6)$$

where i_d, i_q are stator currents; L_d, L_q are stator inductances; λ_d, λ_q are stator flux linkages in the d and q axes; R is the stator resistance; λ_f is the flux linkage due to the rotor permanent magnet (PM) linking the stator. The electric torque generated on the PMSM is given by (Pillay & Krishna, 1989):

$$T_e = \frac{3}{2} P\left[\lambda_f i_q + (L_d - L_q) i_d i_q\right] \qquad (2.7)$$

where P denotes the number of the pole pairs of the PM on the rotor. The mechanical dynamics of the PMSM drive system is given as:

$$T_e = T_L + B\omega_r + J\frac{d\omega_r}{dt} \qquad (2.8)$$

where T_L is the load torque, B is the damping coefficient, ω_r is the rotor angular velocity related to the inverter frequency by $\omega_r = \omega_s / P$, and J is the rotor moment of inertia. The Mathematical model of the PMSM can be summarized in state space representation as:

$$\begin{bmatrix} di_d/dt \\ di_q/dt \\ d\omega_r/dt \end{bmatrix} = \begin{bmatrix} -R/L_d & P\omega_r L_q/L_d & 0 \\ -P\omega_r L_d/L_q & -R/L_q & -P\lambda_f/L_q \\ -1.5P\lambda_q/J & 1.5P\lambda_d/J & -B/J \end{bmatrix} \begin{bmatrix} i_d \\ i_q \\ \omega_r \end{bmatrix} + \begin{bmatrix} v_d/L_d \\ v_q/L_q \\ -T_L/J \end{bmatrix} \qquad (2.9)$$

Taking account of Equations (2.5) and (2.6), Equation (2.9) contains the products of the states such as $\omega_r i_d$, $\omega_r i_q$ and $i_d i_q$. The mathematical model of the PMSM is a multivariable,

nonlinear, strongly coupled system. Its speed and electromagnetic torque are difficult to be controlled by external signal.

One effective control strategy is the *vector control* (Bose, 2001). The vector control algorithm was invented in the 1970s, in which an induction motor can be controlled like a separately excited DC motor. The vector control is also known as decoupling control, orthogonal control, transvector control or field oriented control because of its DC machine-like performance. With the vector control, the state i_d can be governed by a current controller to satisfy the condition $i_d = 0$ so that the PMSM model is decoupled and simplified. The schematic diagram of the vector controlled PMSM (Zambada, 2010) is shown in Figure 10.

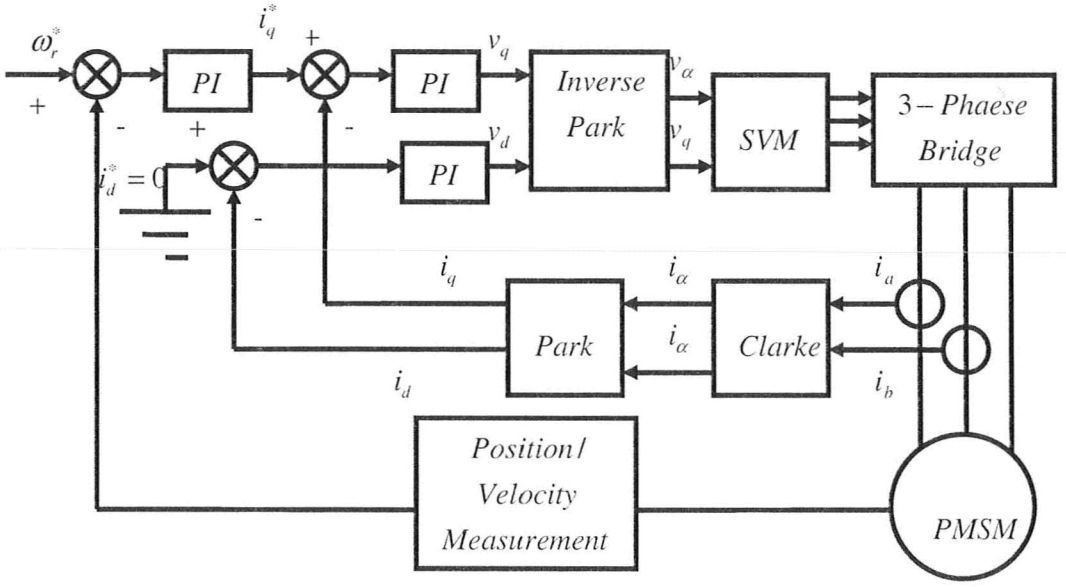

Figure 10 Block Diagram of Vector Control

In Figure 10, ω_r^* is the desired angular velocity of the rotor; i_d^* and i_q^* is the desired current of the direct axis and the quadrature axis, respectively; i_α, i_β, v_α and v_β are currents and voltages in the Clarke transform. There are two control loops, three PI controllers, four transforms and one Space Vector Modulation (SVM) in the vector control process. The outer loop consists of the velocity control. The inner loop consists of the current vector control. The Clarke transform uses three-phase currents i_a, i_b and i_c to calculate currents in the fixed two-phase orthogonal axis i_α and i_β. These currents are transformed to currents i_d and i_q in the rotary two-phase orthogonal axis (Zambada, 2010; Texas Instruments, 1997; Texas Instruments, 1998). The Clarke and Park transforms are shown in shown in Figure 11. The Clarke and Park transforms and their inversions are presented mathematically in Table 1. The SVM is to generate pulse-width modulation signals for 3-pahse motor voltage signals.

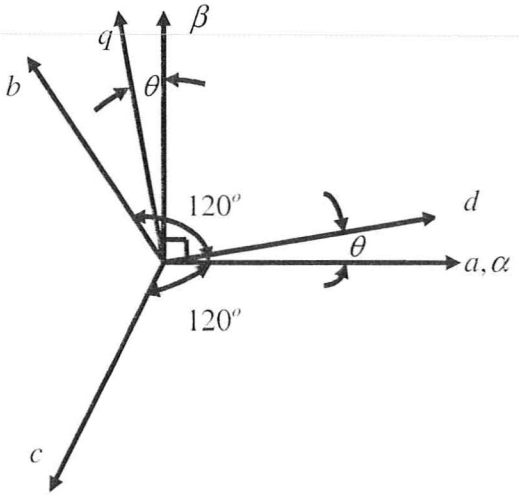

Figure 11 Clarke and Park Transforms

Table 1 *Clarke and Park Transforms and Their Inversions*

Clarke	$i_\alpha = i_a$ $i_\beta = \frac{1}{\sqrt{3}}(i_a + 2i_b)$ $i_a + i_b + i_c = 0$
Inverse Clarke	$i_a = i_\alpha$ $i_b = -\frac{1}{2}i_\alpha + \frac{\sqrt{3}}{2}i_\beta$ $i_c = -\frac{1}{2}i_\alpha - \frac{\sqrt{3}}{2}i_\beta$
Park	$i_d = i_\alpha \cos(\theta) + i_\beta \sin(\theta)$ $i_q = -i_\alpha \sin(\theta) + i_\beta \cos(\theta)$
Inverse Park	$i_\alpha = i_d \cos(\theta) - i_q \sin(\theta)$ $i_\beta = i_d \sin(\theta) + i_q \cos(\theta)$

With the vector control, the current in the direct axis is governed:

$$i_d = 0 \qquad (2.10)$$

The stator flux linkage in *d*-axis of Equation (2.5) becomes:

$$\lambda_d = \lambda_f \qquad (2.11)$$

The developed electromagnetic torque of Equation (2.7) then simplifies to:

$$T_e = 1.5P\lambda_f i_q \qquad (2.12)$$

The resulting mathematical model of the PMSM of Equation (2.9) is subsequently simplified to:

$$\begin{cases} \dfrac{di_q}{dt} = -\dfrac{R}{L_q} i_q - \dfrac{P\lambda_f \omega_r}{L_q} + \dfrac{v_q}{L_q} \\ \dfrac{d\omega_r}{dt} = \dfrac{1.5P\lambda_f i_q}{J} - \dfrac{B\omega_r}{J} - \dfrac{T_L}{J} \end{cases} \qquad (2.13)$$

The equivalent schematic diagram of the speed loop is determined from Equations (2.8) and (2.13) as shown in Figure 12, where V_p is the input voltage signal, K_ω is the gain constant, K_P and K_I are the proportional and the integral gains, respectively.

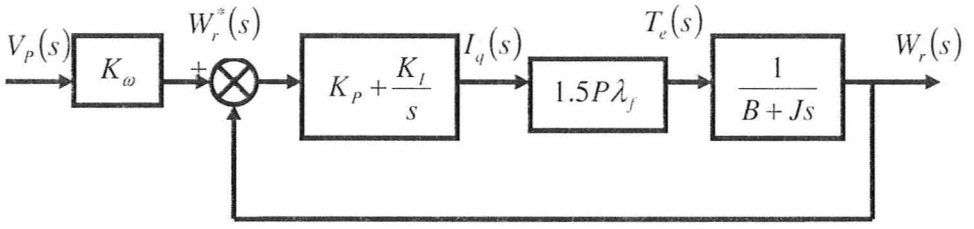

Figure 12 Equivalent Schematic Diagram of the PMSM Speed Loop

The model of the vector controlled PMSM can be obtained as:

$$\frac{W_r(s)}{V_P(s)} = \frac{1.5K_\omega P\lambda_f (K_P s + K_I)}{Js^2 + (B + 1.5K_P P\lambda_f)s + 1.5K_I P\lambda_f} \qquad (2.14)$$

2.2.2 Modeling of the Belt Drive

As stated earlier, the belt drive used in this research consists of a Vee belt and two equal-radius pulleys. The input to the belt drive is the velocity and torque of the drive pulley and the output is the velocity and torque of the driven pulley. The modeling objective is to derive the relationship between the input and the output of the belt drive.

A linear coordinate system, following the moving direction of the belt, can be constructed which decomposes the belt into two spans and two friction interfaces as

shown in Figure 13.

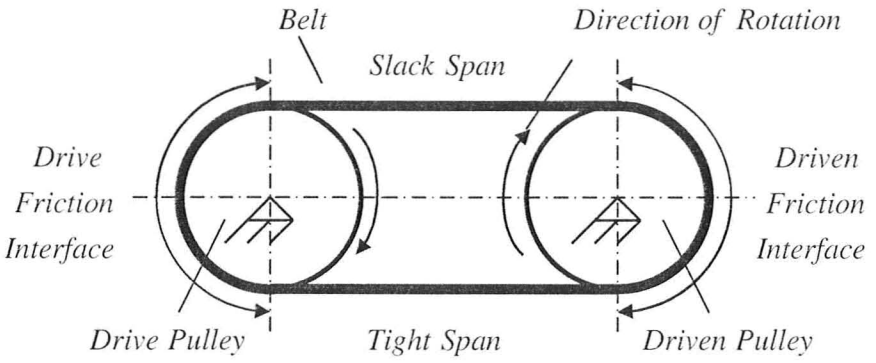

Figure 13 Sketch of a Two-Pulley Belt Drive Structure

The belt is assumed linearly elastic and extensible. The tension of the belt alternates periodically from the lowest value at the *slack span* to the highest value at the *tight span* during each cycle of operation. Sliding of the belt on the pulleys exists occurs along the drive and the driven friction interfaces. For a section of the belt on the pulley, occurrence of slip depends on the difference between the resultant force in the tangential direction and the kinetic friction force as discussed later in this section. There must be one point where the resultant force in the tangential direction equals to the kinetic friction force. This point divides the frictional interface into a slip zone and a non-slip zone. The overall belt is divided into two slip zones, two non-slip zones and two free spans as shown in Figure 14. To model the output velocity and the torque of the belt drive, the tension and the velocity of the belt in different zones need to be determined as follows.

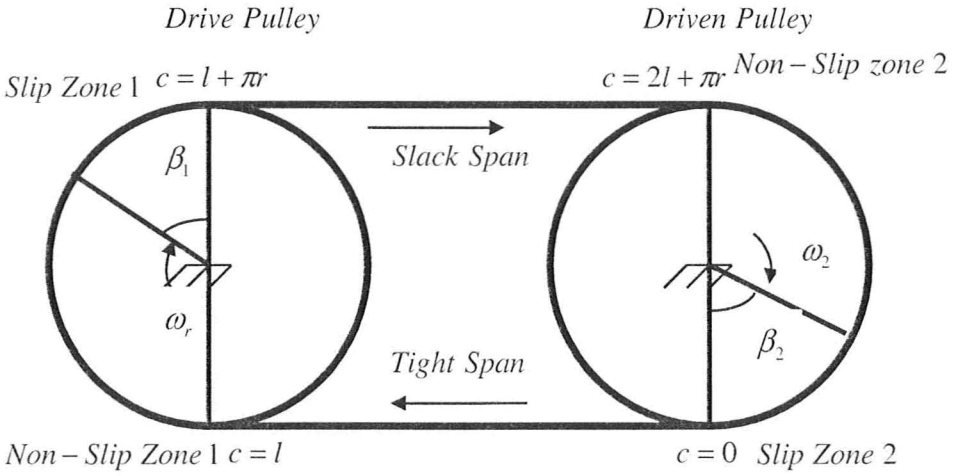

Figure 14 Different Zones and Spans of the Two-Pulley Belt Drive

The pulley radius r, the transmitted torque T_L, the driving pulley angular velocity ω_r, the initial belt tension F_{init}, the belt elastic modulus k_e, and the coefficient of friction μ between the belt and the pulleys are assumed to be known. The unknowns to be determined are the angular velocity of the driven pulley ω_2, the angles β_1 and β_2 over which the belt is slipping on the drive and driven pulleys, the belt tension $F(c)$, the velocity $v(c)$ at all locations along the belt, and the normal and tangential forces per length $n(c)$ and $f(c)$. The letter c denotes the linear coordinate of the belt.

Modeling of the Belt in the Free Spans

The free spans consist of the tight span $(0 \le c \le l)$ and the slack span $(l + \pi r \le c \le 2l + \pi r)$, where l is the center-distance of the drive and the driven pulleys. It is assumed that the weight of the belt in free spans is negligible compared to

the belt tension. The free body diagram of a section of the belt in a free span is shown in Figure 15 (Bechtel, Vohra, Jacob & Carlson, 2000).

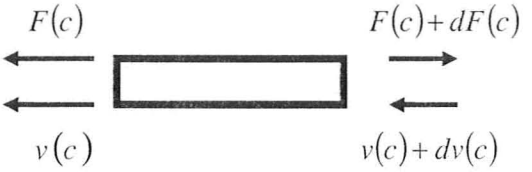

Figure 15 Free Body Diagram of a Section of Belt in a Free Span

Application of the Newton's Second Law to a section of the belt gives the governing equation as:

$$(F(c)+dF(c))-F(c)=\Omega(v(c)+dv(c))-\Omega v(c)$$

The above equation can be simplified as:

$$dF(c)=\Omega dv(c) \qquad (2.15)$$

where $\Omega=\rho A(c)v(c)$ is the mass flow rate, ρ is the mass per volume and $A(c)$ is the cross-sectional area of the belt. Integrating Equation (2.15) gives:

$$F(c)=\Omega v(c)+c_0 \qquad (2.16)$$

where c_0 is the constant of integration corresponding to initial conditions.

For a linear elastic belt, the tension at point c on the belt linearly depends only on the strain of the belt at that point (Bechtel, Vohra, Jacob & Carlson, 2000). So that a model which relates the stretch to the tension and the velocity is obtained as:

$$F(c)=k_e\varepsilon(c) \qquad (2.17)$$

$$v(c)=v_{ref}(1+\varepsilon(c)) \qquad (2.18)$$

where v_{ref} is the velocity of the belt in the reference state[2], k_e is the elastic modulus and ε is the strain of the belt at point c which is determined as:

$$\varepsilon(c) = \frac{dl(c)}{dl_{ref}} - 1 \qquad (2.19)$$

where $dl(c)$ denotes the length of an infinitesimal section of the belt at location c and dl_{ref} denotes the length of the section at the reference state. Combing Equations (2.17) and (2.18) gives:

$$F(c) = \frac{k_e}{v_{ref}} v(c) - k_e \qquad (2.20)$$

Comparing Equations (2.16) and (2.20), for the coefficient of $v(c)$ the condition $\Omega = \frac{k_e}{v_{ref}}$ should hold. This condition can not be realized since the mass flow rate Ω is proportional to the velocity $v(c)$, while $\frac{k_e}{v_{ref}}$ can not be related to the velocity $v(c)$. So a relevant solution of Equation (2.15) in the free span for isothermal motion can be obtained as (Bechtel, Vohra, Jacob & Carlson, 2000):

$$F(c) = Constant \text{ and } v(c) = Constant \qquad (2.21)$$

From Equation (2.21), the tension and the velocity of the belt are constant in the free spans. It is assumed that the bearings are frictionless and the summing moments on either pulley is zero. The tight span tension F_t and the slack span tension F_s can be determined as (Bechtel, Vohra, Jacob & Carlson, 2000):

[2] The reference state is a fictitious state in which the strain of the belt is zero.

$$F_t = F_{init} + \frac{T_L}{2r} \qquad (2.22)$$

$$F_s = F_{init} - \frac{T_L}{2r} \qquad (2.23)$$

The belt first enters the non-slip zones when it exits the free spans. The velocities on the tight span v_t and on the slack span v_s can be determined as (Bechtel, Vohra, Jacob & Carlson, 2000):

$$\begin{aligned} v_t &= r\omega_r \\ v_s &= r\omega_2 \end{aligned} \qquad (2.24)$$

Applying Equation (2.20) to the tight span gives:

$$v_{ref} = \frac{k_e v_t}{k_e + F_t} \qquad (2.25)$$

Modeling of the Belt on the Pulleys

An extendable belt on the pulley of radius r and angular velocity ω is shown in Figure 16.

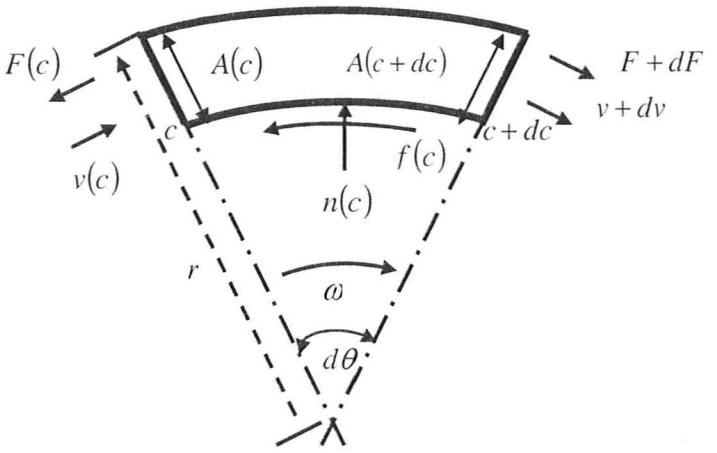

Figure 16 Free Body Diagram of a Section of Belt on the Drive Pulley

It is assumed that the circumferential point c $\left(l \le c \le l+\pi r \text{ or } 2l+\pi r \le c \le 2l+2\pi r\right)$ is fixed in space and the motion is steady. The condition at location c is independent of time. Consider a free body of the belt dc at location c subtending an angle $d\theta$. The belt enters the portion at location c with a tension $F(c)$, a linear velocity $v(c)$ and a sectional area $A(c)$, and exits the portion at location $c+dc$ with a tension $F(c)+dF(c)$, a linear velocity $v(c)+dv(c)$ and a sectional area $A(c+dc)$. Let $f(c)$ and $n(c)$ be the projections of the resultant force per unit length in the tangential and normal directions, respectively. The positive direction of $f(c)$ is opposite to that of the angular velocity of the pulley. The positive direction of $n(c)$ is radially outward.

The momentum projected in the tangential and normal directions are, respectively (Bechtel, Vohra, Jacob & Carlson, 2000):

$$-F(c)\cos\left(\frac{d\theta}{2}\right)+\left(F(c)+dF(c)\right)\cos\left(\frac{d\theta}{2}\right)-f(c)rd\theta=\Omega dv(c)\cos\left(\frac{d\theta}{2}\right) \qquad (2.26)$$

$$-F(c)\sin\left(\frac{d\theta}{2}\right)-(F(c)+dF(c))\sin\left(\frac{d\theta}{2}\right)+n(c)dc=$$
$$-\Omega v\sin\left(\frac{d\theta}{2}\right)-\Omega(v(c)+dv(c))\sin\left(\frac{d\theta}{2}\right) \qquad (2.27)$$

where $\Omega=\rho A(c)v(c)$ denotes the mass flow rate of the belt, ρ is the mass per volume. For steady motion, the conservation of mass requires that the mass flow rate Ω is constant (Bechtel, Vohra, Jacob & Carlson, 2000). For small $d\theta$ with $\frac{d\theta}{2}<<1$, then $\cos\left(\frac{d\theta}{2}\right)\approx 1$ and $\sin\left(\frac{d\theta}{2}\right)\approx\frac{d\theta}{2}$. Equations (2.26) and (2.27) can be simplified, with the products of infinitesimal quantities neglected, to:

$$dF(c)-f(c)dc=\Omega dv(c) \qquad (2.28)$$

$$n(c)=\frac{F(c)-\Omega v(c)}{r} \qquad (2.29)$$

The normal force per unit length $n(c)$ in Equation (2.29) has to be greater or equal to zero to keep the belt in contact with the pulley, that is:

$$F(c)\geq\Omega v(c) \qquad (2.30)$$

Expression (2.30) sets the lower bound on the belt tension $F(c)$.

In practice the magnitude of the tension $F(c)$ is determined by the torque transmitted by the belt drive as discussed latter in this section. The tension $F(c)$ is related to the tangential projection $f(c)$ by Equation (2.28). The tangential projection $f(c)$ has an opposite direction but quantitatively equals to the fractional force generated between the contact surfaces. The tangential projection $f(c)$ increases accordingly following the

increment of the torque transmitted by the belt drive. So does the frictional force until it reaches its upper bound determined by the kinetic frictional force.

For a flat belt, the kinetic frictional force is calculated as $f(c)=\mu n(c)$ with μ denoting a frictional modulus and $n(c)$ denoting the normal force. In case of our Vee belt, the cross-section of the belt is shown in Figure 17.

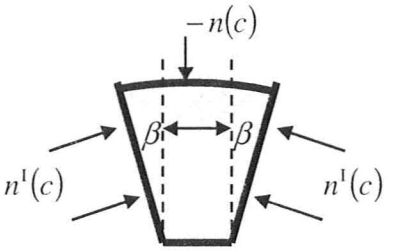

Figure 17 Cross Section of Vee Belt

The equivalent fractional modulus of the Vee belt can be calculated as (Solid Mechanics, 2010):

$$\begin{cases} n(c)=2n^1(c)\sin\beta \\ n(c)\mu_\beta = 2n^1(c)\mu \end{cases} \Rightarrow \mu_\beta = \frac{\mu}{\sin\beta} \qquad (2.31)$$

where β is the wedging angle of the Vee belt, $n^I(c)$ is pressure on the contacting surface and $n(c)$ is the resultant normal force as shown in Figure 17.

As mentioned earlier in this section, there exists a region in the drive (or driven) friction interface where the resultant force in tangential direction $f(c)$ is greater than the kinetic frictional force, the sliding of the belt on the pulley happens; this region is called a slip zone. The drive friction interface is divided into the non-slip zone 1 and the slip zone

1, and the driven pulley is divided into the non-slip zone 2 and the slip zone 2 as shown in Figure 14.

In the non-slip zone 1, the belt attaches to the drive pulley without slip. The tension and the velocity of the belt in this zone are the same as that in the tight span. Since there is no tendency to slip in this zone, the tangential force (frictional force) is zero. The normal force is calculated using Equation (2.29). In summary, the belt tension, the velocity, the tangential force and the normal force in the non-slip zone 1 can be obtained as:

$$F(c)=F_t,\ v(c)=v_t,\ f(c)=0,\ n(c)=\frac{F_t-\Omega v_t}{r} \qquad (2.32)$$

In the slip-zone 1, the belt enters this zone with velocity $v(c)=v_t=r\omega_r$ and exits the zone with velocity $v(s)=v_s=r\omega_2$. The velocity is gradually reduced in this zone. The belt is moving slower than the drive pulley surface. The friction is kinetic with a direction being the same as that of the pulley motion. Rewriting Equations (2.28) and (2.29) with $f(c)=-\mu_\beta n(c)$ gives:

$$\frac{d(F(c)-\Omega v(c))}{F(c)-\Omega v(c)}=-\mu_\beta d\theta \qquad (2.33)$$

Integrating Equation (2.33) over the entire slip-zone 1[3] and solving for β_1 produce:

$$\beta_1=\frac{1}{\mu_\beta}\ln\left(\frac{F_t-\Omega v_t}{F_s-\Omega v_s}\right) \qquad (2.34)$$

---

[3] $\ln(F(c)-\Omega v(c))\Big|_{(F_t,v_t)}^{(F_s,v_s)}=-\mu_\beta\theta\Big|_{\pi-\beta_1}^{\pi}\Rightarrow\ln\left(\frac{F_s-\Omega v_s}{F_t-\Omega v_t}\right)=-\mu_\beta(\pi-(\pi-\beta_1))=-\mu_\beta\beta_1$

Integrating Equation (2.33) from the entering point $c = L + (\pi - \beta_1)r$ to any point in the slip zone 1[4] and combining the result with Equations (2.20) and (2.25) give:

$$F(c) = \frac{k_e + F_t}{k_e + F_t - \Omega v_t}\left(k_e + (F_t - \Omega v_t)e^{-\mu_\beta(\theta(c)-(\pi-\beta_1))}\right) - k_e \qquad (2.35)$$

$$v(c) = \frac{v_t}{k_e + F_t - \Omega v_t}\left(k_e + (F_t - \Omega v_t)e^{-\mu_\beta(\theta(c)-(\pi-\beta_1))}\right) \qquad (2.36)$$

The normal force $n(c)$ and the tangential projection $f(c)$ in the slip zone 1 are determined by substituting Equations (2.35) and (2.36) into Equation (2.29) as:

$$n(c) = \frac{F_t - \Omega v_t}{r} e^{-\mu_\beta(\theta(c)-(\pi-\beta_1))} \qquad (2.37)$$

$$f(c) = -\frac{\mu_\beta(F_t - \Omega v_t)}{r} e^{-\mu_\beta(\theta(c)-(\pi-\beta_1))} \qquad (2.38)$$

In the non-slip zone 2, the belt attaches to the driven pulley without slip. The velocity is v_s and belt tension is F_s. Since there is no tendency to slip in this zone, the tangential force (friction force) is zero. The normal force is calculated using Equation (2.29). In summary, the belt tension, the velocity, the tangential force and the normal force in the non-slip zone 2 can be obtained as:

$$F(c) = F_s,\ v(c) = v_s,\ f(c) = 0,\ n(c) = \frac{F_s - \Omega v_s}{r} \qquad (2.39)$$

In the slip-zone 2, the belt enters this zone with velocity $v(c) = v_s = r\omega_2$ and exits the zone with velocity $v(c) = v_t = r\omega_r$. The velocity is gradually increased in this zone. The belt is moving faster than the driven pulley surface. The friction is kinetic with a direction

[4] $F(c) - \Omega v(c) = (F_t - \Omega v_t)e^{-\mu_\beta(\theta(c)-(\pi-\beta_1))}$

being the same as that of the pulley motion. Rewriting Equations (2.28) and (2.29) with $f(c)=\mu_\beta n(c)$ gives (Bechtel, Vohra, Jacob & Carlson, 2000):

$$\frac{d(F(c)-\Omega v(c))}{F(c)-\Omega v(c)}=\mu_\beta d\theta \qquad (2.40)$$

Integrating Equation (2.40) over the entire slip zone 2 on the driven pulley [5] and solving for β_2 gives:

$$\beta_2=\frac{1}{\mu_\beta}\ln\left(\frac{F_t-\Omega v_t}{F_s-\Omega v_s}\right) \qquad (2.41)$$

Integrating Equation (2.40) from the entering point $c=2l+(2\pi-\beta_2)r$ to any point in the slip zone 2[6] and combining the result with Equations (2.20) and (2.25) gives:

$$F(c)=\frac{k_e+F_t}{k_e+F_t-\Omega v_t}\left(k_e+(F_s-\Omega v_s)e^{\mu_\beta(\theta(c)-(2\pi-\beta_2))}\right)-k_e \qquad (2.42)$$

$$v(c)=\frac{v_t}{k_e+F_t-\Omega v_t}\left(k_e+(F_s-\Omega v_s)e^{\mu_\beta(\theta(c)-(2\pi-\beta_2))}\right) \qquad (2.43)$$

The normal force $n(c)$ and the tangential projection $f(c)$ in the slip zone 2 are determined by substituting Equations (2.42) and (2.43) into Equation (2.29) as (Bechtel, Vohra, Jacob & Carlson, 2000):

$$n(c)=\frac{1}{r}(F_s-\Omega v_s)e^{\mu_\beta(\theta(c)-(2\pi-\beta_2))} \qquad (2.44)$$

$$f(s)=\frac{\mu_\beta}{r}(F_s-\Omega v_s)e^{\mu_\beta(\theta(c)-(2\pi-\beta_2))} \qquad (2.45)$$

---

[5] $\ln(F(c)-\Omega v(c))\big|_{c=2l+(2\pi-\beta_2)r}^{2l+2\pi r}=\mu_\beta\theta\big|_{\theta=2\pi-\beta_2}^{2\pi}\Rightarrow\ln\left(\frac{F_t-\Omega v_t}{F_s-\Omega v_s}\right)=\mu_\beta(2\pi-(2\pi-\beta_2))=\mu_\beta\beta_2$

[6] $F(c)-\Omega v(c)=(F_s-\Omega v_s)e^{\mu_\beta(\theta(c)-(2\pi-\beta_2))}$

In summary, a belt drive consists of two spans, two non-slip zones and two slip zones. In every zone, the belt demonstrates different characteristics which are summarized in Table 2.

Table 2 *Summary of the Belt Drive Characteristics*

	Belt tension $F(c)$	*Belt velocity* $v(c)$	*Belt normal force* $n(c)$	*Belt frictional force* $f(c)$
Tight span	$F_t = F_{init} + \frac{T_L}{2r}$	$v_t = r\omega_r$	0	0
Non slip zone 1	$F(c) = F_t$	$v(c) = v_t$	$n(c) = \frac{F_t - \Omega v_t}{r}$	0
Slip zone 1	*Equation* (2.35)	*Equation* (2.36)	*Equation* (2.37)	*Equation* (2.38)
Slack span	$F_s = F_{init} - \frac{T_L}{2r}$	$v_s = r\omega_2$	0	0
Non slip zone 2	$F(c) = F_s$	$v(c) = v_s$	$n(c) = \frac{F_s - \Omega v_s}{r}$	0
Slip zone 2	*Equation* (2.42)	*Equation* (2.43)	*Equation* (2.44)	*Equation* (2.45)

The torque transmitted to the driven pulley can be calculated by combining Equations (2.22) and (2.23) as:

$$T_L = (F_t - F_s)r \qquad (2.46)$$

The stretching and slipping of the belt on the pulleys cause a portion of velocity loss which means $\omega_2 < \omega_r$. The relationship of the input and the output velocities can be determined by applying Equation (2.20) to the tight and slack spans and combining the results with Equations (2.22), (2.23), as such:

$$\frac{\omega_2}{\omega_r} = \frac{2r(k_e + F_{init}) - T_L}{2r(k_e + F_{init}) + T_L} \qquad (2.47)$$

2.2.3 Modeling of the Overall Belt Drive System

The subsystems of the belt drive system: the electrical subsystem and the belt drive have been modeled in the previous subsections. For a belt drive system consisting of two equal-radius pulleys, the torque applied to the driven pulley comes from the drive pulley. The velocity of the driven pulley is reduced compared to that of the drive pulley. These conclusions infer that the external load added to the belt drive can be directly reflected to the electrical subsystem without changing the dynamics of the overall system except that the velocity of the driven pulley is a reduced version. The block diagram of the overall belt drive system is shown in Figure 18.

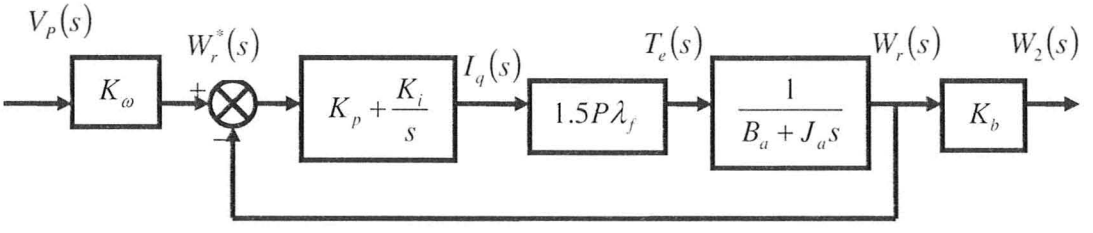

Figure 18 Block Diagram of the Belt Drive System

where $B_a = B + B_L$ is a summation of the damping coefficient of the electric motor and that of the load added on the driven pulley; $J_a = J + J_L$ is a summation of the moment of inertia of the electrical motor and that of load added on the driven pulley; $K_b = \frac{W_2(s)}{W_r(s)} = \frac{2r(k_e + F_{init}) - T_L}{2r(k_e + F_{init}) + T_L}$ is the ratio of the velocity of the driven pulley to that of the drive pulley; and B_L, J_L is the damping coefficient and the moment of inertia of the load on the driven pulley, respectively. The transfer function of the belt drive system is obtained as:

$$\frac{W_2(s)}{V_P(s)} = \frac{1.5K_bK_\omega P\lambda_f(K_P s + K_I)}{J_a s^2 + (B_a + 1.5K_P P\lambda_f)s + 1.5K_I P\lambda_f} \qquad (2.48)$$

Chapter 3 System Identification of the Experimental Belt Drive System

System identification is a process in which a mathematical model of a dynamic system is obtained from a combination of prior knowledge and experimentation involving select measurements. System identification is a structured method (Habibi, 2009). The identification process consists of the following steps: (1) experimental design is made with respect to the intended use of the model and is largely made by a careful selection of input signals such that maximum level of information can be extracted from measurements; (2) a model is obtained by optimization which attempts to provide the best fit to the measurements. The model so generated is known as *empirical or black-box model* as the parameters of the model have no physical significance. The black-box model generated through system identification does not require excessive time, effort and costs typically associated with physical modeling; yet it is usually accurate, and may be directly utilized in many applications including control design.

In the following sections the identification process is applied to our experimental belt drive system.

3.1 System Description and Model Specification

A system is an object with variables that interact and produce observable signals (Ljung, 1987). The observable signals are of interest and are usually called outputs. The system is also affected by external stimuli including inputs, and disturbances. The

disturbances can be divided into measurable ones and non-measurable ones that can only be observed through their influence on the system. The relationship among the observed signals is called a model of the system.

The first step in any system identification process is a consideration of a priori information on the system. This often entails a description of the system under consideration and an initial specification of its model if available. The model specification refers to the determination of the independent variables being included in or excluded from the system model. Non-appropriate inclusion of irrelevant variables or exclusion of relevant variables often results in model specification error. Another factor affecting the model specification is the purpose of the system identification. For example, the identification mainly focuses on low frequency range for the feed-forward control design and on the closed-loop system bandwidth for the closed-loop control design.

In this research, the system identification is applied to an experimental belt drive system as shown in Figure 19.

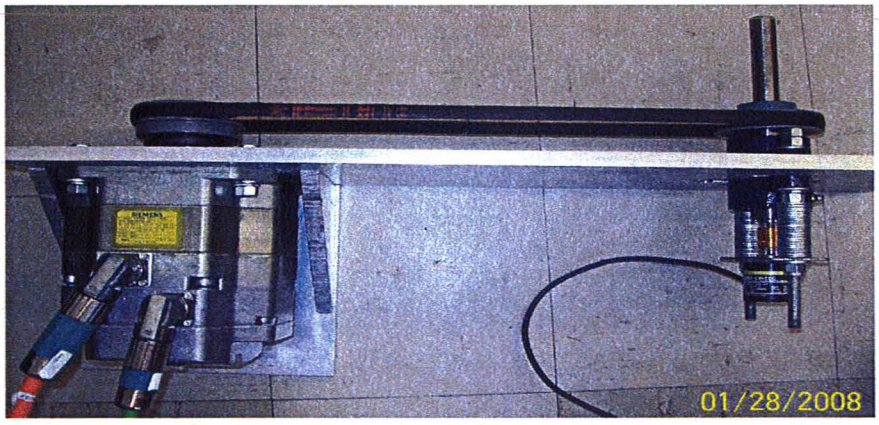

Figure 19 The Experimental Belt Drive System

The purpose of the system identification is to develop a single-input single-output model for closed-loop control design of the belt drive system. The open-loop system configuration is shown in Figure 20.

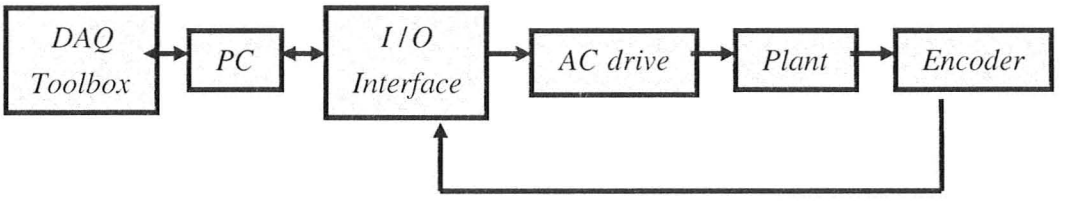

Figure 20 Open-loop Configuration of the Experimental Belt Drive System

The input is a DC voltage signal generated in Matlab's System Identification Toolbox, and sent through an IO interface to a Field Oriented Control (FOC) AC drive. The AC drive amplifies the input signal with a factor of 60.55 to a higher power voltage that actuates a Permanent Magnet Synchronous Motor (PMSM). An encoder is used to measure the angular displacement of the driven pulley of the belt drive system. Taking derivative of the encoder measurement gives the corresponding angular velocity which is treated as the output of the experimental belt drive system. From the input to the output, the overall belt drive system becomes a Single-Input-Single-Output (SISO) dynamical system.

3.2 Initial Tests

The initial tests are carried out for the design of experiments used in system identification and for the determination of the associated elements such as the sampling rate, the type and the magnitude of the input signal, the duration of tests, and the choice of

filters. Improper experimental design at this stage is likely to propagate through the overall identification process and may significantly deteriorate the quality of the identified models. Since repeating experimentation is an undesirable and costly proposition, the initial tests on the system build a solid ground for the careful selection of the parameters related to the final experiment and the data collection. Furthermore, these initial tests are intended to provide insight into the system characteristics by investigating delay, deadzone, drift, nonlinearity, noise level, as well as any other factors which may affect the identification process. The initial tests applied in this research include step responses, sine wave responses, and PRBS tests.

3.2.1 Investigation of Drift

The output of a system may drift subject to internal or external disturbances, even when the input is with zero-amplitude. The drift in the belt drive system is investigated by recording the output signal while the input is set to zero. In this case the drift is constant and can be suppressed by a non-zero bias input. The drift-suppressing bias for the belt drive system is -0.012 volt. The addition of this bias is to cancel out the effect of the drift during all system identification experiments.

3.2.2 Step Response Test

One of the initial tests conducted is the step response. Step inputs varying from 0 to 7 VDC with 1 volt increment are used, and the output is recorded while ensuring the duration is sufficient to allow the system to reach its steady state. The step responses for

the belt drive system are given in Figure 21 and provide several useful pieces of information as follows.

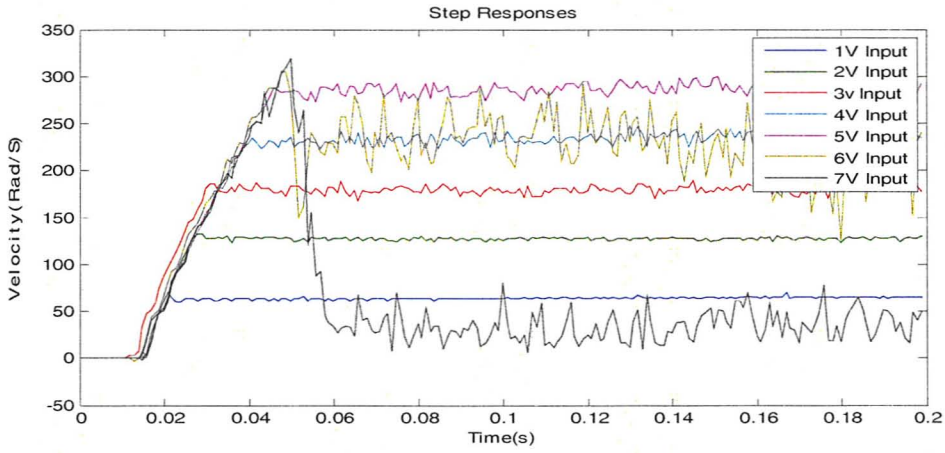

Figure 21 Step Response of the Experimental Belt Drive System

Steady-State Analysis

A linear system has two mathematical characteristics: homogeneity and superposition (Nise, 2004; Smith, 2010). Two properties of linear systems can be visualized by the steady-state plot of step responses. The steady-state outputs versus the amplitudes of the input for the belt drive system are shown in Figure 22. As can be seen, the steady-state gain of the belt drive system is the same for step input amplitudes of up to 5.5 VDC. When the input reaches 5.5 volts, the driven pulley reaches the velocity of 314 radians per second; the incremental encoder reaches its saturation frequency of 200 kHz. Figure 22 suggests that the magnitude of the input should not exceed 5.5 volts in order to retain the belt drive system within a linear operating region. The linearity of the belt drive system will be confirmed by using other methods such as the transient response analysis and the Empirical Transfer Function Estimate (ETFE) (Ljung, 1987).

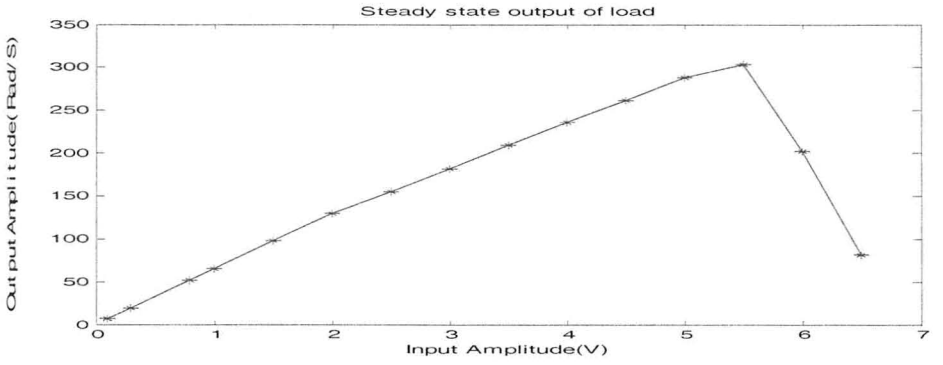

Figure 22 Steady State Outputs versus the Input Amplitudes

Transient Response Analysis

The transient responses of the belt drive system to step inputs of 1 to 7VDC are plotted in Figure 21. The belt drive system demonstrates the characteristic of homogeneity when the amplitude of the input is less than 5.5 VDC. The response characteristics change when the amplitude of the input is greater than 5.5VDC. The transient response analysis supports the linearity range obtained from the steady-state analysis of Figure 22. The transient responses also show that it takes about 15 milliseconds for the system to start to response to the inputs. This may be due to a deadzone or a time delay.

Investigation of the Deadzone

The deadzone of the belt drive system is investigated by gradually increasing the input amplitude until static friction of the system is overcome and a nonzero output is obtained. The plot of the input and the output in Figure 23 shows that a non-zero output is obtained when the input amplitude is greater than 0.02 VDC. The deadzone identified in the belt drive system is 0.02 volts.

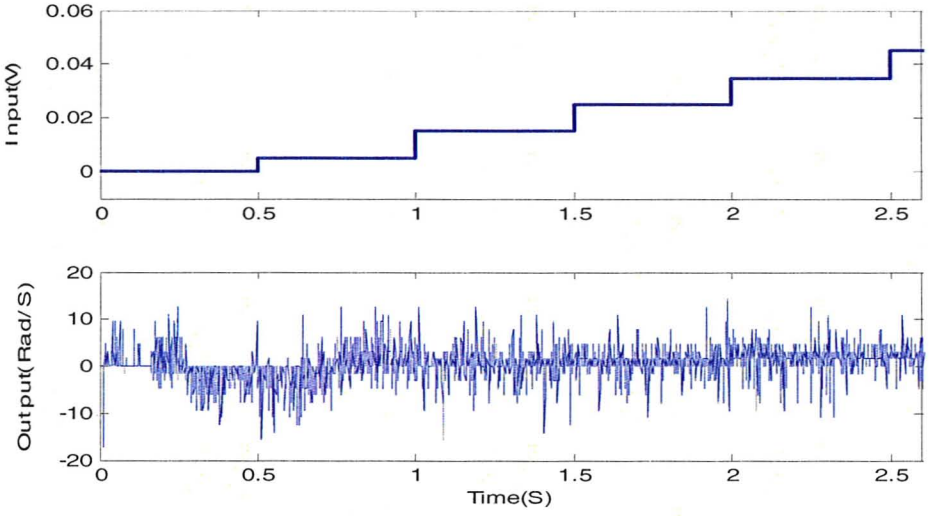

Figure 23 Investigation of the Deadzone

Investigation of the Time Delay

The time delay of the experimental belt drive system is investigated by using the same process as that of testing for the deadzone. A time delay is observed in response to every change in the input. Using the enlarged response of Figure 24, the time delay in the belt drive system is 2 milliseconds.

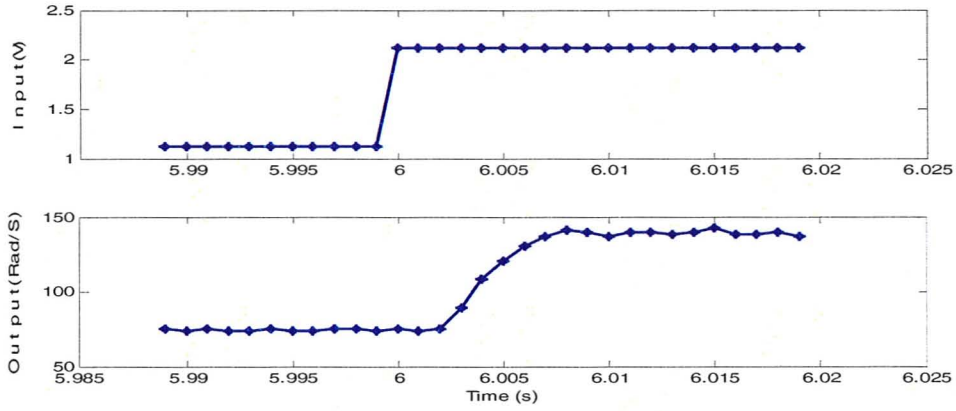

Figure 24 Enlarged Response of the Belt Drive System

Investigation of the Dominant Time Constant

The dominant time constant can be approximately obtained from the settling time or from the exponential decay of the transient response (Distefano, Stubberud & Williams, 1990). For example, for the first and second-order under-damped systems the transient terms have the form $Ae^{-t/\tau}$ and $Ae^{-t/\tau}\cos(\omega_d t+\phi)$, respectively. The decay in each case is governed by the time constant τ. For the system of higher order than two the dominant time constant is approximated by that of the under-damped second-order system. Using the enlarged step response of Figure 25, the settling time of the belt drive system is 0.01 (0.024-0.014) second. The dominant time constant τ of the belt drive system approximated from the settling time is 10 milliseconds.

Based on the time constant τ, the sampling time T is picked as one tenth of the time constant, i.e., 1 millisecond. The period of data recording should be more than 100 times the time constant that is greater than 1 second.

Figure 25 Enlarged Step Response of the Belt Drive System

Noise Analysis

The sources of disturbances or noise include measurement noise and uncontrollable inputs (Ljung, 1987). Measurement noise is due to the sensors. Uncontrollable inputs are disturbances that act as spurious inputs and that are not controllable by the user. By measuring the output corresponding to a constant input, an insight is gained into the system's noise characteristics. An encoder is immune to noise when measuring displacement. However, the quantization noise is introduced when taking derivative of the displacement for velocity. The noise contaminated data is commonly described with *mean* (μ) and *variance* (σ^2) which are calculated, respectively, as (Smith, 2010):

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \qquad (3.1)$$

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu)^2 \qquad (3.2)$$

The mean and variance of the output of the belt drive system (i.e. its velocity) versus the amplitude of the input are listed in Table 3.

