

THE GENERAL TOEPLITZ/OBSERVABILITY SVSF

**THE GENERAL TOEPLITZ/OBSERVABILITY SMOOTH
VARIABLE STRUCTURE FILTER**

By

MOHAMMAD A. AL-SHABI

Master Degree of Science - Mechanical Engineering - Jordan University of Science &
Technology - Jordan

Bachelor Degree of Science - Mechanical Engineering/Mechatronics - Jordan University
of Science & Technology - Jordan

A Thesis Submitted to
College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Mechanical Engineering
McMaster University

© Copyright by Mohammad Al-Shabi, February 2011

Doctor of Philosophy (2011)

McMaster University

(Department of Mechanical Engineering)

Hamilton, Ontario, Canada

TITLE: The General Toeplitz/Observability Smooth Variable Structure Filter

AUTHOR: Mohammad A. Al-Shabi, MSc. (Jordan University of Science & Technology – Jordan)

SUPERVISOR: Professor S. Habibi

NUMBER OF PAGES: 273, xvii

Abstract

Fault detection can be used to improve the reliability and safety of industrial system. An important component of fault detection is filtering for parameter and state estimation. The challenge is to create a filter that is robust and stable in light of modeling uncertainties and parametric changes due to fault conditions.

The Smooth Variable Structure Filter (SVSF) is a recently defined predictor-corrector filter based on the Sliding Mode Control concepts. The SVSF defines a hyperplane and then applies a discontinuous corrective action that forces the estimate to go back and forth cross that plane. The discontinuous action results in chattering. To overcome chattering, the SVSF uses a saturated function with an associated fixed-width smoothing boundary layer. The SVSF is useful in applications that require robustness due to modeling uncertainties.

In this thesis, the SVSF concepts are explored and further investigated. The chattering signals are used to establish a monitoring and reconstructing algorithm that can be used to detect and extract changes and added uncertainties in the system. However, its ability to determine the source of the added uncertainty is limited to cases involving abrupt step changes.

There are limitations to the SVSF due to the use of the Luenberger method in terms of sensitivity to noise and modeling errors. A novel strategy using the Toeplitz and the Observability matrices is proposed to overcome the SVSF's limitations. This strategy is generalized to high order systems with multiple measurements using new proposed the General System Toeplitz and the General Observability matrices. This strategy is linked to the SVSF and improves its performance in terms of robustness and accuracy.

A novel parameter estimation technique referred to as the Iterative Bi-Section/Shooting Method (IBSS) is derived and is linked to the SVSF to estimate model parameters and states for systems in which only the model structure is known.

The benefits of the proposed estimation methods are demonstrated by using example studies involving an electro-hydrostatic actuator proposed in (Habibi, 2007) and a three-degree of freedom mass-spring-damper system.

Acknowledgements

The author expresses his gratitude to his supervisor, Dr. S. Habibi for his guidance, advice and financial support during the research and the writing of this thesis. The committee members; Dr. G. Bone, Dr. S. C. Veldhuis and Dr. S. Sirouspour are also acknowledged for their kindness, encouragement and understanding during regular committee meetings. Special thanks to Philadelphia University, McMaster University committee and the Green Auto Power Train project group for their financial support.

Special thanks must be conveyed to my wife, Amana Foudeh, for her endless moral and material support during my studies, research and writing of the thesis. My heartfelt thanks to my daughter, Shahed, for making these periods enjoyable and for her encouragements.

The author expresses his gratitude to his family, his family in law and his friends especially Andrew Gadsden for their encouragement and supports.

I dedicate this work to my mother, Mayyadah Jabri, who passed away in October, 30th, 2006 during my PhD program.

Table of Contents

Abstract.....	iii
Acknowledgements	iv
Table of Contents	v
List of Figures.....	x
List of Tables	xvi
Chapter One:	
Introduction.....	1
1.1. Introduction to estimation	1
1.2. Historical background	2
1.3. Problem formulation.....	3
1.4. Hypothesis, objectives, contributions and novelties	5
1.4.1. Research objectives and hypothesis	5
1.4.2. Contributions and novelties	6
1.5. Thesis outline	7
Chapter Two:	
Literature Review Part I – The Wiener and Kalman Filters.....	9
2.1 Introduction	9
2.2 The Wiener Filter	9
2.3 The Kalman Filter	10
2.3.1 Improving the KF robustness due to modeling error	12
2.3.1.1 Adding fictitious process noise.....	12
2.3.1.2 Using fading memory	13
2.3.2 Improving the KF performance numerically	13
2.3.2.1 Joseph Stabilized KF.....	14
2.3.2.2 Forcing the covariance matrix to be symmetric	14
2.3.2.3 The Sequential KF.....	14
2.3.2.4 The Information Filter	15
2.3.3 The Kalman Filter with steady state gain.....	16
2.3.4 Modifying the KF to make it applicable to non-linear systems.....	18
2.3.4.1 The Perturbation Kalman Filter	19
2.3.4.2 The Extended Kalman Filter	20
2.3.4.3 The Iterated Extended Kalman Filter	20
2.3.4.4 Higher Order Extended Kalman Filter.....	21
2.3.4.5 The Sigma-point Kalman Filter	22
2.3.5 Combination of the KF with intelligent techniques.....	27

2.3.5.1 Carrasco's Fuzzy Observer/EKF method.....	27
2.3.5.2 Simon's Fuzzy /KF method	28
2.3.5.3 Matia's Fuzzy /KF method	29
2.3.6 Using adaptive techniques.....	29
2.3.6.1 Noise level adjustment	29
2.3.6.2 Noise level switching.....	30
2.3.6.3 Multiple Models (hybrid system)	30
2.3.6.4 Model Switching	32
2.3.6.5 Interacting Multiple Models.....	33
Chapter Three:	
Literature Review Part II - Sliding Mode Observers	35
3.1. Introduction to Variable Structure and Sliding Mode Control and Systems..	35
3.2. The Sliding Mode Control	36
3.3. Sliding Mode Observer	37
3.3.1. Introduction	37
3.3.2. SMO principles.....	38
3.4. Discontinuous observer	40
3.5. Observer with linear and discontinuous injection.....	42
3.5.1. Slotine et al. Observer.....	43
3.5.2. The Walcott and Zak Observer	47
3.5.3. Convex Parameterization/Tan and Edwards's Observer	48
3.6. SMO applications	51
3.6.1. Fault detection and isolation	51
3.6.2. Sensorless control scheme.	52
3.6.3. Parameter estimation.....	53
Chapter Four:	
The Smooth Variable Structure Filter for Systems with Full Ranked Measurement	
Matrix.....	55
4.1 Historical and mathematical background.....	55
4.1.1 The Variable Structure Filter	55
4.1.2 The Extended Variable Structure Filter	57
4.1.3 The Smooth Variable Structure Filter	58
4.2 The SVSF's stability and gain derivations	60
4.2.1 The Lyapunov stability theorem.....	60
4.2.2 The derivation of the SVSF's gain.....	61

4.3	Application of the SVSF to systems with a measurement matrix of full rank.	63
4.3.1	The SVSF's existence subspaces	63
4.3.1.1.	The a priori existence subspace	64
4.3.1.2.	The a posteriori existence subspace	67
4.3.2	The smoothing boundary layer	68
4.3.2.1	The time varying smoothing boundary layer	74
4.3.3	The chattering amplitudes and its information content	82
4.3.3.1	Obtaining the chattering contents for non-stationary input using the average	83
4.3.3.2	Obtaining the chattering contents for stationary random input with zero mean using the covariance	85
4.3.3.3	Obtaining the chattering contents for non-stationary biased input with zero mean over the segment using the cross-correlation	87
4.3.4	Application of the SVSF to an electro-hydrostatic actuator that has a full-rank measurement matrix	89
4.3.4.1.	The SVSF with time-varying smoothing boundary layer	91
4.3.4.1.1.	Simulation setup	91
4.3.4.1.2.	Simulation results	93
4.3.4.2.	The SVSF with tuning process based on information contents in chattering	97
4.3.4.2.1.	Simulation setup	97
4.3.4.2.2.	Simulation results	99

Chapter Five:

	The Toeplitz/Observability Smooth Variable Structure Filters for Systems with Measurement Matrix that is not Full Rank	107
5.1	Historical and mathematical background	107
5.1.1	The SVSF algorithm for linear systems with partially ranked measurement matrices	107
5.1.2	The SVSF algorithm for non-linear systems with partially ranked measurement matrix	110
5.2	The Toeplitz/Observability SVSF	112
5.2.1.	Introduction to the Toeplitz and Observability matrices	112
5.2.2.	The Toeplitz/Observability SVSF	114
5.2.3.	Application of the Toeplitz/Observability SVSF into an electro-hydrostatic actuator	116

5.2.3.1.	Simulation results	117
5.2.3.1.1.	Results using an accurate model (No modeling errors).....	117
5.2.3.1.2.	Results using a non-accurate model	121
5.3	The General Toeplitz/Observability SVSF	125
5.3.1	Introduction to canonical forms representation	125
5.3.2	Discrete observability canonical form.....	126
5.3.3	General Observability Canonical Form.....	128
5.3.3.1.	The General Observability matrix	132
5.3.3.1.1.	Derivation of the General Observability matrix	135
5.3.3.2.	The selection of the General Observability matrix for overlapped information	146
5.3.3.2.1.	The General Observability matrix's conditions.....	149
5.3.3.2.2.	The General Observability matrix's – Selection procedure	150
5.3.3.2.3.	The General Observability matrix's – Modified selection procedure	152
5.3.3.3.	The General System and the General System Noise Toeplitz matrices	162
5.4	Measurement matrix treatment.....	165
5.5	The General Toeplitz/Observability SVSF concepts for systems with measurement matrices that have partial rank	169
5.5.1.	The General Toeplitz/Observability SVSF	169
5.5.2.	The General Toeplitz/Observability SVSF's existence subspaces.....	171
5.5.2.1.	The a priori existence subspace for the General Toeplitz/ Observability SVSF	171
5.5.2.2.	The a posteriori existence subspace for the General Toeplitz /Observability SVSF	172
5.5.3.	The smoothing boundary layer.....	173
5.5.4.	The chattering amplitudes and its information content.....	174
5.5.5.	Application of the General Toeplitz/Observability SVSF to a three degrees of freedom system	179
5.5.5.1.	Simulation setup	179
5.5.5.2.	Simulation results	182
Chapter Six:		
Iterative Bi-Section/Shooting Method for Parameter Estimation.....		212
6.1.	Introduction	212
6.2.	The Bi-Section Method	212

6.3. The Shooting Method.....	214
6.4. The Iterative Bi-Section/Shooting Method	216
6.5. The Iterative Bi-Section/Shooting Method combined with the SVSF and the KF	223
6.5.1. The Iterative Bi-Section/Shooting KF	223
6.5.2. The Iterative Bi-Section/Shooting SVSF.....	224
6.6. Application of the IBSS/SVSF and the IBSS/KF to an electro-hydrostatic actuator	225
6.6.1. Simulation results.....	227
6.6.1.1. The IBSS/KF	227
6.6.1.2. The IBSS/SVSF	234
6.6.1.3. Discussion	241
Chapter Seven:	
Summary and Conclusion Remarks.....	245
Appendix A:	
Nomenclature	249
Bibliography	255

List of Figures

Fig 2.1: The Wiener Filter, (Simon, 2006).	10
Fig 2.2: The Kalman Filter Steps, (Simon, 2006).....	11
Fig 2.3: The Sequential Kalman Filter (Simon, 2006).....	15
Fig 2.4: The Sequential KF Algorithm, [(Simon, 2006) and (Li & Li, 2007)]......	15
Fig 2.5: The information Filtering Algorithm, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]......	16
Fig 2.6: The Perturbation Kalman Filter, (Negenborn, 2003).	19
Fig 2.7: The Extended Kalman Filter, (Simon, 2006).	20
Fig 2.8: The Iterated Kalman Filter, (Simon, 2006).	21
Fig 2.9: Sigma-Points for $n = 2$, (Van Der Merwe & Wan, 2004).....	22
Fig 2.10: (a) The actual system states and their nonlinear measurement (b) The Sigma-Point KF's estimates, (Van Der Merwe & Wan, 2004).....	23
Fig 2.11: The Unscented Kalman Filter, (Simon, 2006).....	24
Fig 2.12: The Sigma-Point Central Difference Kalman Filter, [(Van Der Merwe, 2004) and (Van Der Merwe, 2004)].	27
Fig 2.13: The Carrasco's Fuzzy nonlinear model, (Carrasco, Cipriano, & Carelli, 2005).	28
Fig 2.14: The Fuzzy membership function, (Carrasco, Cipriano, & Carelli, 2005).	28
Fig 2.15: The Carrasco's Fuzzy Kalman Filter, (Carrasco, Cipriano, & Carelli, 2005). ...	28
Fig 2.16: The Multiple Model algorithm, (Bar-Shalom, Li, & Kirubarajan, 2001).	32
Fig 2.17: The Model Switching Approach for $r = 2$ at time k , (Chang & Athans, 1978).	32
Fig 2.18: The Interacting Multiple Model Algorithm, (Li & Bar-Shalom, 1994).	34
Fig 3.1: (a) Mechanical System with Coulomb Friction, (b) Discontinuously Coulomb Friction, (Utkin, Guldner, & Shi, 1999).....	35
Fig 3.2: (a) Ship Model and its states, (b) Sliding Mode Control of Ship's Course, (Utkin, Guldner, & Shi, 1999).....	36
Fig 3.3:(a) Unstable 2nd order Subsystem, (b) SMC effectiveness, (Utkin, Guldner, & Shi, 1999).	36
Fig 3.4: Real Discontinuous Control Signal – Chattering, (Utkin, Guldner, & Shi, 1999).	37
Fig 3.5: Sliding Mode Observer's Phases.	39

Fig 3.6: The Slotine et al. Observer behaviour for $\mathbf{K}_{l,1} = \mathbf{K}_{l,2} = \mathbf{K}_{s,1} = 0$ and $\mathbf{K}_{s,2} \geq \Delta\mu$.	44
Fig 3.7: The Slotine et al. Observer behaviour for $\mathbf{K}_{s,1} \neq 0$.	44
Fig 3.8: The effect of the gain $\mathbf{K}_{s,1}$.	45
Fig 3.9: The effect of the gain $\mathbf{K}_{l,1}$.	46
Fig 3.10: The effect of the gain $\mathbf{K}_{l,2}$.	46
Fig 3.11: Slosh mass example, (Bandyopadhyay, Gandhi, & Kurode, 2009).	53
Fig 4.1: The Definition of the Existence Subspace, (Habibi I, 2008).	64
Fig 4.2: The Existence Subspace versus time.	64
Fig 4.3: The Smoothing Boundary Layer Effects on the Convergence Rate.	71
Fig 4.4: The a Priori Chattering as an indicator of parametric changes.	74
Fig 4.5: The time varying SVSF Structure.	75
Fig 4.6: The SVSFs with time-varying smoothing boundary layer including the effects of modeling errors.	81
Fig 4.7: Multi-step input.	88
Fig 4.8: The components of the EHA proposed in (Habibi & Burton, 2007).	90
Fig 4.9: The prototype of the EHA proposed in (Habibi & Burton, 2007).	90
Fig 4.10: The input signal.	92
Fig 4.11: The first 300 points of (a) x_1 's actual and estimated values, (b) estimation error of x_1 , for the SVSF with time-varying smoothing boundary layer.	94
Fig 4.12: The first 300 points of (a) x_2 's actual and estimated values, (b) estimation error of x_2 , for the SVSF with time-varying smoothing boundary layer.	95
Fig 4.13: The first 300 points of (a) x_3 's actual and estimated values, (b) estimation error of x_3 , for the SVSF with time-varying smoothing boundary layer.	96
Fig 4.14: The input vs. time for the tuning process simulation.	98
Fig 4.15: (a) x_1 's actual and estimated values, (b) estimation error of x_1 for the simulation of the tuning process obtained by using the information contained in the chattering signal.	101
Fig 4.16: (a) x_2 's actual and estimated values, (b) estimation error of x_2 for the simulation of the tuning process obtained by using the information contained in the chattering signal.	102
Fig 4.17: (a) x_3 's actual and estimated values, (b) estimation error of x_3 for the simulation of the tuning process obtained by using the information contained in the chattering signal.	103