Table 3 *Mean and Variance of the Output of the Experimental Belt Drive System*

Input (volts)	Mean (μ) (Rad/S)	Variance (σ^2)
0	-0.0807	6.6719
1	64.8991	1.1664
2	129.7806	3.4942
3	182.1712	19.6462
4	235.6950	29.9154
5	287.8570	41.7282

The noise is further analyzed by considering its Power Spectral Density (PSD). The analysis is performed by first running the system at a constant velocity, recording the output, removing the mean of the output and then taking the power spectral density of the remainder. The noise power spectral densities for the input amplitude of 1 to 5VDC are shown in Figures 26 to 31.

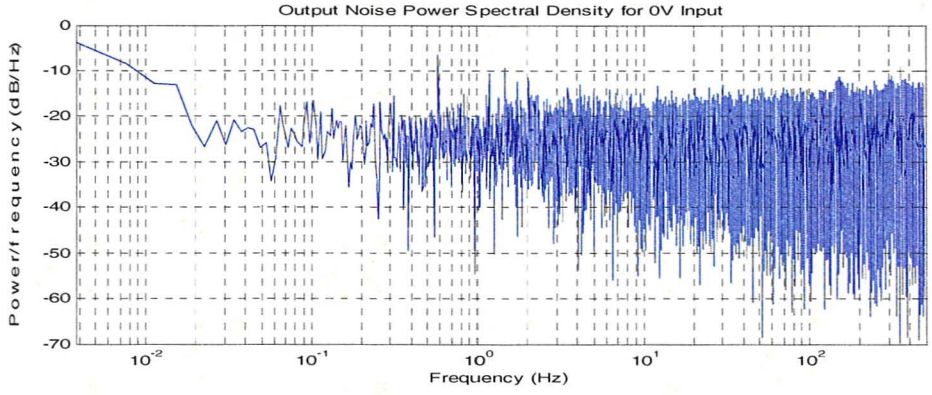

Figure 26 Output Noise Spectral Density for 0 Volt Input

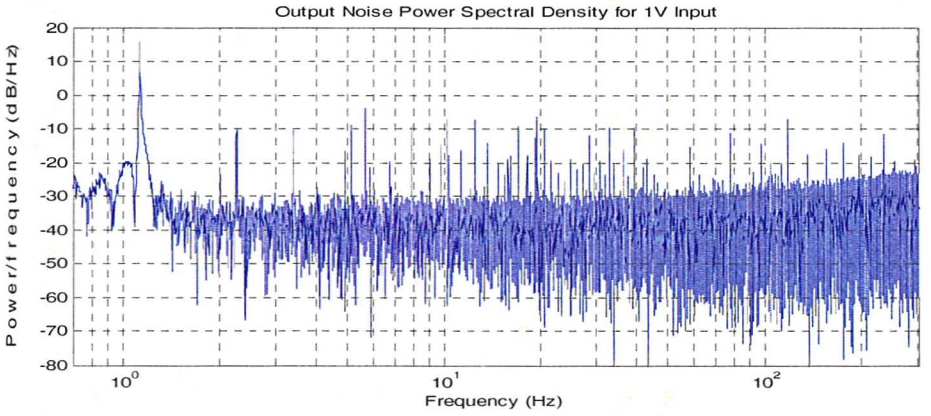

Figure 27 Output Noise Spectral Density for 1 Volt Input

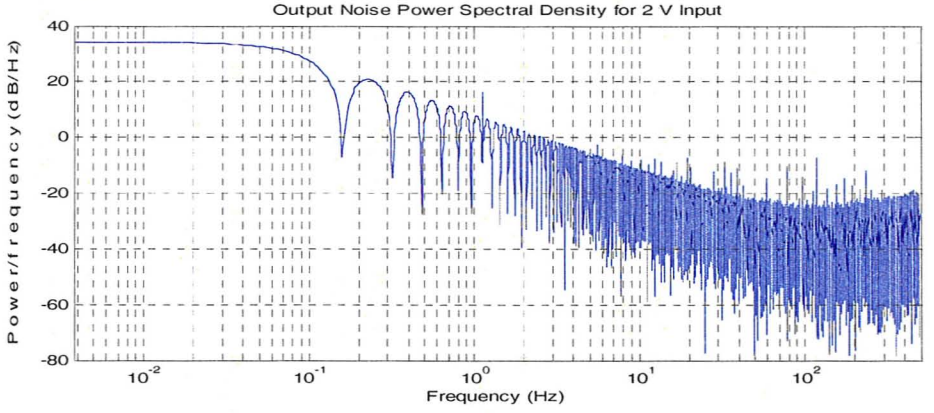

Figure 28 Output Noise Spectral Density for 2 Volts Input

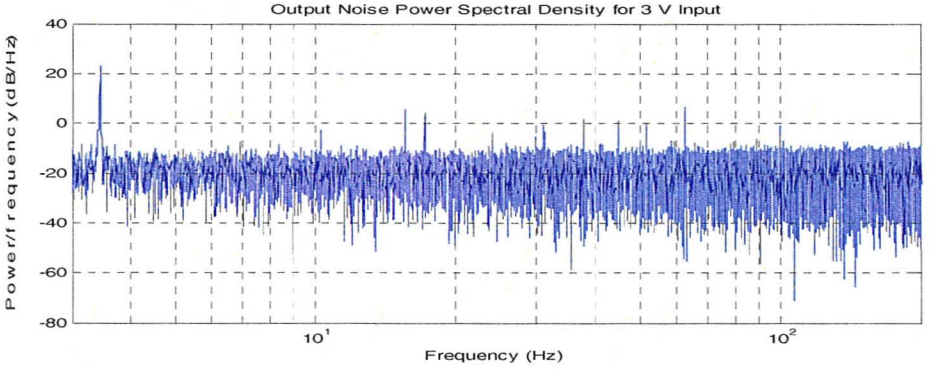

Figure 29 Output Noise Spectral density for 3 Volts Input

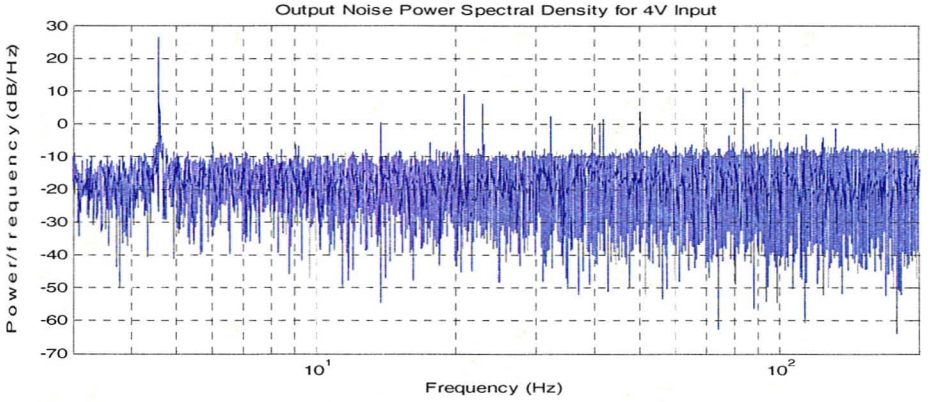

Figure 30 Output Noise Spectral Density for 4 volt Input

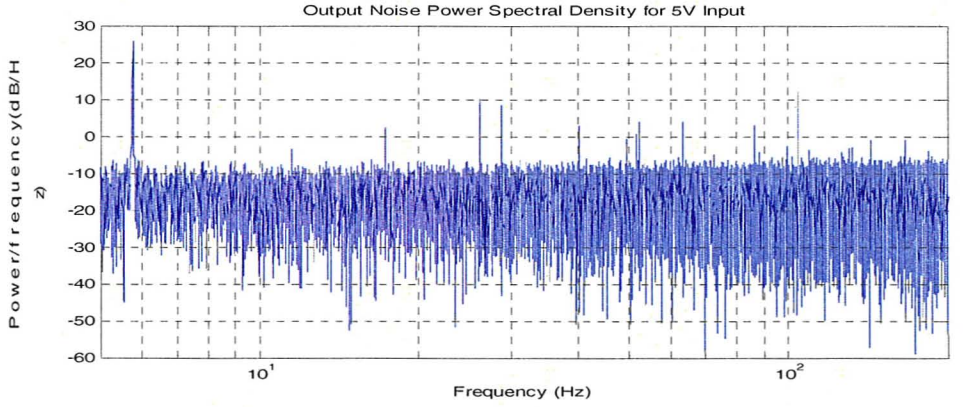

Figure 31 Output Noise Spectral Density for 5 Volts Input

There is a low frequency content in the noise. The frequency increases according to the input level as shown in Table 4. This is caused by a defect in the encoder itself. This low frequency element is within the frequency range of interest and can not be filtered out. It will be present throughout the system identification process and will affect the quality of the identified model.

Table 4 *Low Frequency Content in the Noise*

Amplitude of input(V)	0	1	2	3	4	5
Low Frequency Content (Hz)	0.6	1.2	1.2	3.6	4.5	5.5

3.2.3 Sine-wave Test

A computationally simple method to estimate a frequency function of a system is the frequency response analysis or sine-wave testing. A sine wave with a specific frequency is applied to the system. Neglecting the noise and the transient term, the amplitude and the phase shift of the resulting output signal with respect to the input are measured graphically. The magnitude and the phase of the output sine wave are then calculated. The frequency response function of the system over a frequency range can be obtained by applying input signals with different frequencies and repeating the above calculation. The system bandwidth is estimated by using Bode magnitude and phase plots of the frequency response function of the system. The Bode magnitude plot of the belt drive system is shown in Figure 32. The **open-loop bandwidth** of the belt system is estimated approximately at **70 Hz.**

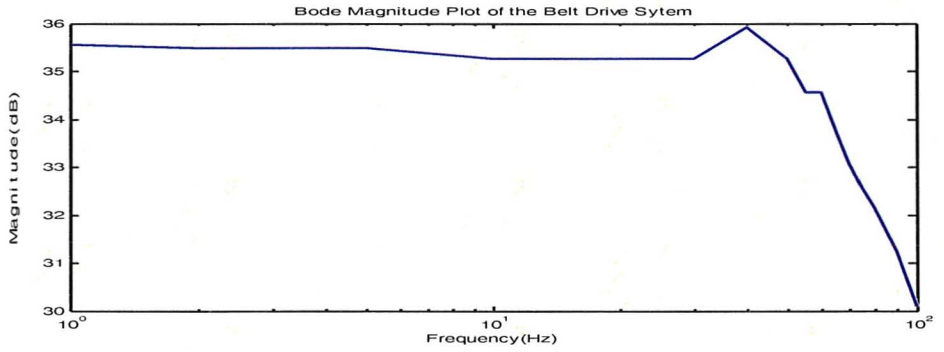

Figure 32 Bode Magnitude Diagram of the Belt Drive System

The disadvantages of the frequency response analysis using sine-wave method are: (1) the test has to be calculated at a range of frequencies making the method very time consuming; (2) it is not always safe to apply pure sinusoidal signals in industrial applications. An alternative strategy is the use of Empirical Transfer Function Estimate (ETFE).

3.2.4 Empirical Transfer Function Estimate (ETFE)

The ETFE is the ratio of the Fourier transforms of the output and the input. It is an estimate of the frequency response of the system. In the Fourier analysis, the discrete Fourier transform (DFT) of the finite sequence of the input u_k, $k = 1, 2, \cdots, N$, is defined as (Ljung, 1987):

$$U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} u(t) e^{-i\omega t}, \ \omega = 2\pi k / N, k = 0, 1, 2, ..., N-1 \qquad (3.3)$$

The DFT maps u_k in time domain into $U_N(\omega)$ in frequency domain. The signal u_k can be recovered by using the Inverse Discrete Fourier Transform (IDFT) as:

$$u_k = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} U_N(\omega) e^{i\omega k}, \ \omega = 2\pi k / N, \ k = 1, 2, \cdots, N \qquad (3.4)$$

The signal u_k is a linear combination of $e^{i\omega k}$ for N different frequencies ω. Using the same definition, the output signal y_k, $k = 1, 2, \cdots, N$ is mapped into and recovered from $Y_N(\omega)$, respectively, as:

$$Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} y_k e^{-i\omega k}, \ \omega = 2\pi k / N, \ k = 1, 2, \cdots, N \qquad (3.5)$$

$$y_k = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} Y_N(\omega) e^{i\omega k}, \ \omega = 2\pi k / N, \ k = 1, 2, \cdots, N \qquad (3.6)$$

The ETFE of the system can be represented as:

$$\hat{\hat{G}}_N(e^{i\omega}) = \frac{Y_N(\omega)}{U_N(\omega)} \qquad (3.7)$$

The ETFE allows a user to approximately estimate the frequency response of a system using experimental data.

The input signals have to contain a wide range of frequencies. The frequency range should ideally be greater than the system bandwidth by more than 5 to 10 times. With an estimated bandwidth of 70 Hz from the initial frequency response analysis, a sampling frequency is of 1500Hz; a frequency range of 233 (700/3) Hz was chosen for the input signal. The power spectrum of the input signal should be flat for this range. As such, the following types of signals can be selected: the random binary sequence (RBS), the random Gaussian sequence (RGS), the chirp signal, or the pseudorandom binary sequence (PRBS). In this research, a PRBS is adopted as the input signal with a frequency of 700 Hz, amplitude of 1volt, giving a flat power spectrum in the range of 0 to 700/3 Hz. The

experiments are conducted using PRBS signals of amplitudes of 1 to 5 volts in 1 volt steps.

The output of the system is obtained by taking the derivative of the encoder signal through the IO interface. The estimated ETFEs are filtered in order to obtain a more usable representation. The filtering is conducted by using a windowing function. There are several types of windowing functions available in both time and frequency domains. In the time domain, the finite sequences $u(t)$ and $y(t)$ are convolved with a time domain window, respectively. The DFTs are taken from the windowed sequences, and the ETFEs are obtained from the ratios of the DFTs. In the frequency domain, the DFTs of the signals $u(t)$and $y(t)$ are taken. The ETFEs are then obtained from the ratios of the DFTs and filtered with a Hamming window in the frequency domain (Ljung, 1987) as follows:

$$\hat{G}_N\left(e^{i\omega_0}\right)=\frac{\int_{-\pi}^{\pi} W_\gamma(\xi-\omega_0)\left|U_N(\xi)\right|^2 \hat{\hat{G}}_N\left(e^{i\xi}\right)d\xi}{\int_{-\pi}^{\pi} W_\gamma(\xi-\omega_0)\left|U_N(\xi)\right|^2 d\xi} \tag{3.8}$$

$$W_\gamma(\omega)=\frac{1}{4\pi}\left(D_\gamma(\omega)+\frac{1}{2}D_\gamma\left(\omega-\frac{\pi}{\gamma}\right)+\frac{1}{2}D_\gamma\left(\omega+\frac{\pi}{\gamma}\right)\right) \tag{3.9}$$

$$D_\gamma(\omega)=\frac{\sin\left(\gamma+\frac{1}{2}\right)\omega}{\sin\omega/2} \tag{3.10}$$

where $W_\gamma(\xi)$ is the window function in frequency domain, which is centered around $\xi=0$, and γ is the "shape parameter" determining the width of the frequency window. The width of the frequency window controls the trade-off between bias and variance: a large value of γ corresponds to a narrow window, and leads to a large

variance of $\hat{G}_N\left(e^{i\omega_0}\right)$ and small bias. In the discrete-time system, the integrations in Equation (3.8) are replaced with summations as:

$$\hat{G}_N\left(e^{i\omega_0}\right)=\frac{\sum_{k=0}^{N-1}W_\gamma\left(\frac{2\pi(k-k_0)}{N}\right)\left|U_N\left(\frac{2\pi k}{N}\right)\right|^2 G_N\left(e^{\frac{i2\pi k}{N}}\right)}{\sum_{k=0}^{N-1}W_\gamma\left(\frac{2\pi(k-k_0)}{N}\right)\left|U_N\left(\frac{2\pi k}{N}\right)\right|^2},\ k_0=0,1,2,\cdots,N-1 \qquad (3.11)$$

The ETFEs and their filtered versions at different amplitudes of the input signal are shown in Figures 33 to 35.

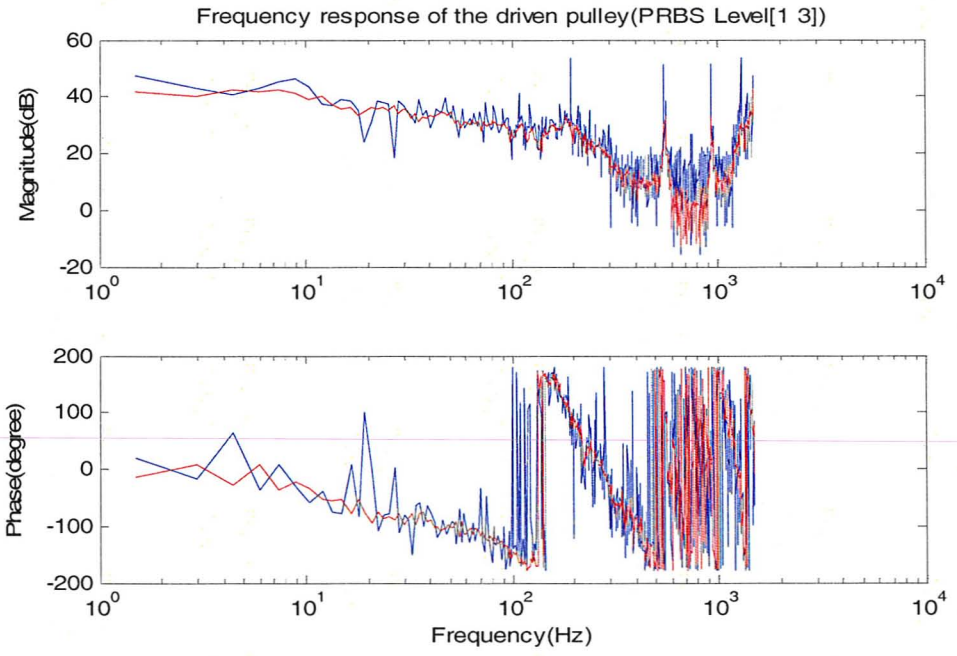

Figure 33 ETFE of the Belt Drive System with Input PRBS Level [1 3]

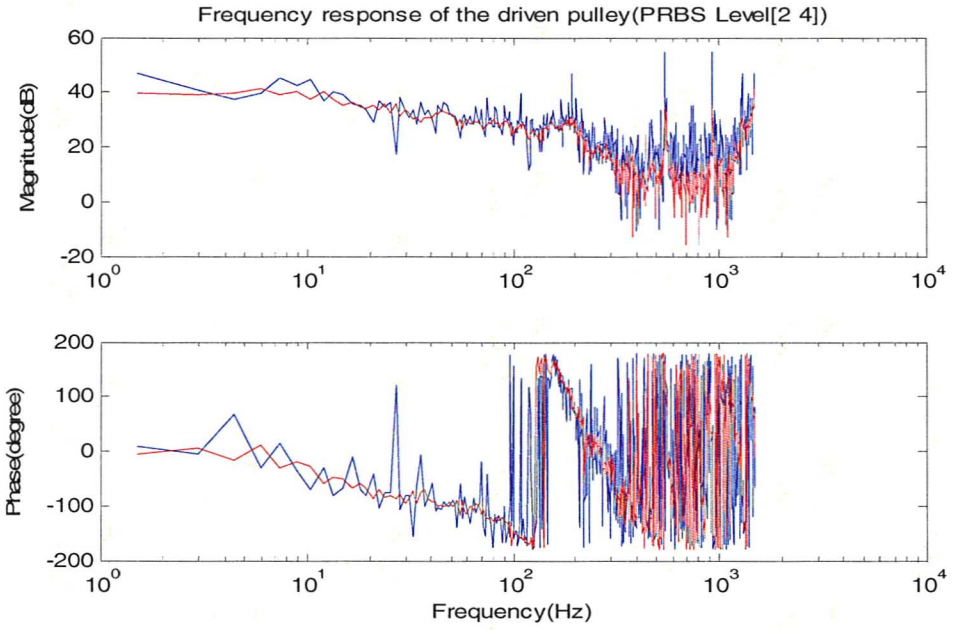

Figure 34 ETFE of the Belt Drive System with Input Level [2 4]

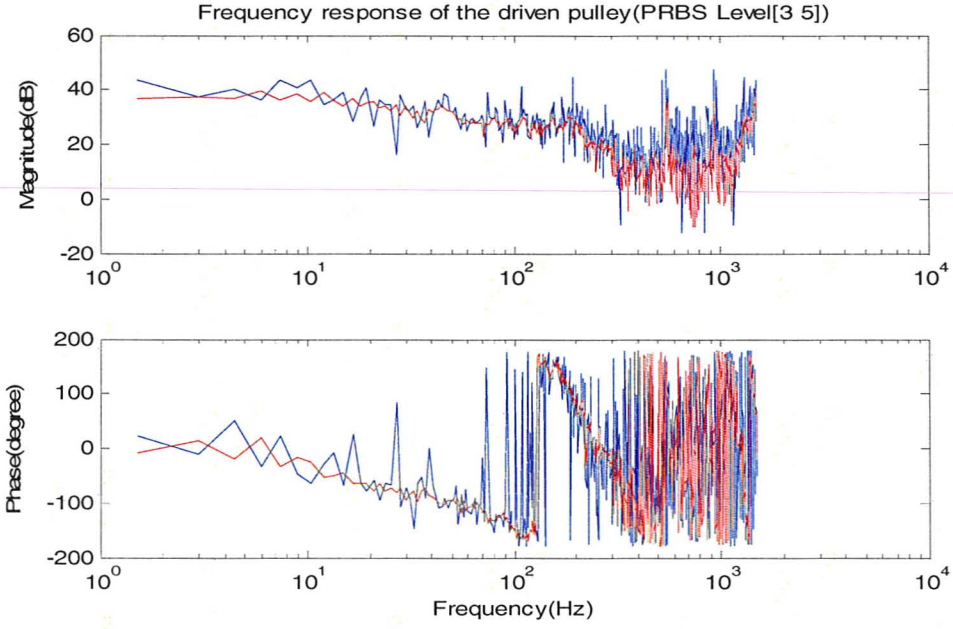

Figure 35 ETFE of the Belt Drive System with Input PRBS Level [3 5]

The objective of the ETFEs is to verify the linearity across the frequency range of interest and with respect to the input level. The ETFEs demonstrate that the bandwidth of the system is approximately 70 Hz which is consistent with the result obtained in the frequency response analysis. Note also that the ETFEs obtained from different levels of inputs are approximately identical. This further supports the earlier results and indicates that the system is linear in the input range of 1 to 5 volts.

3.3 Order Estimation

The order of a linear system can be estimated in different ways. The commonly used methods include the spectral analysis, and the impulse response analysis (Habibi, 2009).

3.3.1 Spectral Analysis Estimate

Spectral analysis uses the Bode plot to provide the DC gain and the break frequencies. By looking the ETFE plots from Figures 33 to 35, there are no positive-slope straight lines in the log-magnitude curves and no positive phase angle contributions in the phase angle curves, which indicate that there are no zeros in the continuous time transfer function. In the decade from 10 to 100Hz the log-magnitude plot is approximately a straight line with a slope greater than -40 dB/decade, which means that the number of poles is not more than two. Furthermore, the phase curves decreasing 180 degrees over the entire frequency range indicate that there are two poles in the system. All of these indications demonstrate that this is a second-order system.

3.3.2 Order Estimation via Impulse Response

Impulse response (IR) is a well-proven method to conduct system identification, and enables an accurate estimation of the system order (Habibi, 2009; Pitstick & Mulholland, 1986; Vacher, 2010). A discrete linear time invariant system can be represented in a state space form (Ogata, 1994) as:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k \\ y_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}u_k \end{aligned} \qquad (3.12)$$

where u_k denotes the input; $\mathbf{x}_k$ denotes the state variable; and y_k denotes the output. $\mathbf{A}$ is the system matrix; $\mathbf{B}$ is the input matrix; $\mathbf{C}$ is the output matrix; and $\mathbf{D}$ is the feed-forward matrix. The transfer function of the system of Equation (3.12) can be written as:

$$G(z) = \mathbf{D} + \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \qquad (3.13)$$

where z denotes the delay operator. Using the geometric series expansion to $(z\mathbf{I} - \mathbf{A})^{-1}$, Equation (3.13) can be further written as:

$$G(z) = \mathbf{D} + \mathbf{C}z^{-1}\left(\mathbf{I} + \mathbf{A}z^{-1} + \mathbf{A}^2z^{-2} + \cdots + \right)\mathbf{B} = \mathbf{D} + \sum_{i=1}^{\infty}\mathbf{h}_i z^{-i} \quad \mathbf{h}_i = \mathbf{C}\mathbf{A}^{i-1}\mathbf{B} \qquad (3.14)$$

Equation (3.14) can also be expressed as a function of matrices $\mathbf{P}$ and $\mathbf{E}$ such that:

$$G(z) = \mathbf{D} + \mathbf{P}\mathbf{E} \qquad (3.15)$$

where $\mathbf{P} = \begin{bmatrix}\mathbf{C} & \mathbf{CA} & \mathbf{CA}^2 & \cdots & \mathbf{CA}^{n-1}\end{bmatrix}$, and $\mathbf{E} = \begin{bmatrix}\mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \cdots & \mathbf{A}^{n-1}\mathbf{B}\end{bmatrix}^T$, are the observability and the controllability matrices (Ogata, 1994), respectively.

For a unit impulse $U(z) = 1$, the system's impulse response equals to the series in Equation (3.14). Equation (3.14) relates the system's impulse response to its transfer

function $G(z)$, in which, $\mathbf{D}$ is the initial value, and $h_i = \mathbf{CA}^{i-1}\mathbf{B}$, $i = 1,2\cdots\infty$ is the impulse response at discrete time i. This provides a very useful way of using the impulse response for identifying the structure of an unknown linear time invariant system (LTI). The transfer function $G(z)$ can be directly obtained from the impulse response of the system (Habibi, 2009; Lim & Phan, 1998; Wang, Lee & Zachery, 1996).

Acquiring of Impulse Response

Ideally the impulse response data would be directly measured from a system which is activated by a unit impulse input signal. This is however not feasible for most real life systems due to the difficulty of generating an impulse input with a high enough amplitude to activate the system.

The second option is to numerically differentiate the step response data. Since an impulse can be generated by differentiating a step input. Given that system identification deals with linear systems, the impulse response of a linear system can be obtained by differentiating its step response.

The third option of estimating the impulse response is by using correlation analysis (System, 2009). The impulse response can be determined by calculating the corresponding correlation function of input and output signals. The impulse response function is derived as:

$$y(t) = \sum_{k=0}^{\infty} g(k)u(t-k) + v(t) \qquad (3.16)$$

where $u(t)$ is the input signal, $y(t)$ is the output signal, $g(k)$ is the impulse response function, and $v(t)$ is the white noise. The equivalent relationship between the input-output cross-correlation and the input auto correlation is represented as:

$$R_{yu}(\tau) = \sum_{k=0}^{\infty} g(k) R_{uu}(\tau - k) \qquad (3.17)$$

where $R_{yu}(\tau) = E\left[y(t+\tau)u(t)\right]$, $R_{uu}(\tau) = E\left[u(t+\tau)u(t)\right]$, in which, $E[\bullet]$ is the expectation operator. In practice, R_{yu} and R_{uu} are replaced with their estimated functions:

$$\hat{R}_{yu}(\tau) = \frac{1}{N} \sum_{t=1-\min(\tau,0)}^{N-\max(\tau,0)} y(t+\tau)u(t),\ \tau = 0, \pm 1, \pm 2, \ldots \qquad (3.18)$$

$$\hat{R}_{uu}(\tau) = \frac{1}{N} \sum_{t=1}^{N-\tau} u(t+\tau)u(t),\ \hat{R}_{uu}(-\tau) = \hat{R}_{uu}(\tau) \qquad (3.19)$$

The estimated impulse response function is determined by solving Equation (3.17) as:

$$\hat{g}(k) = \hat{R}_{yu}\hat{R}_{uu}^{-1} \qquad (3.20)$$

In this research, the numerical differentiation approach is applied to estimate the impulse response from the step response. For the experimental belt drive system, the step and its impulse response obtained by differentiation are provided in Figures 36 and 37. Note that the delays in the signals and outliers have been removed.

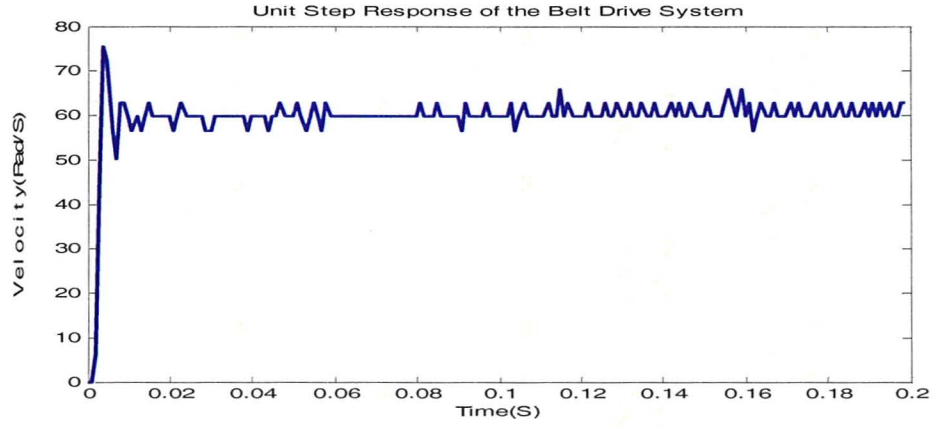

Figure 36 Unit Step Response of the Belt Drive System

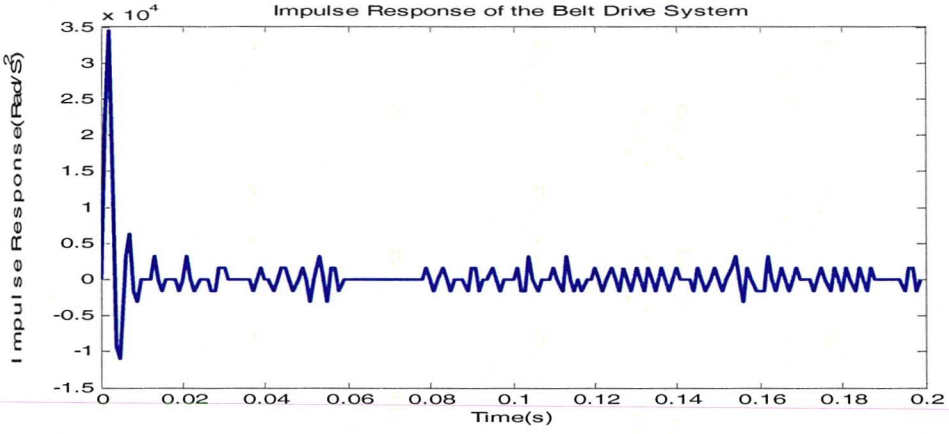

Figure 37 Estimated Impulse Response of the Belt Drive System

Order Estimation

For the state space model of Equation (3.12), the impulse response can be easily found by direct calculation as:

$$h_k = \begin{cases} \mathbf{D}, k = 0 \\ \mathbf{CA}^{k-1}\mathbf{B}, k > 0 \end{cases} \qquad (3.21)$$

Note that in Single-Input-Single-Output (SISO) case, the impulse response consists of scalar outputs. While in Multiple-Input-Multiple-Output (MIMO) cases, every sample of the impulse response is a $p \times q$ matrix where p, q denote the dimensions of the output and input, respectively. In SISO case, the impulse response terms in Equation (3.21) are known as the Markov parameters of the state space model.

A $n_h \times n_h$ matrix can be constructed using the Markov parameters in such a way that each row would contain the Markov parameters shifted in time as:

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 & h_3 & \cdots & h_{n_h} \\ h_2 & h_3 & h_4 & \cdots & h_{n_h+1} \\ h_3 & h_4 & h_5 & \cdots & h_{n_h+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{n_h} & h_{n_h+1} & h_{n_h+2} & \cdots & h_{2n_h-1} \end{bmatrix} \tag{3.22}$$

The matrix $\mathbf{H}$ is patterned in a way that every ascending diagonal contains the same value, and because of this pattern it is called a Hankel matrix. In state space system identification theory, the Hankel matrix plays an important role. For example, a state space model can be obtained through a factorization of the Hankel matrix via the Singular Value Decomposition (SVD) (Lim & Phan, 1998). In the absence of noise, the Hankel matrix derived from a step response contains only the first n linearly independent rows and the remaining rows are linear combinations of the first n for a n^{th} -order system (Habibi, 2009). The rank of the Hankel matrix equals the system order. This provides a way of identifying the model structure using the impulse response data.

However in practice, the impulse response data is contaminated due to perturbations, noise and numerical methods. The calculated rank of the Hankel matrix may be therefore

corrupted. In this case the Singular Value Decomposition (SVD) can be used to calculate the rank of the Hankel matrix. The SVD is a powerful tool dealing with sets of equations or matrices that are either singular or numerically very close to singular (SVD, 2010). It decomposes a matrix into a product of three simpler ones. The SVD is defined as:

$$\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^T \qquad (3.23)$$

where $\mathbf{U}$ is a n-by-n unitary matrix ($\mathbf{U}^T\mathbf{U} = \mathbf{I}$), $\mathbf{V}$ is a n-by-n unitary matrix ($\mathbf{V}^T\mathbf{V} = \mathbf{I}$), the superscript T denotes the matrix transposition; $\sum$ is a n-by-n diagonal matrix of the same dimension as the Hankel matrix $\mathbf{H}$. The matrix Σ consists of the positive singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$ on the diagonal. The order information of the system is recovered by counting the singular values that are significantly higher than others (Habibi, 2009).

In this research, 1587 points of impulse response data are taken to build a 794-by-794 Hankel matrix. After the singular value decomposition, the first 20 singular values of the Hankel matrix are plotted in Figure 38, which indicates that **the experimental belt drive system is a second-order system.** This conclusion is consistent with that drawn by spectral analysis.

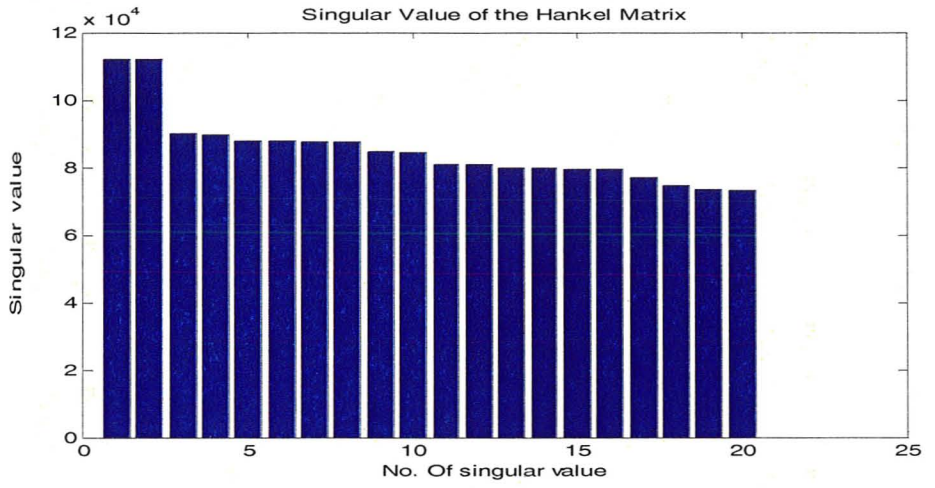

Figure 38 The First 20 Singular Values of the Hankel Matrix

3.4 Main Experiments and Data Processing

Design of the main experimentation is an extremely important aspect of system identification. The input of the system is a voltage signal to the amplifier and the output is the derivative of the position encoder. From the initial sine wave test and the analysis of the ETFEs, it is determined that the system bandwidth is approximately 70Hz. In selecting an input type which would provide an even power spectrum well beyond this frequency range of interest, a Pseudo Random Binary Sequence (PRBS) is chosen. The magnitude of this input is chosen based on the noise characteristics of the system as considered from the step tests, and is generally a trade-off between a strong signal-to-noise ratio and remaining in a piecewise linear region. Also from the steady state gain test and ETFE, the system is linear in the input range of 0 to 5 volts. The input amplitude is hence chosen to be 0 to 5 volts for the main experiments.

The duration of the experiment is chosen as 5 seconds based on the dominant time constant discussed in Section 3.2.2. In determining the sampling rate for these experiments the open loop system bandwidth of approximately 70 Hz is considered. A 1000Hz sampling rate is chosen, that is over 10 times this bandwidth.

Data processing is an important intermediate step. The data must go through a conditioning process. The conditioning process requires not only filtering but also including detection of outliers, removal of low frequency trends, and scaling. To detect outliers, a boundary layer is constructed though time plot of the output data ensuring that the most data points enclosed inside, and the points located outside are considered to be the outliers and replaced with the closest upper or lower bound value. The output data is scaled to the angular velocity of the driven pulley with the unit of rad/s taking into account the pulley radius and the encoder resolution. The low frequency trend is removed by subtracting the input and the output data with their respective mean values.

High frequency noise is removed by using a low-pass digital filter. Finite and infinite impulse response filters always introduce a time delay or a phase lag. They can however be implemented in a *bidirectional filtering mode* to realize zero-phase filtering. This is done by filtering the sequence in the forward direction, then reversing it and passing it through the filter in its reversed sequence. In the forward direction a phase lag is introduced. In the reversed direction a reversed lag (or phase lead) results in zero-phase filtering. According to the ETFEs in Figure 33-35, at frequencies higher than 100Hz, the signal-to-noise ratio is low; the cut-off frequency of the filter is set to 100 Hz. By cutting off the high frequency contents, a better fit in the remaining frequency band is expected.

The type of filter used in this research is the 10^{th} order digital Butterworth filter. The magnitude plot of the digital Butterworth filter is shown in Figure 39.

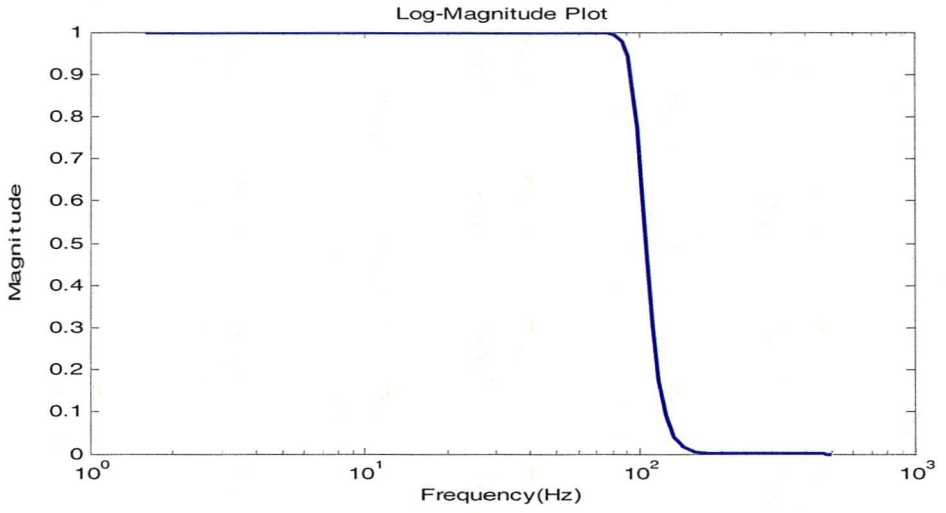

Figure 39 Magnitude Plot of the 10^{th} Order Digital Butterworth Filter

A 5000-point PRBS input with a switching frequency of 500 Hz is used. This establishes a flat spectrum of 500/3Hz for the PRBS signal that is greater than the 100 Hz frequency range of interest. The first 300 points of input and output are shown in Figure 40. The data sets are filtered with the 10^{th} order Butterworth filter to serve as the estimation data. The filtering process results in a reduced frequency range of 0-100 Hz. The first 300 points of the filtered data are shown in Figure 41. An independent validation data set is generated and filtered in the same way. It is used to validate the estimated model.

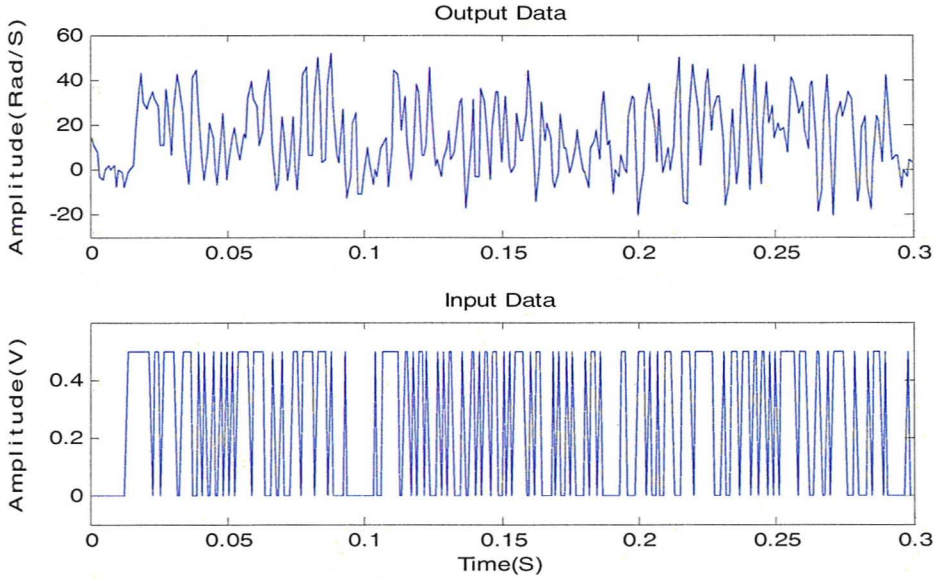

Figure 40 The First 300 Points of the Model Estimation Data

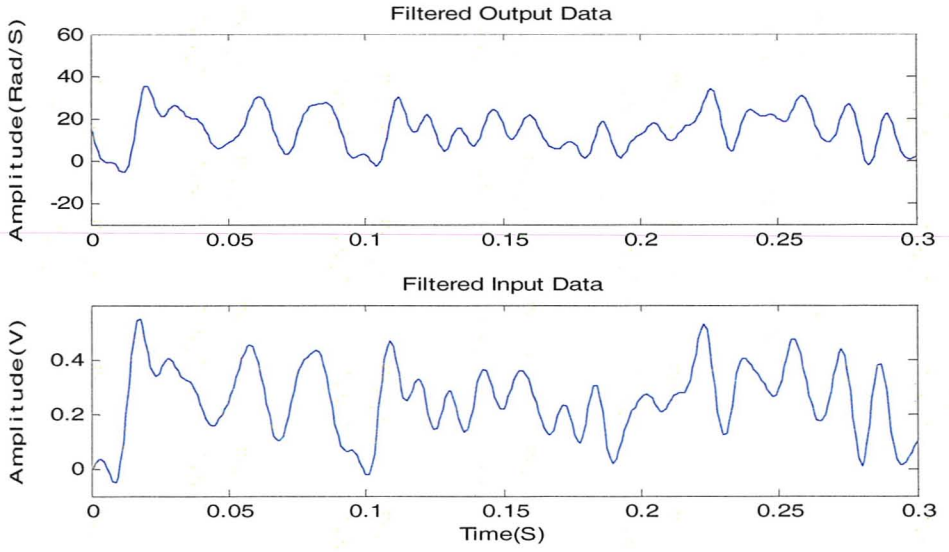

Figure 41 The First 300 Points of the Filtered Data

3.5 Model Estimation

Solving the model estimation problem requires: 1) input-output data from the process; 2) a class of models; 3) a criterion. The model estimation problem can be treated as an optimization problem, in which a best model can be obtained to best fit the data with respect to the given criterion (Astrom & Wittenmark, 1997).

A class of models can be represented in the following generalized form (Ljung, 1987):

$$
\begin{aligned}
&A(z)y_k = \frac{B(z)}{F(z)}u_k + \frac{C(z)}{D(z)}e_k \\
&A(z) = 1 + a_1 z^{-1} + \cdots + a_{n_a} z^{-n_a} \\
&B(z) = b_1 z^{-1} + \cdots + b_{n_b} z^{-n_b} \\
&C(z) = 1 + c_1 z^{-1} + \cdots + c_{n_c} z^{-n_c} \\
&D(z) = d_1 z^{-1} + \cdots + d_{n_d} z^{-n_d} \\
&F(z) = 1 + f_1 z^{-1} + \cdots + f_{n_f} z^{-n_f}
\end{aligned}
\qquad (3.24)
$$

where y_k, u_k and e_k are the output, the input and the zero-mean white noise, respectively; n_a, n_b, n_c, n_d and n_f are the order of the polynomials $A(z)$, $B(z)$, $C(z)$, $D(z)$ and $F(z)$, respectively, and z is the delay operator. The system delay z^{-n_k} enters the model through the polynomial $B(z)$. By adopting different combinations of the polynomials $A(z)$, $B(z)$, $C(z)$, $D(z)$, and $F(z)$ in Equation (3.24), special model forms such as AutoRegressive with Exogenous input (ARX) model, AutoRegressive-Moving Average with eXogenous input (ARMAX) model, Output-Error (OE) model and Box-Jenkins (BJ) model are obtained.

A criterion is intended to measure how well the estimated model fits the experimental data. Hence the criterion can be defined as a cost function of the prediction error. With a given form of model, the model parameters are then determined by minimizing the cost function through optimization methods such as the Least-Squares method or the maximum likelihood method depending on the form of model that is being used.

3.5.1 AutoRegressive with Exogenous Input (ARX) Model

The ARX model is the simplest input-output relationship representation that can be described as the following form:

$$y_k = \frac{B(z)}{A(z)} u_{k-n_k} + \frac{1}{A(z)} e_k \qquad (3.25)$$

The ARX model is diagrammatically illustrated in Figure 42. The ARX model is derived by setting $C(z), D(z)$ and $F(z) = 1$ in Equation (3.24).

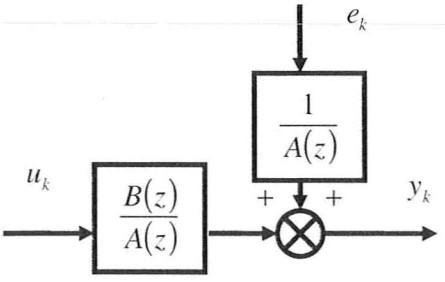

Figure 42 Block Diagram of the ARX Model

The best fit of th second-order ARX model of the experimental belt drive system is obtained as:

$$y_k = \frac{29.62z^{-2} - 24.07z^{-3}}{1 - 1.514z^{-1} + 0.6051z^{-2}} u_k + \frac{1}{1 - 1.514z^{-1} + 0.6051z^{-2}} e_k \qquad (3.26)$$

The ARX model prediction and the measurement are shown in Figure 43.