Fig 4.18: (a) a_{32} 's actual and estimated values, (b) estimation error of a_{32} for the simulation of the tuning process obtained by using the information contained in the chattering signal.	104
Fig 4.19: (a) a_{33} 's actual and estimated values, (b) estimation error of a_{33} for the simulation of the tuning process obtained by using the information contained in the chattering signal.	105
Fig 4.20: (a) b_3 's actual and estimated values, (b) estimation error of b_3 for the simulation of the tuning process obtained by using the information contained in the chattering signal.	106
Fig 5.1: (a) x_1 's actual and estimated values, (b) estimation error of x_1 for the Toeplitz/Observability SVSF and the KF when there are no modeling errors.....	118
Fig 5.2: (a) x_2 's actual and estimated values, (b) estimation error of x_2 for the Toeplitz/Observability SVSF and the KF when there are no modeling errors.....	119
Fig 5.3: (a) x_3 's actual and estimated values, (b) estimation error of x_3 for the Toeplitz/Observability SVSF and the KF when there are no modeling errors.....	120
Fig 5.4: (a) x_1 's actual and estimated values, (b) estimation error of x_1 for the Toeplitz/Observability SVSF and the KF when modeling errors are present.	122
Fig 5.5: (a) x_2 's actual and estimated values, (b) estimation error of x_2 for the Toeplitz/Observability SVSF and the KF when modeling errors are present.	123
Fig 5.6: (a) x_3 's actual and estimated values, (b) estimation error of x_3 for the Toeplitz/Observability SVSF and the KF when modeling errors are present.	124
Fig 5.7: Example 2 – System with three degrees of freedom.	131
Fig 5.8: The modified selection procedure of the General Observability matrix - Step 1: dividing the system matrix into blocks	153
Fig 5.9: The modified selection procedure of the General Observability matrix - Step 2: dividing the state vector into sub-vectors according to step 1 blocks.....	154
Fig 5.10: The modified selection procedure of the General Observability matrix - Step 3: Assigning the measurements sub-vectors	155
Fig 5.11: The modified selection procedure of the General Observability matrix - Step 4: Obtaining the initial measurement segments lengths.....	156
Fig 5.12: The iterative procedure of increasing the segment length in order to obtain the General Observability matrix with full rank.....	159
Fig 5.13: Transforming the measurement matrix to equation (5.53)	168

Fig 5.14: The time steps of the measurement and system noise vectors that are used to obtain the alternative measurement vector at time $k_1 - 1$ compared to the measurement and system noise vectors at time $k - 1$	177
Fig 5.15: (a) x_{11} 's actual and estimated values, (b) estimation error of x_{11} for the General Toeplitz/Observability SVSF's tuning process.....	185
Fig 5.16: (a) x_{12} 's actual and estimated values, (b) estimation error of x_{12} for the General Toeplitz/Observability SVSF's tuning process.....	186
Fig 5.17: (a) x_{21} 's actual and estimated values, (b) estimation error of x_{21} for the General Toeplitz/Observability SVSF's tuning process.....	187
Fig 5.18: (a) x_{22} 's actual and estimated values, (b) estimation error of x_{22} for the General Toeplitz/Observability SVSF's tuning process.....	188
Fig 5.19: (a) x_{31} 's actual and estimated values, (b) estimation error of x_{31} for the General Toeplitz/Observability SVSF's tuning process.....	189
Fig 5.20: (a) x_{32} 's actual and estimated values, (b) estimation error of x_{32} for the General Toeplitz/Observability SVSF's tuning process.....	190
Fig 5.21: (a) a_{21} 's actual and estimated values, (b) estimation error of a_{21} for the General Toeplitz/Observability SVSF's tuning process.....	191
Fig 5.22: (a) a_{22} 's actual and estimated values, (b) estimation error of a_{22} for the General Toeplitz/Observability SVSF's tuning process.....	192
Fig 5.23: (a) a_{23} 's actual and estimated values, (b) estimation error of a_{23} for the General Toeplitz/Observability SVSF's tuning process.....	193
Fig 5.24: (a) a_{24} 's actual and estimated values, (b) estimation error of a_{24} for the General Toeplitz/Observability SVSF's tuning process.....	194
Fig 5.25: (a) a_{25} 's actual and estimated values, (b) estimation error of a_{25} for the General Toeplitz/Observability SVSF's tuning process.....	195
Fig 5.26: (a) a_{26} 's actual and estimated values, (b) estimation error of a_{26} for the General Toeplitz/Observability SVSF's tuning process.....	196
Fig 5.27: (a) a_{41} 's actual and estimated values, (b) estimation error of a_{41} for the General Toeplitz/Observability SVSF's tuning process.....	197
Fig 5.28: (a) a_{42} 's actual and estimated values, (b) estimation error of a_{42} for the General Toeplitz/Observability SVSF's tuning process.....	198
Fig 5.29: (a) a_{43} 's actual and estimated values, (b) estimation error of a_{43} for the General Toeplitz/Observability SVSF's tuning process.....	199
Fig 5.30: (a) a_{44} 's actual and estimated values, (b) estimation error of a_{44} for the General Toeplitz/Observability SVSF's tuning process.....	200

Fig 5.31: (a) a_{45} 's actual and estimated values, (b) estimation error of a_{45} for the General Toeplitz/Observability SVSF's tuning process.....	201
Fig 5.32: (a) a_{46} 's actual and estimated values, (b) estimation error of a_{46} for the General Toeplitz/Observability SVSF's tuning process.....	202
Fig 5.33: (a) a_{61} 's actual and estimated values, (b) estimation error of a_{61} for the General Toeplitz/Observability SVSF's tuning process.....	203
Fig 5.34: (a) a_{62} 's actual and estimated values, (b) estimation error of a_{62} for the General Toeplitz/Observability SVSF's tuning process.....	204
Fig 5.35: (a) a_{63} 's actual and estimated values, (b) estimation error of a_{63} for the General Toeplitz/Observability SVSF's tuning process.....	205
Fig 5.36: (a) a_{64} 's actual and estimated values, (b) estimation error of a_{64} for the General Toeplitz/Observability SVSF's tuning process.....	206
Fig 5.37: (a) a_{65} 's actual and estimated values, (b) estimation error of a_{65} for the General Toeplitz/Observability SVSF's tuning process.....	207
Fig 5.38: (a) a_{66} 's actual and estimated values, (b) estimation error of a_{66} for the General Toeplitz/Observability SVSF's tuning process.....	208
Fig 5.39: (a) b_2 's actual and estimated values, (b) estimation error of b_2 for the General Toeplitz/Observability SVSF's tuning process.....	209
Fig 5.40: (a) b_4 's actual and estimated values, (b) estimation error of b_4 for the General Toeplitz/Observability SVSF's tuning process.....	210
Fig 5.41: (a) b_6 's actual and estimated values, (b) estimation error of b_6 for the General Toeplitz/Observability SVSF's tuning process.....	211
Fig 6.1: The Bi-Section's Principle, (Kaw & Kalu, 2010).....	212
Fig 6.2: The Bi-Section Steps, (Kaw & Kalu, 2010).	213
Fig 6.3: Reducing the solution interval using Bi-Section method.....	214
Fig 6.4: Multiple zeros and interval effects.	214
Fig 6.5: Adjusting the angle of the Canon Muzzle manually to hit a target.	215
Fig 6.6: The Shooting method for one dimensional ordinary differential equation problem.	215
Fig 6.7: The affection of Modeling errors on the extracted parameter, ξ , for a second order system.	217
Fig 6.8: Refining the interval by Bisection Method.	217
Fig 6.9: The IBSS algorithm for second order system.....	222
Fig 6.10: The IBSS/EKF algorithm for second order systems.	224

Fig 6.11: The IBSS/SVSF algorithm.	225
Fig 6.12: The input signal to the IBSS simulation.	227
Fig 6.13: (a) x_1 's actual and estimated values, (b) estimation error of x_1 obtained by using the IBSS/KF.	228
Fig 6.14: (a) x_2 's actual and estimated values, (b) estimation error of x_2 obtained by using the IBSS/KF.	229
Fig 6.15: (a) x_3 's actual and estimated values, (b) estimation error of x_3 obtained by using the IBSS/KF.	230
Fig 6.16: (a) ξ 's actual and estimated values, (b) estimation error of ξ obtained by using the IBSS/KF.	231
Fig 6.17: (a) ω_n 's actual and estimated values, (b) estimation error of ω_n obtained by using the IBSS/KF.	232
Fig 6.18: (a) gain's actual and estimated values, (b) estimation error of gain obtained by using the IBSS/KF.	233
Fig 6.19: (a) x_1 's actual and estimated values, (b) estimation error of x_1 obtained by using the IBSS/SVSF.	235
Fig 6.20: (a) x_2 's actual and estimated values, (b) estimation error of x_2 obtained by using the IBSS/SVSF.	236
Fig 6.21: (a) x_3 's actual and estimated values, (b) estimation error of x_3 obtained by using the IBSS/SVSF.	237
Fig 6.22: (a) ξ 's actual and estimated values, (b) estimation error of ξ obtained by using the IBSS/SVSF.	238
Fig 6.23: (a) ω_n 's actual and estimated values, (b) estimation error of ω_n obtained by using the IBSS/SVSF.	239
Fig 6.24: (a) gain's actual and estimated values, (b) estimation error of gain obtained by using the IBSS/SVSF.	240
Fig 6.25: The error when changes happen within the segment.	243
Fig 6.26: Comparison between the IBSS/KF and the IBSS/SVSF in term of the damping ratio profile.	244

List of Tables

Table 2.1: The differences between the UKF methods, (Simon, 2006).....	25
Table 4.1: The differences between chattering of SMO and the SVSF a priori chattering.	67
Table 4.2: The differences and similarity between chattering of SMO and the SVSF's a posteriori chattering.	68
Table 4.3: The parameters of the EHA proposed in (Habibi & Burton, 2007).....	91
Table 4.4: The root mean square error and the maximum absolute error of the estimated states.....	93
Table 4.5: The beginning and ending of each region for the tuning process simulation. ..	97
Table 4.6: The % of the error between the system and the filter parameters for the tuning process simulation.....	98
Table 4.7: The estimated beginning of each region for the tuning process simulation. ..	100
Table 4.8: The root mean square error of the tuning process simulation.....	100
Table 5.1: Actual and Estimated system parameters.	117
Table 5.2: Comparison between the Toeplitz/Observability SVSF and the KF for a known estimated model.	117
Table 5.3: Comparison between the Toeplitz/Observability SVSF and the KF for model with parameters uncertainties.	121
Table 5.4: The system, input and measurement matrices' of the continuous observer, controller, observability and controllability canonical forms.	125
Table 5.5: The observability and Toeplitz matrices' structures of the observer, controller, observability and controllability canonical forms.....	126
Table 5.6: The rank of the General Observability matrix of Example 5.2 given six states and one measurement; i.e. $\mathbf{H} = \mathbf{H}_1$	146
Table 5.7: The rank of the General Observability matrix for example 2 given six states and two measurements i.e. $\mathbf{H} = [\mathbf{H}_1 \ \mathbf{H}_2]^T$	147
Table 5.8: The rank of the General Observability matrix of example 2 given six states and three measurements i.e. $\mathbf{H} = [\mathbf{H}_1 \ \mathbf{H}_2 \ \mathbf{H}_3]^T$	148
Table 5.9: The General Observability matrix and its rank for example 1 given three states and one measurement.....	149
Table 5.10: The beginning and the ending of the changes.....	181

Table 5.11: The % of the error between the system and the filter parameters for the simulation regions.....	181
Table 5.12: The estimated beginning of each region for the tuning process simulation.	183
Table 5.13: The root mean square error of the estimated states and parameters obtained from the General Toeplitz/Observability SVSF's tuning process.	183
Table 6.1: The new interval boundaries using the Bi-Section method.....	221
Table 6.2: Comparison between the IBSS/EKF and the IBSS/SVSF.....	242
Table A.1 - Nomenclature.....	254

Chapter One:

Introduction

1.1. Introduction to estimation

Generally, a system's dynamics can be captured by physical models that consist of *parameters* that reflect the system's make up, e.g. mass, elasticity, capacitance and resistance. *States* describe the operation and the dynamic behaviour of the system given an input and/or a disturbance. The latter term is referred to as the *system (process) noise*. To study the system, to describe it and to control it, certain knowledge is needed regarding its parameters and states. The source of this knowledge can be *prior knowledge* and *empirical characterization*. The former one is related to the available information about the system such as its order and structure, and its environment (e.g. the system noise distribution). The empirical knowledge is a model generated using measured signals or *measurements*. Due to the sensor limitations, measurements may contain *measurement noise*. The system and measurement noise are stochastic signals that can be characterized by probability distribution functions with their exact values at a particular time being unknown. However, it is possible to extract both states and the parameters from measurements using a process referred to as *estimation*, which combines the prior and the empirical knowledge [(Van Der Heidan, Duin, de Ridder, & Tax, 2004), (Habibi, 2005), (Grewal & Andrews, 2001) and (Bar-Shalom, Li, & Kirubarajan, 2001)]. The estimation process can be divided into three basic categories, (Haykin, 2002):

- 1- Filtering: The filtering process is used to extract the quantity of interest (e.g. state) at a current time (t) from noisy measurements obtained in the past up to and including the current value.
- 2- Smoothing: The smoothing process is used to obtain a better estimated value for the quantity of interest by processing an entire segment of the measurement consisting of past, current and future measured data points (i.e. start at $t - \tau$ and end at $t + \tau$). Therefore, it is not suitable for real-time application.
- 3- Prediction: The aim of the prediction process is to provide information about the quantity of interest at some time in the future ($t + \tau$). This information is obtained by using the past measured data points up to and including the current point.

The estimation process is obtained by an *estimator* that may be model-based. Applying the estimator provides an opportunity to extract more information from a signal. If the signal contains noise, then the estimator needs to filter-out the noise, henceforth referred to as *filter* (Grewal & Andrews, 2001). Filters are widely used in signal

processing, communications, image processing, radar tracking, satellite systems, aerospace applications, weather forecasting, economics, navigation, dynamic positioning tracking systems, fault detection and monitoring, and prediction [(Welch & Bishop, 2006) and (Haykin, 2002)].

Several filtering techniques have been developed in order to cover numerous applications. These techniques share the same purpose of estimating the information content in a signal while reducing noise effects. However, each method uses different criteria to improve the estimate, i.e. optimality, robustness and stability. The creation and/or development of the estimation techniques generally require the following:

- Understanding the application, and identifying and defining the goal to be achieved: for example, radar tracking problems are different than fault detection problems. The main goal of the former is to track the object and predict its trajectory with an acceptable level of accuracy. As such, the filter must have some level of optimality. While in the latter, early detection of faults and stability are the primary concerns. As such, the filter must have some level of robustness.
- Defining the input and the output of the system, and the form of uncertainties.
- Modeling the system, such as finding its transfer function, based on an a priori understanding of the system: the estimators and the filters considered in this thesis are *model-based*. Modeling and model integrity are therefore important considerations.
- Choosing the filter that best fits the application: some filters may not be suitable for fast real-time applications due to their computational complexity. Others may not be applied to systems that change their structure and/or parameters because of lack of robustness and possible instabilities.

1.2. Historical background

“From earliest time, people have been concerned with interpreting observations and making estimates and predictions.” Kailath said in (Kailath, 1974) and continued by “Neugebauer has noted that the Babylonians used a rudimentary form of Fourier series for such purposes.” The beginning of the “theory of estimation” can be tracked back to 1632, when Galileo tried to minimize the error of some functions (Kailath, 1974). In 1795, Gauss introduced and used the method of *least squares* to locate the asteroid Ceres, although Legendre first published it independently in 1805 [(Kailath, 1974) and (Sorenson, 1970)]. These were followed by numerous investigations and studies pertaining to the least squares method, leading to the pioneering work of Wiener in the

1940's [(Kailath, 1974), (Chen, 2003) and (Simon, 2006)]. In 1942, Wiener gave the first explicit solutions for the problem of estimating a stochastic process using the least squares method, referred to as the *Wiener Filter* [(Kailath, 1974), (Chen, 2003), (Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]. In (Kailath, 1974), Kailath gave a brief history of the estimation problem focusing on the period from 1930s to 1960s. In that article, he summarized the history of optimal filters up to the *Kalman Filter* (KF) which was one of the earliest filters that was implemented in a predictor-corrector form.