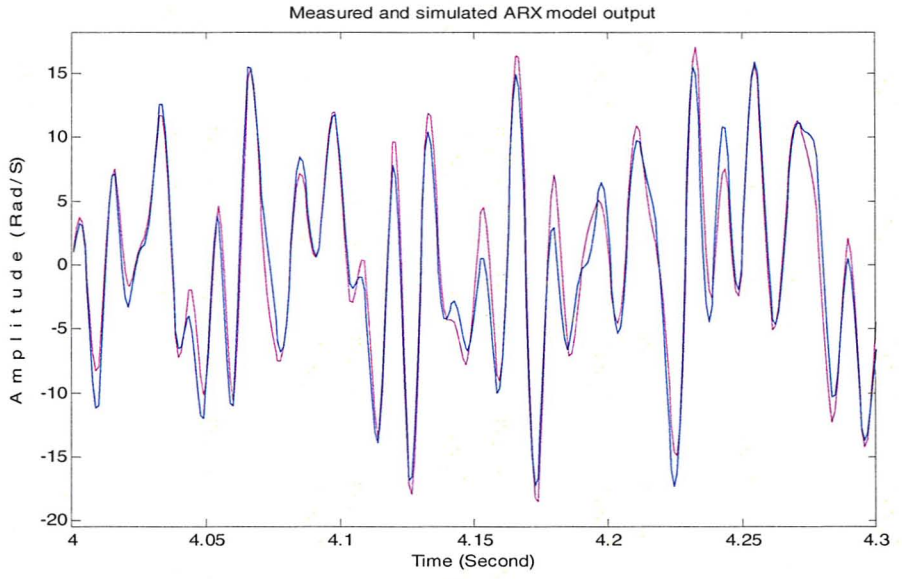

Figure 43 Measured and the ARX Model Output (70.9% Fit)

3.5.2 AutoRegressive-Moving Average with eXogenous input (ARMAX) Model

The main disadvantage of an ARX model structure is its restrictive noise model. More flexibility can be added to the description of noise transfer function by using $C(z)$ as the numerator polynomial. This changes the ARX model into an ARMAX model. The ARMAX model is derived by assigning $D(z), F(z)$ as 1 such that:

$$y_k = \frac{B(z)}{A(z)} u_k + \frac{C(z)}{A(z)} e_k \qquad (3.27)$$

The block diagram of the ARMAX model is shown in Figure 44.

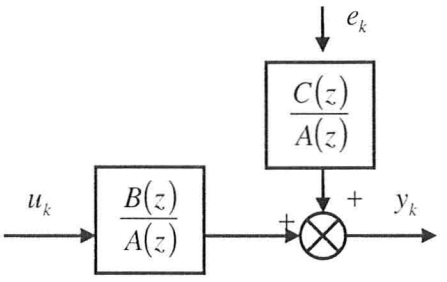

Figure 44 the Block Diagram of the ARMAX Model

The best fit of an ARMAX model for the belt drive system is obtained as:

$$y_k = \frac{29.58z^{-2} - 23.68z^{-3}}{1 - 1.509z^{-1} + 0.6061z^{-2}} u_k + \frac{1 + 1.999z^{-1} + 0.9986z^{-2}}{1 - 1.509z^{-1} + 0.6061z^{-2}} e_k \qquad (3.28)$$

The comparison of the ARMAX model prediction with the measurement is shown in Figure 45.

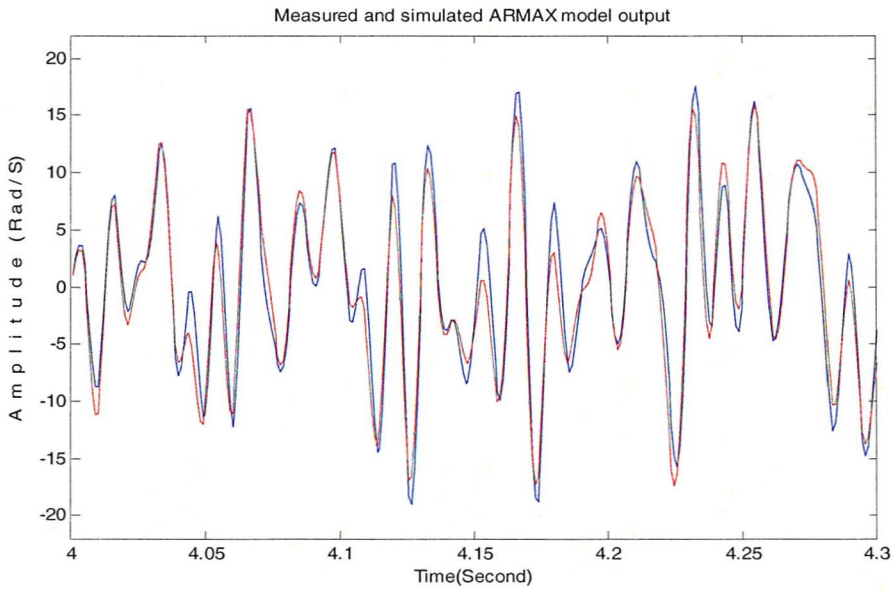

Figure 45 Measured and the ARMAX Model Output (70.72% Fit)

3.5.3 Output Error (OE) Model

In the ARMAX model, the transfer functions $G(z)=\frac{B(z)}{A(z)}$, $H(z)=\frac{C(z)}{A(z)}$ have the same polynomials in denominator; this is not always consistent with most real-life systems. Setting $A(z)=C(z)=D(z)=1$ decouples the noise from the system transfer function and results in the OE model as:

$$y_k=\frac{B(z)}{F(z)}u_k+e_k \qquad (3.29)$$

The output-error model structure is diagrammatically shown in Figure 46.

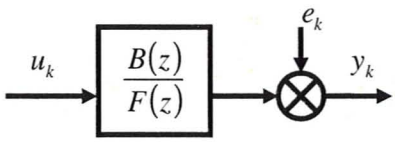

Figure 46 Block Diagram of the OE Model

The best fit of the 2nd order OE model is obtained as:

$$y(t)=\frac{19.2z^{-2}}{1-1.226z^{-1}+0.5311z^{-2}}u(t)+e(t) \qquad (3.30)$$

The comparison of the OE model prediction with the measured output is shown in Figure 47.

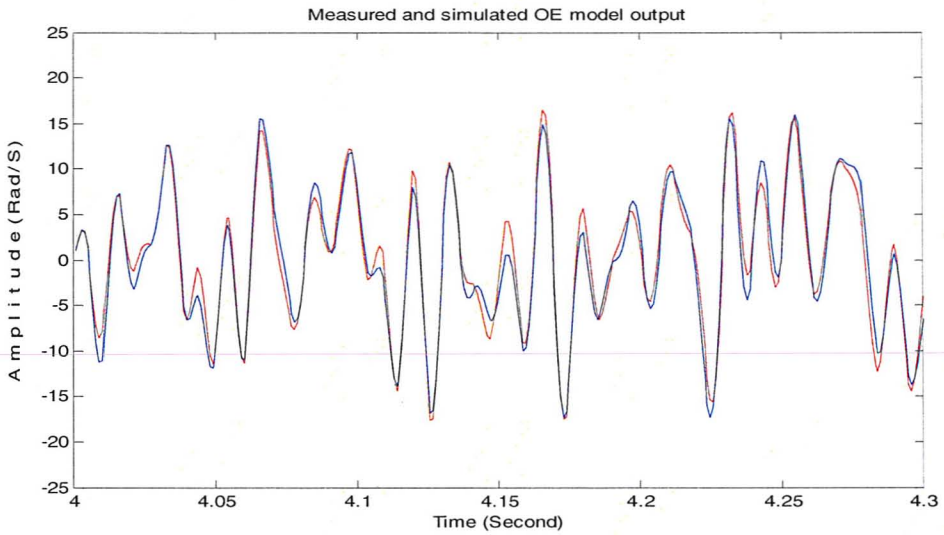

Figure 47 Measured and the OE Model Output (74.9% Fit)

3.5.4 Box-Jenkins (BJ) Model

The BJ model is the most flexible from considered and contains two different transfer functions for the system $G = \frac{B}{F}$ and noise $H = \frac{C}{D}$. Assigning $A(z) = 1$ in Equation (3.24) gives the Box-Jenkins model as:

$$y_k = \frac{B(z)}{F(z)}u(t) + \frac{C(z)}{D(z)}e_k \qquad (3.31)$$

The BJ model is diagrammatically shown in Figure 48.

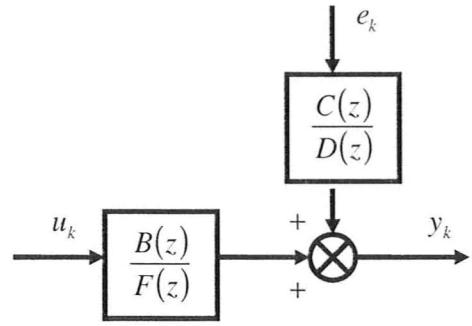

Figure 48 Block Diagram of the BJ Model

The best fit of the 2nd order BJ model is obtained as:

$$y_k = \frac{19.2z^{-2}}{1-1.226z^{-1}+0.5311z^{-2}}u_k + \frac{1+z^{-1}}{1-1.845z^{-1}+0.9187z^{-2}}e_k \qquad (3.32)$$

Note from Equations (3.30) and (3.32) that the OE model and the BJ model have the same system transfer function. The measured and the BJ model output are shown in Figure 49.

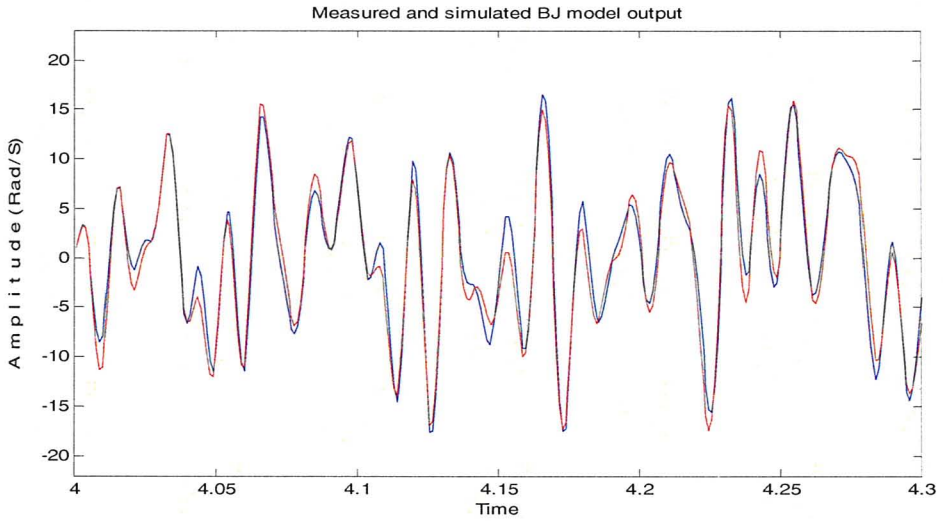

Figure 49 Measured and the BJ Model Output (74.9% Fit)

3.6 Model Validation

The performances of the identified models are further validated by using step, PRBS, and chirp responses. The measured outputs of the step, the PRBS and the chirp responses are all zero-phase filtered by using a 10^{th} order digital Butterworth filter with 100 Hz cut-off frequency. The measured and model outputs are shown in Figure 50. The Root Mean Squared prediction Errors (RMSEs) are listed in Table 5. The performance of the ARX and ARMAX models are very close. The performance of the OE and the BJ Models are very close. The ARX and the ARMAX have better performance on step response. However the OE and the BJ models have better performance when PRBS or Chirp signals are used. The OE model is chosen given its best overall performance.

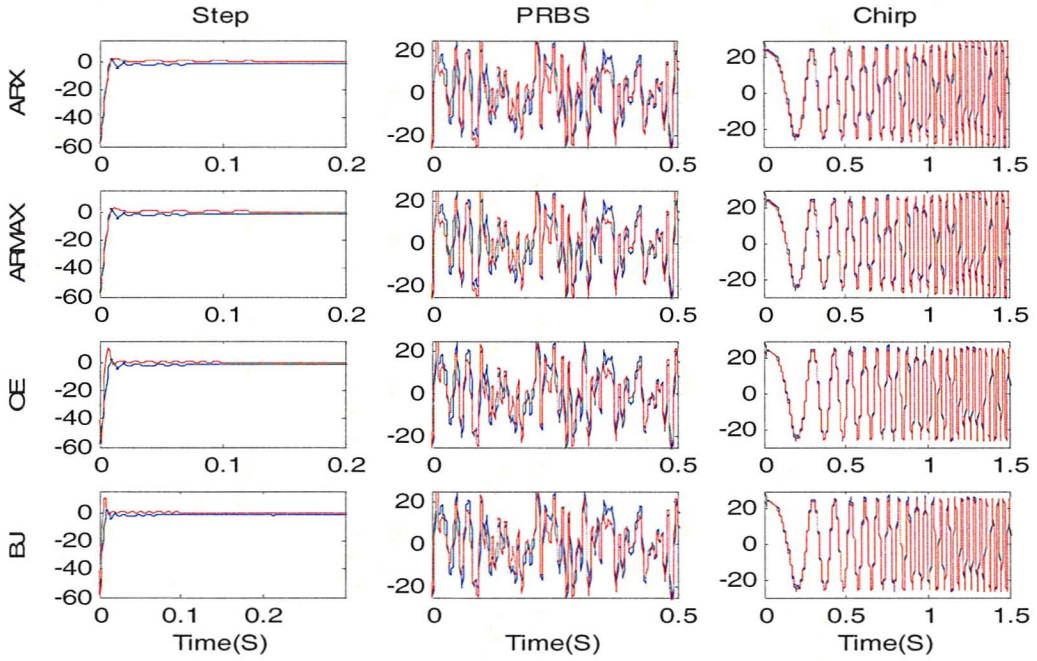

Figure 50 Measured and Simulated Outputs

Table 5 *RMSEs of the Identified Models*

Model	Step	PRBS	Chirp
ARX	0.7843	2.1846	5.7443
ARMAX	0.7837	2.1981	5.7415
OE	0.8729	1.8640	4.5282
BJ	0.8891	1.8797	4.4983

3.7 Conclusions on System Identification

System identification is a very useful way of obtaining an accurate model of a system when it is very difficult to do so through analytical means. The initial tests

provided information of the belt drive system's drift, deadzone, time delay, time constant, linearity, open-loop bandwidth and the order of system as listed in Table 6. Note that the linear region of the system has been reconfirmed as 0 to 1 volt.

Table 6 *Characteristic Parameters of the Belt Drive System*

Drift (V)	Dead-zone(V)	Time Delay(S)	Time Constant(S)	Linear Range(V)	Open-Loop Bandwidth(Hz)	Order of System
-0.015	0.025	0.002	0.01	0-1	70	2

The final black-box model candidate is a 2nd order OE model with 2 ms time delay. The identified transfer function of the belt drive system is obtained as:

$$G(z)=\frac{19.2z^{-2}}{1-1.226z^{-1}+0.5311z^{-2}} \qquad (3.33)$$

This model is used for the design of a robust control strategy.

Chapter 4 Literature Review of Sliding Mode Control Systems

Considerable research has been performed on the control of belt-drives in recent years. Li and Cheng (1994) presented a PID controller with adaptive compensation of inertial force for a belt-driven high speed positioning table. Li and Rehani (1996) designed a PID controller with acceleration and Coulomb friction compensation (PIDAF) for the tracking control of the angular position of the belt drive turning table based on an identified model. Lee and Rutherford (1999) presented the frequency reshaped linear quadratic (FRLQ) control in the development of a low-cost human level performance belt driven robot. In their work, a simple first order filter was chosen for determining a performance index for the linear quadratic optimal controller that would panelize the frequencies at or above the first resonance frequency of the belt drive. Jayawardene, Nakamura and Goto (2003) presented a PID control algorithm with an off-line trajectory planning under maximum velocity and acceleration constraints with compensation of delay dynamics and vibrations for accurate position control of belt drives. The same strategy was utilized for the cooperative control of two industrial robots and belt drive machines (Jayawardene, 2009). Selezneva (2009) implemented a PID algorithm with automatic tuning for the tracking control of the belt-drive systems. A. Sabanovic, O. Sozbilir, G. Goktug & N. Sabanovic (2003) proposed a Sliding Mode Control (SMC) algorithm based on a predefined structure of the time derivative of Lyapunov function candidate for a linear timing-belt drive. The related issues were further investigated by the

same group of the authors (Hace, Jezernik & Sabanovic, 2005; Hace, Jezernik & Sabanovic, 2007; Hace, Jezernik & Terbuc, 2009).

Among those control algorithms mentioned above, the SMC has gradually sought more attention in the control community due to its robustness and simplicity. The SMC is extensively reviewed in this chapter.

4.1 Background of Sliding Mode Control

Variable Structure Systems (VSS) are systems that have discontinuities in the differential equations that describe their dynamic behaviour. Such systems have a state space that is segmented by Hyperplanes. In each segment of the state space, the system may be modeled by a continuous and differentiable function. This continuous characterization of the system changes as the state trajectory cross a discontinuity hyperplane.

The following mathematical definition is commonly used for describing a VSS. A VSS is a discontinuous nonlinear system represented as $\dot{\mathbf{x}} = \Phi(\mathbf{x}, \mathbf{u}, t)$ where $\Phi(\mathbf{x}, t) = [\varphi_1(\mathbf{x}, \mathbf{u}, t), \varphi_2(\mathbf{x}, \mathbf{u}, t), \cdots, \varphi_n(\mathbf{x}, \mathbf{u}, t)]^T : R^{n+1} \rightarrow R^n$ is composed of piecewise continuous nonlinear function with state vector $\mathbf{x} = [x_1, x_2, \cdots, x_n]^T \in R^n$, control $\mathbf{u} \in R^m$ and time variable $t \in R$ (Wikipedia, 2010). The state space is divided into different regions within which the system exhibits continuous behaviour φ_n. At the boundaries of the regions, $\Phi(\mathbf{x}, \mathbf{u}, t)$ is not differentiable and the dynamics of the system switches abruptly, i.e. the structure of the system varies over different regions.

A good characteristic of the VSS is that they can possess new properties which are not inherent in continuous systems. For instance, asymptotic stability can be realized by combining two structures, none of which are asymptotically stable. In variable structure control, a discontinuity hyperplane is artificially introduced in the system by using a control input. A special form of variable structure control (VSC) is obtained when the states are directed toward the discontinuity or switching hyperplane and are forced to slide along it. This special form of VSC is referred to as sliding mode control (SMC) (Utkin, 1977). In SMC, the system dynamics is governed by a reduced order differential equation, and is robust to internal parameter variations and external disturbances.

The design of SMC consists of two tasks. Firstly a switching hyperplane $\mathbf{s}(\mathbf{x}) = \mathbf{0}$ is defined based on a desired dynamic behaviour such that stabilization, tracking and regularization are attained. Note that $\mathbf{s}(\mathbf{x})$ provides a measure of the distance of the state trajectory from the switching hyperplane located at $\mathbf{s}(\mathbf{x}) = \mathbf{0}$. Secondly a discontinuous control law $\mathbf{u}(t) = \begin{cases} \mathbf{u}^+(t) & \mathbf{s}(\mathbf{x}) > 0 \\ \mathbf{u}^-(\mathbf{x}) & \mathbf{s}(\mathbf{x}) < 0 \end{cases}$ is designed to drive the states onto the switching hyperplane in a finite time and to force them to remain and slide along the hyperplane. The control is continuous or smooth in every region, but alters its structure when the dynamics of the state trajectory crosses the boundary defined by the switching hyperplane. SMC input switches from one value or continuous function to another one based on the position of the state trajectory with respect to the switching hyperplane. This switching

can occur at a high frequency. By combining the system model, the switching hyperplane, and the control law, a SMC system is described as $\begin{cases} \dot{\mathbf{x}} = \varphi(\mathbf{x}, \mathbf{u}, t) \\ \mathbf{s}(\mathbf{x}) \\ \mathbf{u} = \begin{cases} \mathbf{u}^+(t) & \mathbf{s}(\mathbf{x}) > 0 \\ \mathbf{u}^-(t) & \mathbf{s}(\mathbf{x}) < 0 \end{cases} \end{cases}$.

There are two phases associated with a SMC system. Phase one is called the *reaching mode* in which the system dynamics is directed to the hyperplane in a finite time. The conditions to guarantee the reachablity are called the *reaching conditions*. For stability, the reaching condition is given as $\begin{cases} \dot{\mathbf{s}} < 0 & when\ \mathbf{s} > \mathbf{0} \\ \dot{\mathbf{s}} > 0 & when\ \mathbf{s} < \mathbf{0} \end{cases}$ or $\mathbf{s}\dot{\mathbf{s}} < 0$ (Gao, 1990). The reaching condition ensures the stability, but does not determine how the sliding surface is reached. The manner in which the sliding surface is reached determines the characteristics of the transient response. There are three commonly used approaches (Gao, 1990): (1) Constant approaching law in which, $\dot{\mathbf{s}} = -\varepsilon \operatorname{sgn}(\mathbf{s}), \varepsilon > 0$ so that the switching function approaches zero with a constant velocity; (2) Proportional approaching law in which $\dot{\mathbf{s}} = -\varepsilon \operatorname{sgn}(\mathbf{s}) - k\mathbf{s}\ \varepsilon > 0, k > 0$ so that the switching function approaches zero with a velocity equalling to the summation of a constant plus an amount proportional to the distance of the state trajectory from the switching hyperplane; (3) Exponential approaching law in which $\dot{\mathbf{s}} = -k|\mathbf{s}|^{\alpha} \operatorname{sgn}(\mathbf{s})\ \ k > 0, 1 > \alpha > 0$ so that the switching function approaches zero with a velocity of an exponential function of the distance of the state trajectory from the switching hyperplane.

Once the sliding surface is reached, the system enters its next phase of sliding mode. When in sliding mode, the system behaviour changes to a reduced order form and robust to disturbances and plant parameter variations.

The exceptional properties of SMC such as order-reduction and robustness to disturbances and parametric variations have drawn significant enthusiasm in the control community both in theoretical development and practical applications.

4.1.1 Brief History of SMC

The SMC theory was originated in the Soviet Union in the early 1930s. Its development has undergone four stages: the originating stage, the early stage, the multi-input linear systems stage, and the advanced stage (Gao, 1990; Hung, Gao & Huang, 1993; Utkin, 1999).

The Originating Stage of VSC (1930-1956)

The origin of Variable Structure Control (VSC) can be traced back to the 1930s in the former USSR (Utkin, 1999). The investigation of systems with discontinuous control actions remained at a high level during this period. As an example, relay or "on-off" regulators played an important role in the design of feedback control systems due to its ease of implementation and high efficiency of hardware. Typical examples included the vibration control in the context of voltage control for a DC generator of an aircraft (Kulebakin, 1932) and the sliding mode relays for controlling the course of a ship (Nikolski, 1934). The concepts of modern control theory such as phase plane, switching line, and even sliding mode, first appeared in the publications in that time.

The Early Stage of the SMC (1957-1970)

The SMC was developed in this stage focusing on the systems in phase conical form. The SMC studied in this period considered systems modeled by high-order, linear differential equations with a single input, such that:

$$x^{(n)}(t)+a_n x^{(n-1)}(t)+a_{n-1}x^{(n-2)}(t)+\cdots+a_2\dot{x}(t)+a_1 x = bu(t)$$

The switching surface was defined as a special second-order function:

$s(\mathbf{x}) = x_1\mathbf{c}^t\mathbf{x} = x_1(c_1x_1 + c_2x_2 + \cdots + c_nx_n)$, and the control law was described as $u = \Psi x_1$

with $\Psi = \begin{cases} \alpha \ when \ s(\mathbf{x}) < 0 \\ \beta \ when \ s(\mathbf{x}) > 0 \end{cases}$ (Hung, Gao & Huang, 1993).

The Stage of Multi-Input Systems (1970-1980)

During this period, the theory of SMC for general multi-input linear systems was firmly established. The general form of a multi-input linear system was represented as $\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}$ with state variable $\mathbf{x} \in R^n$, input $\mathbf{u} \in R^m$, system matrix $\mathbf{A} \in R^{n\times n}$, and input matrix $\mathbf{B} \in R^{n\times m}$. The switching function was depicted as $\mathbf{s}(\mathbf{x}) = \mathbf{Cx} = [s_1(\mathbf{x}), s_2(\mathbf{x}), \cdots, s_m(\mathbf{x})]^T$ with a switching matrix $\mathbf{C} \in R^{m\times n}$. The control law was described as $u_i(\mathbf{x}) = \begin{cases} \Psi_i^+ \ when \ s_i(\mathbf{x}) > 0 \\ \Psi_i^- \ when \ s_i(\mathbf{x}) < 0 \end{cases}$ $i = 1, 2, \cdots, m$, with s_i denoting the i^{th} scalar switching function (Hung, Gao & Huang, 1993). The SMC theory did not draw much attention from the control community during this period mainly due to two reasons: the theory was overshadowed by the popularity of linear control systems design, and the robustness property of the SMC was not fully recognized.

The Advanced Stage of SMC (1980-present)

Since 1980, the research and development of the SMC have been greatly accelerated both in theory and applications. General SMC design methods for complex systems have been proposed taking advantage of SMC's robustness property. The achievements in this period include: the development of SMC for nonlinear systems (Slotine, 1991; Gao, Hung, 1993; Misawa, 1997; Bandyopadhyay & Thakar, 2008), discrete-time systems (Furuta, 1990; Hui & Zak 1999; Chang, 2002), systems with time-delay (H. P. Pang, C. J. Liu & A. Z. Liu, 2006; Qu & Wang, 2006; Pai, 2008), stochastic systems (Y. Shen, W. Shen & J. Gu, 2005; Y. Niu, Z. Wang & X. Wang, 2010), large-scale systems (Chaouki & Moncef, 2007) and infinite dimensional systems (Utkin, 1998). The extension of the objectives of SMC was beyond system stabilization to include motion tracking, model following, model reaching, adaptive and optimal control and state observation. Sliding mode state observation was considered in (Raoufi & Zinober, 2008; Alessandri, Cuneo & Punta, 2009). Additional properties of SMC were explored including the invariance of the sliding mode to system disturbances and parameter variations in (Janardhanan & Bandyopadhyay, 2006), and robustness of the reaching phase, elimination or reduction of chattering in (Loh & Yeung, 2010). The applications of SMC have been reported in numerous industrial problems such as control of electric drives (F. J. Lin, C. H. Lin & Shen, 2002; Bian, Man, Song & Ren, 2006; Rashid & Zidan, 2008; Claudia & Miguel, 2008; Chuang, Haung, Lee, Kao & Fung, 2009), power systems (Samarasinghe & Pahalawaththa, 1997; Fernandes & Alcalde, 2007; Wang & Mao, 2009), robot

manipulators (Hashimoto, Maruyama & Harashima, 1987; Yu & Lloyd, 1997; Pandian, Hanmandlu & Gopal, 1988; Machado & Carvalho, 1988; Habibi & Richards, 1992; Khalal, Mellit, Rahim, Salhi & Guessoum, 2007; Vargas & Ledwich, 2010), mobile drives, autonomous underwater vehicles (Cunha, Costa & Hsu, 1991; Healey & Lienard, 1993; P. M. Lee, Hong, Lim, C. M. Lee, Jeon & Park, 1999; Wang, Liu, Yu & Xu, 2002; Xiong, Bian, Chang & Zou, 2004; Han, Sun & Mo, 2008), automotive applications (Moskwa, 1993; Utkin & Chang, 2002; Uchida, Murata, Yabumi, Morita & Kando, 2006; Mao & Lu, 2008; Kim & Wang, 2009), climate control, aircraft and spacecraft (Young, 1978; Jafarov & Tasaltin, 2000; Shtessel, Buffington & Banda, 2002; Huang, Kuo & Way, 2003; Promtun & Seshagiri, 2009; Ramirez, Thomas & Dwyer, 1986; Ramirez, 1990; Z. B. Li, Wang & J. F. Li, 2004; Hu & Ma, 2005, Hu & Ma, 2006; Jiang, Hu & Ma, 2008).

4.1.2 Literature Review of SMC

Among the abundant publications in SMC field, Utkin's work is notable and outstanding. His early research centered on the design of SMC in phase canonical form and the design principles for multi-input SMC systems. Utkin (1977, 1983) reviewed the historical development of the SMC concept, the role of sliding modes in development of the SMC theory, the stages of SMC development, and the synthesis of SMC using decomposition principle in sliding modes. Another notable survey paper was published by Huang, Gao & Huang (1993). They presented the basic notions, historical development, brief theoretic background of SMC systems, the design of SMC for linear time-invariant multi-input systems and nonlinear multivariable systems. Kaynak, Erbatur & Ertugrul

(2001) surveyed how "intelligence" could be incorporated in sliding mode controllers by using computational intelligence methodologies such as artificial neural networks (ANN), fuzzy logic (FL) and probabilistic reasoning (PR). The focus of their work was to alleviate the well known problems of inadequateness of the traditional analytical techniques for modeling, understanding and predicting the behaviours of complex systems in practical implementations of SMC. Yu & Kaynak (2009) examined the current progress and future perspectives of the integration of SMC with the computational intelligence methodologies.

4.1.3 Lyapunov Stability Analysis in SMC

Stability is one of the most important considerations in the design of a control system. According to (Nise, 2004), a system is stable if the natural response approaches zero as time approaches infinity; a system is unstable if the natural response approaches infinity as time approaches infinity; a system is marginally stable if the natural response neither decays nor grows but remains constant or oscillates. Using total (natural and forced) response, the definition of stability is given as: a system is stable if every bounded input yields a bounded output (BIBO); a system is unstable if any bounded input produces an unbounded output. Unbounded growth of the natural response of an unstable system usually causes damage to that system, to adjacent properties, and can be a danger to human life. The transient response of an unstable system grows infinitely; its total response does not approach a steady-state value.

Determination of stability is extremely important to control systems engineering. For a linear, time-invariant system, there are many stability criteria available such as the Routh's stability criterion (Nise, 2004) and the Nyquist stability criterion (Ogata, 1997). Unfortunately these criteria are only applicable to linear time-invariant systems. If the system is nonlinear, or linear but time-varying, such stability criteria are not directly applicable. The second method of Lyapunov, also called the direct method of Lyapunov, is the most general method for determining the stability of nonlinear, and/or time-varying systems (Ogata, 1997). This method also applies to linear time-invariant systems.

Lyapunov created two methods, called the first method and the second method, for determining the stability of dynamic systems described by ordinary differential equations. The first method of Lyapunov needs an explicit form of solutions of the differential equations for the analysis. The second method of Lyapunov determines the stability of a system without solving the differential equations which is very advantageous due to the fact that the task of solving nonlinear and/or time-varying differential equations is normally difficult. When applied to the determination of stability of nonlinear systems, the second method of Lyapunov usually requires considerable experience and ingenuity, but often succeeds when other methods fail.

Consider a system (Ogata, 1997):

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$$

where $\mathbf{x} \in R^n$ is state vector, and $\mathbf{f} \in R^n$ is a function of $\mathbf{x}$ and time t. It is assumed that the system has a unique solution, starting at the given initial condition denoted as $\phi(t, (\mathbf{x}_0, t_0))$, where $(\mathbf{x}_0, t_0)$ is the initial state and t is time.

Equilibrium State

If the system is linear, time-invariant, i.e., $\mathbf{f}(\mathbf{x},t)=\mathbf{A}\mathbf{x}$, then there exists only one equilibrium state if $\mathbf{A}$ is non-singular and infinitely many equilibrium states if $\mathbf{A}$ is singular. For nonlinear systems, there may be one or more equilibrium states corresponding to the constant solutions of the system. Any isolated equilibrium state can be shifted to the origin of the coordinates, i.e., $\mathbf{f}(\mathbf{0},t)=\mathbf{0}$ by a coordinate translation. So it is only needed to consider the stability analysis of equilibrium at the origin.

Stability in the Sense of Lyapunov

Let a spherical region of radius ς about an equilibrium state $\mathbf{x}_e$ be defined as $\|\mathbf{x}-\mathbf{x}_e\|\le\varsigma$, where $\|\bullet\|$ is the Euclidean norm, and $\|\mathbf{x}-\mathbf{x}_e\|=\left[(x_1-x_{1e})^2+(x_2-x_{2e})^2+\cdots+(x_n-x_{ne})^2\right]^{\frac{1}{2}}$. Let $S(\delta)$consist of all potential initial points $\mathbf{x}_0$ such that $\|\mathbf{x}_0-\mathbf{x}_e\|\le\delta$, and $S(\varepsilon)$ consist of all points such that $\|\phi(t,(\mathbf{x}_0,t_0))-\mathbf{x}_e\|\le\varepsilon$ for all $t>t_0$. An equilibrium state $\mathbf{x}_e$ is called *stable in the sense of Lyapunov* if, for any given ε, there is an $S(\delta)$, such that trajectories starting in $S(\delta)$ would not leave $S(\varepsilon)$ as t grows infinitely (Ogata, 1997).

Asymptotic Stability

An equilibrium state $\mathbf{x}_e$ of the system is called *asymptotically stable* if the equilibrium state $\mathbf{x}_e$ is stable in the sense of Lyapunov and if every solution of $\phi(t,(\mathbf{x}_0,t_0))$ starting within $S(\delta)$ converges to $\mathbf{x}_e$ as t grows to infinity. Asymptotical stability is a local concept. The region in where the system is asymptotically stable is the

domain of attraction. Every trajectory originated in the domain of attraction converges to the equilibrium state $\mathbf{x}_e$ and asymptotically stable. If the domain of attraction covers all of the state space, or in other words, trajectories originated everywhere converge to the equilibrium state $\mathbf{x}_e$, the equilibrium $\mathbf{x}_e$ is *asymptotically stable in the large or globally stable* (Ogata, 1997).

Second Method of Lyapunov

The second method of Lyapunov is based on the fact that if a system has an asymptotic equilibrium state $\mathbf{x}_e$, then the stored energy of the system within the domain of attraction decays as time grows until it reaches its minimum value at the equilibrium state $\mathbf{x}_e$. Lyapunov designed a scalar, fictitious energy function, the Lyapunov function $v(\mathbf{x},t)$. Lyapunov function is continuous, has a unique global minimum at the equilibrium state $\mathbf{x}_e$ with respect to all other states, and is non-increasing for all state trajectories $\mathbf{x}(t)$ in time t. The sign behaviour of the Lyapunov function $v(\mathbf{x},t)$ and its derivative $\dot{v}(\mathbf{x},t)$ give the information about stability, asymptotic stability, and instability without solving the differential equations.

The following theorems apply to the second method of Lyapunov. They are restated here without proof from (Ogata, 1997).

Theorem4.1 For the system: $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathrm{t})$ with $\mathbf{f}(\mathbf{0},\mathrm{t}) = \mathbf{0}$ for all $t > t_0$, if there exists a scalar function $V(\mathbf{x},t)$ having continuous, first partial derivatives and satisfying the

conditions: $V(\mathbf{x},t)$ is positive definite and $\dot{V}(\mathbf{x},t)$ is negative definite, the equilibrium state at the origin is uniformly asymptotically stable.

Theorem4.2 For the system: $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t)$ with $\mathbf{f}(\mathbf{0},t) = \mathbf{0}$ for all $t > t_0$, if there exists a scalar function $V(\mathbf{x},t)$ having continuous, first partial derivatives and satisfying the conditions: $V(\mathbf{x},t)$ is positive definite, $\dot{V}(\mathbf{x},t)$ is negative semi-definite for any initial state $(\mathbf{x}_0(\neq \mathbf{0}),t_0)$, then the equilibrium state at the origin of the system is uniformly asymptotically stable in the large.

Theorem4.3 For the system: $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t)$ with $\mathbf{f}(\mathbf{0},t) = \mathbf{0}$ for all $t > t_0$, if there exists a scalar function $V(\mathbf{x},t)$ having continuous, first partial derivatives and satisfying the conditions: $V(\mathbf{x},t)$ is positive definite in some region about the origin and $\dot{V}(\mathbf{x},t)$ is positive definite in the same region, then the equilibrium state at the origin is unstable.

4.2 Continuous-Time Sliding Mode Control

Continuous-time sliding mode control systems (CSMC) are feedback systems with discontinuous feedback control law. The control law switches the input of the system in order to maintain the state trajectory on a pre-specified switching hyperplane under sliding mode. By introducing the sliding mode into the system, it is possible to achieve stability, robustness to disturbance and parametric variations.

4.2.1 CSMC of Linear Time-Invariant Systems

Consider a linear time invariant system given as:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \qquad (4.1)$$

where $\mathbf{x} \in R^n$ is the state vector, $\mathbf{A} \in R^{n\times n}$ is the system matrix, $\mathbf{B} \in R^{n\times m}$ is the input matrix with $rank(\mathbf{B}) = m$, and $\mathbf{u} \in R^m$ is the input vector. The system is assumed controllable. The design of CSMC can be divided into two steps: (1) design a switching manifold so that the desired dynamics is achieved when the system is confined on it; (2) design a control law to drive the state trajectory onto the switching manifold in finite time and to stay on it thereafter.

Design of the Switching Manifold

Since the input matrix satisfies the condition $rank(\mathbf{B}) = m$, matrix $\mathbf{B}$ can be partitioned by reordering the state vector components as (Utkin, 1999):

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \qquad (4.2)$$

where $\mathbf{B}_1 \in R^{(n-m)\times m}$, $\mathbf{B}_2 \in R^{m\times m}$ with $rank(\mathbf{B}_2) = m$. A non-singular coordinate transformation can be defined as (Utkin, 1999):

$$\mathbf{T} = \begin{bmatrix} \mathbf{I}_{n-m} & -\mathbf{B}_1\mathbf{B}_2^{-1} \\ \mathbf{0} & \mathbf{B}_2^{-1} \end{bmatrix} \qquad (4.3)$$

With the transformation matrix $\mathbf{T}$, the controllable system (4.1) can be converted to the regular form as:

$$\begin{aligned}\dot{\mathbf{x}}_1 &= \mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2 \\ \dot{\mathbf{x}}_2 &= \mathbf{A}_{21}\mathbf{x}_1 + \mathbf{A}_{22}\mathbf{x}_2 + \mathbf{u}\end{aligned} \qquad (4.4)$$

where $\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix} = \mathbf{T}\mathbf{x}$, $\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}$, $\mathbf{T}\mathbf{B} = \begin{bmatrix}\mathbf{0} \\ \mathbf{I}_{m\times m}\end{bmatrix}$. The first equation of (4.4) can be treated as a subsystem with an intermediate input $\mathbf{x}_2$. Since $(\mathbf{A},\mathbf{B})$ is controllable, the pair $(\mathbf{A}_{11},\mathbf{A}_{12})$ is controllable also (Utkin, 1999), the subsystem can be stabilized using state feedback as:

$$\mathbf{x}_2 = -\mathbf{G}\mathbf{x}_1 \qquad (4.5)$$

The closed-loop dynamics can be obtained as:

$$\dot{\mathbf{x}}_1 = (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{G})\mathbf{x}_1 \qquad (4.6)$$

The eigenvalues of matrix $(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{G})$ can be arbitrarily assigned by using the pole-placement technique (Ogata, 1994) to meet the desired dynamic requirements. The switching function can be obtained from Equation (4.5) as:

$$\mathbf{s} = \mathbf{G}\mathbf{x}_1 + \mathbf{x}_2 = \mathbf{0} \qquad (4.7)$$

Design of Control Law

A piecewise linear discontinuous control can be given as (Utkin, 1999):

$$\mathbf{u} = -(\alpha|\mathbf{x}| + \delta)\mathrm{sgn}(\mathbf{s}) \qquad (4.8)$$

where α and δ are positive constants, $\mathrm{sgn}(\mathbf{s})^T = [\mathrm{sgn}(s_1) \;\; \mathrm{sgn}(s_2) \;\; \cdots \;\; \mathrm{sgn}(s_m)]$, and $|\mathbf{x}| = \sum_{i=1}^{n} |x_i|$.

Stability analysis

The stability of the CSMC system is investigated by using the continuous-time Lyapunov stability theory. A Lyapunov function can be defined as:

$$\mathbf{v} = \frac{1}{2}\mathbf{s}^T\mathbf{s} \qquad (4.9)$$

Its derivative can be determined with the system model and the control law of Equation (4.8) as:

$$\begin{aligned}\dot{\mathbf{v}} = \mathbf{s}^T\dot{\mathbf{s}} &= \mathbf{s}^T\left(\mathbf{G}\dot{\mathbf{x}}_1 + \dot{\mathbf{x}}_2\right) = \mathbf{s}^T\left\{\mathbf{G}\left(\mathbf{A}_{11}\mathbf{x}_1 + \mathbf{A}_{12}\mathbf{x}_2\right) + \left[\mathbf{A}_{21}\mathbf{x}_1 + \mathbf{A}_{22}\mathbf{x}_2 - \left(\alpha|\mathbf{x}| + \delta\right)\mathrm{sgn}(\mathbf{s})\right]\right\} \\ &= \mathbf{s}^T\left[\left(\mathbf{GA}_{11} + \mathbf{A}_{21}\right)\mathbf{x}_1 + \left(\mathbf{GA}_{12} + \mathbf{A}_{22}\right)\mathbf{x}_2\right] - \left(\alpha|\mathbf{x}| + \delta\right)|\mathbf{s}| \\ &\le |\mathbf{s}|\left|\left(\mathbf{GA}_{11} + \mathbf{A}_{12}\right)\mathbf{x}_1 + \left(\mathbf{GA}_{12} + \mathbf{A}_{22}\right)\mathbf{x}_2\right| - \left(\alpha|\mathbf{x}| + \delta\right)|\mathbf{s}|\end{aligned} \qquad (4.10)$$

With respect to Expression (4.10), for any given δ, there exists an α in Expression (4.10) that makes the derivative $\dot{\mathbf{v}}$ negative; this guarantees the convergence of the state trajectory to the manifold $\mathbf{s} = \mathbf{0}$. The convergence rate can be made fast by increasing the parameters α and δ in the control law of Equation (4.8).

Taking consideration of disturbances and parametric variations, the linear system of Equation (4.1) can be rewritten as:

$$\dot{\mathbf{x}} = \left(\mathbf{A} + \Delta\mathbf{A}(\mathbf{t})\right)\mathbf{x}(t) + \mathbf{Bu}(t) + \mathbf{Qw}(t) \qquad (4.11)$$

where $\Delta\mathbf{A}(t)$ are parametric variations and $\mathbf{w}(t) \in R^l$ are disturbances. According to Utkin (1999), sliding modes in any manifold are invariant with respect to parametric variations $\Delta\mathbf{A}(t)$ and disturbance vector $\mathbf{w}(t)$ if there exist matrices Λ_A and Λ_Q such that:

$$\begin{aligned}\Delta\mathbf{A} &= \mathbf{B}\Lambda_A \\ \mathbf{Q} &= \mathbf{B}\Lambda_Q\end{aligned} \qquad (4.12)$$

If the conditions (4.12) hold, the system of Equation (4.11) has a regular form similar to Equation (4.4) as:

$$\begin{aligned}\dot{\mathbf{x}}_1 &= \mathbf{A}_{11}\mathbf{x}_1(t)+\mathbf{A}_{12}\mathbf{x}_2(t)\\ \dot{\mathbf{x}}_2 &= \mathbf{A}_{21}\mathbf{x}_1(t)+\mathbf{A}_{22}\mathbf{x}_2(t)+\mathbf{u}(t)+\Lambda_A\mathbf{x}(t)+\Lambda_Q\mathbf{w}(t)\end{aligned} \qquad (4.13)$$

A control law in the form of Equation (4.8) with a manifold of Equation (4.7) leads to a sliding mode equation (4.6) with desired dynamics and invariance property if a additional condition $\delta \geq \mathbf{w}_0$ where $\left|\Lambda_Q\mathbf{w}(t)\right| \leq \mathbf{w}_0$.

4.2.2 CSMC of Nonlinear Systems

Modeling inaccuracies have adverse effects on nonlinear control systems. They can be classified into structural (parametric) uncertainties, and unstructured uncertainties. Structural uncertainties correspond to inaccuracies on the terms included in the model. Unstructured uncertainties correspond to inaccuracies of system order and external perturbations. Two control approaches dealing with the system uncertainties are robust control and adaptive control (Slotine & Li, 1991). A simple approach to robust control is SMC for a nonlinear system as discussed in (Slotine & Li, 1991). Consider a nonlinear single-input system, such that:

$$x^{(n)} = f(\mathbf{x})+b(\mathbf{x})u \qquad (4.14)$$

where $x \in R$ is the output, $u \in R$ is the control input, $\mathbf{x} = \left[\, x \; \dot{x} \cdots x^{(n-1)} \,\right]^T$ is the state vector.

The nonlinear function $f(\mathbf{x})$ is not exactly known but estimated as $\hat{f}(\mathbf{x})$. The estimation error is assumed to be bounded by a known continuous function $F(\mathbf{x})$:

$$\left| f(\mathbf{x}) - \hat{f}(\mathbf{x}) \right| \le F(\mathbf{x}) \qquad (4.15)$$

The control gain $b(\mathbf{x})$ is not exactly known, but is bounded by known, possibly time-varying or state-dependent bounds:

$$0 < b_{\min}(\mathbf{x}) \le b(\mathbf{x}) \le b_{\max}(\mathbf{x}) \qquad (4.16)$$

The control input enters the system dynamics multiplicatively; the estimate $\hat{b}(\mathbf{x})$ of $b(\mathbf{x})$ can be chosen as the geometric mean of the above bounds:

$$\hat{b}(\mathbf{x}) = \sqrt{b_{\min}(\mathbf{x}) b_{\max}(\mathbf{x})} \qquad (4.17)$$

The estimate $\hat{b}(\mathbf{x})$ is also bounded as:

$$\beta^{-1} \le \frac{\hat{b}(\mathbf{x})}{b(\mathbf{x})} \le \beta,\ \beta = \sqrt{b_{\max}(\mathbf{x}) / b_{\min}(\mathbf{x})} \qquad (4.18)$$

The control problem is to get state vector $\mathbf{x}$ to track a pre-specified time varying state $\mathbf{x}_d = \left[x_d \ \dot{x}_d \cdots x_d^{(n-1)} \right]^T$ in the presence of modeling uncertainties on $\hat{f}(\mathbf{x})$ and $\hat{b}(\mathbf{x})$.

The tracking error is defined as $\tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}_d = \left[\tilde{x} \ \dot{\tilde{x}} \cdots x^{(n-1)} \right]^T$, and a time-varying switching surface in R^n state space is defined as:

$$s(\tilde{\mathbf{x}}, t) = \left(\frac{d}{dt} + \lambda \right)^{n-1} \tilde{x} = 0,\ \lambda > 0 \qquad (4.19)$$

With $n = 2$, the switching function becomes $s(\tilde{\mathbf{x}}, t) = \dot{\tilde{x}} + \lambda \tilde{x} = 0$. Similarly, with $n = 3$, the switching function turns into $s(\tilde{\mathbf{x}}, t) = \ddot{\tilde{x}} + 2\lambda \dot{\tilde{x}} + \lambda^2 \tilde{x} = 0$. So the switching function s is simply a weighted sum of the components of the tracking error vector $\tilde{\mathbf{x}}$. Since the

expression of $s(\tilde{\mathbf{x}},t)$ contains $\tilde{x}^{(n-1)}$, one only needs to differentiate $s(\tilde{\mathbf{x}},t)$ once to make the control input $u(t)$ to appear, the original n^{th}-order tracking problem is converted to a simpler first-order stabilization problem as:

$$\dot{s}(\tilde{\mathbf{x}},t)=0 \qquad (4.20)$$

The simplified first-order stabilization problem (4.20) can be solved through finding control input $u(t)$ based on the Lyapunov stability theory. The Lyapunov function for the scalar input system (4.14) is defined as:

$$v(\tilde{\mathbf{x}},t)=\frac{1}{2}s^2(\tilde{\mathbf{x}},t) \qquad (4.21)$$

Taking the derivative of Equation (4.21), and making it negative definite through the selection of the control law will guarantee the asymptotic stability, that is:

$$\dot{v}(\tilde{\mathbf{x}},t)=s\dot{s}<0 \qquad (4.22)$$

For global stability, Equation (4.22) can be redefined as:

$$\dot{v}(\tilde{\mathbf{x}},t)=s\dot{s}\le-\eta|s|,\ \eta>0 \qquad (4.23)$$

where $|\bullet|$ denotes absolute value of $(\bullet)$, the positive constant η determines the convergence rate.

Inequality (4.23) is also called the reaching condition which ensures the existence of sliding mode.

To illustrate the design procedure clearly, a third-order system is adopted as an example, that is, $n=3$. Design of control law is carried out in two steps. In each step only one parametric uncertainty is considered. Firstly a control law is derived by

setting $b(\mathbf{x})=1$. Then the control law is modified to account for uncertainties in the control gain $b(\mathbf{x})$.