The history of estimation was continued in (Chen, 2003) focusing on the period from 1960s to 2000s. He specifically considered the non-linear *Bayesian filtering* problems that include the *Particle Filter*. The Bayesian approach was applied to estimation in stochastic processes in 1964 by Ho and Lee [(Chen, 2003) and (Ho & Lee, 1964)]. They introduced the iterative Bayesian filtering and explored the concepts of “the sequential state estimation problem” (Chen, 2003). Later in the 1980's and the 1990's, the Bayesian filtering was expanded to include the state space structure. During the twentieth century, the “*optimality*” derivatives were formulated specifically in the Bayesian framework. The notion of optimality relies on minimizing a measure referred to as the cost function (Chen, 2003). This type of filter seeks optimality and is discussed in chapter two.

Simultaneously with the development of the optimal filters, another type of filter started to rise and take place in estimation utilising the principles of the *Variable Structure System and Control* as well as the *Sliding Mode Control*. These filters and observers are referred to as *Sliding Mode Observers* (SMO). These observers are used widely in fault detection and signal reconstruction problems due to their robustness and stability against uncertainties. Recently, the sliding mode observer has been formulated in a predictor-corrector form referred to as the *Variable Structure Filter*, (Habibi & Burton, 2002). In 2007, a new form of the Variable Structure Filter, referred to as the *Smooth Variable Structure Filter* (SVSF), was proposed, (Habibi, 2007). This type of filter is robust to uncertainties and is discussed further in chapter three. The SVSF is the core of this research. Its principles and concepts are explored mathematically, and its structure is reformulated to accommodate more general cases.

1.3. Problem formulation

Fault detection is an important application for filtering. This application is based on monitoring system parameters and giving an indication (warning) if they exceed their thresholds. Usually system parameters are not directly measured and they are extracted from the measurements; hence filtering is needed. Monitoring the faults is needed to

minimize the equipment (system) breakdown cost, and to improve the safety of the system. One of the most popular approaches for fault detection is by using model-based filters. By applying the system's input to the filter model and comparing the expected output to the output obtained from the system, the fault can be detected, (Tang & Wang, 2010). However, once a fault occurs, large changes in the system behaviour will cause substantial departure from the expected behaviour captured by the filter model, leading to numerical instability. Therefore, the filter (in such applications) must be robust to uncertainties. One of the most robust filters that can be used for these applications are the SMOs and the SVSF.

The SVSF is a recent estimation technique proposed in (Habibi, 2007) for state and parameter estimation. This filter overcomes the KF limitations in terms of robustness, and stability. In (Habibi, 2007), the gain of the SVSF was derived and the filter stability and robustness were proven. Moreover, the basic principles and properties of the filter were established. However, the derived SVSF has some limitations when applied to systems with a measurement matrix of partial rank. The major limitation is the use of the *Luenberger method* that transforms the estimation problem into a simple algebraic problem involving sub-matrices from the system and input matrices (discussed later in chapter five). The Luenberger method has the following limitations:

- It is compatible with an observer rather than a filter as it assumes that the noise does not exist in the input or the output. Therefore, the estimate of the SVSF becomes sensitive to noise amplitude.
- It extracts the quantity of interest (i.e. hidden state/parameter) by using the inverse of the system model (the inversion is done with respect to the quantity of interest) and the measurement. The inversion process is obtained by taking the inverse of some sub-matrices from the system and input matrices. These sub-matrices may not be square matrices, thus the inversion process may consist of a pseudo-inverse operator which adds some limitations to the Luenberger process.

An inherent element of variable structure control and of the SVSF is chattering. To overcome chattering, the SVSF uses a saturated function with an associated *smoothing boundary layer*. The SVSF as published in (Habibi, 2007) uses a smoothing boundary layer that has a fixed width. In this thesis, a time-varying smoothing boundary layer width is proposed for improving performance.

This thesis considers two applications; the first involves a third order linear system referred to the electro-hydrostatic actuator (Habibi & Goldenberg, 2000), and the

second is a three degrees of freedom mass-spring-damper system. Both systems' parameters are made to change from time to time randomly. When the parameters change, the modeling errors increase and the filter stability is affected. This issue is the major limitation of the KF and most other estimation techniques. Some filters have refinement techniques associated with them such as the SMO where the modeling errors can be reconstructed by using the chattering signal. However, these techniques are limited to special uncertainty structures (as discussed later in chapter three), and require filtering of the chattering signal by a low pass filter. Moreover, the SMOs do not consider noise during their processing, a matter that impacts the estimates and the reconstructed signals, [(Edwards, 2004) and (Edwards, Spurgeon, & Patton, 2000)].

To sum up the problem, modeling errors effects must be reduced and the filter's model needs to be updated once a fault occurs. The filter should be able to detect the instance when the change occurs, track the change in parameters while remaining stable.

1.4. Hypothesis, objectives, contributions and novelties

1.4.1. Research objectives and hypothesis

A target application for this thesis is fault detection that requires tracking and monitoring of system parameters. Therefore, the objective of this research can be defined as the development of a state estimation strategy for fault detection that can:

- Extract the full state vector from noisy measurements.
- Detect parametric changes in the system once they occur. In order to do so, the filter must have a tool that is capable of sensing these changes.
- Self-tune the model according to the reconstructed modeling errors.
- Be robust and stable to parametric uncertainties.
- Have a fast convergence rate and be applicable for online applications.
- Have some extra performance criteria; i.e. minimizing the diagonal elements in the error covariance matrix and reduce the states' sensitivity to noise.

In order to achieve these objectives, **a recently proposed estimation strategy referred to as the Smooth Variable Structure Filter will be reformulated with a model representation referred to as the General Toeplitz and Observability matrices. An indicator of performance in the SVSF is a chattering vector associated with the estimated state vector. The information contents of the chattering vector is investigated. A simple algebraic technique using the chattering vector is developed for the adaptation of the filter model. This will allow the filter to accommodate**

parametric variations and to estimate the system's model when changes occur. Further, the filter's gain is associated with a new smoothing boundary layer vector that has a time-varying width to improve performance.

1.4.2. Contributions and novelties

The novel contributions of this thesis are as follows:

- The SVSF concepts are explored in further detail. It is noted that there are two existence subspaces rather than the one proposed in (Habibi, 2007). This research describes these subspaces and develops mathematical formulas for them.
- The characteristics of the existence subspaces are investigated. The relations between the smoothing boundary layer and the estimation errors are established and mathematically derived. A new time-varying smoothing boundary layer is proposed to improve the SVSF's performance.
- It has been noticed that two types of chattering are present in the SVSF associated each with the a priori and the a posteriori estimates. These chattering signals are mathematically formulated and their amplitudes are linked to the source and level of model uncertainties.
- The chattering equations are used to detect the parametric changes once they occur, and then to reconstruct modeling errors and uncertainties, thereby a method for monitoring and refining the model is established.
- System Toeplitz and Observability matrices for systems with multiple-degrees of freedom and/or multiple measurements are proposed. In this thesis, the developed matrices are referred to as the *General System Toeplitz and the General Observability matrices*. These matrices link the available measurements with the state vector. If the system model is represented in a specific canonical form referred to as the *General Observability Canonical Form* (GOCF), then these matrices become devoid of modeling errors, and the quality of the estimates is improved. Moreover, the estimation process becomes robust and stable.
- The SVSF is combined with the General System Toeplitz and Observability matrices to estimate the full state vector. These matrices provide the SVSF with an *alternative measurement vector* that has a measurement matrix with full rank. This process replaces the Luenberger method in the SVSF, thereby the limitations of the SVSF due to the use of the Luenberger method are eliminated. The combined algorithm between the SVSF and the General system Toeplitz and

Observability matrices is referred to as the **General Toeplitz/Observability SVSF** and it has an algorithm that is similar to the SVSF's algorithm for a measurement matrix with full rank.

- A special case of the General Toeplitz/Observability SVSF referred to as the **Toeplitz/Observability SVSF** is obtained for systems with one measurement that are represented in their Observability canonical form.
- The concepts of both the Toeplitz/Observability SVSF and the General Toeplitz/Observability SVSF are explored. The a priori and the a posteriori existence subspaces are mathematically formulated as well as the chattering in the a priori and the a posteriori estimates.
- A novel monitoring and reconstructing method is developed based on the a priori chattering to detect parametric changes, and to tune the filter's model.
- A novel iterative method is proposed based on two numerical techniques, which are the **Bi-Section** and the **Shooting** methods. This method is used to estimate model parameters for systems in which only the model structure is known. This method is referred to as the **Iterative Bi-Section/Shooting method** (IBSS) and is applicable to systems with small orders.
- The IBSS with the SVSF (IBSS/SVSF) are combined for state and parameter estimation. For comparison, the IBSS is combined with the KF (IBSS/KF). The two methods are compared in terms of robustness and accuracy.

1.5. Thesis outline

This chapter provided a brief introduction to estimation. The contributions and novelties of this work were described.

Due to the numerous estimation algorithms developed in the last century, the literature review is divided into two chapters. In chapter two, optimal strategies are considered. These include the **Wiener Filter**, the **Kalman Filter** (KF) and its variants such as the **Steady State KF**, the **Extended KF**, the **Perturbation KF**, and the **Unscented KF**. The chapter also lists some earlier works developed to overcome the KF limitations such as:

- Integrating the KF with an intelligent algorithm; e.g. **Fuzzy logic**.
- The use of an adaptive technique with the KF; e.g. the **Multiple Model**.
- Improving the robustness of the KF; e.g. adding fictitious process noise to system model and using the fading memory with the KF covariance matrices.

Chapter three discusses robust estimation algorithms. These include several forms of the SMO. The chapter discusses the SMO concepts and their limitations and benefits.

Chapter four reviews the recently proposed SVSF. The chapter links the SVSF to the SMO and describes the previous forms of the SVSF such as the Variable Structure Filter (VSF) and the Extended Variable Structure Filter (EVSF). The concepts of the SVSF applied to linear systems with a full rank measurement matrix are explored, and mathematically formulated. Proof of the SVSF stability is provided. A simple algebraic process is developed to detect and then to extract the modeling errors from the chattering signal. A new time-varying smoothing boundary layer is proposed in order to improve the performance of the SVSF. The new boundary layer minimizes the diagonal elements of the error covariance matrix which results in reducing the estimation error. This improves the performance of the SVSF compared to its corresponding performance obtained with a fixed-width smoothing boundary layer. The features obtained from the chattering and the time-varying smoothing boundary layers are tested by applying the SVSF to an electro-hydrostatic actuator, (Habibi, 2007).

New forms of the system Toeplitz and the Observability matrices are developed in chapter five. These new matrices, which are referred to as the General System Toeplitz and the General Observability matrices, are used in conjunction with the SVSF to extract the full state vector from multiple measurements. The chapter derives the conditions that need to be satisfied for eliminating the effects of modeling errors on the estimates. The resultant existence subspaces of this algorithm are explored and mathematically formulated. Similarly to chapter four, a simple algebraic process is derived for monitoring and reconstructing the modeling errors. This process is applied to a three-degree of freedom mass-spring-damper system.

The IBSS is derived in chapter six as a novel iterative parameter estimation technique. The concepts, algorithm, and limitations of this method are discussed. The IBSS is combined with the SVSF in order to be able to estimate the states and model parameters for systems in which only the model structure is known. This method takes its strength from using the SVSF secondary indicator of performance to detect changes resulting in modeling errors.

Chapter seven contains the concluding remarks. It summarizes the benefits and limitations of using the SVSF with time-varying smoothing boundary layer, the General Toeplitz/Observability SVSF, and the IBSS/SVSF.

Chapter Two:

Literature Review Part I – The Wiener and Kalman Filters

2.1 Introduction

This chapter presents a literature review on optimal state and parameter estimation techniques. The notion of optimality relies on minimizing a measure referred to as the cost function, (Chen Z. , 2003). Common optimal methods used in estimation are as follows, [(Chen, 2003), (Simon, 2006), (Bar-Shalom, Li, & Kirubarajan, 2001), (Barakat, 2005) and (Van Der Heidan, Duin, de Ridder, & Tax, 2004)]:

- 1- **Minimum mean squared error** (MMSE) where the error is usually measured between the estimates and their actual values as temporal functions. The cost function is defined as follows:

$$J = (E[(x_k - \hat{x}_k)^T(x_k - \hat{x}_k) | z_{o\dots k}]) \quad 2.1$$

Where x_k and \hat{x}_k are the state and its estimate at time k , and $z_{o\dots k}$ represents the past measurement vectors up to and including the time k .

- 2- **Maximum a posteriori** (MAP), where the target is to find the posterior probability mode (local and global maxima) that minimizes the following cost function:

$$J = \min(E[1 - I_{x_k: \|x_k - \hat{x}_k\| \leq \delta}(x_k)]) \quad 2.2$$

Where δ is a small threshold value and $I_{x_k: \|x_k - \hat{x}_k\| \leq \delta}(x_k)$ is an indicator function that is defined as follows:

$$I_{x_k: \|x_k - \hat{x}_k\| \leq \delta}(x_k) = \begin{cases} 1 & \|x_k - \hat{x}_k\| \leq \delta \\ 0 & \|x_k - \hat{x}_k\| > \delta \end{cases} \quad 2.3$$

The chapter is organized as follows. In section 2.2 the Wiener Filter is reviewed. Section 2.3 discusses the Kalman Filter principles, limitations and various forms of implementations that improve its performance.

2.2 The Wiener Filter

In 1942, Wiener gave the first explicit solution for the problem of state estimation for a stochastic process using the least squares method, referred to as the **Wiener Filter** (WF) as shown in Fig 2.1, [(Kailath, 1974), (Chen, 2003) and (Simon, 2006)].

The solution of the WF is based on scalar observations and stationary signals with known spectral properties subject to white noise. The target of the WF is to extract information from noisy observations using the frequency domain perspective.

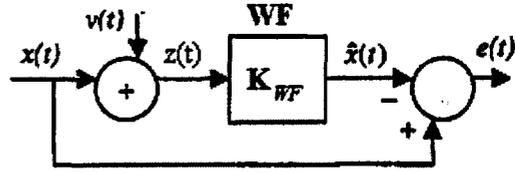


Fig 2.1: The Wiener Filter, (Simon, 2006).

The Wiener Filter algorithm can be summarized as follows, (Simon, 2006).

If the observation $z(t)$ is a function of the noise free signal, $x(t)$, and the noise $v(t)$ such that:

$$z(t) = x(t) + v(t) \quad 2.4$$

then the estimated signal $\hat{x}(t)$ can be obtained as:

$$\hat{x}(t) = K_{WF}(t) * z(t) \quad 2.5$$

Where the terminology $(*)$ is the convolution operator, and $K_{WF}(t)$ is the Wiener gain defined as:

$$K_{WF}(t) = \mathcal{F}^{-1} \left[\frac{S_z - S_v}{S_z} \right] \quad 2.6$$

The terminology (\mathcal{F}^{-1}) means the inverse Fourier Transformation, and S_z and S_v are the Fourier Transforms of the observation and noise autocorrelations (their power spectrums), respectively.

The WF is limited in its application to stationery processes. It was extended to non-stationary processes by Kalman, (Simon, 2006).

2.3 The Kalman Filter

In 1960 Rudolf Kalman presented the *Kalman Filter* (KF), as a recursive, optimal and model based estimator, that falls under the predictor/corrector category for linear systems, [(Bar-Shalom, Li, & Kirubarajan, 2001), (Grewal & Andrews, 2001), (Barker, Brown, & Martin, 1995), (Welch & Bishop, 2006), (Maybeck, 1979) and (Kalman, 1960)]. The KF is an optimal filter for linear Gaussian problems as it minimizes the mean square error between the actual and the estimated state (MMSE). Its gain satisfies two optimality principles; unbiased of the filter (expectation of the state error is zero) and orthogonality between the state's a posteriori error and the measurement's a priori error,

[(Bar-Shalom, Li, & Kirubarajan, 2001), (Anderson & Moore, 1979) and (Arasaratnam, Haykin, & Elliott, 2007)].