Case 1: Control of Systems with Unit Input Gain

By setting $b(\mathbf{x})=1$, the third order differential Equation (4.14) is rewritten as:

$$x^{(3)} = f(\mathbf{x})+u \qquad (4.24)$$

Taking the derivative of $s(\tilde{\mathbf{x}},t)$ and combining it with (4.24) give:

$$\dot{s}(\tilde{\mathbf{x}},t)=\tilde{x}^{(3)}+2\lambda\ddot{\tilde{x}}+\lambda^2\dot{\tilde{x}}=f(\mathbf{x})+u-x_d^{(3)}+2\lambda\ddot{\tilde{x}}+\lambda^2\dot{\tilde{x}} \qquad (4.25)$$

The equivalent control $u_{eq}(t)$ may be defined as the continuous control law that would maintain the sliding mode dynamics without accounting for the presence of modeling uncertainties and external disturbances. In sliding mode, by solving $\dot{s}(\tilde{\mathbf{x}},t)=0$, the equivalent control is obtained as:

$$u_{eq}(t)=x_d^{(3)}-\hat{f}(\mathbf{x})-2\lambda\ddot{\tilde{x}}-\lambda^2\dot{\tilde{x}} \qquad (4.26)$$

where $\hat{f}(\mathbf{x})$ is an estimate of $f(\mathbf{x})$. In practice, the control law has to be discontinuous across the switching surface $s(\tilde{\mathbf{x}},t)=0$ to offset parametric variations and external disturbances. To satisfy the stability condition, a discontinuous term is added to the equivalent control, such that:

$$u(t)=u_{eq}(t)-K\,\mathrm{sgn}(s) \qquad (4.27)$$

where sgn is a signum function, $\begin{cases}\mathrm{sgn}(s)=1, & s>0\\ \mathrm{sgn}(s)=-1, & s<0\end{cases}$, $K>0$. By choosing K to be large enough to overcome uncertainties, the reaching condition of Equation (4.22) can be

satisfied. The magnitude of the constant K must be increased with the extent of parametric uncertainty. Substituting Equation (4.26) into Equation (4.27) gives:

$$u(t) = x_d^{(3)} - \hat{f}(\mathbf{x}) - 2\lambda \ddot{\tilde{x}} - \lambda^2 \dot{\tilde{x}} - K\,\mathrm{sgn}(s) \qquad (4.28)$$

The sliding mode dynamics of Equation (4.24) may be written as:

$$x^{(3)}(t) = f(\mathbf{x}) + x_d^{(3)} - \hat{f}(\mathbf{x}) - 2\lambda \ddot{\tilde{x}} - \lambda^2 \dot{\tilde{x}} - K\,\mathrm{sgn}(s) \qquad (4.29)$$

Combining Equations (4.23), (4.25) and (4.28) gives:

$$s\dot{s} = s\left(f(\mathbf{x}) - \hat{f}(\mathbf{x}) - K\,\mathrm{sgn}(s)\right) = s\left(f(\mathbf{x}) - \hat{f}(\mathbf{x})\right) - K|s| \le -\eta|s| \qquad (4.30)$$

Let $F(\mathbf{x})$ be the upper bound on model uncertainty, the reaching condition (4.30) is satisfied if:

$$K = F(\mathbf{x}) + \eta \qquad (4.31)$$

Case 2: Control of Systems with Uncertain Input Gains

A third-order nonlinear system with arbitrary control input gain is written as:

$$x^{(3)} = f(\mathbf{x}) + b(\mathbf{x})u_m(t) \qquad (4.32)$$

where $u_m(t)$ is the modified control input. As previously mentioned, it is assumed that the control gain $b(\mathbf{x})$ is not exactly known, but has a known sign, and is bounded by some known functions as defined in Equations (4.16), (4.17) and (4.18). The switching surface is defined in (4.19). The objective of design is to modify the control law (4.28) and condition (4.31) to satisfy the reaching condition of Equation (4.23).

Equations (4.24) and (4.32) are identical except that they have different input gains. Let:

$$u_m(t)=\frac{1}{\hat{b}(t)}\left[u_{eq}(t)-K\,\mathrm{sgn}(s)\right] \qquad (4.33)$$

From Equations (4.32) and (4.33), the derivative of the switching function becomes:

$$\dot{s}(\tilde{\mathbf{x}},t)=\tilde{x}^{(3)}+2\lambda\ddot{\tilde{x}}+\lambda^2\dot{\tilde{x}}=f(\mathbf{x})+\frac{b(\mathbf{x})}{\hat{b}(\mathbf{x})}\left[u_{eq}(t)-K\,\mathrm{sgn}(s)\right]-x_d^{(3)}+2\lambda\ddot{\tilde{x}}+\lambda^2\dot{\tilde{x}} \qquad (4.34)$$

Since $f(\mathbf{x})=\hat{f}(\mathbf{x})+\tilde{f}(\mathbf{x})$, Equation (4.34) can be written with Equation (4.26) as:

$$\dot{s}=\tilde{f}(\mathbf{x})+\left(\frac{b}{\hat{b}}-1\right)u_{eq}-\frac{b}{\hat{b}}K\,\mathrm{sgn}(s) \qquad (4.35)$$

Note that time and state variables in Equation (4.35) have been omitted for clarification. Substituting Equation (4.35) into Inequality (4.23) gives:

$$s\dot{s}=s\left(f-\hat{f}\right)+s\left(\frac{b}{\hat{b}}-1\right)u_{eq}-s\frac{b}{\hat{b}}K\,\mathrm{sgn}(s)\le-\eta|s|\Rightarrow\frac{\hat{b}}{b}\left(f-\hat{f}\right)s+\frac{\hat{b}}{b}\eta|s|+\left(1-\frac{\hat{b}}{b}\right)u_{eq}s\le K|s|.$$

From the above inequality with uncertainty bounds (4.15) and (4.18), a conservative value for coefficient K is obtained as:

$$K\ge\beta(F+\eta)+(\beta-1)|u_{eq}| \qquad (4.36)$$

Control law (4.33) with K bounded in Inequality (4.36) satisfies the reaching condition (4.23). However it is discontinuous across the switching surface frequently which results in chattering. Young, Utkin and Ozguner (1999) perceived the chattering phenomenon as motion that oscillates about the sliding manifold due to imperfections in the implementations of the switching curve. They believed that the chattering is produced by two possible mechanisms: (1) the presence of parasitic dynamics such as the fast actuator and sensor dynamics in series with the plant causes a small-amplitude high-frequency oscillation to appear in the vicinity of the sliding manifold; (2) the switching

non-idealities, due to time delays of the switching devices and the microprocessor code executions, cause high frequency oscillations. In general, control chattering is highly undesirable since it involves extremely high control activity that usually causes mechanical actuators to wear out prematurely, or excites un-modeled high-frequency dynamics.

Chattering Elimination

In general, chattering is undesirable, and has to be eliminated for the controller to perform properly. This can be done by setting a thin boundary layer neighbouring the switching function to smooth out the control discontinuity as (Slotine & Li, 1991):

$$\Psi = \left\{\tilde{\mathbf{x}}, \left|s(\tilde{\mathbf{x}},t)\right| \le \varphi\right\},\ \varphi > 0 \qquad (4.37)$$

where φ is the boundary layer thickness. The sign function in the control law $u_m(t)$ of Equation (4.33) is replaced with a saturation function as:

$$u_m(t) = \frac{1}{\hat{b}(t)}\left[u_{eq}(t) - Ksat\left(\frac{s}{\varphi}\right)\right]$$

$$sat\left(\frac{s}{\varphi}\right) = \begin{cases} \dfrac{s}{\varphi}, & \left|\dfrac{s}{\varphi}\right| \le 1 \\ \operatorname{sgn}\left(\dfrac{s}{\varphi}\right), & \left|\dfrac{s}{\varphi}\right| > 1 \end{cases} \qquad (4.38)$$

Note that if $\varphi \to 0$, then $sat(s/\mu) \to \operatorname{sgn}(s)$. In this case, the perfect tracking is achieved in the absence of real sliding mode condition. So the control law (4.33) can be viewed as a particular case of the generalized control law (4.38). Note also that the boundary layer provides a trade-off between the tracking precision and the level of the control activity.

The boundary layer can be time-varying, and through monitoring its width, insight can be gained into uncertainties. The development of control law with variable boundary layer is partially done in Equation (4.38). What remains is the determination of the gain that would satisfy the Lyapunov stability condition. For a system to be stable, the attractiveness of the boundary layer has to be maintained. When the trajectory is outside of the boundary layer, the distance of the switching function to the boundary layer has to decrease always, which is ensured by $\frac{d}{dt}(s-\varphi)\le -\eta$ when $s>\varphi$ and $\frac{d}{dt}(s-(-\varphi))\ge \eta$ when $s<-\varphi$, that is:

$$\begin{cases} \dot{s} \le \dot{\varphi} - \eta, & s > \mu \\ \dot{s} \ge \eta - \dot{\varphi}, & s < -\mu \end{cases} \qquad (4.39)$$

The time derivative of the Lyapunov function of Equation (4.21) with (4.39) can be found as $\dot{v}(\tilde{\mathbf{x}},t) = s\dot{s} \begin{cases} \le s\left(\dot{\varphi}-\eta\right) = |s|\left(\dot{\varphi}-\eta\right), & s>\varphi \\ \le s\left(\eta-\dot{\varphi}\right) = |s|\left(\dot{\varphi}-\eta\right), & s<-\varphi \end{cases}$, i.e. the reaching condition becomes as:

$$s\dot{s} \le \left(\dot{\varphi}-\eta\right)|s|, \quad |s|>\varphi \qquad (4.40)$$

The reaching condition (4.40) indicates that the time derivative of the thickness of the boundary layer is superimposed on to the reaching condition as an additional term; and if the thickness of the boundary layer is time-invariant, the reaching condition is the same as

(4.23); if the thickness increases $\left(\dot{\varphi} > 0\right)$, the boundary layer is getting wider, $s(\tilde{\mathbf{x}},t)$ is more likely trapped inside the boundary layer, the condition (4.40) is less stringent; otherwise, if the thickness decreases $\left(\dot{\varphi} < 0\right)$, the boundary layer is getting narrower, $s(\tilde{\mathbf{x}},t)$ the condition (4.40) needs to be more stringent to attract $s(\tilde{\mathbf{x}},t)$. In other words, the existence condition is more demanding during the contraction of the boundary layer and less demanding during the expansion of the boundary layer.

To satisfy the existence condition (4.40), an additional term $-\dot{\varphi}$ has to be added into the discontinuous control law by modifying the coefficient K into $K-\dot{\varphi}$, such that the control law becomes:

$$u_m(t) = \frac{1}{\hat{b}(t)}\left[u_{eq}(t) - \left(K - \dot{\varphi}\right) sat\left(\frac{s}{\varphi}\right)\right] \qquad (4.40)$$

The s-trajectories inside the boundary layer can be expressed directly from Equations (4.35) and (4.40) as:

$$\dot{s} = \left(f - \hat{f}\right) + \left(\frac{b}{\hat{b}} - 1\right) u_{eq} - \frac{b}{\hat{b}}\left(K - \dot{\varphi}\right)\frac{s}{\varphi} \qquad (4.41)$$

Every term of the right hand side of Equation (4.41) is continuous in $\mathbf{x}$, the solution of Equation (4.41) must be continuous in $\mathbf{x}$ as well. Thus the chattering is eliminated indeed.

4.3 Discrete-Time Sliding Mode Control

Most SMC approaches are based on the continuous-time models with discontinuous control action. In sliding mode, reduced-order dynamics is realized by high frequency switching control action. The switching action is normally achieved with two options: analog implementation and discrete implementation (Utkin, Guldner & Shi, 2009). The analog implementation of SMC needs very fast switching devices such as power transistors. This method is only suitable for systems with voltage control inputs. The discrete implementation is realized by using digital microcontrollers with a finite sampling rate. Direct implementation of a SMC designed for a continuous time system without modifications will introduce chattering due to the finite sampling rate. In order to eliminate chattering, one can theoretically increase the sampling rate, which is subject to limitations of the converting devices and can excite the unmodeled, undesirable high frequency dynamics within them. Design of SMC based directly on discrete-time systems is preferred.

Discrete-time sliding mode control (DSMC) can be categorized into two approaches. In the first, the focus is on mapping CSMC systems into DSMC counterparts (Sarpturk, Istcfanopulos & Kaynak, 1987; Gao, Wang & Homaifa, 1995). In the second, a new control algorithm is developed based on the discrete formulation of the Lyapunov stability condition (Furuta, 1990; Bartolini, Ferrara & Utkin, 1995; Young, Utkin & Ozguner, 1999). Application of DVSC can be found in a wide spectrum of system types, including nonlinear systems, multi-input-multi-output (MIMO) systems, large-scale, infinite dimensional systems and stochastic systems (Kao, 2005). The early stage of

DVSC were mainly implemented on SISO systems, linear systems with state feedback and switching surfaces defined in a special quadratic form (Ignaciuk & Bartoszewicz, 2009). Then, the property of robustness of DVSC systems with respect to system perturbations and disturbances was developed (Kao, 2005).

Compared to CSMC systems, DSMC systems have the following attributes (Gao, Wang & Homaifa, 1995): (1) starting from any initial state, the s-trajectory moves monotonically towards the switching plane and crosses it in finite time; (2) once the s-trajectory crosses the switching hyperplane for the first time, it will cross the hyperplane again at every successive sampling time which results in a zigzag motion about the switching plane; (3) the size of each crossing step is non-increasing and bounded in a fixed width in a subspace around the switching hyperplane.

A DSMC system satisfying Attributes (2) and (3) is called a *Quasi-Sliding-Mode (QSM).*The specified band containing the QSM is called the *Quasi Sliding Mode Band (QSMB)* which is represented as $\{\mathbf{x}|-\varphi < s(\mathbf{x}) < \varphi\}$. The width of QSMB is 2φ. A QSM is an *Ideal-Sliding-Mode (ISM)* if the width of the QSMB becomes zero. A DSMC system having all three attributes (1), (2) and (3) is said to be satisfying the reaching condition (Gao, Wang & Homaifa, 1995).

4.3.1 Design of Discrete-Time Switching Functions

Design of a switching function is the first step in the DSMC design. The techniques of designing switching functions are covered in (Hwang, 1992; Spurgeon, 1992; Gao, Wang & Homaifa, 1995; Chen & Chang, 2000; Bandyopadhyay & Thakar, 2008). The

method used for the design of switching function in continuous time is still applicable for discrete time systems. One method for the design of a switching function has been presented in Section 4.2.1. Another method addressed by Chen and Chang (2000) is presented as follows.

A discrete linear time invariant (LTI) system is given as:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \qquad (4.42)$$

where $\mathbf{x}_k \in R^n$ is the state vector, $\mathbf{A} \in R^{n\times n}$ is the system matrix, and $\mathbf{B} \in R^n$ is the input matrix, $\mathbf{u}_k \in R^m$ is the input vector. It is assumed that the pair $(\mathbf{A},\mathbf{B})$ is controllable. Design of a switching function is carried out in two stages. In stage one, a feedback gain matrix $\mathbf{K}$ for the closed-loop system is designed using pole assignment technique. In stage two, a switching function is determined based on the feedback gain matrix $\mathbf{K}$.

Since the pair $(\mathbf{A},\mathbf{B})$ is controllable, a feedback gain matrix $\mathbf{K} \in R^{m\times n}$ can be obtained by assigning n eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ for matrix $\mathbf{A}-\mathbf{BK}$ with the following three conditions (Chen & Chang, 2000): (1) the matrix $\mathbf{A}-\mathbf{BK}$ has no common eigenvalues with matrix $\mathbf{A}$; (2) the eigenvalues of $\mathbf{A}-\mathbf{BK}$ are so chosen that the closed-loop feedback system is stable and valued as $\left\{\lambda_1, \lambda_2, \cdots, \lambda_{n-m}, \underbrace{\lambda \quad \cdots \quad \lambda}_{m\lambda s}\right\}$ with $\lambda \neq \lambda_i, i = 1, 2, \cdots, n-m, 0 < \lambda < 1$; (3) the matrix $(\mathbf{A}-\mathbf{BK})$ is diagonalizable with m repeated eigenvalues. Under these conditions, a switching function can be defined as:

$$\mathbf{s}_k = \mathbf{G}\mathbf{x}_k = \mathbf{0} \qquad (4.43)$$

where $\mathbf{G} \in R^{m \times n}$ is the switching matrix which is determined by using the feedback gain matrix $\mathbf{K}$ with the following important properties (Chen & Chang, 2000), such that:

$$\begin{aligned} &\mathbf{G} = \mathbf{K}(\mathbf{A} - \lambda \mathbf{I}_n)^{-1} \\ &\mathbf{GB} = \mathbf{I}_m \\ &\left[\mathbf{I}_n - \mathbf{B}(\mathbf{GB})^{-1}\mathbf{G}\right]\mathbf{A} \text{ has the eigenvalues } \{\lambda_1, \lambda_2, \cdots, \lambda_{n-m}, 0, \cdots, 0\} \end{aligned} \quad (4.44)$$

Substituting Equation (4.44) into Equation (4.43) gives the switching function as:

$$\mathbf{s}_k = \mathbf{K}(\mathbf{A} - \lambda \mathbf{I}_n)^{-1}\mathbf{x}_k = \mathbf{0} \quad (4.45)$$

4.3.2 DSMC of Linear Time-Invariant Systems

There are plenty of articles published on DSMC of linear time invariant (LTI) systems. They cover a wide range from single-input single-output (SISO) to multi-input-multi-output (MIMO) systems. The DSMC algorithms are implemented by using either state or output feedbacks.

DSMC of Single Input LTI systems

Reaching Law Approach (RLA) (Gao, Wang & Homaifa, 1995) is a SMC method first proposed for continuous VSC systems. The RLA specifies an autonomous reaching condition, and then a SMC law is derived from the reaching condition combined with a known model of the plant given known bounds no perturbations. For the SMC of a single input discrete system, the reaching law is specified as:

$$s_{k+1} = (1 - qT)s_k - \varepsilon T \operatorname{sgn}(s_k),\ q > 0,\ \varepsilon > 0,\ 0 < 1 - qT < 1 \quad (4.46)$$

where T is the sampling interval. From the definition of the QSM, the state trajectory s_{k+1} must have a sign that is opposite to that of s_k. The first term of the right

hand side of Equation (4.46) has a sign that is the same as that of s_k. To make the sign of s_{k+1} to be opposite to that of s_k, the following condition has to be satisfied by the reaching law, that is:

$$s_{k+1} = (1-1qT)s_k - \varepsilon T \operatorname{sgn}(s_k) = \begin{cases} (1-qT)s_k - \varepsilon T < 0, & s_k > 0 \\ (1-qT)s_k + \varepsilon T > 0, & s_k < 0 \end{cases} \Rightarrow$$

$$\begin{cases} s_k < \dfrac{\varepsilon T}{1-qT}, & s_k > 0 \\ s_k > -\dfrac{\varepsilon T}{1-qT}, & s_k < 0 \end{cases}$$

From the above expression, the state trajectory satisfies the following condition:

$$|s_k| < \frac{\varepsilon T}{1-qT} \qquad (4.47)$$

The QSMB is determined from Expression (4.47) as:

$$\varphi = \frac{\varepsilon T}{1-qT} \qquad (4.48)$$

The reaching law (4.46) satisfies the reaching condition and guarantees the stability of the DSMC system. Given a discrete single input linear system as:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k \qquad (4.49)$$

the control law is derived based on the reaching law of Equation (4.46) and the switching function of Equation (4.43) as:

$$s_{k+1} - s_k = -qTs_k - \varepsilon T \operatorname{sgn}(s_k) = \mathbf{GA}\mathbf{x}_k + \mathbf{GB}u_k - s_k \Rightarrow$$
$$u_k = -(\mathbf{GB})^{-1}(\mathbf{GA}\mathbf{x}_k - s_k + qTs_k + \varepsilon T \operatorname{sgn}(s_k))$$

From above equation, the control law is obtained as:

$$u_k = -(\mathbf{GB})^{-1}[\mathbf{GA} - (1-qT)\mathbf{G}]\mathbf{x}_k - (\mathbf{GA})^{-1}\varepsilon T \operatorname{sgn}(s_k) \qquad (4.50)$$

Substituting (4.50) into (4.49) gives the closed-loop dynamics of the discrete SMC system as:

$$\mathbf{x}_{k+1} = (1-qT)\mathbf{x}_k - \mathbf{B}(\mathbf{GB})^{-1}\varepsilon T\,\mathrm{sgn}(s_k) \qquad (4.51)$$

Consider the system (4.49) with perturbations:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \Delta\mathbf{A}\mathbf{x}_k + \mathbf{B}u_k + \mathbf{w}_k \qquad (4.52)$$

Applying the reaching law (4.46) to Equation (4.52), the control law is then modified as:

$$u_k = -(\mathbf{GB})^{-1}[\mathbf{GA} - (1-qT)\mathbf{G}]\mathbf{x}_k - (\mathbf{GA})^{-1}\varepsilon T\,\mathrm{sgn}(s_k) - (\mathbf{GB})^{-1}\delta_k \qquad (4.53)$$

where $\delta_k = \mathbf{G}\Delta\mathbf{A}\mathbf{x}_k + \mathbf{G}\mathbf{w}_k$. With these perturbations, the structure of QSM is unchanged and the system's invariance property holds. However the control law (4.53) is non-implementable due to the lack of knowledge of the perturbations. It is assumed that $|\delta_k| < \gamma T$, *for* $k = 0,1,\cdots,\infty$. To make (4.53) implementatable, δ_k is replaced with ς, such that:

$$u_k = -(\mathbf{GB})^{-1}[\mathbf{GA} - (1-qT)\mathbf{G}]\mathbf{x}_k - (\mathbf{GA})^{-1}\varepsilon T\,\mathrm{sgn}(s_k) - (\mathbf{GB})^{-1}\varsigma$$

Combining the above equation with Equations (4.52) and (4.43) gives:

$$s_{k+1} = (1-qT)s_k - \varepsilon T\,\mathrm{sgn}(s_k) + \delta_k - \varsigma \qquad (4.54)$$

In Equation (4.54), the term δ_k is out of control. The term ς helps the term $\varepsilon T\,\mathrm{sgn}(s_k)$ to make s_{k+1} to have a sign that is opposite to that of s_k to ensure the QSM. This can be achieved by setting $\varsigma = \gamma T\,\mathrm{sgn}(s_k)$ where $|\delta_k| < \gamma T$, *for* $k = 0,1,\cdots,\infty$. The control law (4.50) becomes:

$$u_k = -(\mathbf{GB})^{-1}[\mathbf{GA} - (1-qT)\mathbf{G}]\mathbf{x}_k - (\mathbf{GA})^{-1}(\varepsilon+\gamma)T\,\mathrm{sgn}(s_k) \qquad (4.55)$$

With the control law of Equation (4.55), the closed-loop dynamics of the system of Equation (4.52) becomes:

$$\mathbf{x}_{k+1} = (1 - qT)\mathbf{x}_k - \mathbf{B}(\mathbf{GB})^{-1}(\varepsilon + \gamma)T\,\mathrm{sgn}(s_k) \qquad (4.56)$$

DSMC of Multi-Input-Multi-Output LTI System

Chang (2002) addressed a DSMC for Multi-Input-Multi-Output (MIMO) LTI systems which uses the switching function of Equation (4.43) to design a DSMC law that can result in a given rate of convergence. Consider a discrete LTI system model with matched uncertainty, such that:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}(\mathbf{u}_k + \mathbf{d}_k) \qquad (4.57)$$

where $\mathbf{x}_k \in R^n$, $\mathbf{u}_k \in R^m$, and $rank(\mathbf{B}) = m$. The matched uncertainty satisfies:

$$\|\mathbf{d}_k\| \le \sigma_k = d_0 + d_1\|\mathbf{x}_k\| \qquad (4.58)$$

where $\|\bullet\|$ represent the Euclidean norm, d_0 and d_1 are positive constants. The switching surface is defined from Equation (4.43) as:

$$\mathbf{s}_k = \mathbf{G}\mathbf{x}_k = \mathbf{K}(\mathbf{A} - \lambda\mathbf{I}_n)^{-1}\mathbf{x}_k = 0, \quad \mathbf{G} \in R^{m\times n} \qquad (4.59)$$

where $\mathbf{G} = \mathbf{K}(\mathbf{A} - \lambda\mathbf{I}_n)^{-1}$, $\mathbf{GB} = \mathbf{I}_m$, and where $\left[\mathbf{I}_n - \mathbf{B}(\mathbf{GB})^{-1}\mathbf{G}\right]\mathbf{A}$ has the eigenvalues $\{\lambda_1, \lambda_2, \cdots, \lambda_{n-m}, 0, \cdots, 0\}$. For the controller design, the following conditions need to be satisfied (Chang, 2002):

Condition C1: The matrix $\mathbf{A} - \mathbf{BK}$ has no common eigenvalues with matrix $\mathbf{A}$.

Condition C2: The eigenvalues of $\mathbf{A}-\mathbf{BK}$ are chosen such that the closed-loop feedback system is stable and valued as $\{\lambda_1, \lambda_2, \cdots, \lambda_{n-m}, \lambda, \cdots, \lambda\}$ with $\lambda \neq \lambda_i, .\ \ i = 1,2,\cdots,m$, $0 < \lambda < 1$.

Condition C3: The matrix $\mathbf{A}-\mathbf{BK}$ can be diagonalized with m repeated eigenvalues.

Further to conditions **C1** to **C**3, the dynamics of the equivalent motion[7] becomes:

$$\mathbf{x}_{k+1} = \left(\mathbf{I}_n - \mathbf{BK}(\mathbf{A} - \lambda \mathbf{I}_n)^{-1}\right)\mathbf{A}\mathbf{x}_k \qquad (4.60)$$

For the system of Equation (4.60), the **L**yapunov function is defined as $\mathbf{v}(\mathbf{x}_k) = \|\mathbf{s}_k\|$. The global asymptotic stability of system of Equation (4.60) is guaranteed by:

$$\|\mathbf{s}_{k+1}\| < \|\mathbf{s}_k\| \qquad (4.61)$$

A smoothing boundary layer is designed with Expression (4.58) and condition **C2**, such that:

$$\Psi = \left\{ \mathbf{x}_k \middle| \|\mathbf{s}_k\| \le \varphi_k = \frac{\sigma_k + \varepsilon}{\lambda} \right\},\ \varepsilon > 0 \qquad (4.62)$$

The control law is in turn defined as:

$$\mathbf{u}_k = -\mathbf{K}\mathbf{x}_k - \lambda \varphi_k sat(\mathbf{s}_k),\ sat(\mathbf{s}_k) = \begin{cases} \dfrac{\mathbf{s}_k}{\|\mathbf{s}_k\|}, \|\mathbf{s}_k\| > \varphi_k \\ \dfrac{\mathbf{s}_k}{\varphi_k}, \|\mathbf{s}_k\| < \varphi_k \end{cases} \qquad (4.63)$$

---

[7] Form the nominal system of Equation (4.57) and the switching function of Equation (4.59), the equivalent control can be obtained as:
$\mathbf{s}_{k+1} = \mathbf{G}\mathbf{x}_{k+1} = \mathbf{0} \Rightarrow \mathbf{G}\mathbf{A}\mathbf{x}_k + \mathbf{G}\mathbf{B}\mathbf{u}_k = \mathbf{0} \Rightarrow \mathbf{u}_{eq.k} = -(\mathbf{G}\mathbf{B})^{-1}\mathbf{G}\mathbf{A}\mathbf{x}_k$ Substituting the equivalent control into the nominal system of Equation (4.57), with $\mathbf{G} = \mathbf{K}(\mathbf{A} - \lambda \mathbf{I}_n)^{-1}$, gives the equivalent dynamics.

where **K** is the feedback gain matrix; λ is one of the m repeated eigenvalues. The dynamics of the s-trajectory by combining Equations (4.57), (4.59), and (4.63) becomes:

$$\mathbf{s}_{k+1} = \lambda \mathbf{s}_k - \lambda \varphi_k sat(\mathbf{s}_k) + \mathbf{d}_k \qquad (4.64)$$

When the s-trajectory is outside of the boundary layer, i.e., $\|\mathbf{s}_k\| > \varphi_k$, such that:

$$sat(\mathbf{s}_k) = \frac{\mathbf{s}_k}{\|\mathbf{s}_k\|}$$

Substituting the above equation into Equation (4.64), the s-trajectory becomes:

$$\mathbf{s}_{k+1} = \frac{\lambda \mathbf{s}_k}{\|\mathbf{s}_k\|}(\|\mathbf{s}_k\| - \varphi_k) + \mathbf{d}_k \qquad (4.65)$$

Taking the Euclidean norm of (4.65) with Expression (4.58) and the boundary layer defined in Expression (4.62) gives:

$\|\mathbf{s}_{k+1}\| \le \lambda(\|\mathbf{s}_k\| - \varphi_k) + \sigma_k = \lambda\|\mathbf{s}_k\| + (-\lambda\varphi_k + \sigma_k) = \lambda\|\mathbf{s}_k\| - \varepsilon \le \lambda\|\mathbf{s}_k\|$. This condition $\|\mathbf{s}_{k+1}\| < \lambda\|\mathbf{s}_k\|$ is satisfied with $0 < \lambda < 1$ from above. The s-trajectory shrinks to the boundary layer of Expression (4.62) with the shrinking rate governed by the positive constant λ. When the s-trajectory gets inside the boundary layer θ_k, i.e., $\|\mathbf{s}_k\| < \varphi_k$, the s-trajectory from Equations (4.57), (4.59) and Expression (4.63) as:

$$\mathbf{s}_{k+1} = \mathbf{d}_k \le \sigma_k \qquad (4.66)$$

The above analysis indicates that the s-trajectory approaches the boundary layer in a finite number of steps and finally stays inside the sliding region.

Adaptive DSMC of LTI system

Song and Huang (2007) proposed a DSMC algorithm based on the reaching law approach with on-line adaptive estimation of uncertainty. Consider a single input, uncertain discrete-time dynamical system:

$$\mathbf{x}_{k+1} = (\mathbf{A} + \Delta\mathbf{A})\mathbf{x}_k + (\mathbf{B} + \Delta\mathbf{B})u_k + \mathbf{w}_k \qquad (4.67)$$

where $\mathbf{x} \in R^n$ is the state vector; $\mathbf{A} \in R^{n\times n}$ is the system matrix; $\mathbf{B} \in R^n$ is the input vector; $\Delta\mathbf{A}$, $\Delta\mathbf{B}$ are the corresponding uncertainties on in matrix $\mathbf{A}$ and the vector $\mathbf{B}$, respectively; u is scalar input; $\mathbf{w}(\bullet)$ is external disturbances.

The sliding surface is defined as

$$s_k = \mathbf{G}\mathbf{x}_k = 0 \qquad (4.68)$$

where $\mathbf{G} \in R^{1\times n}$ is the switching surface vector. A Lyapunov function is defined as $v_k = |s_k|$. A control law is determined using RLA to ensure the stability such that $v_{k+1} \le v_k$, as:

$$s_{k+1} = \mu s_k - \rho_k \varepsilon T \operatorname{sgn}(s_k) \qquad (4.69)$$

where $\varepsilon > 0$, T is the sampling period, $0 < \mu < 1$ and $\rho_k = \begin{cases} 1, & |s_k| > \dfrac{\varepsilon T}{1+\mu} \\ \dfrac{2\mu|s_k|}{\varepsilon T}, & |s_k| \le \dfrac{\varepsilon T}{1+\mu} \end{cases}$. The reaching law itself can ensure the s-trajectory to reach the sliding surface of Equation (4.68) asymptotically. The control law is derived by combining the system model of Equation (4.67), the sliding surface of Equation (4.68) and the reaching law of Equation (4.69) as:

$$u_k = -(\mathbf{GB})^{-1}\left(\mathbf{GA}\mathbf{x}_k - \mu s_k + \rho_k \varepsilon T \operatorname{sgn}(s_k) + \hat{d}_k\right) \qquad (4.70)$$

where $d_k = \mathbf{G}(\Delta\mathbf{A}\mathbf{x}_k + \Delta\mathbf{B}u_k + \mathbf{w}_k)$ is the uncertain part of the s-trajectory, and $\hat{d}_k$ is the estimated value of d_k determined through a recursive process as:

$$\hat{d}_k = \hat{d}_{k-1} + \lambda(\hat{d}_{k-1} - \hat{d}_{k-2}) + \tilde{d}_{k-1} + \lambda(\tilde{d}_{k-1} - \tilde{d}_{k-2}) \qquad (4.71)$$

where $|\lambda| < 1$, $\tilde{d} = d - \hat{d}$ is the estimation error on s-trajectory between the measured s-value and the reaching law as:

$$\tilde{d}_{k-1} = s_k - (\mu s_{k-1} - \rho_{k-1}\varepsilon T \operatorname{sgn}(s_{k-1})) \qquad (4.72)$$

A modified reaching law can be obtained from the estimation error (4.72) as:

$$s_{k+1} = \mu s_k - \rho_k \varepsilon T \operatorname{sgn}(s_k) + \tilde{d}_k \qquad (4.73)$$

Equation (4.73) can be interpreted as the modified reaching law consisting of the s-dynamics and the estimation error dynamics. The estimation error dynamics is decoupled from the s-dynamics, that is, it is reasonable to represent the reaching law as:

$$s_{k+1} = s_{1.k+1} + s_{2.k+1},\ s_{1.k+1} = \mu s_{1.k} - \rho_k \varepsilon T \operatorname{sgn}(s_{1.k}),\ s_{2.k+1} = \tilde{d}_k \qquad (4.74)$$

With the definition $d_k = \hat{d}_k + \tilde{d}_k$, it can be determined from (4.71) and (4.74) that:

$$s_{2.k+1} = \tilde{d}_k = d_k - (\hat{d}_{k-1} + \tilde{d}_{k-1}) - \lambda\left[(\hat{d}_{k-1} + \tilde{d}_{k-1}) - (\hat{d}_{k-2} + \tilde{d}_{k-2})\right] = d_k - d_{k-1} - \lambda(d_{k-1} - d_{k-2}),\ \text{so:}$$

$$s_{2.k+2} = \tilde{d}_{k+1} = d_{k+1} - d_k - \lambda(d_k - d_{k-1}) \Rightarrow |\tilde{d}_{k+1}| \le |d_{k+1} - d_k| + \lambda|d_k - d_{k-1}| \qquad (4.75)$$

If the condition $|d_k - d_{k-1}| < m$ holds for all k and some positive constant m, then $s_{2.k+1} = |\tilde{d}_k| < 2m$ holds for all k. The amplitude of m is dependent on varying rate of the estimation error dynamics. The estimation error converges to zero asymptotically if the uncertainty is constant or slow varying.

From Equations (4.71) and (4.74), if $s_{1.k} > \frac{\varepsilon T}{1+\mu}$ then:

$$\begin{aligned} s_{1.k+1} + s_{1.k} &= (1+\mu)s_{1.k} - \varepsilon T > (1+\mu)\frac{\varepsilon T}{1+\mu} - \varepsilon T = 0 \\ s_{1.k+1} - s_{1.k} &= (\mu-1)s_{1.k} - \varepsilon T < (\mu-1)\frac{\varepsilon T}{1+\mu} - \varepsilon T < 0 \end{aligned} \tag{4.76}$$

If $s_{1.k} < -\frac{\varepsilon T}{1+\mu}$ then from Equations (4.71) and (4.74):

$$\begin{aligned} s_{1.k+1} + s_{1.k} &= (1+\mu)s_{1.k} + \varepsilon T < (1+\mu)\left(-\frac{\varepsilon T}{1+\mu}\right) + \varepsilon T = 0 \\ s_{1.k+1} - s_{1.k} &= (\mu-1)s_{1.k} + \varepsilon T > (\mu-1)\left(-\frac{\varepsilon T}{1+\mu}\right) + \varepsilon T > 0 \end{aligned} \tag{4.77}$$

If $-\frac{\varepsilon T}{1+\mu} \le s_{1.k} < 0$ then from Equations (4.71) and (4.74):

$$\begin{aligned} s_{1.k+1} + s_{1.k} &= (1+\mu)s_{1.k} - 2\mu|s_{1.k}|\mathrm{sgn}(s_{1.k}) = (1-\mu)s_{1.k} < (1-\mu)0 = 0 \\ s_{1.k+1} - s_{1.k} &= (\mu-1)s_{1.k} - 2\mu|s_{1.k}|\mathrm{sgn}(s_{1.k}) = -(1+\mu)s_{1.k} > -(1+\mu)0 = 0 \end{aligned} \tag{4.78}$$

If $0 \le s_{1.k} \le \frac{\varepsilon T}{1+\mu}$ then from Equations (4.71) and (4.74):

$$\begin{aligned} s_{1.k+1} + s_{1.k} &= (1+\mu)s_{1.k} - 2\mu|s_{1.k}|\mathrm{sgn}(s_{1.k}) = (1-\mu)s_{1.k} \ge (1-\mu)0 = 0 \\ s_{1.k+1} - s_{1.k} &= (\mu-1)s_{1.k} - 2\mu|s_{1.k}|\mathrm{sgn}(s_{1.k}) = -(1+\mu)s_{1.k} \le -(1+\mu)0 = 0 \end{aligned} \tag{4.79}$$

All the expressions from (4.76) to (4.79) imply $|s_{1.k+1}|^2 \le |s_{1.k}|^2$ which means that as $k \to \infty$, s_1 converges to zero. From the analysis of both the s-dynamics and the estimation error dynamics, it can be concluded that the system of Equation (4.67) with control of Equation (4.70) and the recursive uncertainty estimation of Equation (4.71) is asymptotically stable.

DSMC of MIMO LTI Systems Using Multi-Rate Output Feedback

A state observer can be used when all of the states are not available. Theoretically the state estimation error from an observer asymptotically approaches zero in finite time. Multirate output feedback (MROF) can be used in place of a state observer to compute the state. In MROF, the measurement of the output and the computation of the control input take place at different rates. The output is measured at a faster rate than the control input. The computed state error from MROF is zero at the next control input sample time (Waterman & Nonami, 2010). Janardhanan and Bandyopadhyay proposed a control algorithm using the MOF strategy for achieving quasi-sliding mode in discrete-time LTI systems with bounded unmatched uncertainties.

Consider the discrete-time system representation for a sampling rate of τ second:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}_\tau \mathbf{x}_k + \mathbf{B}_\tau \mathbf{u}_k + \mathbf{D}_\tau \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k \end{aligned} \qquad (4.80)$$

where $\mathbf{x} \in R^n$ is the state vector, $\mathbf{u} \in R^m$ is the input vector, $\mathbf{y} \in R^p$ is the output vector, and $\mathbf{w} \in R^q$ is bounded disturbance vector, with matrices $\mathbf{A}_\tau \in R^{n\times n}, \mathbf{B}_\tau \in R^{n\times m}$, $\mathbf{D}_\tau \in R^{n\times q}$ and $\mathbf{C} \in R^{p\times n}$. It is assumed that the pair $(\mathbf{A}_\tau, \mathbf{B}_\tau)$ is controllable and the pair $(\mathbf{A}_\tau, \mathbf{C})$ is observable. The matching condition $rank(\mathbf{B}_\tau) = rank(\mathbf{B}_\tau \mid \mathbf{D}_\tau)$ may not be satisfied. The switching function is defined as $\mathbf{s}_k = [s_{1k}\ s_{2k} \cdots s_{mk}]^T = \mathbf{G}\mathbf{x}_k = \mathbf{0}, \mathbf{s} \in R^m, \mathbf{G} \in R^{m\times n}$ and $\mathbf{G}\mathbf{B}_\tau$ is assumed non-singular and the disturbance is bounded:

$$\begin{aligned}
&\mathbf{d}_l \le \mathbf{d}_k = \mathbf{G}\mathbf{D}_d\mathbf{w}_k \le \mathbf{d}_u \\
&\mathbf{d}_0 = \frac{\mathbf{d}_l + \mathbf{d}_u}{2} \\
&\mathbf{d}_1 = \frac{\mathbf{d}_u - \mathbf{d}_l}{2}
\end{aligned} \tag{4.81}$$

where $\mathbf{d}_l = \left[d_{l_1}\ d_{l_2}\ \cdots\ d_{l_m}\right]^T$, $\mathbf{d}_u = \left[d_{u_1}\ d_{u_2}\ \cdots\ d_{u_m}\right]^T$. The objective of the control is to achieve quasi-sliding mode by steering the system states to the vicinity of the sliding mode $\mathbf{s}_k = 0$ such that once the states enter the neighbourhood of the switching surface, stay thereafter.

For state observation a fast output sampling is performed (Bandyopadhyay & Thakar 2008) and the system of Equation (4.80) is resampled at $\Delta = \tau / N$, where the integer $N \ge o$, and o is the observability index[8] of the system $(\mathbf{A}, \mathbf{B}, \mathbf{C})$, such that:

$$\begin{aligned}
&\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{D}\mathbf{w}_k \\
&\mathbf{y}_k = \mathbf{C}\mathbf{x}_k
\end{aligned} \tag{4.82}$$

The output measurements are taken at time instants $t = l\Delta,\ l = 1,2,\cdots,N-1$. The control input $\mathbf{u}(t),\ k\tau \le t \le (k+1)\tau$ is constructed as a linear combination of the last N output measurements, that is:

$$\mathbf{y}_k = \begin{bmatrix} \mathbf{y}(k\tau - \tau) \\ \mathbf{y}(k\tau - \tau + \Delta) \\ \vdots \\ \mathbf{y}(k\tau - \Delta) \end{bmatrix} \tag{4.83}$$

---

[8] The observability index o of a linear time-invariant discrete system of Equation (4.82) is the smallest natural number for which is satisfied that $rank(\mathbf{O}_o) = rank(\mathbf{O}_{o+1})$ where:

$\mathbf{O}_o = \left[\mathbf{C}\quad \mathbf{CA}\quad \mathbf{CA}^2\quad \cdots\quad \mathbf{CA}^{o-1}\right]$

The multirate system with the input sampling time of τ and the output sampling time of Δ is denoted as:

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{A}_\tau \mathbf{x}_k + \mathbf{B}_\tau \mathbf{u}_k + \mathbf{D}_\tau \mathbf{w}_k \\ \mathbf{y}_{k+1} &= \mathbf{C}_0 \mathbf{x}_k + \mathbf{D}_0 \mathbf{u}_k + \mathbf{E}_0 \mathbf{w}_k\end{aligned} \qquad (4.84)$$

$$\text{where } \mathbf{C}_0 = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{N-1} \end{bmatrix} \in R^{(pN)\times n}, \mathbf{D}_0 = \begin{bmatrix} \mathbf{0} \\ \mathbf{CB} \\ \mathbf{CAB}+\mathbf{CB} \\ \vdots \\ \mathbf{C}\sum_{i=0}^{N-2}\mathbf{A}^i\mathbf{B} \end{bmatrix} \in R^{(pN)\times m}, \mathbf{E}_0 = \begin{bmatrix} \mathbf{0} \\ \mathbf{CD} \\ \mathbf{CAD}+\mathbf{CD} \\ \vdots \\ \mathbf{C}\sum_{i=0}^{N-2}\mathbf{A}^i\mathbf{D} \end{bmatrix} \in R^{(pN)\times q}.$$

The system state $\mathbf{x}_k$ can be approximately constructed from the past output $\mathbf{y}_k$ by using the second equation of Equation (4.84):

$$\mathbf{x}_k = \left(\mathbf{C}_0^T\mathbf{C}_0\right)^{-1}\mathbf{C}_0^T\mathbf{y}_{k+1} - \left(\mathbf{C}_0^T\mathbf{C}_0\right)^{-1}\mathbf{C}_0^T\mathbf{D}_0\mathbf{u}_k - \left(\mathbf{C}_0^T\mathbf{C}_0\right)^{-1}\mathbf{C}_0^T\mathbf{E}_0\mathbf{w}_k \qquad (4.85)$$

Substituting (4.85) into the first equation of Equation (4.84) gives the relationship between $\mathbf{x}_k$ and $\mathbf{y}_k$ as:

$$\mathbf{x}_k = L_y\mathbf{y}_k + L_u\mathbf{u}_{k-1} + L_w\mathbf{w}_{k-1} \qquad (4.86)$$

where $L_y = \mathbf{A}_\tau\left(\mathbf{C}_0^T\mathbf{C}_0\right)^{-1}\mathbf{C}_0^T$, $L_u = \mathbf{B}_\tau - \mathbf{A}_\tau\left(\mathbf{C}_0^T\mathbf{C}_0\right)^{-1}\mathbf{C}_0^T\mathbf{D}_0$, $L_d = \mathbf{D}_\tau - \mathbf{A}_\tau\left(\mathbf{C}_0^T\mathbf{C}_0\right)^{-1}\mathbf{C}_0^T\mathbf{E}_0$. In Equation (4.86) $\mathbf{w}_{k-1}$ is not known, but its bounds are assumed known. The bounds on $L_w\mathbf{w}_{k-1}$ can be determined as $\mathbf{l}_l \le L_w\mathbf{w}_{k-1} \le \mathbf{l}_u$, $\mathbf{l}_0 = \frac{\mathbf{l}_u + \mathbf{l}_l}{2}$, $\mathbf{l}_1 = \frac{\mathbf{l}_u - \mathbf{l}_l}{2}$. The estimation of the state $\mathbf{x}_k$ is calculated as:

$$\hat{\mathbf{x}}_k = L_y\mathbf{y}_k + L_u\mathbf{u}_{k-1} + \mathbf{l}_0 \qquad (4.87)$$

The relationship between $\mathbf{x}_k$ and $\hat{\mathbf{x}}_k$ is derived with (4.86) and (4.87) as:

$$\mathbf{x}_k = \hat{\mathbf{x}}_k - \mathbf{l}_0 + L_w \mathbf{w}_{k-1} \qquad (4.88)$$

The estimated sliding manifold can be defined as:

$$\hat{\mathbf{s}}_k = \mathbf{G}\hat{\mathbf{x}}_k = \mathbf{G}L_y \mathbf{y}_k + \mathbf{G}L_u \mathbf{u}_{k-1} + \mathbf{G}\mathbf{l}_0 \qquad (4.89)$$

Following the state estimation of Equation (4.88), a quasi-sliding mode control law can be proposed as stated in the following theorem (Janardhanan & Bandyopadhyay, 2006).

Theorem4.4: A multirate output feedback based control law is defined as:

$$\mathbf{u}_k = -(\mathbf{G}\mathbf{B}_\tau)^{-1}\left[\mathbf{G}(\mathbf{A}_\tau - \mathbf{I})\hat{\mathbf{x}}_k + \mathbf{K}\mathbf{f}_{sat}(\hat{\mathbf{s}}_k, \phi) + \mathbf{d}_0\right] \qquad (4.90)$$

$$\begin{cases} \mathbf{f}_{sat}(\hat{\mathbf{s}}_k, \phi) = \left[sat\left(\dfrac{\hat{s}_{1k}}{\phi_1}\right) \quad sat\left(\dfrac{\hat{s}_{2k}}{\phi_2}\right) \quad \cdots \quad \dfrac{\hat{s}_{mk}}{\phi_1} \right]^T \\ sat\left(\dfrac{\hat{s}_{ik}}{\phi_i}\right) = \begin{cases} \mathrm{sgn}(\hat{s}_{ik}), & |\hat{s}_{ik}| > \phi_i \\ \dfrac{\hat{s}_{ik}}{\phi_i}, & |\hat{s}_{ik}| \le \phi_i \end{cases} \quad i = 1,2,\cdots,m \end{cases} \qquad (4.91)$$

where ϕ_i is the component of the vector ϕ which sets the widths of the quasi-sliding mode smoothing boundary layers, and $\mathbf{K} = diag(k_1, k_2, \cdots, k_m)$ with:

$$k_i > (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i \qquad (4.92)$$

and:

$$\begin{cases} \mathbf{a}_1 = \dfrac{\mathbf{a}_u - \mathbf{a}_l}{2} \\ \mathbf{a}_0 = \dfrac{\mathbf{a}_u + \mathbf{a}_l}{2} = \mathbf{G}\mathbf{A}_\tau \mathbf{l}_0 \\ \mathbf{a}_l \le \mathbf{a}_k = \mathbf{G}\mathbf{A}_\tau L_w \mathbf{w}_k \le \mathbf{a}_u \end{cases} \qquad (4.93)$$

$$\begin{cases} \mathbf{b}_1 = \dfrac{\mathbf{b}_u - \mathbf{b}_l}{2} \\ \mathbf{b}_0 = \dfrac{\mathbf{b}_u + \mathbf{b}_l}{2} = \mathbf{G}\mathbf{l}_0 \\ \mathbf{b}_l \le \mathbf{b}_k = \mathbf{G}L_w\mathbf{w}_k \le \mathbf{b}_u \end{cases} \qquad (4.94)$$

$$0 < 2\phi_i - k_i - (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i \qquad (4.95)$$

The controller of Equation (4.90) achieves quasi-sliding mode for the discrete-time system of Equation (4.84); that is, for any initial state $\mathbf{x}_0$, there exists an index k^* such that $\mathbf{x}_k \in \overline{\Omega} \subset \Omega$ for all $k > k^*$ where Ω is the quasi-sliding mode defined as $\Omega = \{\mathbf{x} \| \mathbf{s}_k | < \phi\}$.