The KF is referred to as a predictor/corrector filter because it uses a mathematical model of the system represented in equation (2.7) to obtain an a priori estimate of the state in a period referred to as the prediction step. The a priori estimates represent the noise-free state/parameter at that time step. The a priori estimation errors are due to uncertainties as well as noise effects. The KF then uses the measurements and an optimal gain to refine the a priori estimates to an a posteriori form in what is referred to as an update step. The KF process and equations are given in Fig 2.2.

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}u_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{z}_k &= \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \end{aligned} \quad 2.7$$

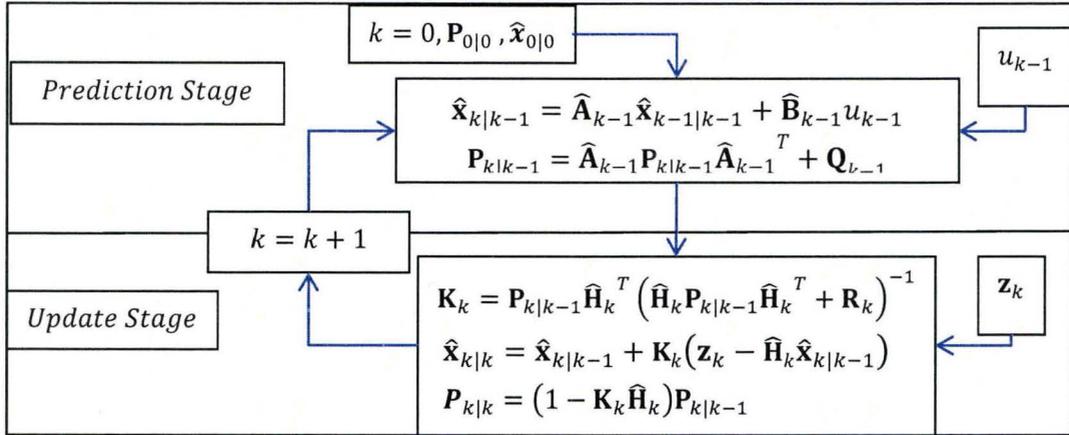


Fig 2.2: The Kalman Filter Steps, (Simon, 2006).

Where \mathbf{A}_{k-1} , \mathbf{B}_{k-1} and \mathbf{H}_k are the system, input and measurement matrices, respectively and their estimated matrices used by the filter are $\hat{\mathbf{A}}_{k-1}$, $\hat{\mathbf{B}}_{k-1}$ and $\hat{\mathbf{H}}_k$, respectively. \mathbf{w}_{k-1} and \mathbf{v}_k are the process (system) and measurement noise vectors, respectively. The subscripts $k|k-1$ and $k|k$ represents the quantity's a priori and the a posteriori value, respectively. \mathbf{K}_k is the KF gain and \mathbf{P} , \mathbf{Q} and \mathbf{R} are the error, process noise and the measurement noise covariance matrices, respectively.

The KF assumes that the system model is known and is linear, system and measurement noise are white, and the states have initial conditions that are modeled as random variables with known means and variances, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]. However, these assumptions do not always hold in real

applications. If one of these assumptions is violated, the KF performance becomes sub-optimal and could potentially become unstable. The KF's limitations can be summarized as follows, [(Simon, 2006), (Fitzgerald, 1971) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

1. It provides only an implicit mechanism for dealing with modeling uncertainties.
2. It assumes that noise is white.
3. It is prone to instability due to numerical errors (i.e., round-off errors and inversions of ill-conditioned matrices).

In the last few decades, many researchers have proposed improvements to the KF. These are presented in the following subsections.

2.3.1 Improving the KF robustness due to modeling error

Modeling errors add uncertainties and impact the stability and optimality of the KF. Taking into account modeling uncertainties, the a priori covariance matrix equation differs from its original equation. Assuming the input u_{k-1} is zero, the a priori and the a posteriori covariance matrices can be obtained as follows, (Bar-Shalom, Li, & Kirubarajan, 2001):

$$\mathbf{P}_{k|k-1} = \left[\begin{array}{c} \overbrace{\widehat{\mathbf{A}}_{k-1} \mathbf{P}_{k-1|k-1} \widehat{\mathbf{A}}_{k-1}^T + \mathbf{Q}_{k-1} + \mathbf{A}_{k-1} \boldsymbol{\sigma}_{x_{k-1}, x_{k-1}} \mathbf{A}_{k-1}^T}^{\text{original } P_{k|k-1}} \\ + \widehat{\mathbf{A}}_{k-1} \boldsymbol{\sigma}_{x_{k-1}, e_{x_{k-1}|k-1}} \Delta \mathbf{A}_{k-1}^T + \Delta \mathbf{A}_{k-1} \boldsymbol{\sigma}_{x_{k-1}, e_{x_{k-1}|k-1}} \widehat{\mathbf{A}}_{k-1}^T \end{array} \right] \quad 2.8$$

and

$$\mathbf{P}_{k|k} = \left[\begin{array}{c} \overbrace{(\mathbf{I} - \mathbf{K}_k \widehat{\mathbf{H}}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \widehat{\mathbf{H}}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T}^{\text{original } P_{k|k}} \\ + \mathbf{K}_k \Delta \widehat{\mathbf{H}}_k \boldsymbol{\sigma}_{x_{k-1}, x_{k-1}} \Delta \widehat{\mathbf{H}}_k^T \mathbf{K}_k^T - (\mathbf{I} - \mathbf{K}_k \widehat{\mathbf{H}}_k) \boldsymbol{\sigma}_{x_{k-1}, e_{x_{k-1}|k-1}} \Delta \widehat{\mathbf{H}}_k^T \mathbf{K}_k^T \\ - \mathbf{K}_k \Delta \widehat{\mathbf{H}}_k \boldsymbol{\sigma}_{x_{k-1}, x_{k-1}} (\mathbf{I} - \mathbf{K}_k \widehat{\mathbf{H}}_k)^T \end{array} \right] \quad 2.9$$

Where $\boldsymbol{\sigma}_{\mathbf{a}, \mathbf{b}}$ is the cross-correlation between \mathbf{a} and \mathbf{b} .

To reduce the effects of modeling errors, various techniques have been proposed. These are outlined in sub-sections (2.3.1.1) and (2.3.1.2).

2.3.1.1 Adding fictitious process noise.

A *fictitious process* noise is added to tune the process noise covariance matrix and adapt the differences between the filter and system models. By setting the process noise

covariance matrix to be $\mathbf{Q}_{new_{k-1}}$ defined in equation (2.10), the filters performance is improved, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)].

$$\mathbf{Q}_{new_{k-1}} \geq \left[\mathbf{Q}_{k-1} + \mathbf{A}_{k-1} \boldsymbol{\sigma}_{\mathbf{x}_{k-1}, \mathbf{x}_{k-1}} \mathbf{A}_{k-1}^T + \widehat{\mathbf{A}}_{k-1} \boldsymbol{\sigma}_{\mathbf{x}_{k-1}, \mathbf{e}_{\mathbf{x}_{k-1}|k-1}} \Delta \mathbf{A}_{k-1}^T + \Delta \mathbf{A}_{k-1} \boldsymbol{\sigma}_{\mathbf{x}_{k-1}, \mathbf{e}_{\mathbf{x}_{k-1}|k-1}} \widehat{\mathbf{A}}_{k-1}^T \right] \quad 2.10$$

The tuning is done by trial and error because the modeling errors are usually unknown. One of this method's limitations is that if there is a modeling error in the input matrix and the input is valid, the resultant covariance matrices depend on the input properties which may violate the assumption of Gaussian noise, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)].