Proof: Define a Lyapunov function $v_{ik} = \hat{s}_{ik}^2$, $i = 1,2,\cdots,m$. The control law (4.90) shall decrease the Lyapunov function monotonically, that is, the reaching condition is:

$$v_{ik+1} = \hat{s}_{ik+1}^2 < v_{ik} = \hat{s}_{ik}^2 \Rightarrow (\Delta\hat{s}_{ik} + 2\hat{s}_{ik})\Delta\hat{s}_{ik} < 0, \forall \hat{s}_{ik} \neq 0 \qquad (4.96)$$

where $\Delta\hat{s}_{ik}$ is the i^{th} component of the vector $\Delta\hat{\mathbf{s}}_k$. $\Delta\hat{\mathbf{s}}_k$ can be calculated from (4.88) as:

$$\Delta\hat{\mathbf{s}}_k = \hat{\mathbf{s}}_{k+1} - \hat{\mathbf{s}}_k = \mathbf{G}(\hat{\mathbf{x}}_{k+1} - \hat{\mathbf{x}}_k) \qquad (4.97)$$

Substituting Equations (4.84), (4.88) and (4.90) into Equation (4.97):

$$\begin{aligned} \Delta\hat{\mathbf{s}}_k &= (\mathbf{G}\mathbf{A}_\tau L_w\mathbf{w}_{k-1} - \mathbf{G}\mathbf{A}_\tau\mathbf{l}_0) + (\mathbf{G}\mathbf{l}_0 - \mathbf{G}L_w\mathbf{w}_k) + (\mathbf{d}_k - \mathbf{d}_0) - \mathbf{K}\mathbf{f}_{sat}(\hat{\mathbf{s}}_k, \phi) \\ &= (\mathbf{a}_{k-1} - \mathbf{a}_0) + (\mathbf{b}_0 - \mathbf{b}_k) + (\mathbf{d}_k - \mathbf{d}_0) - \mathbf{K}\mathbf{f}_{sat}(\hat{\mathbf{s}}_k, \phi) \end{aligned} \qquad (4.98)$$

Rearranging Equation (4.98):

$$\Delta\hat{s}_{ik} = (\mathbf{a}_{k-1} - \mathbf{a}_0)_i + (\mathbf{b}_0 - \mathbf{b}_k)_i + (\mathbf{d}_k - \mathbf{d}_0)_i - k_i sat\left(\frac{\hat{s}_{ik}}{\phi_i}\right) \qquad (4.99)$$

Substituting from Equation (4.91):

$$\Delta\hat{s}_{ik} = (\mathbf{a}_{k-1} - \mathbf{a}_0)_i + (\mathbf{b}_0 - \mathbf{b}_k)_i + (\mathbf{d}_k - \mathbf{d}_0)_i - \begin{cases} k_i; & \hat{s}_{ik} > \phi_i \\ k_i \dfrac{\hat{s}_{ik}}{\phi_i}; & |\hat{s}_{ik}| \le \phi_i \\ -k_i; & \hat{s}_{ik} < -\phi_i \end{cases} \quad (4.100)$$

From Equations (4.81), (4.93) and (4.94), then:

$$-(\mathbf{a}_1+\mathbf{b}_1+\mathbf{d}_1)_i < (\mathbf{a}_{k-1} - \mathbf{a}_0)_i + (\mathbf{b}_0 - \mathbf{b}_k)_i + (\mathbf{d}_k - \mathbf{d}_0) < (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i \quad (4.101)$$

To determine if the stability and the reaching condition (4.96) are satisfied, the three cases associated with the saturation function, that is, $\hat{s}_{ik} > \phi_i$, $\hat{s}_{ik} < -\phi_i$ and $|\hat{s}_{ik}| \le \phi_i$, have to be considered as follows:

(1) For $\hat{s}_{ik} > \phi_i$, from Equation (4.100):

$$\Delta\hat{s}_{ik} = (\mathbf{a}_{k-1} - \mathbf{a}_0)_i + (\mathbf{b}_0 - \mathbf{b}_k)_i + (\mathbf{d}_k - \mathbf{d}_0)_i - k_i$$

and from Expression (4.101):

$$-(\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i - k_i \le \Delta\hat{s}_{ik} \le (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i - k_i < 0$$

From the left half of the above expression:

$$2\hat{s}_{ik} + \Delta\hat{s}_{ik} > 2\phi_i - (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i - k_i > 2\phi_i - 2\phi_i = 0,$$

these imply that:

$$(\Delta\hat{s}_{ik} + 2\hat{s}_{ik})\Delta\hat{s}_{ik} < 0 \quad when\ \hat{s}_{ik} > \phi_i \quad (4.102)$$

The stability condition is satisfied.

(2) For $\hat{s}_{1k} < -\phi_i$, from (4.100):

$$\Delta\hat{s}_{ik} = (\mathbf{a}_{k-1} - \mathbf{a}_0)_i + (\mathbf{b}_0 - \mathbf{b}_k)_i + (\mathbf{d}_k - \mathbf{d}_0)_i + k_i$$

and from (4.101):

$$0 < k_i - (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i < \Delta\hat{s}_{ik} < k_i + (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i$$

From the right half of the above expression and Expression (4.95):

$$2\hat{s}_{ik} + \Delta\hat{s}_{ik} < -2\phi_i + (\mathbf{a}_1 + \mathbf{b}_1 + \mathbf{d}_1)_i + k_i < 0.$$

These imply:

$$(\Delta\hat{s}_{ik} + 2\hat{s}_{ik})\Delta\hat{s}_{ik} < 0 \quad when\ \hat{s}_{ik} < -\phi_i \qquad (4.103)$$

The stability condition is satisfied.

(3) For $|\hat{s}_{ik}| \le \phi_i$, from (4.100):

$$\hat{s}_{ik+1} = \left(1 - \frac{k_i}{\phi_i}\right)\hat{s}_{ik} + (\mathbf{a}_{k-1} - \mathbf{a}_0)_i + (\mathbf{b}_0 - \mathbf{b}_k)_i + (\mathbf{d}_k - \mathbf{d}_0)_i$$

From Equations (4.88), (4.89) and (4.94), the above equation can be rewritten, with $\mathbf{s}_k = \mathbf{G}\mathbf{x}_k$, as:

$$\begin{aligned} &s_{ik+1} = \left(1 - \frac{k_i}{\phi_i}\right)\hat{s}_{ik} + (\mathbf{a}_{k-1} - \mathbf{a}_0)_i + (\mathbf{d}_k - \mathbf{d}_0)_i \Rightarrow \\ &|s_{ik+1}| \le |\phi_i - k_i|\left|\frac{\hat{s}_{ik}}{\phi_i}\right| + |(\mathbf{a}_{k-1} - \mathbf{a}_0)_i| + |(\mathbf{d}_k - \mathbf{d}_0)_i| \le |\phi_i - k_i| + (\mathbf{a}_1 + \mathbf{d}_1)_i \end{aligned} \qquad (4.104)$$

If $\phi_i > k_i$, with (4.92) the expression (4.104) becomes:

$$|s_{ik+1}| \le \phi_i - k_i + (\mathbf{a}_1 + \mathbf{d}_1)_i < (\phi - \mathbf{b}_1)_i < \phi_i \qquad (4.105)$$

If $\phi_i < k_i$, with (4.95) the expression (4.104) becomes:

$$|s_{ik+1}| \le k_i - \phi_i + (\mathbf{a}_1 + \mathbf{d}_1)_i < (2\phi - \mathbf{b} - \phi)_i = (\phi - \mathbf{b}_1)_i < \phi_i \qquad (4.106)$$

If $\phi_i = k_i$, with (4.92) the expression (4.104) becomes:

$$|s_{ik+1}| \le (\mathbf{a}_1 + \mathbf{d}_1)_i < (\phi - \mathbf{b}_1)_i < \phi_i \qquad (4.107)$$

From Expressions (4.105), (4.106) and (4.107), it can be concluded that the condition $|s_{ik+1}| < \phi_i$ is satisfied. This proves that a quasi-sliding mode is obtained in the system using multirate output feedback.

4.3.3 DSMC of Non-Linear Systems

Most plants in industry have nonlinearities. Sliding mode control of nonlinear systems has therefore drawn great attention within control community.

DSMC of Single Input Nonlinear System

Misawa (1997) presented a DSMC method with a boundary layer generated by a saturation function without requiring either matched uncertainties or smooth functions. The boundary layer is attractive. The thickness of the layer is bounded with mild conditions so that global stability is guaranteed.

Consider a single-input nonlinear system:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u(t) + \Delta\mathbf{f} \qquad (4.108)$$

where $\mathbf{f}(\mathbf{x}) \in R^n$ is a known function, $\mathbf{g}(\mathbf{x}) \in R^n$, $\Delta\mathbf{f} \in R^n$, $u(t)$ is the scalar input, $\mathbf{x}(t) \in R^n$ is the state vector. The uncertainty $\Delta\mathbf{f}$ is assumed to be bounded by constants or time-varying bounds of functions of the states, the known inputs, or by explicitly known functions of time. The input gain vector is defined as $\mathbf{g}(\mathbf{x}) = \hat{\mathbf{g}}(\mathbf{x})\Delta g$ with $\hat{\mathbf{g}}(\mathbf{x}) \in R^n$ being a known function. The scalar uncertainty Δg is assumed to be bounded such that:

$$1/\beta < \Delta g < \beta \ \ for\ \beta \geq 1 \qquad (4.109)$$

The objective of the tracking control is to find a control law such that the system trajectory $\mathbf{x}(t)$ would follow the desired trajectory $\mathbf{x}_d(t)$. The tracking error is defined as:

$$\tilde{\mathbf{x}}(t) = \mathbf{x}_d(t) - \mathbf{x}(t) \qquad (4.110)$$

The sliding manifold is defined as:

$$s(\tilde{\mathbf{x}}(t)) = 0 \qquad (4.111)$$

It should be chosen such that $s(\tilde{\mathbf{x}}(t))$ is a function of all of the elements of the error vector $\tilde{\mathbf{x}}(t)$, and the equivalent dynamics, $\dot{\mathbf{x}}(t) = \left[\mathbf{I} - \mathbf{g}(\mathbf{x})\left[\frac{\partial s}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\right]^{-1}\frac{\partial s}{\partial \mathbf{x}}\right]\mathbf{f}(\mathbf{x})$[9] is stable.

The smoothing boundary layer is defined as:

$$\psi = \left\{\tilde{\mathbf{x}} \middle\| s(\tilde{\mathbf{x}})\right| \le \varphi\} \qquad (4.112)$$

The difference in reference trajectory between consecutive samplings is defined as:

$$\Delta\mathbf{x}_{dk} = \mathbf{x}_{dk+1} - \mathbf{x}_{dk} \qquad (4.113)$$

It is assumed that the reference trajectory $\mathbf{x}_{dk}$ is always finite and the Euclidean norm of the rate of change of the reference trajectory between consecutive sampling $|\Delta\mathbf{x}_d|$ is bounded, and the following norm is bounded also:

$$\left|\mathbf{p}_k \frac{\Delta\mathbf{x}_{dk}}{\Delta t}\right| = \left|\mathbf{p}_k \frac{\mathbf{x}_{dk+1} - \mathbf{x}_{dk}}{\Delta t}\right| \le \gamma_{xd} \qquad (4.114)$$

---

[9] The equivalent dynamics is obtained by:

$$s(\tilde{\mathbf{x}}(t)) = 0 \Rightarrow \dot{s}(\tilde{\mathbf{x}}(t)) = 0 \Rightarrow \frac{\partial s}{\partial \mathbf{x}}\dot{\mathbf{x}} = 0 \Rightarrow \frac{\partial s}{\partial \mathbf{x}}[\mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})\mathbf{u}(t)] = 0 \Rightarrow u_{eq}(t) = -\left[\frac{\partial s}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\right]^{-1}\frac{\partial s}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x})$$

$$\Rightarrow \dot{\mathbf{x}}(t) = \left[\mathbf{I} - \mathbf{g}(\mathbf{x})\left[\frac{\partial s}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x})\right]^{-1}\frac{\partial s}{\partial \mathbf{x}}\right]\mathbf{f}(\mathbf{x})$$

$\mathbf{p}_k$ is the partial derivative of the sliding manifold and defined as:

$$\mathbf{p}_k = \frac{\partial s}{\partial \tilde{\mathbf{x}}}\Big|_{\tilde{\mathbf{x}}_k} = \left[\frac{\partial s}{\partial \tilde{x}_1}\frac{\partial s}{\partial \tilde{x}_2}\cdots\frac{\partial s}{\partial \tilde{x}_n}\right]\Big|_{\tilde{\mathbf{x}}_k} \qquad (4.115)$$

The discrete-time sliding surface is obtained by approximating the continuous time counterpart of Equation (4.111) in its tangent plane using Taylor series expansion:

$$s(\tilde{\mathbf{x}}_{k+1}) = s(\tilde{\mathbf{x}}_k) + \mathbf{p}_k(\tilde{\mathbf{x}}_{k+1} - \tilde{\mathbf{x}}_k) + \sigma_k \qquad (4.116)$$

where σ_k is the approximation error and is assumed to be bounded as:

$$|\sigma| \le \gamma_\sigma \qquad (4.117)$$

The discrete-time state vector model is obtained by using the Euler forward difference as:

$$\dot{x}_{ik} = \frac{x_{ik+1} - x_{ik}}{\Delta t} + \delta_{ik} \quad for\ i = 1,2,\cdots,n \qquad (4.118)$$

where Δt is the sampling period and δ_{ik} represent the numerical approximation errors which are assumed to be of the order of Δt, that is:

$$|\delta_{ik}| \le \xi \Delta t \quad for\ 0 < \xi < \infty \qquad (4.119)$$

The discrete-time model can be obtained by discretizing the continuous-time counterpart of Equation (4.108) by using the Euler forward difference of Equation (4.118):

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta t\,\dot{\mathbf{x}} = \mathbf{x}_k + \Delta t\mathbf{f}(\mathbf{x}_k) + \Delta t\mathbf{g}(\mathbf{x}_k)u_k + \Delta t\alpha_k \qquad (4.120)$$

where α_k is the lumped uncertainty with $\alpha_k^T = [\alpha_{1k}\ \alpha_{2k}\cdots\alpha_{nk}]$, and is a function of time, exogenous inputs, states and is defined as:

$$\alpha_{ik} = \Delta f_{ik} - \delta_{ik} \quad for\ i = 1,2,\cdots,n \qquad (4.121)$$

It is assumed that α_k is bounded as:

$$\left|\mathbf{p}_k \alpha_k\right| \le \gamma_{Dh} \qquad (4.122)$$

It is further assumed that the system is controllable and the measurement of all state variables is available or otherwise obtained using state observers.

The discrete-time control law is derived based on the discrete-time Lyapunov stability theory. The Lyapunov function is chosen to be $\mathbf{v}_k = s_k^2$. The reaching condition of sliding mode is $v_{k+1} < v_k$, that is, $s_{k+1}^2 < s_k^2$, and can be further written with $\Delta s_k = s_{k+1} - s_k$ as (Misawa, 1997):

$$\left(\Delta s_k + 2s_k\right)\Delta s_k < \mathbf{0} \Rightarrow \Delta s_k^2 < -2s_k \Delta s_k \Rightarrow \begin{cases} -2s_k < \Delta s_k < 0 \ for \ s_k > 0 \\ 0 < \Delta s_k < -2s_k \ for \ s_k < 0 \end{cases} \qquad (4.123)$$

Now Δs_k is calculated by combining Equations (4.110), (4.116) and (4.120):

$$\Delta s_k = \mathbf{p}_k \Delta \mathbf{x}_{dk} - \Delta t \mathbf{p}_k \left(\mathbf{f} + \mathbf{g} u_k + \alpha_k\right) + \sigma_k = \mathbf{p}_k \Delta \mathbf{x}_{dk} - \Delta t \mathbf{p}_k \mathbf{f} - \Delta t \mathbf{p}_k \mathbf{g} u_k + \upsilon_k \qquad (4.124)$$

where $\upsilon_k = \sigma_k - \Delta t \mathbf{p}_k \alpha_k$, and argument $\mathbf{x}_k$ has been dropped for clarity. Substituting from Expression (4.123) and Equation (4.124) with $\mathbf{g}(\mathbf{x}) = \hat{\mathbf{g}}(\mathbf{x})\Delta g$ and assuming $\mathbf{p}_k \mathbf{g}(\mathbf{x}) > 0$, then rearranging Equation (4.123), for $s_k > 0$:

$$\frac{\mathbf{p}_k \Delta \mathbf{x}_{dk}}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} - \frac{\mathbf{p}_k \mathbf{f}}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} < u_k < \frac{\mathbf{p}_k \Delta \mathbf{x}_{dk}}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} - \frac{\mathbf{p}_k \mathbf{f}}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{2s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} \qquad (4.125)$$

For $s_k < 0$:

$$\frac{\mathbf{p}_k \Delta \mathbf{x}_{dk}}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} - \frac{\mathbf{p}_k \mathbf{f}}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{2s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} < u_k < \frac{\mathbf{p}_k \Delta \mathbf{x}_{dk}}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} - \frac{\mathbf{p}_k \mathbf{f}}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} \qquad (4.126)$$

The control law is defined as the combination of a continuous part and a switching part:

$$u_k = \hat{u}_k + \tilde{u}_k \tag{4.127}$$

where $\hat{u}_k$ is model based and is further defined as:

$$\hat{u}_k = \frac{\mathbf{p}_k \Delta \mathbf{x}_{dk}}{\Delta t \mathbf{p}_k \hat{\mathbf{g}}} - \frac{\mathbf{p}_k \mathbf{f}}{\mathbf{p}_k \hat{\mathbf{g}}} \tag{4.128}$$

Substituting Equations (4.127) and (4.128) into Expressions (4.125) and (4.126), respectively, and with $\upsilon_k = \sigma_k - \Delta t \mathbf{p}_k \mathbf{h}_k$ gives, for $s_k > 0$:

$$\begin{aligned} &\left(\frac{1}{\Delta g} - 1\right)\hat{u}_k + \frac{1}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g}\left(\frac{\sigma_k}{\Delta t} - \mathbf{p}_k \mathbf{h}_k\right) < \tilde{u}_k < \\ &\left(\frac{1}{\Delta g} - 1\right)\hat{u}_k + \frac{1}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g}\left(\frac{\sigma_k}{\Delta t} - \mathbf{p}_k \mathbf{h}_k\right) + \frac{2 s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} \end{aligned} \tag{4.129}$$

for $s_k < 0$:

$$\begin{aligned} &\left(\frac{1}{\Delta g} - 1\right)\hat{u}_k + \frac{1}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g}\left(\frac{\sigma_k}{\Delta t} - \mathbf{p}_k \mathbf{h}_k\right) + \frac{2 s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} < \tilde{u}_k < \\ &\left(\frac{1}{\Delta g} - 1\right)\hat{u}_k + \frac{1}{\mathbf{p}_k \hat{\mathbf{g}} \Delta g}\left(\frac{\sigma_k}{\Delta t} - \mathbf{p}_k \mathbf{h}_k\right) \end{aligned} \tag{4.130}$$

The switching control law $\tilde{u}_k$ has to be chosen such that asymptotic stability is guaranteed in the presence of uncertainties, that is, Expressions (4.129) and (4.130) both have to be satisfied to ensure the stability of the uncertain system of Equation (4.120). Using the left half of Expression (4.129) and the right half of Expression (4.130), the switching control law $\tilde{u}_k$ can be determined as (Misawa, 1997):

$$\tilde{u}_k = q + \frac{\beta\kappa}{\mathbf{p}_k\hat{\mathbf{g}}} sat\left(\frac{s_k}{\varphi}\right),\ \kappa \geq \eta = \frac{\gamma_\sigma}{\Delta t} + \gamma_{Dh} \geq \left|\frac{\sigma_k}{\Delta t} - \mathbf{p}_k\mathbf{h}_k\right| \qquad (4.131)$$

The new parameter q is defined as:

$$q = \begin{cases} q^+ \geq \left(\frac{1}{\Delta g} - 1\right)\hat{u}_k, & s_k > 0 \\ q^- \leq \left(\frac{1}{\Delta g} - 1\right)\hat{u}_k, & s_k < 0 \end{cases} \qquad (4.132)$$

The expression (4.132) can be rewritten as:

$$q^+ = \sup\left[\left(\frac{1}{\Delta g} - 1\right)\hat{u}_k\right] = \begin{cases} (\beta - 1)\hat{u}_k, & \hat{u}_k > 0 \\ \left(\frac{1}{\beta} - 1\right)\hat{u}_k, & \hat{u}_k < 0 \end{cases} \qquad (4.133)$$

$$q^- = \inf\left[\left(\frac{1}{\Delta g} - 1\right)\hat{u}_k\right] = \begin{cases} \left(\frac{1}{\beta} - 1\right)\hat{u}_k, & \hat{u}_k > 0 \\ (\beta - 1)\hat{u}_k, & \hat{u}_k < 0 \end{cases} \qquad (4.134)$$

where sup and inf are the supremum and the infimum, respectively. Equations (4.133) and (4.134) can be satisfied by selecting q as:

$$q = \left(\frac{\beta - 1}{2}\right)\left[\hat{u}_k + |\hat{u}_k|\mathrm{sgn}(s_k)\right] + \left(\frac{1 - \beta}{2\beta}\right)\left[\hat{u}_k - |\hat{u}_k|\mathrm{sgn}(s_k)\right] \qquad (4.135)$$

Combining (4.127), (4.128), (4.131) gives the control law as:

$$u_k = \left(\frac{\beta^2 + 1}{2\beta}\right)\hat{u}_k + \left(\frac{\beta^2 - 1}{2\beta}\right)|\hat{u}_k|\mathrm{sgn}(s_k) + \frac{\beta\kappa}{\mathbf{p}_k\hat{\mathbf{g}}} sat\left(\frac{s}{\varphi}\right), \qquad (4.136)$$

where $\hat{u}_k = \dfrac{\mathbf{p}_k\Delta\mathbf{x}_{dk}}{\Delta t\mathbf{p}_k\hat{\mathbf{g}}} - \dfrac{\mathbf{p}_k\mathbf{f}}{\mathbf{p}_k\hat{\mathbf{g}}}$.

Note that the control law of Equation (4.136) is determined by using Expressions (4.129) and (4.130) with the width φ of the boundary layer in (4.112) undefined.

Expressions (4.129) and (4.130) can also be used to determine the width φ of the boundary layer as follows. For $s_k > \varphi$:

$$u_k < \frac{1}{\Delta g}\hat{u}_k + \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{2s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} \Rightarrow u_k - \frac{1}{\Delta g}\hat{u}_k - \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} < \inf\left(\frac{2s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g}\right) = \frac{2\varphi}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g}$$

Solving the above expression by substituting from Equation (4.136) gives:

$$\varphi > \frac{\Delta t}{2}\left\{\mathbf{p}_k \hat{\mathbf{g}}\left[\left(\frac{\Delta g}{\beta}\frac{\beta^2+1}{2} - 1\right)\hat{u}_k + \frac{\Delta g}{\beta}\frac{\beta^2-1}{2}\left|\hat{u}_k\right|\right] + \eta + \Delta g \beta \kappa\right\} \qquad (4.137)$$

For $s_k < -\varphi$:

$$u_k > \frac{1}{\Delta g}\hat{u}_k + \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} + \frac{2s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} \Rightarrow u_k - \frac{1}{\Delta g}\hat{u}_k - \frac{\upsilon_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g} > \sup\left(\frac{2s_k}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g}\right) = -\frac{2\varphi}{\Delta t \mathbf{p}_k \hat{\mathbf{g}} \Delta g}$$

Solving the above expression by substituting from Equation (4.136) gives:

$$\varphi > \frac{\Delta t}{2}\left\{\mathbf{p}_k \hat{\mathbf{g}}\left[\left(1 - \frac{\Delta g}{\beta}\frac{\beta^2+1}{2}\right)\hat{u}_k + \frac{\Delta g}{\beta}\frac{\beta^2-1}{2}\left|\hat{u}_k\right|\right] + \eta + \Delta g \beta \kappa\right\} \qquad (4.138)$$

Expressions (4.137) and (4.138) can be combined into:

$$\begin{aligned} &\varphi \geq \frac{\Delta t}{2}\left\{\mathbf{p}_k \hat{\mathbf{g}}\left[\left(\frac{\Delta g}{\beta}\frac{\beta^2+1}{2} - 1\right)\hat{u}_k \operatorname{sgn}(s_k) + \frac{\Delta g}{\beta}\frac{\beta^2-1}{2}\left|\hat{u}_k\right|\right] + \eta + \Delta g \beta \kappa\right\} \Rightarrow \\ &\varphi \geq \frac{\Delta t}{2}\left\{\mathbf{p}_k \hat{\mathbf{g}}\left[\beta_{\max} + \frac{\beta^2-1}{2}\left|\hat{u}_k\right|\right] + \eta + \Delta g \beta \kappa\right\} \end{aligned} \qquad (4.139)$$

where

$$\beta_{\max} = \sup\left[\left(\frac{\Delta g}{\beta}\frac{\beta^2+1}{2}-1\right)\hat{u}_k \operatorname{sgn}(s_k)\right]$$
$$= \begin{cases} \dfrac{\beta^2-1}{2}|\hat{u}_k| \ when\ \hat{u}_k \operatorname{sgn}(s_k) > 0 \\ \dfrac{1-\beta^2}{2\beta^2}\left(-|\hat{u}_k|\right) = \dfrac{\beta^2-1}{2\beta^2}|\hat{u}_k| \ when\ \hat{u}_k \operatorname{sgn}(s_k) < 0 \end{cases}$$
$$= \frac{\beta^2-1}{2}\frac{|\hat{u}_k|+\hat{u}_k \operatorname{sgn}(s_k)}{2} + \frac{\beta^2-1}{2\beta^2}\frac{|\hat{u}_k|-\hat{u}_k \operatorname{sgn}(s_k)}{2}$$

With the positive constant κ defined in Equation (4.131), without loss of generality, selection of $\kappa = \eta + 2\varepsilon$, $\varepsilon > 0$, the boundary layer width φ in Expression (4.139) can be determined as:

$$\varphi \geq \frac{\Delta t}{2}\left\{\frac{\mathbf{p}_k \hat{\mathbf{g}}\left(\beta^2-1\right)}{4\beta^2}\left[\left(3\beta^2+1\right)|\hat{u}_k| + \left(\beta^2-1\right)\hat{u}_k \operatorname{sgn}(s_k)\right] + \left(1+\beta^2\right)\eta + 2\beta^2\varepsilon\right\} \qquad (4.140)$$

With control law (4.136) and the smoothing boundary layer thickness of Expression (4.140), the stability condition, $s_{k+1}^2 < s_k^2$, is satisfied outside of the boundary layer and the smoothing boundary layer with thickness φ is attractive. All trajectories in the state space will converge to the smoothing boundary layer despite uncertainties in the system. However this convergence does not imply the system is stable. For instance, the system trajectory is bounded inside the boundary layer, but the thickness of the boundary grows unbounded.

As discussed above, the control law of Equation (4.136) and the boundary layer thickness of Expression (4.140) only guarantee the system trajectory converging to the boundary layer. In order to make the system asymptotically stable, boundedness of the thickness of the boundary layer has to be realized.

For the system of Equation (4.108) with all assumptions mentioned above, further assuming that the system is bounded, such that:

$$\mathbf{p}_k\mathbf{f} \le \gamma_{pf},\ 0 < \gamma_{pg}^{-} \le \mathbf{p}_k\hat{\mathbf{g}} \le \gamma_{pg}^{+} \qquad (4.141)$$

Taking the absolute value of (4.128) and substituting from Equation (4.114) and Expression (4.138):

$$|\hat{u}_k| = \left|\frac{\mathbf{p}_k\Delta\mathbf{x}_{d.k}}{\Delta t\mathbf{p}_k\hat{\mathbf{g}}} - \frac{\mathbf{p}_k\mathbf{f}}{\mathbf{p}_k\hat{\mathbf{g}}}\right| \le \left|\frac{1}{\mathbf{p}_k\hat{\mathbf{g}}}\right|\left(\left|\mathbf{p}_k\frac{\Delta\mathbf{x}_{d.k}}{\Delta t}\right| + |\mathbf{p}_k\mathbf{f}|\right) \le \frac{1}{\gamma_{pg}^{-}}\left(\gamma_{xd} + \gamma_{pf}\right) \qquad (4.142)$$

The infimum of the thickness of the smoothing boundary layer from Expression (4.140) can be obtained as:

$$\varphi_{\inf} = \frac{\Delta t}{2}\left\{\frac{\mathbf{p}_k\hat{\mathbf{g}}(\beta^2-1)}{4\beta^2}\left[(3\beta^2+1)|\hat{u}_k| + (\beta^2-1)\hat{u}_k\,\mathrm{sgn}(s_k)\right] + (1+\beta^2)\eta + 2\beta^2\varepsilon\right\},\ \varepsilon > 0$$

Taking the absolute value of the above equation and substituting from Expressions (4.141) and (2.142) give:

$$|\varphi_{\inf}| \le \frac{\Delta t}{2}\left\{\frac{|\mathbf{p}_k\hat{\mathbf{g}}|(\beta^2-1)}{4\beta^2}\left[(3\beta^2+1)|\hat{u}_k| + (\beta^2-1)\hat{u}_k\,\mathrm{sgn}(s_k)\right] + (1+\beta^2)\eta + 2\beta^2\varepsilon\right\} \Rightarrow$$
$$|\varphi_{\inf}| \le \frac{\Delta t}{2}\left[(\beta^2-1)\frac{\gamma_{pg}^{+}}{\gamma_{pg}^{-}}\left(\gamma_{xd} + \gamma_{pf}\right) + (1+\beta^2)\eta + 2\beta^2\varepsilon\right] \qquad (4.143)$$

It is proved that the thickness of the smoothing boundary layer is bounded provided that the assumptions of (4.109), (4.114), (4.131) and (4.141) are held. The thickness of the smoothing boundary layer (4.143) indicates that the sampling period, the bounds of modeling errors and the command signal are factors affecting the tracking quality of the discrete–time sliding mode controller. Conversely, smaller sampling period and bettering

modeling can reduce the thickness of the boundary layer and improve the tracking accuracy.

If the input gain vector $\mathbf{g}(\mathbf{x})$ in Equation (4.108) is exactly known, so that $\Delta g = 1$ and β in Expression (4.109) is set to one. The control law of Equation (4.136) and the thickness of the smoothing boundary layer (4.143) can be simplified as:

$$\begin{cases} u_k = \hat{u}_k + \dfrac{\eta + 2\varepsilon}{\mathbf{p}_k \hat{\mathbf{g}}} sat\left(\dfrac{s_k}{\varphi}\right) \\ \hat{u}_k = \dfrac{\mathbf{p}_k \Delta \mathbf{x}_{d.k}}{\Delta t \mathbf{p}_k \hat{\mathbf{g}}} - \dfrac{\mathbf{p}_k \mathbf{f}}{\mathbf{p}_k \hat{\mathbf{g}}} \\ \varphi = \Delta t(\eta + \varepsilon) \end{cases} \tag{4.144}$$

Inside the boundary layer, from (4.124):

$$s_{k+1} - s_k = \mathbf{p}_k \Delta \mathbf{x}_{d.k} - \Delta t \mathbf{p}_k \mathbf{f} - \Delta t \mathbf{p}_k \mathbf{g} u_k + \upsilon_k$$

Substituting from Equation (4.141) into the above equation gives:

$$\begin{aligned} & s_{k+1} = -\frac{\varepsilon}{\eta + \varepsilon} s_k + \Delta t\left(\frac{\sigma}{\Delta t} - \mathbf{p}_k \mathbf{h}_k\right) \Rightarrow |s_{k+1}| \le \frac{\varepsilon}{\eta + \varepsilon}|s_k| + \Delta t\left|\frac{\sigma}{\Delta t} - \mathbf{p}_k \mathbf{h}_k\right| \Rightarrow \\ & s_{k+1} \le \frac{\varepsilon}{\eta + \varepsilon}|s_k| + \Delta t \eta \end{aligned} \tag{4.145}$$

The above expression clearly verifies the boundary layer is bounded.

DSMC Based on Sliding Mode Prediction

Xiao, Su & Chu (2007) proposed a sliding mode prediction based DSMC algorithm by introducing a predictive control switching function into the design of DSMC for nonlinear systems. In their work, a sliding mode prediction model (SMPM) was created first, and the error between the output of the SMPM model and the practical sliding mode value was used to make feedback corrections for the SMPM. Then a performance index

was constructed based on the SMPM and the control law. Finally the DSMC law was derived which ensured that the performance index was minimized.

Consider the following nonlinear coupled system, for example, $n-$ joint rigid robotic manipulator:

$$\begin{cases} x_{2i-1k+1} = x_{2ik} \\ x_{2ik+1} = f_i(\mathbf{x}_k) + g_i(\mathbf{x}_k)u_{ik} + w_{ik} \end{cases} \quad i = 1, \cdots, n \qquad (4.146)$$

where $\mathbf{x} \in R^{2n}$ is the state vector and is assumed measurable, $f_i(\bullet)$ and $g_i(\bullet)$are nonlinear scalar functions, u_i is scalar input. w_{ik} is uncertainty which may include parametric uncertainties and external disturbances.

The design objective is to construct a no-switching DSMC law $\mathbf{u}_k$ which guarantees that the state trajectory of Equation (4.146) asymptotically converges to the sliding mode. The design task is fulfilled in two steps. Firstly a sliding surface is designed such that the system in the sliding mode demonstrates the desired performance. Secondly, a control law is designed to drive the states to the sliding surface and maintain on or in the neighbourhood of the sliding surface thereafter.

The sliding surface is designed as:

$$s_{ik} = x_{2ik} + \sigma_i x_{2i-1k} = 0, \; i = 1, \cdots, n \qquad (4.147)$$

where $-1 < \sigma_i < 0$ is chosen to guarantee the stability and dynamic performance of ideal sliding mode motion of the system of Equation (4.146).

A recursive switching function is designed based the sliding surface of Equation (4.147) as:

$$s_{mik+1} = s_{ik+1} + \gamma_i s_{ik}, \quad i = 1, \cdots, n \qquad (4.148)$$

where $0 < \gamma_i \leq 1$. Substituting from Equation (4.146), Equation (4.148) can be rewritten as:

$$s_{mik+1} = f_i(\mathbf{x}_k) + g_i(\mathbf{x}_k)u_{ik} + \sigma_i x_{2ik} + \gamma_i s_{ik} \qquad (4.149)$$

The recursive switching function of Equation (4.149) can be viewed as a prediction of the state s_{ik+1}, such that it is called a sliding mode prediction model.

Due to time variance, nonlinearity, and external disturbance, the output s_{mik+1} of the SMPM will not be the same as the real switching function s_{ik+1}. The error can be reduced by using feedback correction. The error between the practical switching function $s_{i.k}$ and the SMPM s_{imk} is feedback to correct the SMPM. The corrected SMPM output becomes

$$\hat{s}_{mik+1} = s_{mik+1} + \xi_i(s_{ik} - s_{mik}) \qquad (4.150)$$

where $\xi_i \in R$ is the weighting factor, and $0 \leq \xi_i \leq 1$. With the representation $\tilde{s}_{ik} = s_{ik} - s_{mik}$, Equation (4.150) is rewritten as:

$$\hat{s}_{mik+1} = s_{mik+1} + \xi_i \tilde{s}_{ik} \qquad (4.151)$$

Since state $x_{i.k}$ is measurable, $s_{i.k}$ can be calculated from Equation (4.147), and $s_{mi.k}$ can be determined from Equation (4.148), then $\hat{s}_{mi.k+1}$ can be computed as:

$$\hat{s}_{mik+1} = h_{ik} + g_i(\mathbf{x}_k)u_{ik} \qquad (4.152)$$

where $h_{ik} = f_i(\mathbf{x}_k) + \sigma_i x_{2ik} + \gamma_i s_{ik} + \xi_i \tilde{s}_{ik}$.

A performance index is constructed as a summation of the squared error of the SMPM and the weighted squared control input (Xiao, Su & Chu, 2007), such that:

$$J_i = (\hat{s}_{mik+1} - s_{ri})^2 + \lambda_i u_{ik}^2, \quad i = 1, \cdots, n \qquad (4.153)$$

where s_{ri} is sliding mode reference value. The objective of the DSMC is to keep the states on the sliding surface, that is, $s_{ri} = 0$. In this case, the performance index is reduced to:

$$J_i = \hat{s}_{mik+1}^2 + \lambda_i u_{ik}^2, \quad i = 1, \cdots, n \qquad (4.154)$$

where λ_i is a weighting factor adjusting the relation between the closed loop output of SMPM and the input signal.

The control objective is to find a control $u_{i.k}$ which minimizes the performance index. The control law can be solved by setting $\frac{\partial J_i}{\partial u_i} = 0$ and solving for an optimal $u_{i.k}$ as:

$$u_{i.k} = -\frac{h_{ik} g_i(\mathbf{x}_k)}{g_i^2(\mathbf{x}_k) + \lambda_i} \quad i = 1, \cdots, n \qquad (4.155)$$

From Equation (4.154), decreasing of λ_i weakens the contribution of $u_{i.k}$ in the performance index, such that, for the same value of performance index, more control energy is needed. By setting $\lambda_i = 0$, the non-optimal control is derived as:

$$u_{i.k} = -\frac{h_{ik}}{g_i(\mathbf{x}_k)} \quad i = 1, \cdots, n \qquad (4.156)$$

The closed-loop dynamics of the DSMC system can be determined as follows. Substituting from the control law of Equation (4.156) and the system model of Equation (4.146), the switching function of Equation (4.147) is rewritten as:

$$s_{ik+1} = -\gamma_i s_{ik} - \xi_i \tilde{s}_{ik} + w_{ik} \qquad (4.157)$$

Combining Equations (4.149), (4.146) and (4.147), then:

$$\tilde{s}_{ik} = s_{ik} - s_{mik} = -\gamma_i s_{ik-1} + w_{ik-1} \qquad (4.158)$$

Substituting from Equation (4.158), Equation (4.157) can be rewritten as:

$$s_{ik+1} = \gamma_i\left(-s_{ik} + \xi_i s_{ik-1}\right) + \left(w_{ik} - \xi_i w_{ik-1}\right) \qquad (4.159)$$

Because the polynomial $\left(-1+\xi_i z^{-1}\right)$ has a root $z = \xi_i$, $0 < \xi_i \le 1$, and $0 < \gamma_i \le 1$, both terms on the right hand side of Equation (4.159) are stable, that is:

$$\left|s_{i.k+1}\right| \le \gamma_i\left|-s_{i.k} + \xi_i s_{i.k-1}\right| + \left|w_{i.k} - \xi_i w_{i.k-1}\right| \le \varsigma_i + \zeta_i \qquad (4.160)$$

where ς_i and ζ_i are the bounds of s-dynamics and the uncertainty dynamics, respectively. The practical switching function will converge to a $\varsigma_i + \zeta_i$ vicinity of the sliding surface and stay thereafter.

Chapter 5 Control of the Belt Drive System

In this section, a discrete-time Proportional-Integral-Derivative (PID) controller and a Discrete-time Sliding Mode Controller (DSMC) are designed and implemented to track the desired output of the belt drive system. DSMC is a model based control algorithm with velocity feed forward compensation and trajectory tracking. The trajectory is specified in terms of desired state trajectories. Tracking the desired output is realized by tracking the states.

5.1 Design of Discrete-time PID Controller Using the SISOTool

Root locus is a graphical representation of the position and movement of the closed-loop poles as a system parameter or a gain is varied. Root locus provides a qualitative and a quantitative indication of the control system's performance such as percent overshoot, settling time, peak time as well as stability. It is used to solve problems for high order single-input-single-output systems (Nise, 2004). The root locus method was originally developed for continuous systems. It can be extended to discrete systems by using the z-transform in the z-domain.

The Proportional-Integral-Derivative (PID) controller is widely used in industrial applications. Block diagrams of the continuous-time and discrete-time PID controllers are given in Figures 51 and 52, respectively.

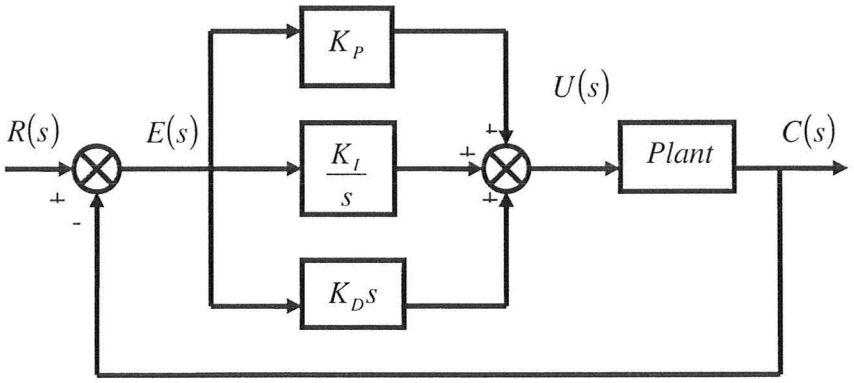

Figure 51 Block Diagram of the Continuous-time PID Controller

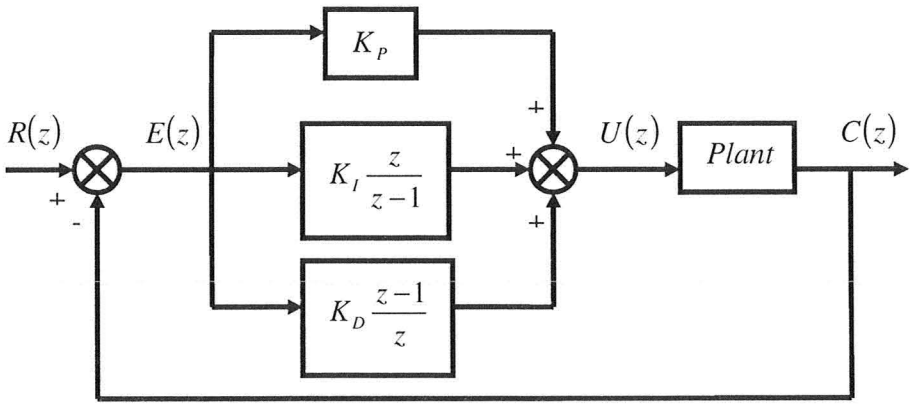

Figure 52 Block Diagram of the Discrete-time PID Controller

The transfer function of the discrete-time PID controller is as follows:

$$G(z) = K_P + K_I\left(\frac{z}{z-1}\right) + K_D\left(\frac{z-1}{z}\right) \qquad (5.1)$$

where K_P, K_I and K_D are the proportional, integral and derivative gains, respectively. The proportional term provides an overall control action proportional to the error through an all-pass gain factor. The integral term can remove steady-state error through low frequency compensation by using an integrator. The derivative term improves transient response by using a differentiator. Equation (5.1) can be re-written as:

$$G_{PID}(z)=\frac{(K_P+K_I+K_D)z^2-(K_P+2K_D)z+K_D}{z(z-1)} \qquad (5.2)$$

The PID controller adds two poles, one at the origin and the other at $z=1$, and two zeros to be positioned according to the required performance specifications. Matlab's SISO design tool combines the root locus and frequency response techniques to be used concurrently for the design of compensators. By using the SISO design tool, a DPID compensator can be easily designed by adding two poles, two zeros and simply adjusting of the compensator gain. With the fixed compensator poles, the compensator gains, K_P, K_I and K_D, solely depend on the positions of the added zeros and the amplitude of the compensator gain. There are many zero-gain combinations which result in various choices of the gains K_P, K_I and K_D. The final choice of zeros and gain should satisfy the required control performance.

Further to Chapter 3, the transfer function of the belt drive system of Equation (3.33) is rewritten as follows:

$$G(z)=\frac{19.2}{z^2-1.226z+0.5311} \qquad (5.3)$$

The root locus, the open-loop and the closed-loop Bode plots of the uncompensated belt drive system are shown in Figure 53.

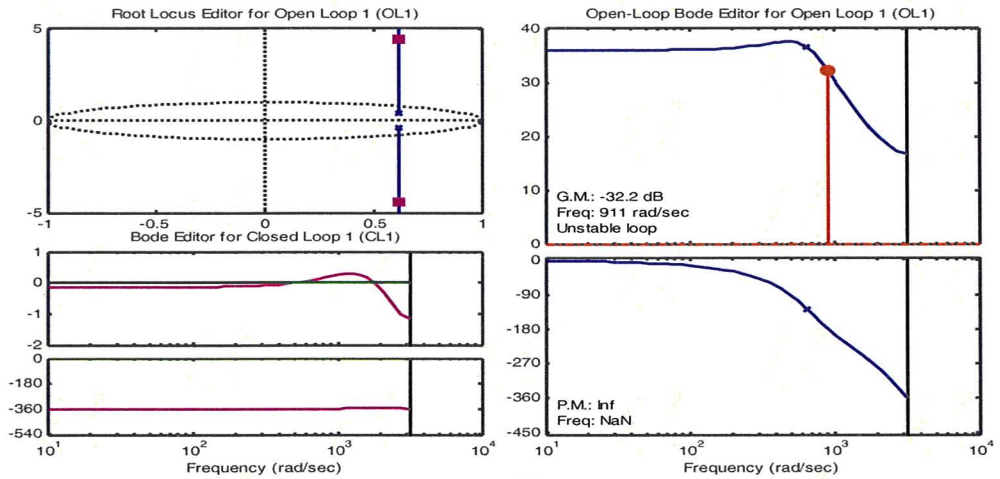

Figure 53 The Root Locus and Bode Plots of the Uncompensated Belt Drive System

Using the SISO design tool the transfer function of the DPID controller is obtained as:

$$G_{PID}(z)=\frac{0.0044\left(z^2-0.4286z+0.1349\right)}{z(z-1)} \qquad (5.4)$$

The DPID controller of Equation (5.4) can be written in a parallel form as:

$$G_{PID}(z)=0.0007+0.0031\frac{z}{z-1}+0.0006\frac{z-1}{z} \qquad (5.5)$$

The transfer function of the closed-loop DPID controlled belt drive system can be written as:

$$G_c(z)=\frac{G_{PID}(z)G(z)}{1+G_{PID}(z)G(z)}=\frac{0.0845z^2-0.03622z+0.0114}{z^4-2.226z^3+1.842z^2-0.5673z+0.0114} \qquad (5.6)$$

The final pole and zero locations for the closed-loop control system are obtained within the unit circle and as follows:

- Poles: $\begin{cases} z = 0.7247 + 0.4182i \\ z = 0.7247 - 0.4182i \\ z = 0.7550 \\ z = 0.0216 \end{cases}$

- Zeros: $\begin{cases} z = 0.2143 + 0.2983i \\ z = 0.2143 - 0.2983i \end{cases}$

The root locus and the Bode plots of the DPID controlled belt drive system are shown in Figure 54.

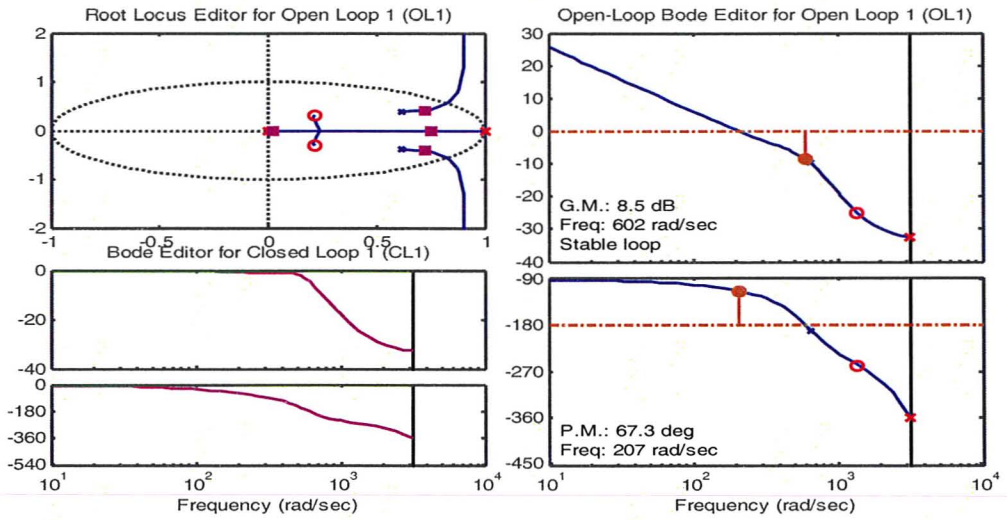

Figure 54 The Root Locus and the Bode Plots of the DPID Controlled Belt Drive System

The simulated step response of the PID controlled system is shown in Figure 55. The controlled system produces 2.53% percent overshoot, 5 ms rise time, 15 ms settling time.

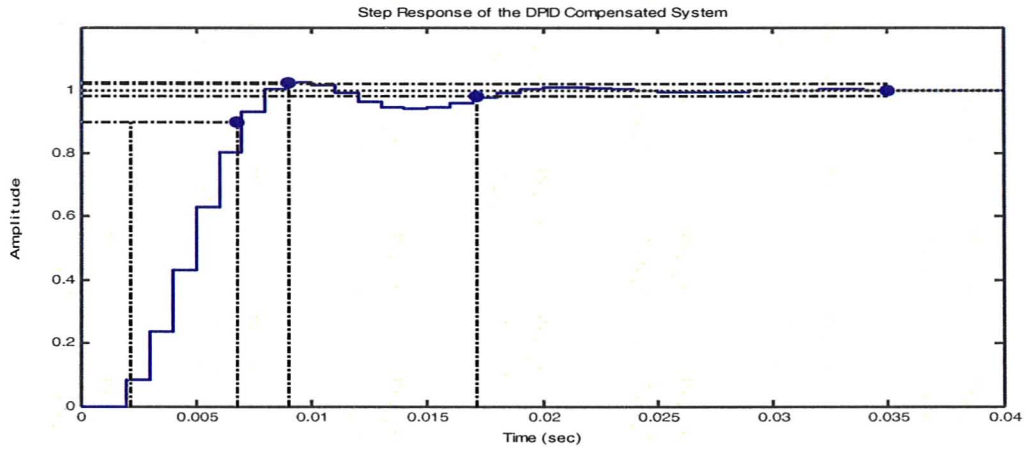

Figure 55 The Simulated Step Response of the PID Controlled Belt Drive System

5.2 State-Space Representation of the Belt Drive System

Normally the DSMC algorithm is implemented in the state space model of the system to be controlled. The transfer function of the belt drive system obtained through system identification has to be converted to a state space model. Given a discrete transfer function $\frac{Y(z)}{U(z)} = \frac{b_0 z^n + b_1 z^{n-1} + \cdots + b_n}{z^n + a_1 z^{n-1} + \cdots + a_n}$, the controllable canonical form state-space model can be obtained as (Ogata, 1994):

$$\begin{bmatrix} x_{1k+1} \\ x_{2k+1} \\ \vdots \\ x_{n-1k+1} \\ x_{nk+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_n & -a_{n-1} & -a_{n-2} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{n-1k} \\ x_{nk} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u_k \qquad (5.7)$$

$$y_k = \begin{bmatrix} b_n - a_n b_0 & b_{n-1} - a_{n-1} b_0 & b_{n-2} - a_{n-2} b_0 & \cdots & b_1 - a_1 b_0 \end{bmatrix} \begin{bmatrix} x_{1.k} \\ x_{2.k} \\ \vdots \\ x_{n-1.k} \\ x_{n.k} \end{bmatrix} + b_0 u_k \quad (5.8)$$

Equations (5.7) and (5.8) are the system and the output equations, respectively. For the identified model of Equation (5.3), the state-space model becomes:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k \\ y_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}u_k \end{aligned} \quad (5.9)$$

where $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -0.5311 & 1.226 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{C} = \begin{bmatrix} 19.2 \\ 0 \end{bmatrix}^T$, $\mathbf{D} = 0$.

5.2.1 Controllability

According to Ogata (1994), a system is completely controllable if the system at any arbitrary state can be moved to any other desired arbitrary state by some unconstrained input signal in a finite time period. In other words, if the state variable is independent of the control signal, then it is impossible to control the system state and the system is uncontrollable. Most practical systems are controllable. Knowing the condition under which the system is controllable is very important in control systems engineering.

Consider the state space model of the second-order belt drive system given by Equation (5.9) as:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k$$

where u_k is a scalar input, $\mathbf{x}_k \in R^2$ is the state vector, $\mathbf{A} \in R^{2\times 2}$ is the system matrix, $\mathbf{B} \in R^{2\times 1}$ is the input matrix. The above system is completely controllable if the rank of its controllability matrix equals to the order of the system. The matrix $\left[\mathbf{B} \vdots \mathbf{BA} \vdots \cdots \vdots \mathbf{A}^{n-1}\mathbf{B}\right]$ is called the controllability matrix, and the condition for complete controllability can be stated as:

$$rank\left[\mathbf{B} \vdots \mathbf{BA} \vdots \cdots \vdots \mathbf{A}^{n-1}\mathbf{B}\right] = n \qquad (5.10)$$

For the belt drive system given in Equation (5.9), the controllability matrix is obtained as:

$$\mathbf{M} = \left[\mathbf{B} \vdots \mathbf{AB}\right] = \begin{bmatrix} 0 & 1 \\ 1 & 1.226 \end{bmatrix} \qquad (5.11)$$

Since the controllability matrix $\mathbf{M}$ has full rank, the belt drive system is completely controllable.

5.2.2 Observability

According to Ogata (1994), a system is completely observable if every initial state $\mathbf{x}_0$ can be determined from the observation of y_k over a finite number of sampling periods. In other words, if every transition of the state eventually affects every element of the output vector, the system is completely observable. For a discrete-time linear time invariant system, the observability matrix is defined as:

$$\mathbf{N} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} \qquad (5.12)$$

To be able to uniquely extract the state vector from the measurements, the rank of the observability matrix must be equal to the order of the system as follows:

$$rank\left(\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix}\right) = n \qquad (5.13)$$

For our belt drive system given in Equation (5.9), the observability matrix can be obtained as:

$$\mathbf{N} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \end{bmatrix} = \begin{bmatrix} 19.2 & 0 \\ 0 & 19.2 \end{bmatrix} \qquad (5.14)$$

The observability matrix has the rank of 2 for the 2^{nd}-order belt drive system. It is completely observable.

5.3 Definition of Desired State Trajectories

The objective of trajectory tracking control is to minimize the tracking error of a state following a desired trajectory. Normally the trajectories to be followed are low frequency signals such as ramp or and sinusoidal signals. In this research, the trajectory to be followed is a velocity profile defined as a low frequency sine wave. The other desired state trajectories are then derived from the desired velocity profile.

The state space model of the belt drive system is given in Equation (5.9) as:

$$\begin{bmatrix} x_{1k+1} \\ x_{2k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -0.5311 & 1.226 \end{bmatrix} \begin{bmatrix} x_{1k} \\ x_{2k} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_k$$

$$y_k = 19.2 x_{1k}$$

The desired output trajectory is given as:

$$y_{dk} = B + A\sin(\omega kT),\ k = 0,1,\cdots,N+1 \qquad (5.15)$$

where A is the amplitude of the sine wave, B is the bias to keep the output in the linear region , ω is the angular frequency in radians/second, T is the sampling time. The desired state trajectories can be calculated from Equation (5.15) and the state space model as:

$$x_{1dk} = 0.0521[B + A\sin(\omega kT)],\ k = 0,1,\cdots,N \qquad (5.16)$$

$$x_{2dk} = 0.0521\{B + A\sin[\omega(k+1)T]\},\ k = 0,1,\cdots,N \qquad (5.17)$$

Noticed from Equations (5.15), (5.16) and (5.17), the relationships between the state and output trajectories can be shown in Figure 56, where $X_{1d}(z)$, $X_{2d}(z)$ and $Y_d(z)$ are z-transforms of the desired trajectory x_{1d} , x_{2d} and y_d , respectively.

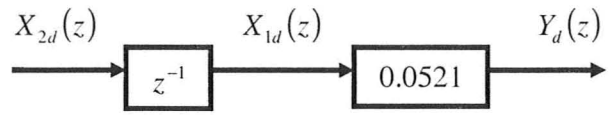

Figure 56 Relationship between the State and the Output Trajectories

With $A = 15$, $B = 16$ and $\omega = \pi$, the desired output and state trajectories are shown in Figure 57.

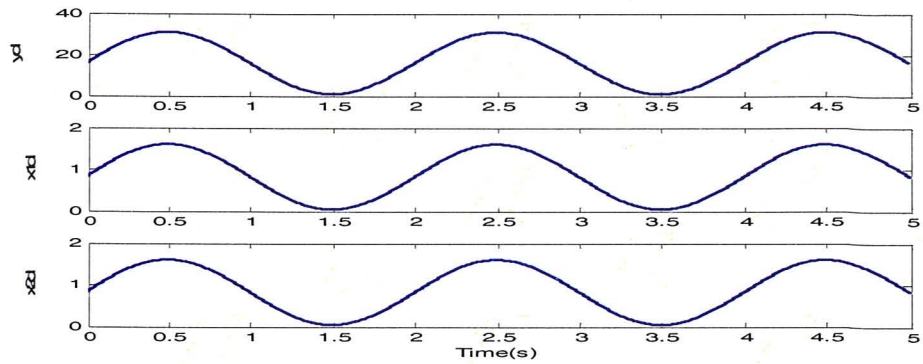

Figure 57 Desired Output and State Trajectories

5.4 State Observers

The design of SMC is based on the assumption that the state vector of the system to be controlled is available for measurement. In practice, only a few output quantities are available for measurement. Fortunately, the state vector of a linear system can be reconstructed from observations of the system output through a state observer. The state observer is a subsystem in the control system that performs an estimation of state variables based on the measurements of output. The state observer can be designed if and only if the observability condition is satisfied, that is, the observability matrix has full rank (Ogata, 1994). This condition is satisfied in the experimental belt drive system.

The block diagram of a DSMC system with a state observer is shown in Figure 58, where y_d is the desired trajectory, y is the measured output, $\mathbf{x}_d$ is the desired state trajectory, $\hat{\mathbf{x}}$ is the estimated state trajectory, $\tilde{\mathbf{x}}$ is the state tracking error, u is the output of the DSMC.

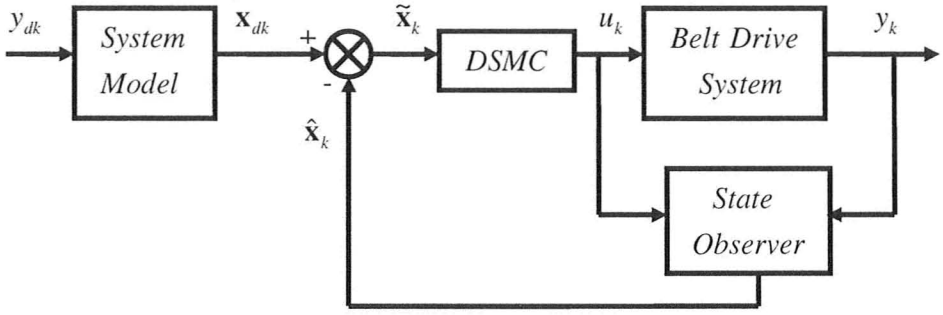

Figure 58 DSMC System with a State Observer

The generation of the desired state trajectory $\mathbf{x}_d$ from the desired output y_d using the system model is contained in the previous subsection. The design of the state observer is addressed as follows.

There are three types of state observers available depending on the order of the state observers: (1) full-order state observer, (2) reduced-order state observer, and (3) minimum-order state observer (Ogata, 1994). Full-order state observer estimates all n state variables regardless of whether some state variables are available for direct measurement. Reduced-order state observer estimates all unmeasurable state variables plus some (but not all) of the measurable state variables. Minimum-order state observer estimates only the state variables that are not measured directly.

5.4.1 Full-Order State Observer

Consider the state space model of the belt drive system given in Equation (5.9) as:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}u_k \end{aligned} \tag{5.18}$$

The full-order state observer has the same order as that of the system. The full-order state observer can be designed based on the model of the original system with minor modifications as shown in Figure 59. The input of the state observer consists of the control input u_k and the measured output y_k. The output of the state observer is the estimated state $\hat{\mathbf{x}}_k$.

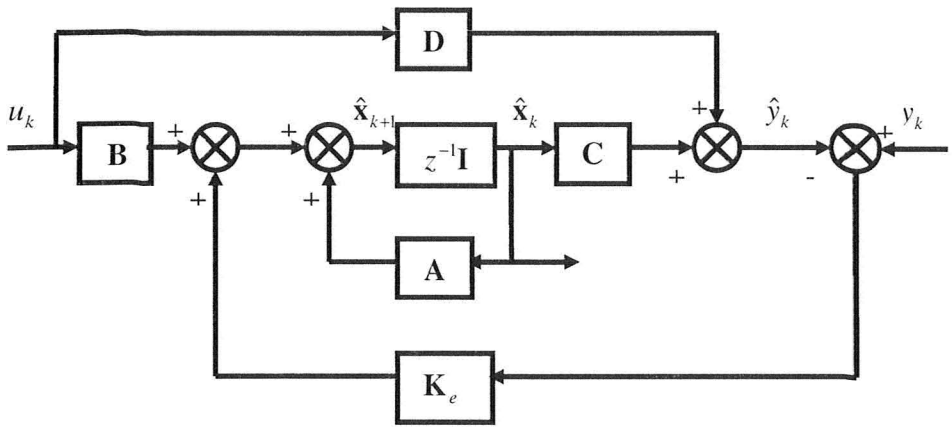

Figure 59 Block Diagram of the Full-Order State Observer

The full-order state observer can be categorized into the prediction observer or the current observer. In the former, the observed state $\hat{\mathbf{x}}_{k+1}$ is obtained from the measured output and the input up to time step k. In the latter, the state $\hat{\mathbf{x}}_{k+1}$ is obtained from the measured output and the input up to the step $k+1$.

Prediction Observer

The prediction observer is built on the previous step input and output information. The estimated output $\hat{y}_k$ can be obtained as:

$$\hat{y}_k = \mathbf{C}\hat{\mathbf{x}}_k + \mathbf{D}u_k \qquad (5.19)$$

The state of the system is estimated as:

$$\hat{\mathbf{x}}_{k+1} = \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}u_k + \mathbf{K}_e(y_k - \hat{y}_k) \qquad (5.20)$$

where $\mathbf{K}_e$ is the state observer gain serving as a weighting matrix. Substituting Equation (5.19) into Equation (5.20) gives:

$$\hat{\mathbf{x}}_{k+1} = (\mathbf{A} - \mathbf{K}_e\mathbf{C})\hat{\mathbf{x}}_k + (\mathbf{B} - \mathbf{K}_e\mathbf{D})u_k + \mathbf{K}_e y_k \qquad (5.21)$$

The state estimate $\hat{\mathbf{x}}_{k+1}$ is one sampling period ahead of the measurement y_k, so the state observer given by Equation (5.21) is called the prediction observer and the eigenvalues of the matrix $(\mathbf{A} - \mathbf{K}_e\mathbf{C})$ are commonly called the observer poles (Ogata, 1994). If $\hat{\mathbf{x}}_k = \mathbf{x}_k$, Equation (5.21) becomes:

$$\hat{\mathbf{x}}_{k+1} = \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}u_k$$

which is identical to the state space model of the system and the response of the state observer is identical to the response of the original system.

Subtracting the actual state in Equation (5.18) from estimations in Equation (5.21) gives the estimation error dynamics as:

$$\mathbf{e}_{k+1} = (\mathbf{A} - \mathbf{K}_e\mathbf{C})\mathbf{e}_k \qquad (5.22)$$

where $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ is the estimation error. The dynamic behaviour of the estimation error is governed by the eigenvalues of matrix $(\mathbf{A} - \mathbf{K}_e\mathbf{C})$. If matrix $(\mathbf{A} - \mathbf{K}_e\mathbf{C})$ is stable, the error vector will converge to zero for any initial error $\mathbf{e}_0$; that is, $\hat{\mathbf{x}}_k$ will converge to $\mathbf{x}_k$ for any initial values of $\mathbf{x}_0$ and $\hat{\mathbf{x}}_0$. Since the original system is completely observable, an arbitrary placement of the eigenvalues of the matrix $(\mathbf{A} - \mathbf{K}_e\mathbf{C})$ is possible.

The dynamic behaviour of the estimation error can be made adequately fast by locating the eigenvalues of the matrix in $(\mathbf{A}-\mathbf{K}_e\mathbf{C})$ in proper positions so that the estimation error tends to zero with an adequately fast speed. One way to obtain a fast response is to use deadbeat response by choosing all eigenvalues of the matrix $(\mathbf{A}-\mathbf{K}_e\mathbf{C})$ to be zero.

Feedback Gain Matrix $\mathbf{K}_e$

There are a few of approaches available for determining the feedback gain matrix $\mathbf{K}_e$ for the state observer. For the full-order prediction observer given in Equation (5.21), a general approach to determine the feedback gain matrix $\mathbf{K}_e$ is discussed as follows (Ogata, 1994).

First a transformation is defined as:

$$\mathbf{Q} = \left(\mathbf{W}\mathbf{N}^T\right)^{-1} \qquad (5.23)$$

where $\mathbf{N}$ is the observability matrix defined as:

$$\mathbf{N} = \left[\mathbf{C}^T \vdots \mathbf{A}^T\mathbf{C}^T \vdots \cdots \vdots \left(\mathbf{A}^T\right)^{n-1}\mathbf{C}^T\right] \qquad (5.24)$$

and:

$$\mathbf{W} = \begin{bmatrix} a_{n-1} & a_{n-2} & \cdots & a_1 & 1 \\ a_{n-2} & a_{n-3} & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \qquad (5.25)$$

where a_1, a_2, …, a_{n-1} are coefficients in the characteristic equation of the original system of Equation (5.18), that is, $|z\mathbf{I}-\mathbf{G}| = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = 0$. Then the state vector is defined as:

$$\mathbf{x}_k = \mathbf{Q}\xi_k \qquad (5.26)$$

where $\xi_k \in R^{n\times 1}$. The original system of Equation (5.18) can be rewritten with Equation (5.26) as:

$$\begin{aligned} \xi_{k+1} &= \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}\xi_k + \mathbf{Q}^{-1}\mathbf{B}u_k \\ y_k &= \mathbf{C}\mathbf{Q}\xi_k + \mathbf{D}u_k \end{aligned} \qquad (5.27)$$

with [Ogata1994]:

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 \end{bmatrix} \qquad (5.28)$$

$$\mathbf{C}\mathbf{Q} = [0 \quad 0 \quad \cdots \quad 0 \quad 1] \qquad (5.29)$$

The estimated state vector is defined as:

$$\hat{\mathbf{x}}_k = \mathbf{Q}\hat{\xi}_k \qquad (5.30)$$

where $\hat{\xi}_k \in R^n$. Substituting Equation (5.30) into the full-order state observer of Equation (5.21) gives:

$$\hat{\xi}_{k+1} = \mathbf{Q}^{-1}(\mathbf{A}-\mathbf{K}_e\mathbf{C})\mathbf{Q}\hat{\xi}_k + \mathbf{Q}^{-1}(\mathbf{B}-\mathbf{K}_e\mathbf{D})u_k + \mathbf{Q}^{-1}\mathbf{K}_e y_k \qquad (5.31)$$

Subtracting Equation (5.27) from Equation (5.31) gives the estimation dynamics as:

$$\mathbf{e}_{k+1} = \mathbf{Q}^{-1}(\mathbf{A}-\mathbf{K}_e\mathbf{C})\mathbf{Q}\mathbf{e}_k \qquad (5.32)$$

where $\mathbf{e}_k = \xi_k - \hat{\xi}_k$ is the estimation error. From Equation (5.23), a new vector $\delta \in R^n$ is defined as:

$$\mathbf{Q}^{-1}\mathbf{K}_e = \mathbf{W}\mathbf{N}^T\mathbf{K}_e = \begin{bmatrix} a_{n-1} & a_{n-2} & \cdots & a_1 & 1 \\ a_{n-2} & a_{n-3} & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_1 & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-2} \\ \mathbf{CA}^{n-1} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_{n-1} \\ k_n \end{bmatrix} = \begin{bmatrix} \delta_n \\ \delta_{n-1} \\ \vdots \\ \delta_2 \\ \delta_1 \end{bmatrix} \qquad (5.33)$$

where $\mathbf{K}_e = \begin{bmatrix} k_1 & k_2 & \cdots & k_{n-1} & k_n \end{bmatrix}^T$. Substituting Equations (5.28), (5.29) and (5.33) into Equation (5.32) gives the tracking error dynamics as:

$$\mathbf{e}_{k+1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_n - \delta_n \\ 1 & 0 & \cdots & 0 & -a_{n-1} - \delta_{n-1} \\ 0 & 1 & \cdots & 0 & -a_{n-2} - \delta_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_1 - \delta_1 \end{bmatrix} \mathbf{e}_k \qquad (5.34)$$

The characteristic equation $\left| z\mathbf{I} - \mathbf{Q}^{-1}(\mathbf{A} - \mathbf{K}_e\mathbf{C})\mathbf{Q} \right|$ becomes:

$$\begin{vmatrix} z & 0 & \cdots & 0 & a_n + \delta_n \\ -1 & z & \cdots & 0 & a_{n-1} + \delta_{n-1} \\ 0 & -1 & \cdots & 0 & a_{n-2} + \delta_{n-2} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & z + a_1 + \delta_1 \end{vmatrix} = 0$$

Using minors and cofactors for the determinant, the above equation can be rewritten as:

$$z^n + (a_1 + \delta_1)z^{n-1} + (a_2 + \delta_2)z^{n-2} + \cdots + (a_n + \delta_n) = 0 \qquad (5.35)$$

It is assumed that the desired characteristic equation for the error dynamics is given as:

$$(z - \mu_1)(z - \mu_2)\cdots(z - \mu_n) = z^n + \alpha_1 z^{n-1} + \alpha_2 z^{n-2} + \cdots + \alpha_n = 0 \qquad (5.36)$$

The desired eigenvalues μ_i or the locations of the poles determine how fast the observed state converges to the actual state of the plant. By comparing the coefficients of the equal powers of z in Equations (5.35) and (5.36), the following relation holds true:

$$\begin{bmatrix} \delta_n \\ \delta_{n-1} \\ \vdots \\ \delta_2 \\ \delta_1 \end{bmatrix} = \begin{bmatrix} \alpha_n - a_n \\ \alpha_{n-1} - a_{n-1} \\ \vdots \\ \alpha_2 - a_2 \\ \alpha_1 - a_1 \end{bmatrix} \qquad (5.37)$$

Substituting Equation (5.37) into Equation (5.33) gives:

$$\mathbf{K}_e = \left(\mathbf{W}\mathbf{N}^T\right)^{-1} \begin{bmatrix} \alpha_n - a_n \\ \alpha_{n-1} - a_{n-1} \\ \vdots \\ \alpha_2 - a_2 \\ \alpha_1 - a_1 \end{bmatrix} \qquad (5.37)$$

The block diagram of the full-order state observer with general form of feedback gain matrix $\mathbf{K}_e$ is shown in Figure 60.

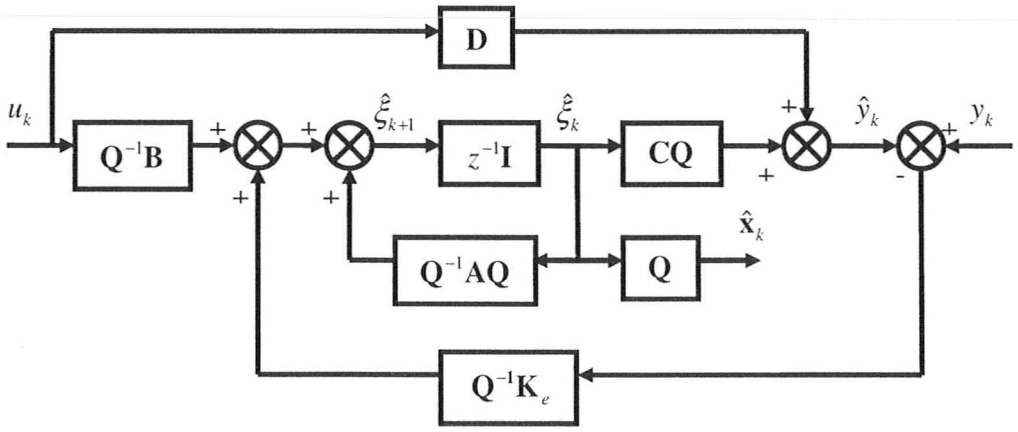

Figure 60 Full-Order State Observer with a General Form of Feedback Gain Matrix

The desired eigenvalues of the estimation error dynamics can be assigned in such a way that the observer responds at least four to five times faster than the closed-loop system (Ogata, 1994). In some applications, deadbeat response may be desired. Hence the characteristic equation becomes $z^n = 0$, which results in a feedback gain $\mathbf{K}_e$ as:

$$\mathbf{K}_e = \left(\mathbf{W}\mathbf{N}^T\right)^{-1}\begin{bmatrix} -a_n \\ -a_{n-1} \\ \vdots \\ -a_2 \\ -a_1 \end{bmatrix} \qquad (5.38)$$

For a low order system, the observer feedback gain matrix $\mathbf{K}_e$ can be determined by equating the coefficients of the same powers of z in the characteristic equation of Equation (5.22), $\left|z\mathbf{I} - \left(\mathbf{A} - \mathbf{K}_e\mathbf{C}\right)\right| = 0$, and those of the desired characteristic equation given by:

$$(z-\mu_1)(z-\mu_2)\cdots(z-\mu_n) = z^n + \alpha_1 z^{n-1} + \cdots + \alpha_{n-1}z + \alpha_n = 0$$

where $\mu_i, i = 1,\cdots,n$ are desired eigenvalues of the matrix $\left(\mathbf{A} - \mathbf{K}_e\mathbf{C}\right)$. For our belt drive system of Equation (5.9), $\mathbf{K}_e = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$, the characteristic equation, $\left|z\mathbf{I} - \left(\mathbf{A} - \mathbf{K}_e\mathbf{C}\right)\right| = 0$, can be obtained as:

$$z^2 + (19.2k_1 - 1.226)z - 23.54k_1 + 19.2k_2 + 0.5311 = 0 \qquad (5.39)$$

The desired characteristic equation can be assigned as:

$$z^2 - (\mu_1 + \mu_2)z + \mu_1\mu_2 = 0 \qquad (5.40)$$

By equating the coefficients of the like powers of z in Equations (5.39) and (5.40), the full-order prediction observer gain matrix can be obtained as:

$$k_1 = \frac{1.226-(\mu_1+\mu_2)}{19.2}$$
$$k_2 = \frac{\mu_1\mu_2-1.226(\mu_1+\mu_2)+0.972}{19.2} \qquad (5.41)$$

By setting $\mu_1 = 0.5+0.5i$, $\mu_2 = 0.5-0.5i$, the observer gain matrix is obtained as $\mathbf{K}_e = [0.0118 \quad 0.0128]^T$ and the performance of the full-order prediction observer is shown in Figure 61.

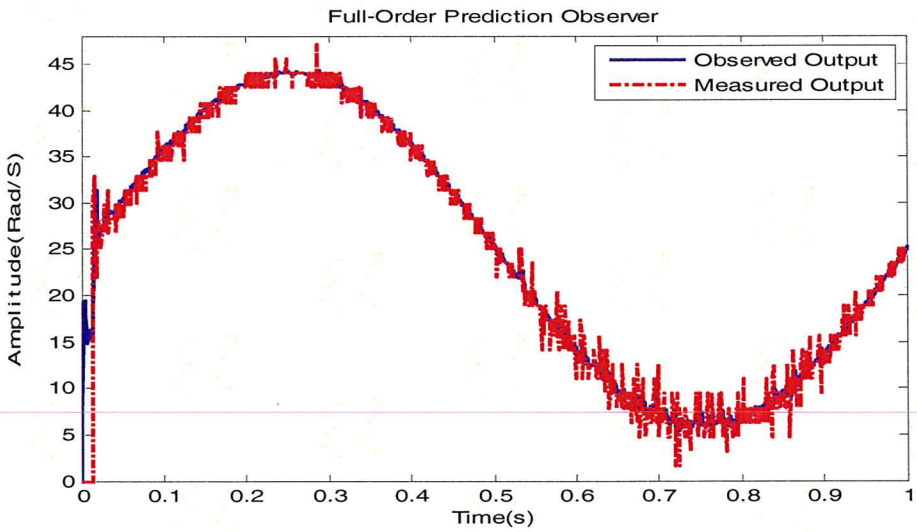

Figure 61 Performance of the Full-Order Prediction Observer

Current Observer

A current observer utilizes the current measured output y_k to estimate the state $\mathbf{x}_k$. The observation process consists of two steps: (1) determination of $\mathbf{z}_{k+1}$, an approximation

of state $\mathbf{x}_{k+1}$, based on the observed state $\hat{\mathbf{x}}_k$ and input u_k, (2) improving $\mathbf{z}_{k+1}$ using the measurement y_{k+1} (Ogata, 1994).

For the completely observable system given in Equation (5.9) as:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k$$
$$y_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}u_k$$

The current observer equations are given (Ogata, 1994) as:

$$\hat{\mathbf{x}}_{k+1} = \mathbf{z}_{k+1} + \mathbf{K}_e\left(y_{k+1} - \left(\mathbf{C}\mathbf{z}_{k+1} + \mathbf{D}u_{k+1}\right)\right) \qquad (5.42)$$

$$\mathbf{z}_{k+1} = \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}u_k \qquad (5.43)$$

Equation (5.43) predicts the $\mathbf{z}_{k+1}$ based on the current observed state $\hat{\mathbf{x}}_k$ and input u_k at time step k. Equation (5.42) improves the prediction $\mathbf{z}_{k+1}$ based on the measurement y_{k+1} to obtain the improved state $\hat{\mathbf{x}}_{k+1}$.

The estimation error dynamics can be obtained by subtracting the observer Equations (5.42) and (5.43) from the original system of Equation (5.9) as:

$$\mathbf{e}_{k+1} = \left(\mathbf{A} - \mathbf{K}_e\mathbf{C}\mathbf{A}\right)\mathbf{e}_k \qquad (5.44)$$

The observability matrix of Equation (5.44) becomes:

$$\begin{bmatrix} (\mathbf{CA})\mathbf{A} \\ (\mathbf{CA})\mathbf{A}^2 \\ \vdots \\ (\mathbf{CA})\mathbf{A}^{n-2} \\ (\mathbf{CA})\mathbf{A}^{n-1} \end{bmatrix} = \begin{bmatrix} \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{n-2} \\ \mathbf{CA}^{n-1} \end{bmatrix} \mathbf{A} \qquad (5.45)$$

To make it possible to arbitrarily place the eigenvalues of the matrix $\left(\mathbf{A} - \mathbf{K}_e\mathbf{C}\mathbf{A}\right)$, the observability matrix in Equation (5.45) has to have a rank of n. It is the product of the

observability matrix of the original system and the process matrix $\mathbf{A}$. It is assumed that the original system is completely observable; its observability matrix is full rank. The rank of matrix in Equation (5.45) solely depends on the rank of the process matrix $\mathbf{A}$. If the matrix $\mathbf{A}$ is full rank, the matrix (5.45) has a rank of n. The feedback matrix $\mathbf{K}_e$ can be determined using approaches presented in the previous section. For a non-singular matrix $\mathbf{A}$, the complete solution for determining the feedback gain matrix $\mathbf{K}_e$ can be found in (Ogata, 1994). The block diagram of the full-order current observer is shown in Figure 62.

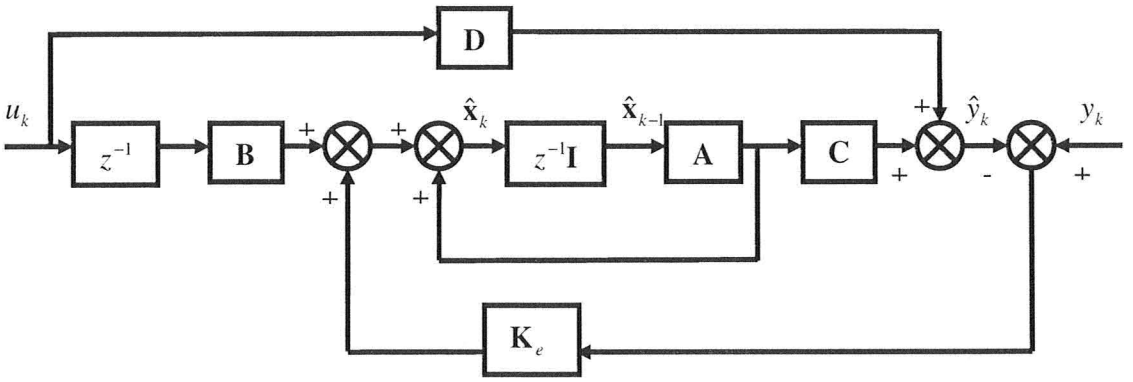

Figure 62 Block Diagram of the Full-Order Current Observer

For a low order system, the observer feedback gain matrix $\mathbf{K}_e$ can be determined by equating the coefficients of the same powers of z in the characteristic equation of Equation (5.44), $|z\mathbf{I}-(\mathbf{A}-\mathbf{K}_e\mathbf{C}\mathbf{A})|=0$, and those of the desired characteristic equation given by:

$$(z-\mu_1)(z-\mu_2)\cdots(z-\mu_n)=z^n+\alpha_1 z^{n-1}+\cdots+\alpha_{n-1}z+\alpha_n=0$$

where $\mu_i, i = 1, \cdots, n$ are desired eigenvalues of the matrix $(\mathbf{A} - \mathbf{K}_e\mathbf{CA})$. For our belt drive system of Equation (5.9), with $\mathbf{K}_e = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$, the characteristic equation, $|z\mathbf{I} - (\mathbf{A} - \mathbf{K}_e\mathbf{CA})| = 0$, can be obtained as:

$$z^2 + (19.2k_2 - 1.226)z - 10.2k_1 + 0.5311 = 0 \quad (5.46)$$

The desired characteristic equation can be assigned as:

$$z^2 - (\mu_1 + \mu_2)z + \mu_1\mu_2 = 0 \quad (5.47)$$

By equating the coefficients of the like powers of z in Equations (5.46) and (5.47), the observer gain matrix can be obtained as:

$$k_1 = \frac{0.5311 - \mu_1\mu_2}{10.2}$$
$$k_2 = \frac{1.226 - (\mu_1 + \mu_2)}{19.2} \quad (5.48)$$

By setting $\mu_1 = 0.5 + 0.5i$, $\mu_2 = 0.5 - 0.5i$, the full-order current observer gain matrix can be obtained as $\mathbf{K}_e = [0.003 \quad 0.0118]^T$, and the performance of the state observer is shown in Figure 63.

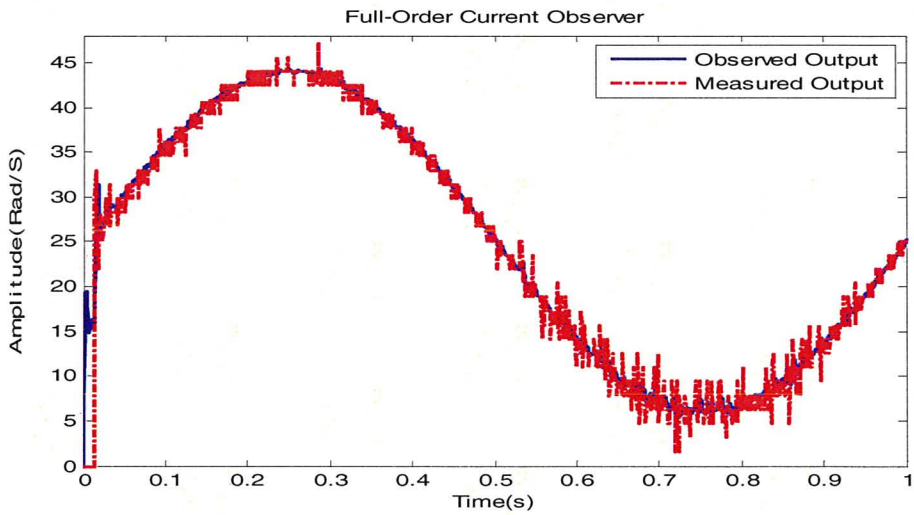

Figure 63 Performance of the Full-Order Current Observer

5.4.2 Minimum-Order State Observer

Full-order state observers are designed to reconstruct all state variables. In practice, some state variables may be accurately measured; there is no need for estimation of the measurable variables. An observer that estimates the number of state variables fewer than the dimension of the state is called reduced-order observer. A minimum-order observer is a reduced-order observer that has a minimum-order (Ogata, 1994). It is assumed that the state vector has the dimension of n and there are m accurately measured state variables. The minimum-order state observer is a $(n-m)^{th}$ -order observer. The block diagram of a DSMC system with a minimum-order state observer is shown in Figure 64. In the case of the experimental belt drive system, the measurement quality is poor, making the choices of the reduced-order and the minimum-order observers inappropriate.

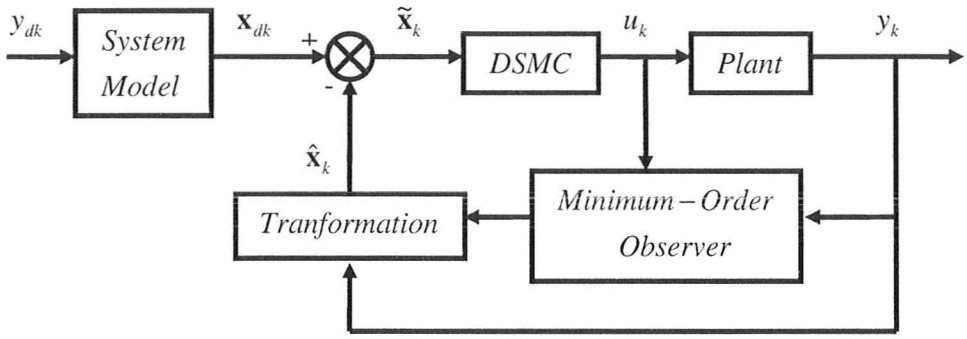

Figure 64 Minimum-Order State Observer

Effects of Observer on a Closed-Loop System

In practice, the true state may not be available for measurement. The estimated state $\hat{\mathbf{x}}_k$ is used in the control systems. The effects of using the estimated states in place of the real states in a control system are investigated with a state feedback control system as an example.

A completely controllable and observable system is given as:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k$$
$$y_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}u_k$$

The state feedback control based on the estimated state $\hat{\mathbf{x}}_k$ is given as:

$$u_k = -\mathbf{K}\hat{\mathbf{x}}_k$$

where $\mathbf{K}$ is the feedback control gain matrix. The dynamics of the closed-loop system is obtained from the above two equations as:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k - \mathbf{B}\mathbf{K}\hat{\mathbf{x}}_k \qquad (5.49)$$

With the estimation error $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$ and its dynamics given in Equation (5.22), Equation (5.49) can be rewritten as:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{e}_{k+1} \end{bmatrix} = \begin{bmatrix} \mathbf{A}-\mathbf{BK} & \mathbf{BK} \\ \mathbf{0} & \mathbf{A}-\mathbf{K}_e\mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{e}_k \end{bmatrix} \qquad (5.50)$$

The characteristic equation for the closed-loop system of Equation (5.50) is given as:

$$\begin{vmatrix} \mathbf{A}-\mathbf{BK} & \mathbf{BK} \\ \mathbf{0} & \mathbf{A}-\mathbf{K}_e\mathbf{C} \end{vmatrix} = 0 \text{ or}$$

$$|\mathbf{A}-\mathbf{BK}||\mathbf{A}-\mathbf{K}_e\mathbf{C}| = 0 \qquad (5.51)$$

The closed-loop poles of the observed state feedback control system consist of the poles of the feedback control design and the poles of the observer design. The control design and the observer design are independent of each other and can be considered separately, then combined to form an observed state feedback control system.

The closed-loop poles of the feedback control system are chosen in such a way that the control system demonstrates the desired performance. The poles of the observer are chosen so that the response of the observer is much faster than that of the system. A rule of thumb is the response of the observer is at four to ten times faster than the original system. The observer is programmed in the computer. Its response speed can be increased up to deadbeat response so that the observed state quickly converges to the real state. The only limiting factor on the response speed of the observer is sensitivity to noise and disturbances in the system.

5.5 The Discrete Time Kalman Filter

The state observers presented in the previous section work well in the absence of disturbance and noise. In practice, in the case of the belt drive system, the measurement signals are often contaminated by process and measurement noise. Although the feedback gain matrix $\mathbf{K}_e$ of the state observer can be properly chosen to trade-off speed of response for robustness to disturbance and noise, noise greatly degrades the observer performance. A digital filter can be used in conjunction with the observer to suppress the noise at the expense of introducing a phase shift that is highly undesirable in real-time control applications. The Kalman filter provides an alternative observation strategy (Haykin, 1996; Welch & Bishop, 2006).

The Kalman filter was invented by R. E. Kalman in 1960s (Welch & Bishop, 2006). A Kalman filter has two distinctive features: (1) the mathematical formulation of a Kalman filter is a predictor-corrector method described in terms of state space representation, (2) the solution of a Kalman filter is computed recursively. Each state estimate is calculated from the previous estimate and the current input data. This feature results in two advantages: (1) only the previous estimate is required, minimizing the storage requirements; (2) the Kalman filtering is computationally very efficient. The Kalman filter recursively estimates the state of a process contaminated with Gaussian noise by minimizing the mean of the squared error.

Consider the second-order discrete belt drive system given in Equation (5.9). The equations are augmented to account for noise as follows:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k + \mathbf{w}_k \\ y_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}u_k + v_k \end{aligned} \qquad (5.52)$$

where u_k is the scalar input, y_k is the scalar output, $w_k \in R^2$ is the process noise vector, v_k is the scalar measurement noise, $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -0.5311 & 1.226 \end{bmatrix}$ is the system matrix, $\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is the input matrix, $\mathbf{C} = \begin{bmatrix} 19.2 \\ 0 \end{bmatrix}^T$ is the output matrix and $\mathbf{D} = 0$ is the direct feed through matrix.

The noise is assumed to be uncorrelated, zero mean with normal probability distributions, i.e. $E(\mathbf{w}_k) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $E(\mathbf{w}_k \mathbf{w}_j^T) = \begin{cases} \mathbf{Q}_k & k = j \\ \mathbf{0} & k \neq j \end{cases}$, $E(v_k) = 0$ and $E(v_k v_j^T) = \begin{cases} R_k & k = j \\ 0 & k \neq j \end{cases}$. It is further assumed that the process noise covariance matrix $\mathbf{Q}$ and the measurement noise covariance matrix R are constant for each iteration step. $\mathbf{x}_k$ is an unknown 2×1 state vector to be estimated through output observation y_k and input u_k.

The Kalman filtering problem is stated as using the entire observed data consisting of $y_1, y_2, \cdots, y_k$ to find for each $j \geq 1$ the minimum mean squared error estimate of the states $\mathbf{x}_j$. The problem is called a filtering problem if $j = k$, a prediction problem if $j > k$ and a smoothing problem if $1 \leq j < n$ (Haykin, 1996).

The Kalman filter estimates a process by using a form of feedback control, that is, it estimates the process state at some time and then obtains feedback in the form of measurements. The Kalman filter equations fall into two groups: the time-update

equations and the measurement-update equations (Welch & Bishop, 2006). Each of these groups works separately in two individual steps in a single operation cycle: the prediction step and the correction step.

In the prediction step, the Kalman filter uses the time-update equations to predict the state of the process based on the previous state estimation and the process model. The state and the error covariance matrix obtained in this step are the *a priori* estimation for step k. In the correction step, the Kalman gain is first calculated. Then the measurement update equations are used to incorporate a new measurement into the *a priori* estimation to obtain an improved *a posteriori* estimation.

Let $\hat{\mathbf{x}}_{k|k-1}$ to be the *a priori* state estimate at step k given the knowledge of the process prior to step k, $\hat{\mathbf{x}}_{k|k}$ to be the *a posteriori* state estimate at step k given measurement y_k, $\mathbf{e}_{k|k-1}$ to be the *a priori* estimation error defined as:

$$\mathbf{e}_{k|k-1} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \tag{5.53}$$

Let $\mathbf{e}_{k|k}$ to be the *a posteriori* estimation error defined as:

$$\mathbf{e}_{k|k} = \mathbf{x}_k - \hat{\mathbf{x}}_{k|k} \tag{5.54}$$

The *a priori* and the *a posteriori* estimation error covariance matrices are obtained as:

$$\mathbf{P}_{k|k-1} = E\left[\mathbf{e}_{k|k-1}\mathbf{e}_{k|k-1}^T\right] \tag{5.55}$$

$$\mathbf{P}_{k|k} = E\left[\mathbf{e}_{k|k}\mathbf{e}_{k|k}^T\right] \tag{5.56}$$

The *a priori* state estimation is obtained from the process equation with the previous *a posteriori* state estimation and input as:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}u_{k-1} \qquad (5.57)$$

The *a posteriori* state estimation $\hat{\mathbf{x}}_{k|k}$ is computed as a linear combination of the *a priori* state estimation $\hat{\mathbf{x}}_{k|k-1}$ and a weighted difference between the actual measurement y_k and the measurement prediction $\hat{y}_{k|k-1}$ as:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left(y_k - \hat{y}_{k|k-1}\right) \qquad (5.58)$$

$$\hat{y}_{k|k-1} = \mathbf{C}\hat{\mathbf{x}}_{k|k-1} + \mathbf{D}u_k \qquad (5.59)$$

The difference $y_k - \hat{y}_{k|k-1}$ in Equation (5.58) is called the filtering *innovation* or *residual*, which reflects the discrepancy between the actual measurement y_k and its prediction value $\hat{y}_{k|k-1}$. The 2×1 matrix $\mathbf{K}_k$ is the Kalman gain aiming to minimize the *a posteriori* estimation error covariance matrix of Equation (5.56). Substituting Equations (5.59), (5.58), (5.53) and (5.52) into Equation (5.55), then the *a posteriori* error covariance matrix $\mathbf{P}_{k|k}$ becomes:

$$\begin{aligned}
\mathbf{P}_{k|k} &= E\left\{\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k\left(y_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1} - \mathbf{D}u_k\right)\right]\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k\left(y_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1} - \mathbf{D}u_k\right)\right]^T\right\} \\
&= E\left\{\begin{array}{l}\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k\left(\mathbf{C}\mathbf{x}_k + \mathbf{D}u_k + \mathbf{v}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1} - \mathbf{D}u_k\right)\right] \\ \times\left[\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k\left(\mathbf{C}\mathbf{x}_k + \mathbf{D}u_k + \mathbf{v}_k - \mathbf{C}\hat{\mathbf{x}}_{k|k-1} - \mathbf{D}u_k\right)\right]^T\end{array}\right\} \\
&= E\left\{\begin{array}{l}\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right) - \mathbf{K}_k\left(\mathbf{C}\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right) + \mathbf{v}_k\right)\right] \\ \times\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right) - \mathbf{K}_k\left(\mathbf{C}\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right) + \mathbf{v}_k\right)\right]^T\end{array}\right\} \\
&= E\left\{\left[\mathbf{e}_{k|k-1} - \mathbf{K}_k\left(\mathbf{C}\mathbf{e}_{k|k-1} + \mathbf{v}_k\right)\right]\left[\mathbf{e}_{k|k-1} - \mathbf{K}_k\left(\mathbf{C}\mathbf{e}_{k|k-1} + \mathbf{v}_k\right)\right]^T\right\} \\
&= E\left\{\left[\mathbf{e}_{k|k-1} - \mathbf{K}_k\mathbf{C}\mathbf{e}_{k|k-1} + \mathbf{K}_k\mathbf{v}_k\right]\left[\mathbf{e}_{k|k-1}^T - \mathbf{e}_{k|k-1}^T\mathbf{C}^T\mathbf{K}_k{}^T + \mathbf{v}_k{}^T\mathbf{K}_k{}^T\right]\right\} \\
&= \mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{C}^T\mathbf{K}_k{}^T - \mathbf{K}_k\mathbf{C}\mathbf{P}_{k|k-1} + \mathbf{K}_k(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R})\mathbf{K}_k{}^T
\end{aligned} \qquad (5.60)$$

The optimal Kalman gain $\mathbf{K}_k$, which minimizes the *a posteriori* error covariance matrix, can be obtained by taking the derivative of Equation (5.60) with respect to the Kalman gain $\mathbf{K}_k$ and setting the term to zero. This is equivalent to minimizing the trace of the *a posteriori* error covariance matrix, i.e. setting the trace derivative (Petersen & Pedersen, 2008) of Equation (5.60) to zero, noting that $\mathbf{P}_k$ and R are symmetric matrices, that is:

$$\begin{aligned}\frac{\partial}{\partial \mathbf{K}_k} Tr(\mathbf{P}_{k|k}) &= \frac{\partial}{\partial \mathbf{K}_k} Tr\left[\mathbf{P}_{k|k-1} - \mathbf{P}_{k|k-1}\mathbf{C}^T\mathbf{K}_k^T - \mathbf{K}_k\mathbf{C}\mathbf{P}_{k|k-1} + \mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + \mathbf{R}\right)\mathbf{K}_k^T\right] \\ &= -\mathbf{P}_{k|k-1}\mathbf{C}^T - \left(\mathbf{C}\mathbf{P}_{k|k-1}\right)^T + \left[\mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + R\right)^T + \mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + R\right)\right] \\ &= -2\mathbf{P}_{k|k-1}\mathbf{C}^T + 2\mathbf{K}_k\left(\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + R\right) \\ &= 0\end{aligned} \tag{5.61}$$

From Equation (5.61), the optimal Kalman gain is obtained as:

$$\mathbf{K}_k = \frac{\mathbf{P}_{k|k-1}\mathbf{C}^T}{\mathbf{C}\mathbf{P}_{k|k-1}\mathbf{C}^T + R} \tag{5.62}$$

Substituting Equation (5.62) into Equation (5.60) gives the *a posteriori* error covariance as:

$$\mathbf{P}_{k|k} = \left(\mathbf{I} - \mathbf{K}_k\mathbf{C}\right)\mathbf{P}_{k|k-1} \tag{5.63}$$

The *a priori* error covariance matrix is calculated from Equation (5.55) as:

$$\begin{aligned}\mathbf{P}_{k|k-1} &= E\left[\mathbf{e}_{k|k-1}\mathbf{e}_{k|k1}^T\right] = E\left[\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right)\left(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}\right)^T\right] \\ &= E\left[\left(\mathbf{x}_k - \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} - \mathbf{B}u_{k-1}\right)\left(\mathbf{x}_k - \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} - \mathbf{B}u_{k-1}\right)^T\right] \\ &= E\left[\left(\mathbf{A}\mathbf{e}_{k-1|k-1} + \mathbf{w}_{k-1}\right)\left(\mathbf{A}\mathbf{e}_{k-1|k-1} + \mathbf{w}_{k-1}\right)^T\right] \\ &= E\left(\mathbf{A}\mathbf{e}_{k-1|k-1}\mathbf{e}_{k-1|k-1}{}^T\mathbf{A}^T + \mathbf{A}\mathbf{e}_{k-1}\mathbf{w}_{k-1}^T + \mathbf{w}_k\mathbf{e}_{k-1}^T\mathbf{A}^T + \mathbf{w}_{k-1}\mathbf{w}_{k-1}^T\right) \\ &= \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^T + \mathbf{Q}\end{aligned} \tag{5.64}$$

The overall Kalman filtering operation can be viewed as an iterative prediction-correction process as shown in Figure 65. After each time-update and measurement-update cycle, the filtering process is repeated by using the process model with the current *a posteriori* estimate and input to predict the new *a priori* estimate. This recursive feature makes the Kalman filter computationally efficient.

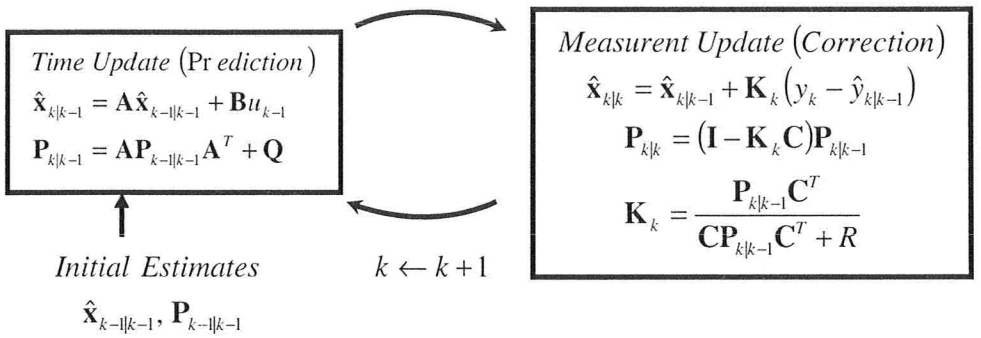

Figure 65 The Ongoing Kalman Filter Cycle

The Kalman Filter Parameters Tuning

To implement the Kalman filter algorithm, the initial conditions such as the initial estimate, including the state $\hat{\mathbf{x}}_{0|0}$ and the error covariance matrix $\mathbf{P}_{0|0}$, the process noise covariance matrix $\mathbf{Q}$ and the measurement noise covariance R are needed. However, the initial conditions of the process may not be known precisely.

In the absence of any observed data, the initial state may be chosen from its mean value which produces an unbiased estimate (Haykin, 1996). The initial estimation error covariance matrix $\mathbf{P}_{0|0}$ can be chosen as $C\mathbf{I}$ where C is a large constant (C should not be too large as it can cause premature convergence with the Kalman gain going and remaining at zero) and $\mathbf{I}$ is an 2×2 identity matrix.

The measurement noise covariance R is usually measured prior to the operation of the filter. Some off-line measurements, for example, a constant input can be taken to determine the variance of the measurement noise.

The determination of the process noise covariance matrix $\mathbf{Q}$ is more difficult since it can not be observed directly. If the process measurement is reliable, selection of matrix $\mathbf{Q}$ with enough uncertainty can produce acceptable results even with a poor model (Welch & Bishop, 2006).

Superior filter performance can be achieved by tuning the filter parameter $\mathbf{Q}$ and R. The tuning is usually carried out off-line. The filter innovation used to correct the *a priori* estimate also provides the starting point for checking the filter operation. A necessary and sufficient condition for a Kalman filter to be optimal is that the residual is zero mean and white. The Kalman gain of Equation (5.62) with the *a priori* error covariance matrix of Equation (5.64) provide guidance to the tuning of the Kalman filter parameters $\mathbf{Q}$ and R.

In conditions where $\mathbf{Q}$ and R are constant, the estimation covariance and the Kalman gain will stabilize quickly and then remain constant (Welch & Bishop, 2006). If the noise covariance matrix $\mathbf{Q}$ is too large, the Kalman gain $\mathbf{K}_k$ will be too large, and as a result the estimate has a tendency to follow the measurements "too much", and the estimate will bounce a lot. R has an opposite effect to that of $\mathbf{Q}$, that is, with a larger R the estimate will be more dependent on the model. The parameters $\mathbf{Q}$ and R can be adjusted to make the filter tighter or more relaxed. A tighter filter is resulted from a large R and small $\mathbf{Q}$,

and the estimated state is weighted more to the model. On the other hand, a relaxed filter with a small R and large $\mathbf{Q}$ will more closely follow the measurement (Cadet, 2010).

For our belt drive system, the performance of the Kalman filter with the covariance matrices tuned, $\mathbf{Q} = \begin{bmatrix} 0.0001 & 0 \\ 0 & 0.0001 \end{bmatrix}$ and $R = 2$, is shown in Figure 66.

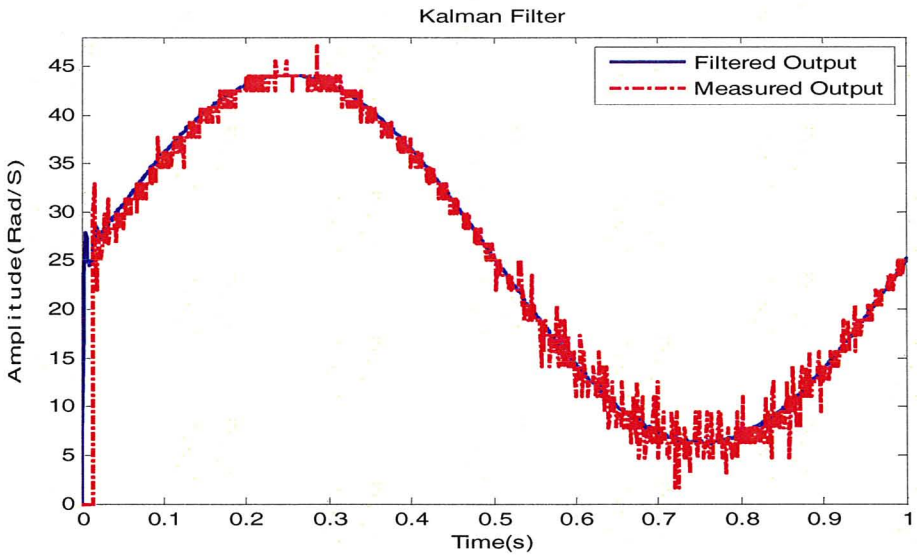

Figure 66 The Performance of the Kalman Filter

5.6 Design of a DSMC Tracking Controller

Consider the discrete-time state space model of the belt drive system given in Equation (5.52)(162) and written as follows:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k + \mathbf{w}_k \\ y_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}u_k + v_k \end{aligned} \qquad (5.65)$$

where w_k is 2×1 process noise vector, v_k is a scalar measurement noise, $\mathbf{A}=\begin{bmatrix} 0 & 1 \\ -0.5311 & 1.226 \end{bmatrix}$, $\mathbf{B}=\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{C}=\begin{bmatrix} 19.2 \\ 0 \end{bmatrix}^T$, $\mathbf{D}=0$. As stated earlier, the system is completely controllable and completely observable. A DSMC is designed by using Misawa's DSMC strategy [Misawa1997a] as follows. The design is carried out in two steps: (1) a switching function is designed in such a way that the dynamics of the system demonstrates a desired performance; (2) a DSMC law is designed that forces the state trajectory of the system to reach the switching hyperplane defined by the switching function and slide on it thereafter.

5.6.1 Design of Switching Function

The system equation in (5.65) is the controllable canonical form. An advantage of the controllable canonical form is that it is convenient for the design of a switching function. In the controllable canonical form, a state space model can be partitioned so that the control input is disconnected from some of the state variables. These state variables (independent of the control input u_k) form an unforced subsystem. A switching function is then easily designed for the unforced system by using the pole-assignment technique.

For stabilization of the nominal system of Equation (5.65) a switching function is defined as:

$$s_k = \mathbf{G}\mathbf{x}_k = \begin{bmatrix} g_1 & g_2 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} = 0 \qquad (5.66)$$

where $\mathbf{G} = [g_1 \quad g_2]$ is the switching vector. By setting $g_1 = 1$, Equation (5.66) can be rewritten as:

$$x_{2k} = -\frac{g_1}{g_2} x_{1.k} = -\frac{x_{1k}}{g_2} \tag{5.67}$$

The nominal system of Equation (5.65) can be partitioned as:

$$\begin{bmatrix} [x_{1k+1}] \\ [x_{2k+1}] \end{bmatrix} = \begin{bmatrix} [0] & [1] \\ [-0.5311] & [1.226] \end{bmatrix} \begin{bmatrix} [x_{1k}] \\ [x_{2k}] \end{bmatrix} + \begin{bmatrix} [0] \\ [1] \end{bmatrix} u_k \tag{5.68}$$

Using of the internal square brackets in Equation (5.68), although unnecessary for the single-input second-order system, is to reflect that the individual elements may become proper dimensional vectors or matrices for multiple-input higher order systems. Equation (5.68) is decomposed as:

$$\begin{aligned} x_{1k+1} &= x_{2k} \\ x_{2k+1} &= -0.5311 x_{1k} + 1.226 x_{2k} + u_k \end{aligned} \tag{5.69}$$

Substituting Equation (5.67) into Equation (5.69) gives:

$$x_{1k+1} = -\frac{x_{1k}}{g_2} \tag{5.70}$$

Since the pair $(\mathbf{A}, \mathbf{B})$ is completely controllable, the closed-loop system of Equation (5.70) is completely controllable also and its pole can be placed at a desired location by means of state feedback through an appropriate constant g_2. If g_2 satisfies the condition $0 < -\frac{1}{g_2} < 1$, that is $g_2 < -1$, the system of Equation (5.70) is stable. There are many hyperplanes available based on different choices of the constant g_2. The hyperplanes guaranteeing the stability of the belt system are straight lines crossing the

origin with slopes between 0 and 1 as shown in Figure 67. By setting $g_2 = -2$, the switching vector becomes $\mathbf{G} = [1 \quad -2]$ and the switching hyperplane becomes $x_2 = 0.5x_1$.

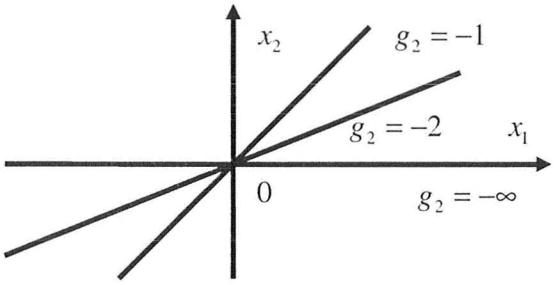

Figure 67 Hyperplanes Guaranteeing the Stability of the Belt Drive System

5.6.2 The Equivalent Dynamics of the Belt Drive System

The equivalent dynamics can be obtained by combining the switching function with the model of the nominal system. Note that the matrix $(\mathbf{GB})$ is non-singular. The equivalent control law (Utkin, 1977) of the belt drive system is obtained as:

$$\begin{aligned} s_{k+1} &= \mathbf{G}\mathbf{x}_{k+1} = 0 \\ \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}u_k \end{aligned} \Rightarrow u_{eqk} = -(\mathbf{GB})^{-1}\mathbf{GA}\mathbf{x}_k \qquad (5.71)$$

Substituting Equation (5.71) into the nominal system gives the equivalent dynamics of the belt drive system as:

$$\mathbf{x}_{k+1} = \left(\mathbf{A} - \mathbf{B}(\mathbf{GB})^{-1}\mathbf{GA}\right)\mathbf{x}_k = \begin{bmatrix} 0 & 1 \\ 0 & 0.5 \end{bmatrix}\mathbf{x}_k \qquad (5.72)$$

The closed-loop poles of Equation (5.72) are obtained by solving the characteristic equation $\left|z\mathbf{I} - \begin{bmatrix} 0 & 1 \\ 0 & 0.5 \end{bmatrix}\right| = 0$ as $z_1 = 0$ and $z_2 = 0.5$.

5.6.3 Design of the Tracking Control Law

For the tracking control of the system in Equation (5.65), the switching function of Equation (5.66) for the stabilization is modified as:

$$s_k = \mathbf{G}\tilde{\mathbf{x}}_k = 0 \qquad (5.73)$$

where $\tilde{\mathbf{x}}_k = \mathbf{x}_{d.k} - \mathbf{x}_k = \begin{bmatrix} x_{1dk} - x_{1k} \\ x_{2dk} - x_{2k} \end{bmatrix} = \begin{bmatrix} \tilde{x}_{1k} \\ \tilde{x}_{2k} \end{bmatrix}$ is the state tracking error. The switching vector $\mathbf{G}$ designed for the stabilization problem is still applicable to the tracking problem at hand. The switching hyperplane becomes a function of the state tracking error variables $\tilde{x}_{1k}$ and x_{2k}.

A DSMC law is designed to ensure the switching function is satisfied, so that the system state $\mathbf{x}_k$ reaches the desired state $\mathbf{x}_{dk}$ in a finite time and slide on it thereafter. The tracking error $\tilde{\mathbf{x}}_k$ is driven to zero and kept as small as possible.

Following the procedures in (Misawa, 1997; Misawa, 1997a), a switching surface is redefined as:

$$\Sigma = \left\{ \tilde{\mathbf{x}}_k \middle| s_k = \mathbf{G}\tilde{\mathbf{x}}_k = 0 \right\} \qquad (5.74)$$

A boundary layer in the state tracking error space neighbouring the switching surface of Equation (5.74) is defined as:

$$\Psi = \left\{ \tilde{\mathbf{x}}_k \middle| |s_k| = |\mathbf{G}\tilde{\mathbf{x}}_k| \le \varphi \right\} \qquad (5.75)$$

where φ is the boundary layer thickness.

The DSMC law can be derived by using discrete-time Lyapunov stability theory with a Lyapunov function candidate defined as $V_k = s_k^2$. To ensure the stability, the Lyapunov function must satisfy $V_{k+1} < V_k$, i.e. (Misawa, 1997):

$$s_{k+1}^2 < s_k^2 \qquad (5.76)$$

Expression (5.76) can be rewritten as:

$$\Delta s_k^2 < -2 s_k \Delta s_k \qquad (5.77)$$

where $\Delta s_k = s_{k+1} - s_k$. Expression (5.77) implies:

$$\begin{cases} -2s_k < \Delta s_k < 0, & s_k > 0 \\ 0 < \Delta s_k < -2s_k, & s_k < 0 \end{cases} \qquad (5.78)$$

Further to Equations (5.65) and (5.73), Δs_k is calculated as:

$$\begin{aligned} \Delta s_k &= s_{k+1} - s_k = \mathbf{G}\tilde{\mathbf{x}}_{k+1} - s_k = \mathbf{G}(\mathbf{x}_{dk+1} - \mathbf{x}_{k+1}) - s_k \\ &= \mathbf{G}\mathbf{x}_{dk+1} - \mathbf{G}\mathbf{A}\mathbf{x}_k - \mathbf{G}\mathbf{B}u_k - \mathbf{G}\mathbf{w}_k - s_k \end{aligned} \qquad (5.79)$$

Substituting Equation (5.79) into Expression (5.78) gives:

$$\begin{cases} -s_k < \mathbf{G}\mathbf{x}_{dk+1} - \mathbf{G}\mathbf{A}\mathbf{x}_k - \mathbf{G}\mathbf{B}u_k - \mathbf{G}\mathbf{w}_k < s_k, & s_k > 0 \\ s_k < \mathbf{G}\mathbf{x}_{dk+1} - \mathbf{G}\mathbf{A}\mathbf{x}_k - \mathbf{G}\mathbf{B}u_k - \mathbf{G}\mathbf{w}_k < -s_k, & s_k < 0 \end{cases} \qquad (5.80)$$

Since $(\mathbf{GB})^{-1} = \left(\begin{bmatrix} 1 & -2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)^{-1} = -0.5 < 0$, Expression (5.80) can be rewritten as:

$$\begin{cases} (\mathbf{GB})^{-1}(\mathbf{G}\mathbf{x}_{dk+1} - \mathbf{G}\mathbf{A}\mathbf{x}_k - \mathbf{G}\mathbf{w}_k + s_k) < u_k \\ \qquad < (\mathbf{GB})^{-1}(\mathbf{G}\mathbf{x}_{dk+1} - \mathbf{G}\mathbf{A}\mathbf{x}_k - \mathbf{G}\mathbf{w}_k - s_k), & s_k > 0 \\ (\mathbf{GB})^{-1}(\mathbf{G}\mathbf{x}_{dk+1} - \mathbf{G}\mathbf{A}\mathbf{x}_k - \mathbf{G}\mathbf{w}_k - s_k) < u_k \\ \qquad < (\mathbf{GB})^{-1}(\mathbf{G}\mathbf{x}_{dk+1} - \mathbf{G}\mathbf{A}\mathbf{x}_k - \mathbf{G}\mathbf{w}_k + s_k), & s_k < 0 \end{cases} \qquad (5.81)$$

The control law u_k can be defined as the combination of a continuous part and a switching part as:

$$u_k = \overline{u}_k + \widetilde{u}_k \qquad (5.82)$$

where the continuous part $\overline{u}_k$ is model based and defined as:

$$\overline{u}_k = (\mathbf{GB})^{-1}(\mathbf{G}\mathbf{x}_{dk+1} - \mathbf{GA}\mathbf{x}_k) \qquad (5.83)$$

The reaching condition of Expression (5.81) can be modified as:

$$(\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k + s_k) < \widetilde{u}_k < (\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k - s_k),\ s_k > 0 \qquad (5.84)$$

$$(\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k - s_k) < \widetilde{u}_k < (\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k + s_k),\ s_k < 0 \qquad (5.85)$$

The switching control law $\widetilde{u}_k$ has to be chosen such that asymptotic stability is guaranteed in the presence of uncertainties, that is, Expressions (5.84) and (5.85) both have to be satisfied to ensure the stability of the uncertain system of Equation (5.75). Using right half of Expression (5.84) and left half of Expression (5.85), the switching control law can be chosen as (Misawa, 1997a):

$$\widetilde{u}_k = (\mathbf{GB})^{-1}\left[-s_k + Ksat\left(\frac{s_k}{\varphi}\right)\right] \qquad (5.86)$$

$$sat\left(\frac{s_k}{\varphi}\right) = \begin{cases} +1, & \frac{s_k}{\varphi} > 1 \\ \frac{s_k}{\varphi}, & \left|\frac{s_k}{\varphi}\right| \le 1 \\ -1, & \frac{s_k}{\varphi} < -1 \end{cases} \qquad (5.87)$$

where $K = \gamma + 2T\varepsilon$ is the controller gain, φ is the boundary layer thickness referred to Expression (5.75) and is further defined by $\varphi \ge \gamma + T\varepsilon$, ε is a positive constant, T is the

sampling time and γ is positive constant satisfying $|\mathbf{G}\mathbf{w}_k| < \gamma$. In the case of the belt drive system, the sampling time $T = 0.001$ second, the constant ε is selected as $\varepsilon = 10\gamma$, $\varphi = \gamma + T\varepsilon$ and γ is tuned on-line.

With the switching control law of Equation (5.86), the boundary layer of Equation (5.75) is attractive. The attractiveness of the boundary layer can be verified in the following three cases:

Case1. $s_k > \varphi$, *the value of the switching function is positive and is located outside of the boundary layer. Expression (5.84) should be satisfied.*

Verification: Since $s_k > \varphi$, Equation (5.86) can be rewritten as:

$$\tilde{u}_k = -(\mathbf{GB})^{-1} s_k + (\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) \qquad (5.88)$$

With Equation (5.88), $|\mathbf{G}\mathbf{w}_k| < \gamma$ and $(\mathbf{GB})^{-1} < 0$, so:

$$\begin{aligned} |\mathbf{G}\mathbf{w}_k| < \gamma &\Rightarrow -\gamma < -\mathbf{G}\mathbf{w}_k < \gamma \\ &\Rightarrow (\mathbf{GB})^{-1}\gamma < -(\mathbf{GB})^{-1}\mathbf{G}\mathbf{w}_k < -(\mathbf{GB})^{-1}\gamma \\ &\Rightarrow -(\mathbf{GB})^{-1} s_k + (\mathbf{GB})^{-1}\gamma < -(\mathbf{GB})^{-1} s_k - (\mathbf{GB})^{-1}\mathbf{G}\mathbf{w}_k \\ &\Rightarrow -(\mathbf{GB})^{-1} s_k + (\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) < -(\mathbf{GB})^{-1} s_k - (\mathbf{GB})^{-1}\mathbf{G}\mathbf{w}_k \\ &\Rightarrow \tilde{u}_k < (\mathbf{GB})^{-1}(-s_k - \mathbf{G}\mathbf{w}_k) \end{aligned} \qquad (5.89)$$

From Expression (5.81), the right half of Expression (5.84) is satisfied. Since $s_k > \varphi \geq \gamma + T\varepsilon$, the left hand side of Expression (5.84) can be derived as:

$$\begin{aligned} (\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k + s_k) < -(\mathbf{GB})^{-1}\gamma + (\mathbf{GB})^{-1} s_k < -(\mathbf{GB})^{-1}\gamma + (\mathbf{GB})^{-1}(\gamma + T\varepsilon) \\ \Rightarrow (\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k + s_k) < (\mathbf{GB})^{-1}(T\varepsilon) \end{aligned} \qquad (5.90)$$

From Equation (5.88) and $s_k > \varphi \geq \gamma + T\varepsilon$:

$$\begin{aligned}\tilde{u}_k &= -(\mathbf{GB})^{-1}s_k + (\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) > -(\mathbf{GB})^{-1}(\gamma + T\varepsilon) + (\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) \\ &= (\mathbf{GB})^{-1}(T\varepsilon) \\ &\Rightarrow \tilde{u}_k > (\mathbf{GB})^{-1}(T\varepsilon)\end{aligned} \tag{5.91}$$

Combining Expressions (5.90) and (5.91) gives:

$$\tilde{u}_k > (\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k + s_k) \tag{5.92}$$

From Expression (5.92), the left half of Expression (5.84) is satisfied. With Expressions (5.89) and (5.92), the switching control law satisfy the condition (5.84).

Case2. $s_k < -\varphi$, *the value of the switching function is negative and is located outside of the boundary layer. Expression (5.85) should be satisfied.*

Verification: Since $s_k < -\varphi$, Equation (5.86) can be rewritten as:

$$\begin{aligned}\tilde{u}_k &= -(\mathbf{GB})^{-1}s_k - (\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) \\ &< -(\mathbf{GB})^{-1}(-\varphi) - (\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) \\ &\le -(\mathbf{GB})^{-1}[-(\gamma + T\varepsilon)] - (\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) = -(\mathbf{GB})^{-1}T\varepsilon \Rightarrow \\ \tilde{u}_k &< -(\mathbf{GB})^{-1}T\varepsilon\end{aligned} \tag{5.93}$$

The right hand side of Expression (5.85) can be rewritten as:

$$\begin{aligned}(\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k + s_k) &> (\mathbf{GB})^{-1}\gamma + (\mathbf{GB})^{-1}s_k \\ &> (\mathbf{GB})^{-1}\gamma + (\mathbf{GB})^{-1}(-\varphi) \\ &\ge (\mathbf{GB})^{-1}\gamma + (\mathbf{GB})^{-1}[-(\gamma + T\varepsilon)] \\ &= -(\mathbf{GB})^{-1}T\varepsilon\end{aligned} \tag{5.94}$$

Combining Expressions (5.93) and (5.94) gives:

$$\tilde{u}_k < (\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k + s_k) \tag{5.95}$$

The right half of Expression (5.85) is satisfied. The left half of Expression (5.85) can be rewritten as:

$$\begin{aligned}(\mathbf{GB})^{-1}(-\mathbf{G}\mathbf{w}_k - s_k) &< -(\mathbf{GB})^{-1}\gamma - (\mathbf{GB})^{-1}s_k \\ &< -(\mathbf{GB})^{-1}(\gamma + 2T\varepsilon) - (\mathbf{GB})^{-1}s_k = \tilde{u}_k\end{aligned} \tag{5.96}$$

The left half of Expression (5.85) is satisfied. With Expressions (5.95) and (5.96), the condition (5.85) is satisfied.

Case3. $-\varphi \le s_k \le \varphi$, *the value of the switching function is located inside the boundary layer.*

For the convenience of discussion, it is assumed, without loss of generality, that $\varphi = \gamma + T\varepsilon$. Since $-\varphi \le s_k \le \varphi$, the saturation function in Equation (5.86) can be replaced by a linear function and the switching function is obtained as:

$$\tilde{u}_k = -(\mathbf{GB})^{-1}s_k + (\mathbf{GB})^{-1}\frac{(\gamma + 2T\varepsilon)}{\gamma + T\varepsilon}s_k \tag{5.97}$$

Substituting Equation (5.97), (5.83) and (5.82) into Equations (5.65) and (5.66) gives the s-function inside the boundary:

$$s_{k+1} = \frac{T\varepsilon}{\gamma + T\varepsilon}s_k - \mathbf{G}\mathbf{w}_k \tag{5.98}$$

The parameters ε and γ may be chosen such that there a low pass filtering through the s-function (Misawa, 1997a). So long as $\varepsilon > 0$ and $\gamma > |\mathbf{G}\mathbf{w}_k|$, the parameters ε and γ can be regarded as the design parameters. In the absence of modeling errors, the s-function asymptotically approaches zero inside the boundary layer. The convergence rate may be arbitrarily chosen by the selection of the ratio ε / γ and can be made arbitrarily fast at the expense of the boundary layer thickness.

The control law can be obtained by substituting Equations (5.86) and (5.83) into (5.82) as:

$$u_k = (\mathbf{GB})^{-1}(\mathbf{G}\mathbf{x}_{dk+1} - \mathbf{GA}\mathbf{x}_k) - (\mathbf{GB})^{-1} s_k + (\mathbf{GB})^{-1} Ksat\left(\frac{s_k}{\varphi}\right) \qquad (5.99)$$

This control law guarantees that the state of the system converges to the desired state with a predefined convergence rate.

5.7 Experiment Results

Based on the model of the plant to be controlled, a discrete-time PID controller and a DSMC controller have been designed in the previous subsections. The states of the plant are estimated through either a state observer or a Kalman filter. The system model is obtained through system identification with only one driven pulley. Modeling uncertainties are later introduced by adding extra load, including two pulleys and one eccentric flywheel, on the driven shaft. The PID and the DSMC are both applied. For comparison, the PID control is implemented on the system with and without modeling uncertainties. The output is filtered with a first-order Butterworth filter before feeding it back into the PID controller. The order of the filter is kept low to reduce the undesirable phase shift. The DSMC is implemented with the system states estimated using the state observer presented in Section 5.4 and the Kalman filter presented in Section 5.5, respectively. Both the filter and the state observer were designed and tested off-line using data generated from the plant. The Kalman filter was tuned off-line by adjusting the process noise covariance matrix $\mathbf{Q}$ and measurement noise covariance R . The

experimental results are evaluated by using the Root-Mean-Square of the tracking Error (RMSE), the Root-Mean-Square of the Estimation Error (RMSEE), the Root-Mean-Square of the control Error (RMSCE), Root-Mean-Square of the control signal (RMSU), the Maximum-Absolute value of the tracking Error (MAE), and the Maximum-Absolute value of the control signal (MAU) to demonstrate the effectiveness and robustness of the DSMC algorithm.

The specifications of the tracking performance of the controllers are defined as follows. With N denoting the number of data points taken in the experiment, r_i denoting the i^{th} point of the reference output, y_i denoting the i^{th} measured output, $\hat{y}_i$ is denoting the i^{th} estimated or filtered output, e_i denoting the i^{th} point of the tracking error and u_i denoting the i^{th} point of the control signal, the Mean-Square-Root of the tracking Error (RMSE) is defined as:

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(r_i - y_i)^2} \qquad (5.100)$$

The Mean-Square-Root of the Estimation Error (RMSEE) is defined as:

$$RMSEE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (5.101)$$

The Mean-Square-Root of the Control Error is defined as:

$$RMSCE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(r_i - \hat{y}_i)^2} \qquad (5.102)$$

The Mean-Square-Root of the control signal (RMSU) is defined as:

$$RMSU = \frac{1}{N}\sqrt{\sum_{i=1}^{N} u_i^2} \qquad (5.103)$$

The Maximum-Absolute-value of the tracking Error (MAE) is defined as:

$$MAE = \max(abs(e)) \qquad (5.104)$$

The Maximum-Absolute-value of the control signal (MAU) is defined as:

$$MAU = \max(abs(u)) \qquad (5.105)$$

The outputs of the DSMC with state observer of the belt drive system without model uncertainty are shown in Figure 68. Tracking error of the DSMC with state observer without model uncertainty is shown in Figure 69.

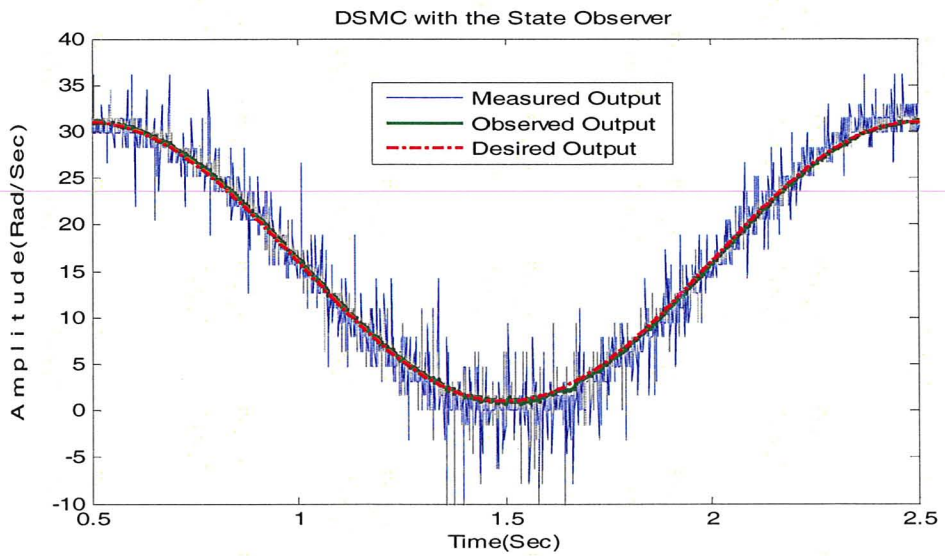

Figure 68 DSMC with State Observer of the Belt Drive System without Model Uncertainty

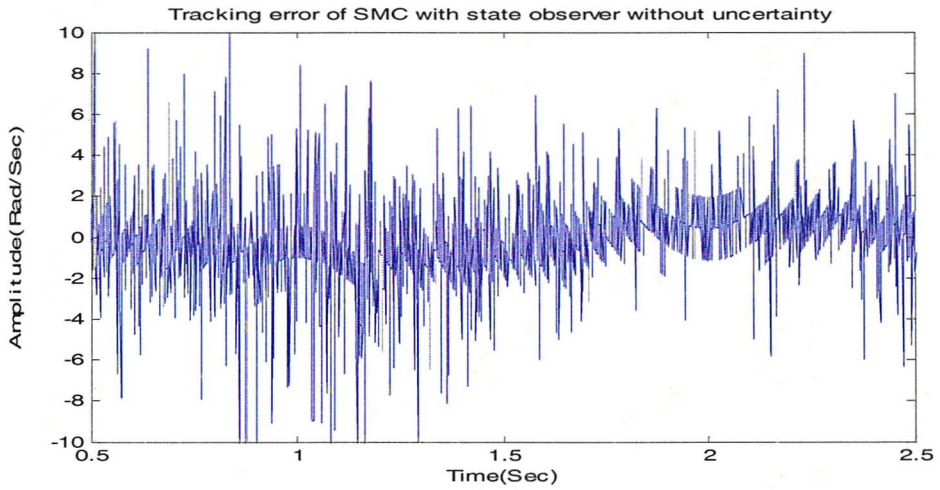

Figure 69 Tracking Error of the DSMC with the State Observer without Model Uncertainty

The outputs of the DSMC with state observer of the belt drive system with model uncertainty are shown in Figure 70. Tracking error of the DSMC with state observer with model uncertainty is shown in Figure 71.

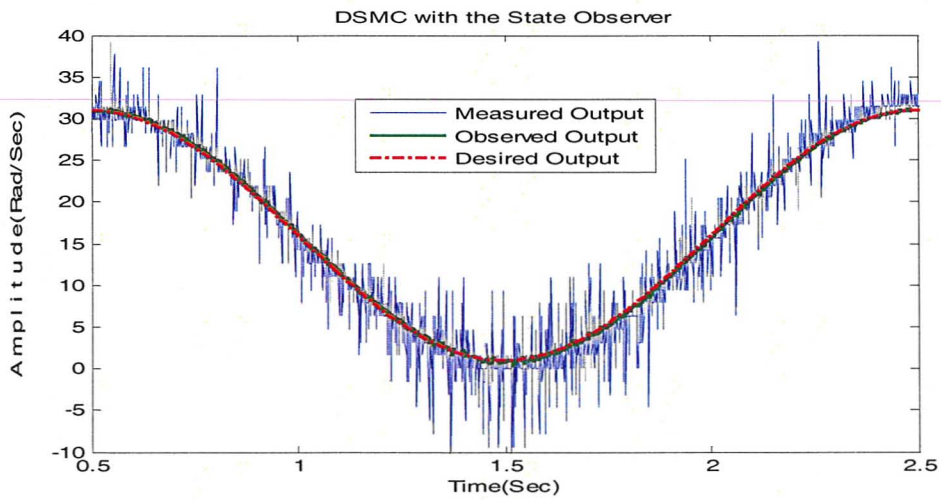

Figure 70 DSMC with the State Observer of the Belt drive System with Model Uncertainty

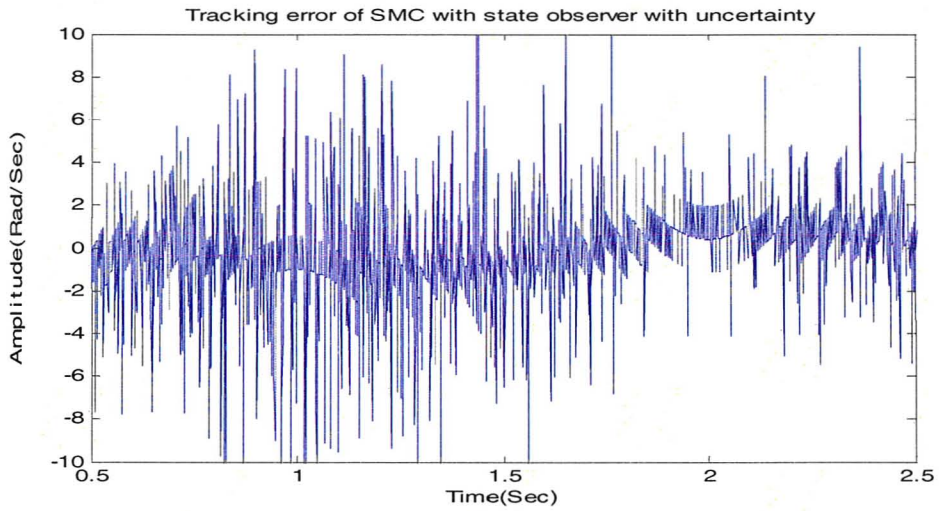

Figure 71 Tracking Error of the DSMC with State Observer of the Belt Drive System with Uncertainty

The outputs of the DSMC with the Kalman filter of the belt drive system without model uncertainty are shown in Figure 72. Tracking error of the DSMC with the Kalman filter of the belt drive system without model uncertainty is shown in Figure 73.

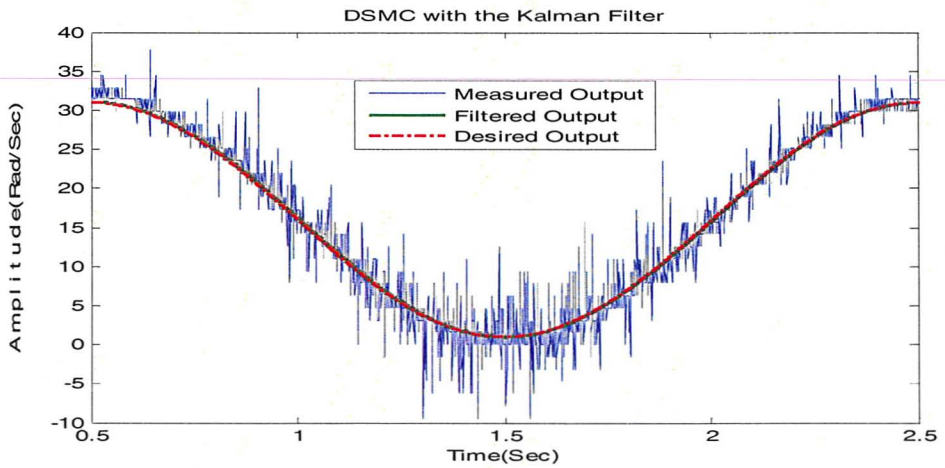

Figure 72 DSMC with the Kalman Filter of the Belt Drive System without Model Uncertainty

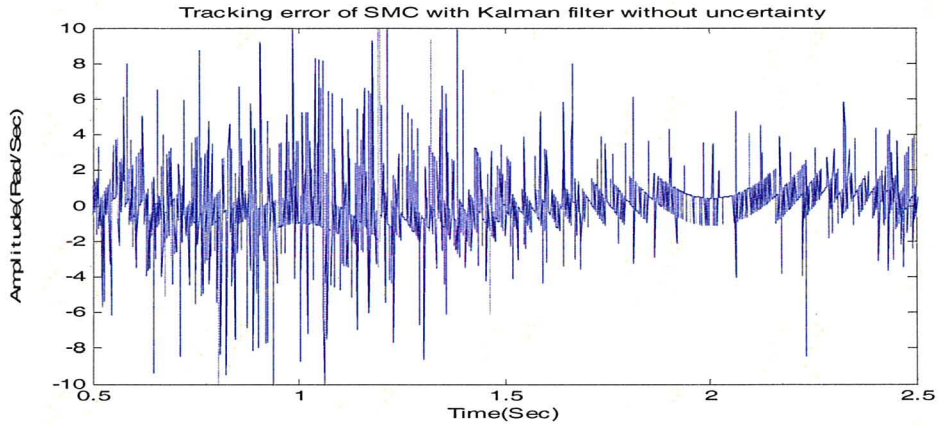

Figure 73 Tracking Error of the DSMC with the Kalman Filter of the Belt Drive System without Model Uncertainty

The outputs of the DSMC with the Kalman filter of the belt drive system with model uncertainty are shown in Figure 74. Tracking error of the DSMC with the Kalman filter of the belt drive system with model uncertainty is shown in Figure 75.

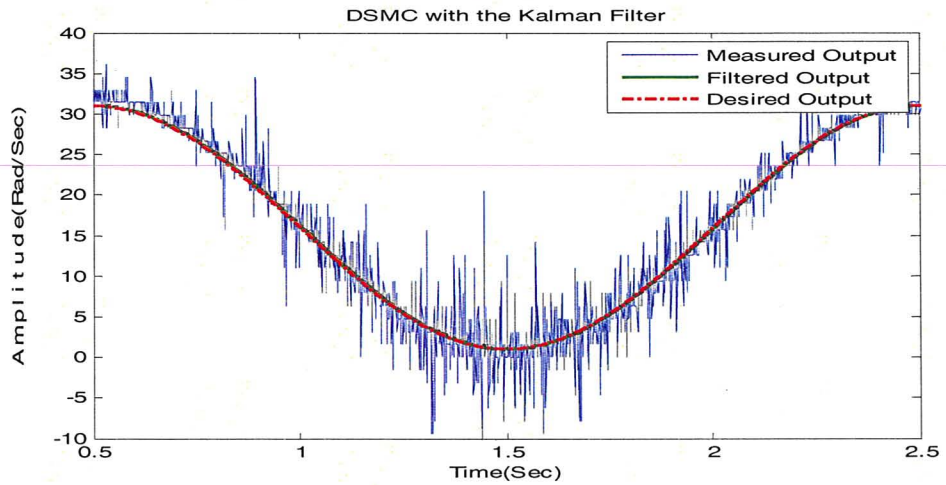

Figure 74 DSMC with the Kalman Filter of the Belt Drive System with Model Uncertainty

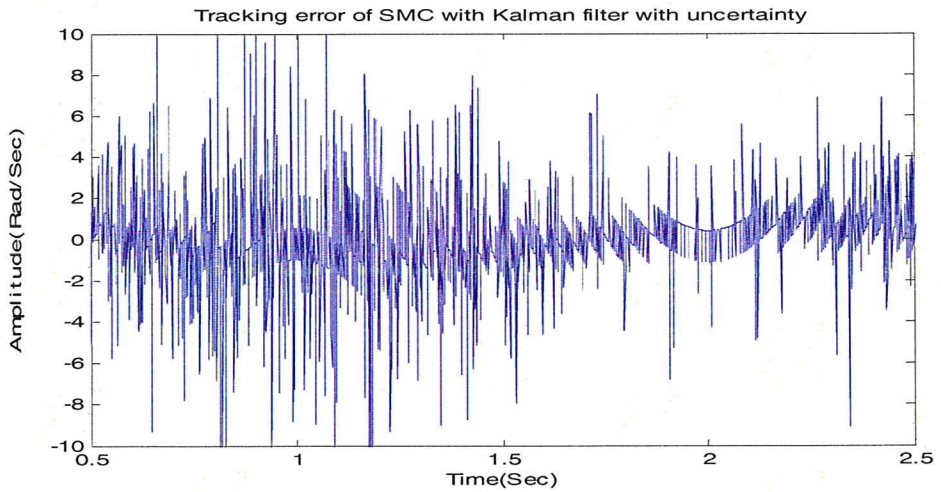

Figure 75 Tracking Error of the DSMC with the Kalman Filter of the Belt Drive System with Model Uncertainty

The variation of the Kalman gain is shown in Figure 76.

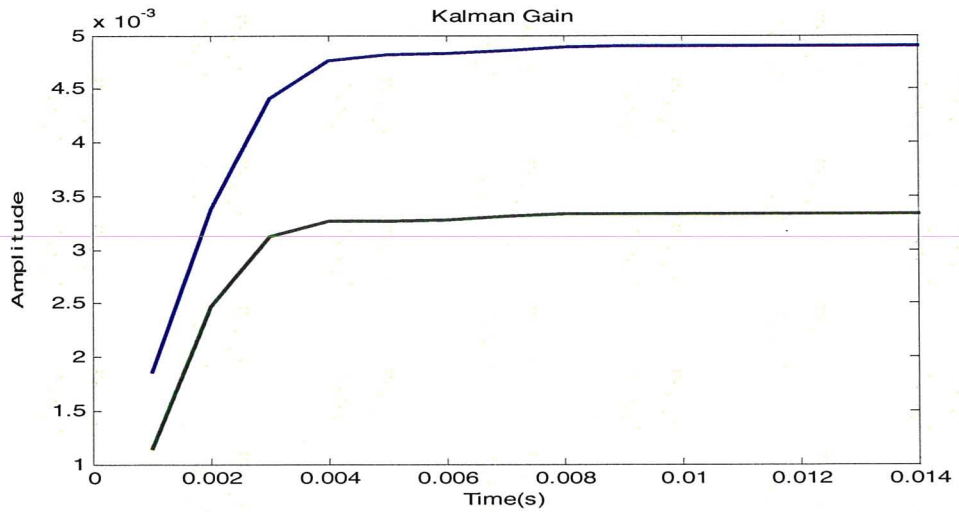

Figure 76 Variation of the Kalman Gain

The outputs of the PID control of the belt drive system without model uncertainty are shown in Figure 77. Tracking error of the PID control of the belt drive system without model uncertainty is shown in Figure 78.

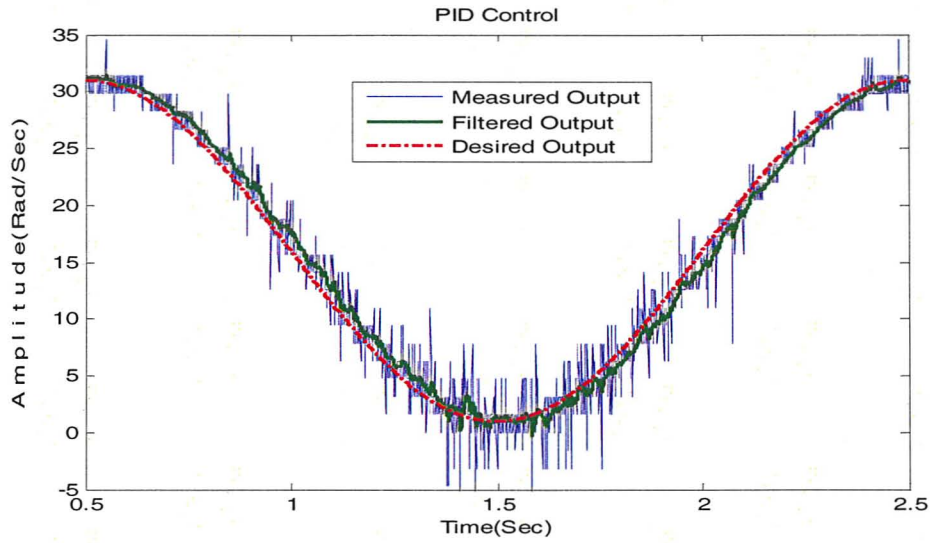

Figure 77 PID Control of the Belt Drive System without Model Uncertainty

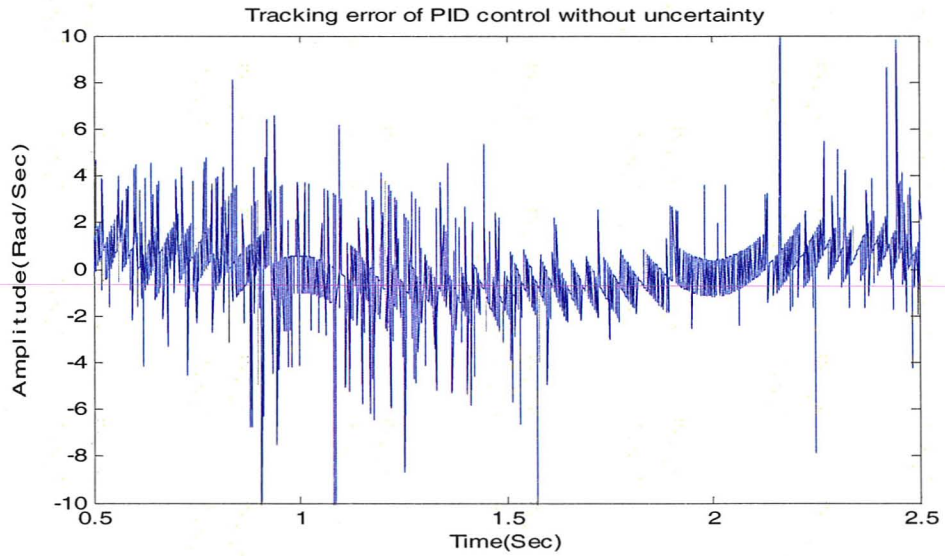

Figure 78 Tracking Error of the PID Control of the Belt Drive System without Model Uncertainty

The outputs of the PID control of the belt drive system model uncertainty are shown in Figure 79. Tracking error of the PID control of the belt drive system model uncertainty is shown in Figure 80.

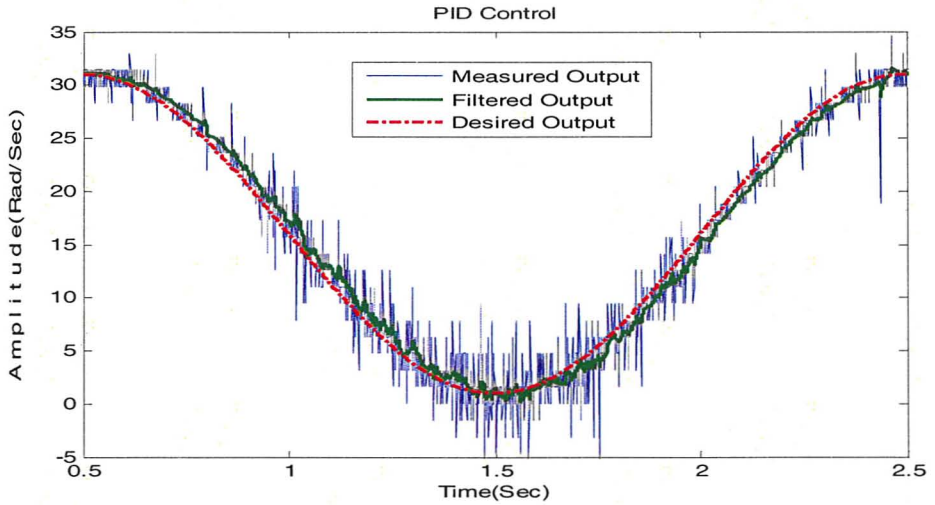

Figure 79 PID Control of the Belt Drive System with Model Uncertainty

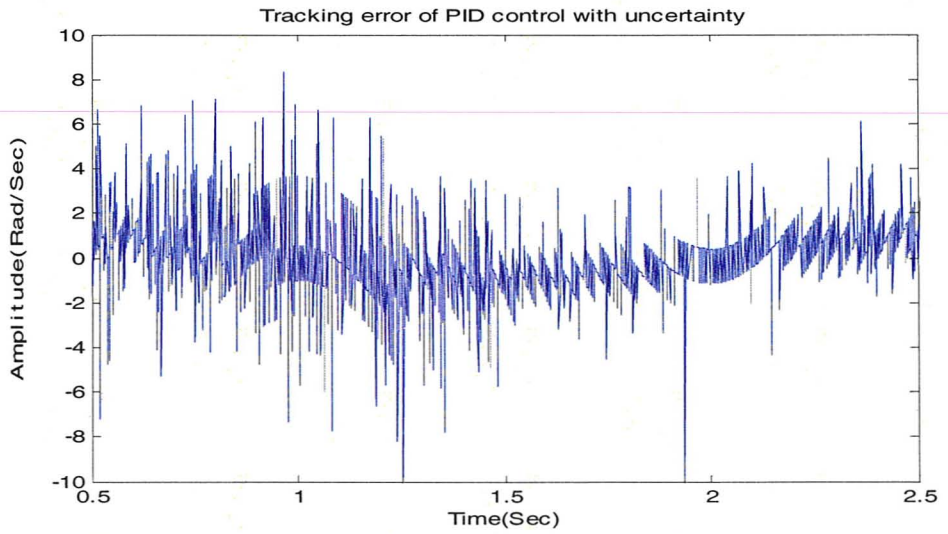

Figure 80 Tracking Error of the PID Control of the Belt Drive System with Model Uncertainty

The RMSE, the RMSEE, the RMSCE, the RMSU, the MAE and the MAU are listed in Tables 7, 8, 9, 10, 11 and 12, respectively.

Table 7 *RMSE of the Controlled Belt Drive System*

Controller	Uncertainty	RMSE(Rad/Sec)
DSMC with Kalman Filter	No	0.0303
	Yes	0.0336
DSMC with State Observer	No	0.0335
	Yes	0.0353
PID Control	No	0.0239
	Yes	0.0253

Table 8 *RMSEE of the Controlled Belt Drive System*

Controller	Uncertainty	RMSEE(Rad/Sec)
DSMC with the Kalman Filter	No	0.0308
	Yes	0.0342
DSMC with the State Observer	No	0.0329
	Yes	0.0349
PID Control with the Butterworth Filter	No	0.0216
	Yes	0.0231

Table 9 *RMSCE of the Controlled Belt Drive System*

Controller	Uncertainty	RMSCE(Rad/Sec)
DSMC with the Kalman Filter	No	0.0019
	Yes	0.0020
DSMC with the State Observer	No	0.0040
	Yes	0.0040
PID Control with the Butterworth Filter	No	0.0146
	Yes	0.0148

Table 10 *RMSU of the Controlled Belt Drive System*

Controller	Uncertainty	RMSU(Volt)
DSMC with Kalman Filter	No	0.0046
	Yes	0.0046
DSMC with State Observer	No	0.0046
	Yes	0.0046
PID Control	No	0.0045
	Yes	0.0046

Table 11 *MAE of the Controlled Belt Drive System*

Controller	Uncertainty	MAE(Rad/Sec)
DSMC with Kalman Filter	No	13.5465
	Yes	13.9629
DSMC with State Observer	No	17.1950
	Yes	17.9337
PID Control	No	11.6058
	Yes	12.2469

Table 12 *MAU of the Controlled Belt Drive System*

Controller	Uncertainty	MAU(Volt)
DSMC with the Kalman Filter	No	0.4884
	Yes	0.4891
DSMC with the State Observer	No	0.4907
	Yes	0.4908
PID Control with the Digital Butterworth Filter	No	0.4747
	Yes	0.4808

The robustness of DSMC of the belt drive system is verified by experiments. The experiments involved running the system without added uncertainty compared to cases when unmodeled two extra pulleys and one eccentric flywheel were added to the system.

Compared to the other control algorithm, the DSMC with Kalman filter produced a better performance mainly as follows: (1) a better transient response was achieved in the DSMC with the Kalman filter; (2) a smoother filtered signal was obtained by using the Kalman filter; (3) phase shift was avoided by using the Kalman filter as shown in Figure 81. In the sense of tracking error, the PID control with digital Butterworth filter gives the best performance. This unusual result is mainly because that the mathematical model used in the DSMC is not accurate.

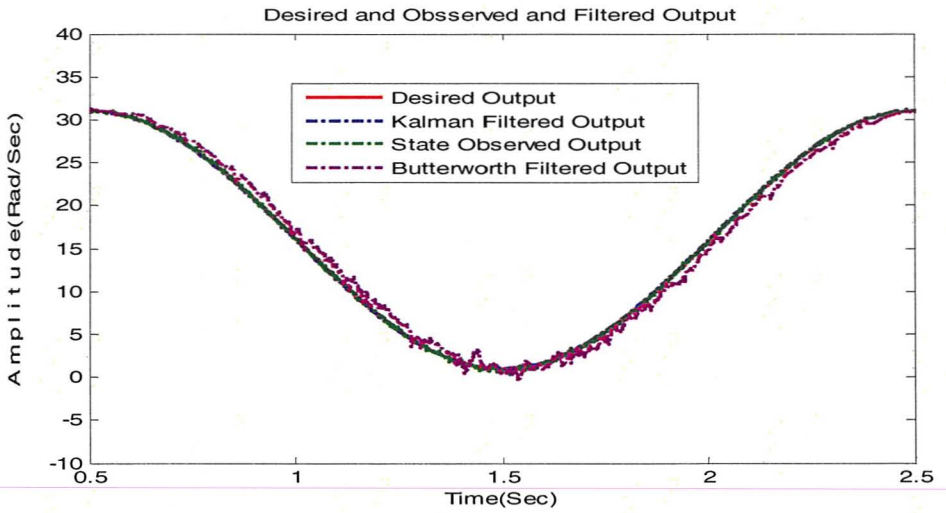

Figure 81 Desired, Filtered and Observed Outputs

Chapter 6 Conclusions

A mathematical model has been proposed and applied to a laboratory-scale experimental prototype belt drive system. This model is used as a priori information to support experimental verification in system identification.

A transfer function of the belt drive system has been obtained using physical modeling even though the parameters involved are left unknown. The transfer function has provided pieces of useful information needed for system identification such as the order of the system. The system order information is re-obtained empirically using the impulse response data during the initial tests of the system identification. The system order information obtained through different methods has been consistent with each other.

Experiments on the prototype have shown that correct determination of the linear region of the belt drive system is the most important factor ensuring the success of the system identification process. The system identification is expected to produce a linear time-invariant dynamic model which is dependent on the operating point. The model obtained using system identification should not include any discontinuities or non-linearities. It has been determined experimentally that the saturation of the sensors is the bottleneck, restricting the width of the linear region of the system.

Different linear regions have been identified through different sensors. The linear region determined through the encoder (Encoder *A*) originally mounted inside the electric motor covers the entire input region of the AC drive, namely between 0 to 7 volts with the consideration of only unidirectional operation. Two main drawbacks of the encoder

(Encoder *B*) on the driven pulley have been identified: (1) a low saturation point narrows the linear region of the system; (2) it is hard to concentrically align the encoder with the driven pulley as a result the low frequency noises can not be avoided. The linear region identified through Encoder *B* is in the range of 0 to 1 volt due to its saturation.

The second drawback of Encoder *B*, namely eccentric installation, has shown a profound influence on the quality of the system identification and the control process. Low frequency content has been identified in the measurement noise, its frequency has been found positively correlated to the amplitude of the input, and is located in the frequency range of interest for the system identification and the control process. Hence the low frequency content can not be filtered out and impacts the identification and control process. Due to sensor B, an acceptable transfer function could only be obtained for input amplitudes in the range of 0 to 1 volt. In this narrow region, the combination of low input amplitude and noise adversely affects signal to noise ratio, restricting the system operation, system identification and control in the low velocity region. The output of the system is the velocity signal calculated from the encoder measurement. When calculating the velocity from an encoder signal at low velocity, the results become even noisier due to numerical differentiation. The eccentricity of Encoder *B* has greatly degraded the overall quality of system identification and control.

The second objective of this research project has been to explore the application and the robustness of the sliding mode control algorithm. A DSMC tracking controller has been designed and implemented on the belt drive system with the state variables estimated using a state observer and the Kalman filter, respectively. In comparison, a

discrete PID controller has also been designed and implemented. A comparison of the controllers is presented under the following three categories: (1) the performance of the observer and the filters including the Kalman filter and the digital Butterworth filter is compared using the error between the estimated or filtered and the measured outputs with respect to the Mean-Square-Root of the Estimation Error (RMSEE); (2) the controllers, namely the DSMC with the state observer, the DSMC with the Kalman filter, and the PID with the digital Butterworth filter, are compared with respect to the Mean-Square-Root of the Control Error (RMSCE) and the tracking Error (RMSE); and the Mean-Square-Root of the control signal, the Maximum-Absolute value of the tracking Error (MAE) and the Maximum-Absolute value of the control signal (MAU) are also compared.

The digital Butterworth filter applied with the PID controller is easy to design and implement and can effectively remove high frequency contents from a signal of interest. The drawback is that it introduces a phase-shift that is highly undesirable in real-time control applications. In this research project, the output of the system is a velocity signal which is calculated from an encoder resulting in noise in the frequency range of interest for control, and phase-shift due to digital filtering.

Model based filters and observers are capable of estimating the state of the system without introducing a phase-shift. The most well-known model based filter is the Kalman filter. It has been successfully used for estimating the states of the experimental system. The Kalman filter has been designed based on the assumptions that both the process noise and the measurement noise are white. These assumptions are not satisfied in this research project and the performance of the Kalman filter has been adversely affected as expected.

However compared to a state observer, a properly tuned Kalman filter provides a smoother state estimate.

Overall, the smallest Mean-Square-Root of the filtering error (RMSEE) was obtained by using the digital Butterworth filter. The Kalman filter and the state observer did not introduce any phase-shift into the system. But their RMSEEs are higher than the digital Butterworth filter. The state observer produced the highest RMSEE.

Amongst all the controllers implemented, the discrete sliding mode control (DSMC) with the Kalman filters provided the smallest Mean-Square-Root of the Control Error (RMSCE). The best tracking performance was achieved using the DSMC with the Kalman filter. Its steady state response was much better than that of the PID control but slightly poorer than that of the DSMC with the Kalman filter. The PID controller produced the worst steady-state response among all of the controllers.

The verification of the robustness of the DSMC algorithm was carried out by adding two extra pulleys and one eccentric flywheel to the system without adjusting the controller parameters. All of the controllers performed equally well. This is probably because the overall load including the unmodeled uncertainty of the system is too small compared to the power that the electric motor can supply to cause any one of the controller to fail.

References

Abrate, S. Vibrations of belts and belt drives. *Mechanism and Machine Theory, 27,* 645– 659, 1992

Adjustable-speed drives. (2007). Retrieved October 26, 2007, from http://www.emerson-ept.com/eptroot/public/schools/adjsddrv.pdf/

Alessandri, A. Cuneo, M. & Punta, E. Sliding mode state observers for multi-output nonlinear systems with bounded noise on dynamics and measurements. *Proceedings of the 48th IEEE Conference on Decision and Control,* 7729-7734, 2009

Aoki, M. & Huddle, J. Estimation of State vector of a Linear Stochastic System with a Constrained Estimator, *IEEE Transactions on Automatic Control, 12(4),* 432-433, 1967

Astrom, K. J. & Wittenmark, B. *Computer-Controlled Systems: Theory and Design.* Prentice Hall, 1997

Bandyopadhyay B. & Thakar, V. K. Discrete time output feedback sliding mode control for nonlinear MIMO systems: A step motor case. *International Journal of Systems Science, 39(1),* 89–104, 2008

Bang-bang control. (2010). Retrieved January 5, 2010, from http://en.wikipedia.org/wiki/Bang–bang_control

Barker, C. R., Oliver, L. R. & Brieg, W. F. Dynamic Analysis of Belt Drive Tension Forces during Rapid Engine Acceleration, *SAE (Society of Automotive Engineers) Transactions, 100(6),* 898-913, 1991

Bartolini, G., Ferrara, A. & Utkin, V. I. Adaptive Sliding Mode Control in Discrete-time Systems. *Automatica, 31(5),* 769-773, 1995

Bechtel, S. E., Vohra, S., Jacob, K. I. & Carlson, C. D. The stretching and slipping of belts and fibbers on pulleys, *Journal of Applied Mechanics, 67,* 197-206, 2000

Beikmann, R. S., Perkins, N. C. & Ulsoy, A. G. Free Vibration of Serpentine Belt Drive Systems, *Journal of Vibration and Acoustics, Transactions of the ASME, 118(3),* 406-413, 1996

Belt and Chain Drives. (2009). Retrieved August 19, 2009, from http://www.emerson-ept.com/eptroot/public/schools/**belt**chan.pdf

Belt Drive. (2009). Retrieved August 18, 2009, from http://www.srmuniv.ac.in/downloads/**belt**_drive.doc

Bian, C., Man, Y., Song C. & Ren, S. Variable Structure Control of Switched Reluctance Motor and Its Application, *Proceedings of the World Congress on Intelligent Control and Automation (WCICA), 1,* 2490-2493, 2006

Bose, B. K. *Modern Power Electronics and AC Drives*. Prentice-Hall Inc., 2001

Buja, G. S. & Souliaev, A. A Variable Structure Controller. *IEEE Transactions on Automatic Control, 33(2),* 206-209, 1988

Cadet, O. (2010). *Introduction to Kalman Filter and Its Use in Dynamic Positioning Systems*. Retrieved November 20, 2010, from http://www.dynamic-positioning.com/dp2003/design_cadet.pdf

Cepon, G. & Boltezar, M. Dynamics of a belt-drive system using a linear complimentarity problem for the belt-pulley contact description. *Journal of Sound and Vibration, 319(3-5),* 1019-1035, 2009

Chang, J. L. Discrete Sliding-Mode Control of MIMO Linear Systems. *Asian Journal of Control, 4(2),* 217-222, 2002

Chaouki, M. & Moncef, G. Robust Decentralized Sliding Mode Control for Large Scale Uncertain Systems. *Control Conference 2007,* 76-81, 2007

Chen Y. P. & Chang, J. L. A new method for constructing sliding surfaces of linear time-invariant systems. *International Journal of Systems Science, 31(4),* 417- 420, 2000

Chuang, C. W., Haung, C. L., Lee, C. D., Kao, C. C. & Fung, R. F. Synchronization and tension control of dual motor systems via MIMO discrete pseudo model following integral variable structure control. *Mechanism and Machine Theory, 44(2),* 499-510, 2009

Claudia, P. & Miguel, S. Speed Control of DC Motor by Using Fuzzy Variable Structure Controller. *Proceedings of the 27th Chinese Control Conference, CCC,* 311-315, 2008

Chakraborty, N., Berard, S., Akella, S. & Trinkle, J. (2009). *An Implicit Time-Stepping Method for Multibody Systems with Intermittent Contact.* Retrieved August 25, 2009, from http://www.cs.rpi.edu/~sakella/papers/rss07-130-CBAT.pdf

Clark, R. N. *Control System Dynamics.* Cambridge University Press, 1996

Cunha, J. P.V.S. D., Costa, R. R. & Hsu, L. Input/Output Variable Structure Position Control of a Remotely Operated Underwater Vehicle. *Fifth International Conference on Advanced Robotics,* 1305-1310, 1991

DeCarlo, R. A., Zak, S. H. & Matthews, G. P. Variable Structure Control of Nonlinear Multivariable Systems: A Tutorial. *Proceedings of the IEEE, 76(3),* 212-232, 1988

Dehkordi, A. B., Gole, A. M. & Maguire, T. L. (2010). Permanent Magnet Synchronous Machine Model for Real-time Simulation. Retrieved October 30, 2010, from http://www.ipst.org/TechPapers/2005/IPST05_Paper159.pdf

Distefano, J., Stubberud, A. R. & Williams, I. J. *Schaum's Outlines of Feedback and Control Systems.* Second Edition, the Mcgraw-Hill Companies, Inc., 1990

Dufva, K., Kerkkanen, K. S., Maqueda, L. G. & Shabana, A. A. Nonlinear dynamics of three-dimensional belt drives using the finite-element method. *Nonlinear Dynamics, 48*, 449-466, 2007

Edwards, C. & Spurgeon, S. K. *Sliding Mode Control: Theory and Applications*. Taylor and Francis Ltd., 1998

Engineering Statistics Handbook. (2009). Retrieved July 15, 2009, from http://www.itl.nist.gov/div898/handbook/pmd/section4/pmd431.htm

Fawcett, J. N. Chain and Belt Drives-A Review. *Shock and Vibration Digest, 13(5),* 5-12, 1981

Fehr, R. E. (2010). *A Novel Approach for Understanding Systematical Components and Sequence Networks of Three-Phase Power Systems.* Retrieved December 8, 2010, from http://people.virginia.edu/~bjb7f/Old%20Stuff/drawings/E-Journal/Failure%20and%20Reliability%20Blood%20Pumps/Fault%20Tolerance/Pillay_Modeling%20PM%20drive.pdf

Fernandes, A. A. & Alcalde, V.H. C. Serious Compensation Using Variable Structure and Lyapunov Function Controls for Stabilization Multimachine Power System. *Mediterranean Conference on Control and Automation, MED,* 2007

Film Belts. (2009). Retrieved August 18, 2009, from http://news.thomasnet.com/fullstory/560695

Film Belts. (2009a). Retrieved August 18, 2009, from http://www.beltingindustries.com/page15.htm

Firbank, T. C. Mechanics of the Belt Drive. *International Journal of Mechanical Sciences,*

12(12), 1053-1063, 1970

Flat Belts. (2009). Retrieved August 17, 2009, from http://www.directindustry.com/industrial- manufacturer/flat-belt-60976.html

Furuta, K. Sliding mode control of a discrete system. *Systems & Control Letters, 14(2),* 145-152, 1990

Gao, W. B. *Foundation of Variable Structure Control.* Beijing: China Press of Science and Technology (in Chinese), 1990

Gao, W. B. & Hung, J. C. Variable Structure Control of Nonlinear Systems: A New Approach. *IEEE Transactions on Industrial Electronics, 40,* 45-55, 1993

Gao, W. B., Wang, Y. & Homaifa, A. Discrete-time Variable Structure Control Systems. *IEEE Transactions on Industrial Electronics, 42(2),* 117-122, 1995

Garcia, J. P. F., Ribeiro, J. M. S., Silva, J. J. F. & Martins, E.S. Continuous-time and discrete-time sliding mode control accomplished using a computer. *IEE Proceedings: Control Theory and Applications, 152(2),* 220-228, 2005

Habibi, S. R. & Richards, R. J. Sliding mode control of an electrically powered industrial robot. *IEE Proceedings D: Control Theory and Applications, 139(2),* 207-225, 1992

Habibi, S. R. *System identification lecture notes.* (2009). Hamilton, ON: McMaster University

Hace, A., Jezernik, K. & Terbuc, M. VSS motion control for a laser-cutting machine. *Control Engineering Practice, 9(1),* 67-77, 2001

Hace, A., Jezernik, K. & Šabanovic, A. A New Robust Position Control Algorithm for a Linear Belt Drive. *Proceedings of the IEEE International Conference on Mechatronics, 3(5),* 358-363, 2004

Hace, A., Jezernik, K. & Sabanovic, A. Improved Design of VSS Controller for a Linear Belt-driven Servomechanism. *IEEE/ASME TRANSACTIONS ON MECHATRONICS, 10(4),* 385-390, 2005

Hace, A., Jezernik, K. & Sabanovic, A. SMC with Disturbance Observer for a Linear Belt Drive. *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, 54(6),* 3402-3412, 2007

Hace, A., Jezernik, K. & Terbuc, M. (2009). *VSS motion control for a laser cutting machine. Control engineering practice*. Retrieved November 29, 2009, from

https:// www.ro.feri.unimb.si/~ales/VSSmotionControlForLaserCuttingMachine.pdf

Han, Y., Sun, Y. & Mo, H. Dynamic inversion control based on variable structure theory for underwater high-speed vehicle. *Chinese Control and Decision Conference,* 3761-3764, (in Chinese), 2008

Hashimoto, H., Maruyama, K. & Harashima, F. A Microprocessor-Based Robot Manipulator Control with Sliding Mode. *IEEE Transactions on Industrial Electronics, 34(1),* 1987

Haykin, S. *Adaptive filter theory.* 3rd ed. Prentice Hall, Upper Saddle River, New Jersey, 1996

Healey A. J. & Lienard, D. Multivariable Sliding Mode Control for Autonomous Diving and Steering of Unmanned Underwater Vehicles. *IEEE Journal of Oceanic Engineering, 18(3),* 327-339, Jul 1993

Heemels, W. P. M. H., Camhbel, M. K. & Schumacher, J. M. A time-stepping method for relay systems. *Decision and Control, 5,* 4461-4466, 2000

Hu, Q. & Ma, G. Variable structure control and active vibration suppression of flexible spacecraft during attitude maneuver. *Aerospace science and technology, 9(4),* P. 307-317, 2005

Hu, Q. & Ma, G. Control of Three-axis Stabilized Flexible Spacecraft Using Variable Structure Strategies Subject to Input Nonlinearities. *Journal of Vibration and Control, 12(2),* 659-681, 2006

Huang, J. Y., Gao, W. B. & Huang, J. C. Variable Structure Control: A Survey. *IEEE Transactions on Industrial Electronics, 40(1),* 2-22, 1993

Huang, Y. J., Kuo, T. C. & Way, H. K. Robust vertical takeoff and landing aircraft control via integral sliding mode. *IEE Proceedings-Control Theory and Applications, 150(4),* 383-388, 2003

Hui, S. & Zak, S. H. On discrete-time variable structure sliding mode control. *Systems & Control Letters, 38,* 283-288, 1999

Hung, J. Y., Gao, W. B. & Hung, J. C. Variable Structure Control: A Survey. *IEEE Transactions on Industrial Electronics, 40(1),* 2-22, 1993.

Hwang, C. L. Design of Servocontroller via the Sliding Mode Technique. *IEE Proceedings D (Control Theory and Applications), 139(5),* 439-46, 1992

Hwang, S. J. Perkins, N. C., Ulsoy, A. G. & Mechstroth, R. J. Rotional Response and Slip Prediction of Serpentine Belt Drive Systems. *Journal of Vibration and Acoustics, Transactions of the ASME, 116(1),* 71-78, 1994

Ignaciuk, P. & Bartoszewicz, A. Linear Quadratic Optimal Sliding Mode Flow Control for Connection-Oriented Communication Networks. *International Journal of Robust and Nonlinear Control, 19,* 442-461, 2009

Jafarov, E. M. & Tasaltin, R. Robust Sliding-Mode Control for the Uncertain MIMO Aircraft Model F-18. *IEEE Transactions on Aerospace and Electronic Systems, 36(4),* 1127-1141, 2000

Janardhanan, S. & Bandyopadhyay, B. Discrete sliding mode control of systems with unmatched uncertainty using multirate output feedback. *IEEE Transactions on Automatic Control, 51(6),* 1030-1035, 2006

Jayawardene, T. S. S., Nakamura, M. & Goto, S. Accurate Control Position of Belt Drives under Acceleration and Velocity Constraints. *International Journal of Control, Automation, and Systems, 1(4),* 2003

Jayawardene, T. S. S. (2009). *Trajectory Planners for Cooperative Control of Two Industrial Robots and Belt Drives*. Retrieved November 15, 2009, from https://dlwww.dl.saga-u.ac.jp/contents/diss/GI00000874/th_Thesis.pdf

Jean, M. The non-smooth contact dynamics method. *Computer Methods in Applied Mechanics and Engineering, 177,* 235-257, 1999

Jiang, Y., Hu, Q. & Ma, G. Design of Robust Adaptive Integral Variable Structure Attitude Controller with Application to Flexible Spacecraft. *International Journal of Innovative Computing, Information and Control, 4(9),* 2431-2440, 2008

Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME-Journal of Basic Engineering, 82, Series D,* 35-45, 1960

Kao, C. C. Surveys of Discrete-time Variable Structure Control. *IEEE ICSS2005 International Conference on Systems & Signals,* 2005

Kaynak, O., Erbatur, K. & Ertugrul, M. The Fusion of Computationally Intelligent Methodologies and Sliding Mode Control-A Survey. *IEEE Transactions on Industrial Electronics, 48(1),* 4-17, 2001

Kerkkanen, K. S. Vallejo, D. G. & Mikkola, A. M. Modeling of belt-drives using a large deformation finite element formulation. *Nonlinear Dynamics, 43,* 239-256, 2006

Kerkkänen, K. (2009). *Dynamic Analysis of Belt-Drives Using the Absolute Nodal Coordinate Formulation.* Retrieved February 17, 2009, from https://oa.doria.fi/handle/10024/31226

Khalal, O., Mellit, A., Rahim, M., Salhi, H. & Guessoum, A. Robust Control of Manipulator Robot by Using the Variable Structure Control with Sliding Mode. *2007 Mediterranean on Control and Automation,* 2007

Kim G. W. & Wang, K. W. Switching sliding mode force tracking control of piezoelectric-hydraulic pump-based friction element actuation systems for automotive transmissions. *Smart Materials and Structures, 18(8),* 2009

Kong, L. & Parker, R. G. Microslip friction in flat belt drives. *Proceedings of the institution of mechanical engineers, Part C: Journal of Mechanical Engineering Science, 219(10),*1097-1106, 2005

Kraver, T. C., Fan W. G. & Shah, J. Complex modal analysis of a flat belt pulley system with belt damping and Coulomb damped tensioner. *Computers in Engineering, Proceedings of the International Conference and Exhibit, 2,* 589-595, 1994

Kulebakin, V. On Theory of Vibration Controller for Electric Machines (in Russian). *Theoretical and Experimental Electronics, 4,* 1932

Lai, N. O., Edwards C. & Spurgeon, S. K. On Discrete Dynamic Output Feedback Min-Max Controllers. 43rd *IEEE Conference on Decision and Control, 2,* 1836-41, 2004

Lai, N. O., Edwards C. & Spurgeon, S. K. On Output Tracking Using Dynamic Output Feedback Discrete-Time Sliding Mode Controllers. 44th *IEEE Conference on Decision and Control, and the European Control Conference 2005 Seville, Spain,* 12-15, 2005

Lai, N. O., Edwards C. & Spurgeon, S. K. An implementation of an output tracking dynamic discrete-time sliding mode controller on an aircraft simulator. *Proceedings of the 2006 international workshop on variable structure systems*, Alghero, Italy, 35-40, 2006

Leamy, M. J. & Perkins, N. C. Nonlinear Periodic Response of Engine Accessory Drives with Dry Friction Tensioners. *ASME J. Vibration and Acoustics, 120(4),* 909–916, 1998

Leamy, M. J. Barber J. R. & Perkins, N. C. Distortion of a harmonic elastic wave reflected from a dry friction support. *Transactions of the ASME. Journal of Applied Mechanics*, *65(4),* 851-857, 1998

Leamy, M. J. & Wasfy, T. M. Transient and steady-state dynamic finite element modeling of belt-drives. *Journal of Dynamic Systems, Measurement, and Control, 124* 575-581, 2002

Leamy, M. J. & Wasfy, T. M. Analysis of Belt-Drive Mechanics Using a Creep-Rate-Dependent Friction Law. *Journal of Applied Mechanics, Transactions ASME, Vol. 69(6),* 763-771, 2002a

Lee, K. M. & Rutherford, C. Frequency reshaped quadratic control of a Low-Cost human level performance belt driven robot. *Mechatronics, 9(1),* 95-110, 1999

Lee, P. M., Hong, S. W., Lim, Y. K., Lee, C. M., Jeon, B. H. & Park, J. W. Discrete-Time Quasi-Sliding Mode Control of an Autonomous Underwater Vehicle. *IEEE Journal of Oceanic Engineering, 24(3),* 388-395, 1999

Li, W. & Cheng, X. Adaptive high precision control of positioning tables-theory and experiment. *IEEE Transactions on Control Systems Technology, 2(3),* 265-270, 1994

Li, W. & Rehani, M. Modeling and Control of a Belt-drive Positioning Table. *International Conference on Industrial Electronics, control, and Instrumentation, 3,* 1984-1989, 1996

Li, Z. B., Wang, Z. L. & Li, J. F. A hybrid control scheme of adaptive and variable structure for flexible spacecraft. *Aerospace Science and Technology, 8(5),* 423-430, 2004

Lim, R. K. & Phan, M. Q. (1998). *State-Space System Identification with Identified Hankel Matrix.* Department of Mechanical and Aerospace Engineering Technical Report No.3045, Princeton University, Princeton, NJ, 1998

Lin, F. J., Lin C. H. & Shen, P. H. Variable Structure Control for Linear Synchronous Motor Using Recurrent Fuzzy Neural Network. *IECON Proceedings (Industrial Electronics Conference), 3,* 2108-2113, 2002

Liu, M., Cai, Z., Cheng, X. & Ouyang, M. Adaptive Position Servo Control of Permanent Magnet Synchronous Motor. *Proceedings of the 2004 American Control Conference,* Boston, Massachusetts, 2004

Ljung, L. *System Identification: Theory for the User.* PRT Prentice Hall, Englewood Cliffs, New Jersey, 1987

Loh, A. M. & Yeung, L. F. (2010). *Chattering Reduction in Sliding Mode Control: an Improvement for Nonlinear Systems.* Retrieved November 4, 2010, from Http://www.wseas.us/e-library/conferences/venice2004/papers/476-181.pdf

Machado, J. A. T. & Carvalho, J. L. M. de. A New Variable Structure Controller for Robot Manipulators. *Third International Symposium on Intelligent Control*, 441-446, 1988

Mao, X. & Lu, K. Research on the Nonlinear Governor of Diesel Engine with Variable Structure Control Theory. *2008 IEEE International Conference on Robotics, Automation and Mechatronics,* 537-542, 2008

Misawa, E. A. Discrete-Time Sliding Mode Control for Nonlinear Systems with Unmatched Uncertainties and Uncertain Control Vector. *Journal of Dynamic Systems, Measurement, and Control, 119(3),* 503-512, 1997

Misawa, E. A. Discrete-Time Sliding Mode Control: The Linear Case. *Journal of Dynamic Systems, Measurement, and Control, 119(4),* 819-821, 1997a

Monsees, G. (2010). *Discrete-time Sliding Mode Control.* Retrieved May 24, 2010, from http://www.dcsc.tudelft.nl/Research/PublicationFiles/publication-5390.pdf

Moskwa, J. J. Sliding Mode Control of Automotive Engines. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME, 115(4),* 687-693, 1993

Muñoz, D. & Sbarbaro, D. An adaptive sliding-mode controller for discrete nonlinear systems. *IEEE transactions on industrial electronics, 47(3),* 574-581, 2000

Nikolski, G. On Automatic Stability of a Ship on a Given Course (in Russian). *Proceedings of the Central Communication Laboratory, 1,* 34-75, 1934

Nise, N. S. Control Systems Engineering. 4th ed. John Willy and Sons, 2004.

Niu, Y., Wang, Z. & Wang, X. Robust sliding mode design for uncertain stochastic systems based on H_∞ control method, *Optimal Control Applications and Methods, 31(2),* 93–104, 2010

Nouri, A. S., Abdennour, R. B. A Sliding Mode Control for Multivariable Systems. *International Journal on Sciences and Techniques of Automatic Control, 1(1),* 29-42, 2007

Ogata, K. *Discrete-time control systems.* 2nd ed. Prentice-Hall, 1994

Ogata, K. *Modern control engineering.* 3rd ed. Prentice-Hall, 1997

Pai, M. C. Robust Control of Uncertain Time-Delay Systems via Sliding Mode Control and LMI- H_∞ Technique. *Journal of Marine Science and Technology, 16(1),* 1-7,

2008

Pandian, S. R., Hanmandlu, M. & Gopal, M. A decentralized Variable Structure Model Following Controller for Robot Manipulators. *Robotics and Automation, IEEE International Conference on*, 1324-1328, 1988

Pang, H. P., Liu, C. J. & Liu, A. Z. Sliding Mode Control for Time-Delay Systems with Its Application to Networked Control Systems. *Sixth International Conference on Intelligent Systems Design and Applications, 16,*192-197, 2006

Passino, K. M. & Yurkovich, S. *Fuzzy control.* Addison-Wesley Longman, Inc., 1997

Perruquetti, W., Borne, P. & Richard, J. P. A Generalized Regular Form or Sliding Mode Stabilization of MIMO Systems. *Proceedings of the IEEE Conference on Decision and Control, 2,* 957-961, 1997

Perruquetti, W. & Barbot, J. P. *Sliding Mode Control in Engineering*. Marcel Dekker, Inc., New York, 2002

Petersen, K. B. & Pedersen, M. S. (2010). *The Matrix Cookbook*. Retrieved June19, 2010, from http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3274/pdf/imm3274.pdf

Pillay, P. & Krishna, R. Modeling, Simulation, and Analysis of Permanent-Magnet Motor Drives, Part I: The Permanent-Magnet Synchronous Motor Drive. *IEEE Transactions on Industrial Applications, 25(2),* 1989

Pitstick, G. M., Mulholland, R. J. Approximate Realization Algorithms for Truncated Impulse Response Data. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 34(6),*1583-1588, 1986

Promtun, E. & Seshagiri, S. Siding Mode Control of Pitch-Rate of an F-16 Aircraft. *International Journal of Applied Science, Engineering and Technology, 5(1),* 2009

Qu, S. & Wang, Y. Robust Control of Uncertain Time-Delay Systems: a Novel Sliding Mode Control Design via LMI. *Journal of Systems Engineering and Electronics, 17(3),* 624-628, 2006

Ramirez, H. S., Thomas A. & Dwyer, W. Variable Structure Control of Spacecraft Reorientation Maneuvers. *Proceedings of American Institute of Aeronautics and Astronautics Guidance, Navigation, and Control Conference, 19(21),* 88-96, 1986

Ramirez, H. S. A Variable Structure Control Approach to the Problem of Soft Landing on Planet. *Control-Theory and Advanced Technology, 6(1),* 53-73, 1990

Raoufi, R. & Zinober, A. S. I. Sliding Mode State Observation for Time-Delay Uncertain Nonlinear Systems. *American Control Conference*, 5248-5253, 2008

Rashid K. & Zidan, H. Variable Structure Controller with Prescribed Transient Response to Control the Position of the Induction Motor Drives. *International Journal of Advanced Manufacturing Technology, 39(7-8),* 744-754, 2008

Round Belt. (2009). Retrieved August 17, 2009, from http://www.directindustry.com/prod/habasit/round-belt-5857-46507.html

Saarakkala, S. (2009). Design of Position and Anti-swing Control for Linear Belt Drive. Retrieved January 4, 2009, from www.ee.lut.fi/en/learn/courses/Sa2731300/DCD_Demo_Lecture2.pdf

Sabanovic, A., Sozbilir, O., Goktug, G. & Sabanovic, N. Sliding mode control of timing-belt servo-system. *IEEE international symposium on industrial electronics, 2,* 684-689, 2003

Samarasinghe, V. G. D. C., & Pahalawaththa, N. C. Stabilization of a multi-machine power system using nonlinear robust variable structure control. *Electric Power Systems Research, 43(1),* 11-17, 1997

Sarpturk, S. Z., Istcfanopulos, Y. & Kaynak, O. On the stability of Discrete-Time Sliding Mode Control Systems. *IEEE Transactions on Automatic Control, 32(10),* 930-932, 1987

Schur2010: ------, "Further Robustness of the LQR", [online], http://ocw.mit.edu/NR/rdonlyres/Mechanical-Engineering/2-154Fall-2004/B1E001D8-AE8B-4CE3-B618-562CBD8157AE/0/lec25.pdf, accessed on April o9, 2010

Schulte, H. Approximate Modeling of a Case of Nonlinear Oscillators using Takagi-Sugeno Fuzzy Systems and Its Application to Control Design. *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference,* 3387-3392, 2005

Selezneva, A. (2009). *Modeling and Synthesis of Tracking Control for the Belt Drive System*. Retrieved October 20, 2009, from https://oa.doria.fi/bitstream/handle/10024/.../TMP.objres.638.pdf?...1

Shahat, A. E. & Shewy, H. E. Permanent Magnet Synchronous Motor Dynamic Modeling with Genetic Algorithm Performance Improvement. *International Journal of Engineering, Science and Technology, 2(2),* 93-106, 2010

Shen, Y., Shen, W. & Gu, J. Sliding Mode Control for Uncertain Stochastic Systems with Time Delay. *IEEE International Conference on Mechatronics and Automation, 3,* 1224-1229, 2005

Shtessel, Y., Buffington, J. & Banda, S. Tailless Aircraft Flight Control Using Multiple Time Scale Reconfigurable Sliding Modes. *IEEE Transactions on Control Systems Technology, 10(2),* 288 – 296, 2002

Slotine, J. E. & Li, W. *Applied Nonlinear Control.* Prentice Hall, New Jersey, 1991

Smith, S. W. (2010). *The scientist and engineer's guide to digital signal processing.* Retrieved May 1, 2010, from http://www.dspguide.com/

Solid Mechanics Dynamics Tutorial-Pulley Drives Systems. (2010). Retrieved April 20, 2010 from http://www.freestudy.co.uk/dynamics/pulleys.pdf

Song, L. & Huang, P. Adaptive discrete-time sliding mode control of brushless DC servo motors. *Second IEEE Conference on Industrial Electronics and Applications,* 1101-1103, 2007

Spurgeon, S. K. Sliding Mode Control Design for Uncertain Discrete-Time System *Proceedings of the 30th Conference on Decision and Control,* 1991

Spurgeon, S. K. Hyperplane design techniques for discrete-time variable structure control systems. *International Journal of Control, 55,* 445-456, 1992

Stewart, D. & Trinkle, J. An implicit time-stepping scheme for rigid body dynamics with inelastic collisions and coulomb friction. *International Journal for Numerical Methods in Engineering, 39,* 2673-2691, 1996

System identification. (2009). Retrieved April 15, 2009, from http://mazsola.iit.uni-miskolc.hu/ftp/tempus/QUAN-MOD/chap263.doc

SVD Tutorial. (2010). Retrieved October 15, 2010, from http://alias-i.com/lingpipe/demos/tutorial/svd/read-me.html

Takagi T. & Sugeno, M. Fuzzy identification of systems and its applications to modeling and control. *IEEE Trans. Syst., Man, Cybernetics, SMC-15,* 116-132, 1985

Tang, C. Y. & Misawa, E. A. Discrete Variable Structure Control for Linear Multivariable Systems: The State Feedback Case. *Proceedings of the 1998 American Control Conference, 1(21-26),* 114 – 118, 1998

Tang, C. Y. & Misawa, E. A. Discrete Variable Structure Control for Linear Multivariable Systems. *Journal of Dynamic Systems, Measurement, and Control, 122(4),* 783-793, 2000

Texas Instruments. (1997). *Clarke and Park Transforms on the TMS320C2xx.* Retrieved December 12, 2010, from http://web.cecs.pdx.edu/~strom/ev2/articles/foc_clark_park_xfrm_ti.PDF

Texas Instruments. (1998). *Field Oriented Control of 3-Phase AC-Motors*. Retrieved December 11, 2010, from http://focus.ti.com/lit/an/bpra073/bpra073.pdf

Timing belt drives and their advantages regarding engine efficiency and NVH characteristics. Retrieved August 19, 2009, from http://www.gates.com/file_save_common.cfm?...Gates%Fdocuments...

Uchida, M., Murata, R., Yabumi, T., Morita Y. & Kando, H. Sliding Mode Servo Control for Electromagnetic Engine Valve. *International Joint Conference, 18(21),* 3568-3663, 2006

Utkin, V. Variable Structure Systems with Sliding Modes. *IEEE Transactions on Automatic Control, 22(2),* 212-222, 1977

Utkin, V. Variable Structure Systems: Present and Future. *Automation and Remote Control, 44(9),* 1105-1120, 1983

Utkin, V. *Sliding Mode in Control Optimisation.* Springer-Verlag, 1992

Utkin, V. Adaptive Discrete-Time Sliding Mode Control of Infinite Dimensional Systems. *Proceedings of the 37th IEEE Conference on Decision and Control, 4,* 4033-4038, 1998

Utkin, V. *Sliding Mode Control in Electromechanical Systems*. Taylor & Francis, 1999

Utkin, V. & Chang H. Sliding Mode Control on Electromechanical Systems. *Mathematical problems in engineering, 8(4-5),* 451-473, 2002

Utkin, V., Guldner, J. & Shi, J. *Sliding Mode Control in Electro-Mechanical Systems*. 2nd ed. Taylor and Francis Group, 2009

Vacher, P. (2010). *Three Approaches for System Identification by Impulse Response Fitting*. Retrieved May 10, 2010, from http://www.cert.fr/dcsd/idco/perso/Vacher/Publications/Identification/SYSID09.pdf

Vargas, J. F. & Ledwich, G. Variable Structure Control for Power Systems Stabilization.

Electric Power and Energy Systems, 32(2), 101-107, 2010

Wang, S. H., Lee, T.F. & Zachery, R. System identification via singular value decomposition. *IEEE International Conference on Acoustics, Speech, and Signal Processing, 5,* 2638-2641, 1996

Wang, L., Liu, J., Yu, H. & Xu, Y. Sliding Mode Control of an Autonomous Underwater Vehicle. *Proceedings of 2002 International Conference on Machine Learning and Cybernetics, 1,* 247-251, 2002

Wang J., Brussel, H. V. & Swevers, J. Robust Perfect Tracking Control with Discrete Sliding Mode Controller. *Journal of Dynamic Systems, Measurement and Control, 125*, 2003

Wang, B. & Mao, Z. Nonlinear variable structure excitation and steam valving controller for power system stability. *Journal of Control Theory and Applications, 7(1),* 97-102, 2009

Wang, A., Xu, W. & Liu, C. T. On-Line PI Self-Turning Based on Inertia Identification for Permanent Magnet Synchronous Motor Servo System. *International Conference on Power Electronics and Drive Systems, 2,* 1406-1410, 2009

Waterman, D. & Nonami, K. (2010). *Multirate Output Feedback Sliding Mode Control of a Hydraulic Manipulator End Effector.* Retrieved Nov.6, 2010, from http://www.icgst.com/aras/Volume9/Issue1/P1130941891.pdf

Welch, G. & Bishop, G. (2006). *An Introduction to the Kalman Filter.* Retrieved June 17, 2010, from http://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf

Wikipedia. (2010). *Variable Structure Systems.* Retrieved February 03, 2010, from http://en.wikipedia.org/wiki/Variable_structure_system

Xiao, L., Su, H. & Chu, J. Sliding mode prediction based control algorithm for discrete-time nonlinear uncertain coupled systems. *International journal of control, 80(10),* 1616-1625, 2007

Xiong, H., Bian, X., Chang, Z. & Zou, H. Integral Variable Structure Controller for Autonomous Underwater Vehicle. (in Chinese). *Proceedings of the World Congress on Intelligent Control and Automation (WCICA), 6,* 4976-4979, 2004

Young, K. K. D. Design of Variable Structure Model-Following Control Systems. *IEEE Transactions on Automatic Control, 23(6),* 1079-1085, 1978

Young, K. D., Utkin, V. I. & Ozguner, U. A Control Engineer's Guide to Sliding Mode

Control. *IEEE Transactions on Control Systems Technology, 7(3),* 1999

Yu, H. & Lloyd, S. Variable structure adaptive control of robot manipulator. *IEE proceedings on control theory and applications, 144(2),* 167-176, 1997

Yu, X. & Kaynak, O. Sliding-Mode Control with Soft Computing. *IEEE Transactions on Industrial Electronics, 56(9),* 3275-3285, 2009

Zambada, J. (2010). *Sensorless Field Oriented Control of PMSM Motors*. Retrieved December 10, 2010, from http://ww1.microchip.com/downloads/en/AppNotes/01078A.pdf

Zheng, Y., Dimirovski, G. M., Jing, Y. & Yang, M. Discrete-Time Sliding Mode Control of Nonlinear Systems. *American Control Conference, 9(13),* 3825 – 3830, 2007