2.3.1.2 Using fading memory

The a priori covariance equation in the KF is modified by multiplying the term that includes $\mathbf{P}_{k-1|k-1}$ with positive constant, δ , that is slightly larger than one, as follows, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

$$\mathbf{P}_{k-1|k-1} = \delta \widehat{\mathbf{A}}_{k-1} \mathbf{P}_{k-1|k-1} \widehat{\mathbf{A}}_{k-1}^T + \mathbf{Q}_{k-1} \quad 2.11$$

This results in a larger a priori covariance matrix and subsequently a higher gain value that gives more emphasis to the current measurement and less attention to the old measurement, (Simon, 2006). This improves the filter performance in terms of robustness against modeling error. The idea of this method is to use a temporal correction factor instead of a constant one as the previous method in section (2.3.1.), (Bar-Shalom, Li, & Kirubarajan, 2001).

2.3.2 Improving the KF performance numerically

The KF has some numerical limitations that could be summarized as follows, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

- 1- The covariance matrix should be positive definite to guarantee the convergence of the filter. However, the covariance matrix may not be positive definite because of round off errors further impacting numerical accuracy of the matrix inversion used in its iterative calculation as shown in Fig 2.2.
- 2- The inversion operator may result in ill-conditioned matrices which would cause instability.

Recovering the covariance matrix properties (i.e. symmetry and positive definiteness) and increasing the arithmetic precision (i.e. reducing the round off error) have received attention by many researchers to improve the filter stability and recover its optimality. This has resulted in different versions of the Kalman Filter, such as the forms presented in sub-sections (2.3.2.1), (2.3.2.2), (2.3.2.3) and (2.3.2.4).

2.3.2.1 Joseph Stabilized KF

The a posteriori covariance equation is rearranged in a form called *Joseph stabilized* form to ensure the symmetry of the error covariance matrix, [(Simon, 2006), (Bar-Shalom, Li, & Kirubarajan, 2001) and (Grewal & Andrews, 2001)]. The advantages of this form are that the subtraction part in the a posteriori covariance equation is squared and written in a way that stores it as a symmetric matrix as follows, (Grewal & Andrews, 2001):

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \hat{\mathbf{H}}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \hat{\mathbf{H}}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \quad 2.12$$

2.3.2.2 Forcing the covariance matrix to be symmetric

The covariance matrix is forced to be symmetric by simple methods, such as, (Simon, 2006):

- Taking the average of the covariance matrix and its transpose at each time step as follows:

$$\mathbf{P} = (\mathbf{P} + \mathbf{P}^T)/2 \quad 2.13$$

- Forcing the terms under the diagonal to have the same values as the terms above the diagonal.
- Using an appropriate initial covariance matrix that will not experience large changes later.

2.3.2.3 The Sequential KF

The *sequential KF* is a version of the KF that is applicable to systems with multiple measured states. The KF equations are rewritten in a way that uses the a priori covariance matrix and states as initial starting condition. These are then updated several times, instead of once, according to the number of measurements, m , as shown in Fig 2.3, [(Simon, 2006) and (Li & Li, 2007)], where $\hat{\mathbf{x}}_{k|k}(r)$ and $\mathbf{P}_{k|k}(r)$ are the a posteriori estimate and the error covariance matrix after r updates, respectively. R_{rr_k} is the r^{th} diagonal element in \mathbf{R} , $\hat{\mathbf{H}}_k(r)$ is the r^{th} row in \mathbf{H} , and \mathbf{z}_{r_k} is the r^{th} row in \mathbf{z}_k .

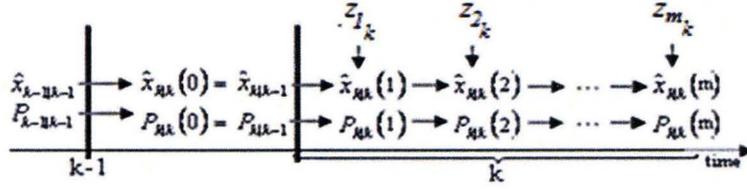


Fig 2.3: The Sequential Kalman Filter, (Simon, 2006).

At each update session, the output from the previous session, $\hat{x}_{k|k}(r-1)$, is considered as the a priori estimate and it is updated using one of the measurements. The a posteriori estimate from the last session is the filter a posteriori estimate as shown in Fig 2.4. By taking a single measurement at each iteration, the KF equations deal with scalar values instead of matrices which overcome the matrix inversion problems in terms of computational time, existence and complexity.

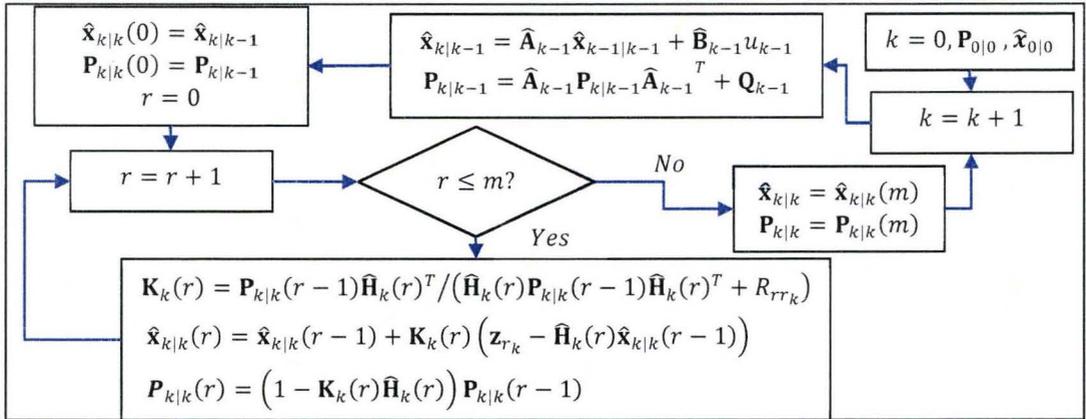


Fig 2.4: The Sequential KF Algorithm, [(Simon, 2006) and (Li & Li, 2007)].

The limitation of this method is that the measurement noise covariance matrix, \mathbf{R} , must be diagonal. If it is not a diagonal matrix then a transform is done to the KF equations to make it diagonal. However, this method is insufficient if \mathbf{R} is not diagonal or constant as it would result in a prohibitively large computational time.

2.3.2.4 The Information Filter

In the *Information Filter*, the covariance calculation is replaced by the information matrix calculation, \mathfrak{I} , which is the inverse of the covariance matrix ($\mathfrak{I} = \mathbf{P}^{-1}$). By rearranging the KF equations according to the information matrix, as shown in Fig 2.5, the subtraction operator in the a posteriori equation is converted to an addition

operator; thereby the positive definiteness of the covariance matrix, \mathbf{P} , is reserved, [(Simon, 2006), (Tylavsky & Sohie, 1986) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

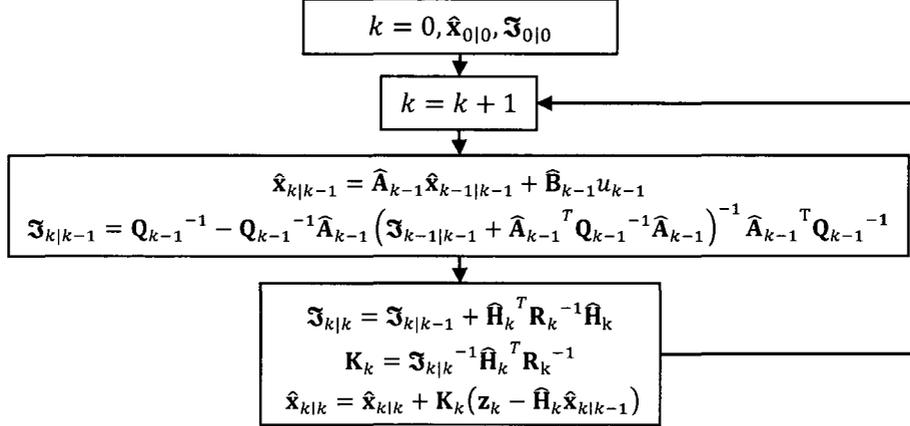


Fig 2.5: The information Filtering Algorithm, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)].

On the other hand, the computational time depends on the noise covariance matrices. If these matrices are constant then the computational time of this method becomes significantly faster, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)].

2.3.3 The Kalman Filter with steady state gain

If the system model and the system and measurement noise covariance matrices are time invariant, then the steady state KFs (SSKFs) can be used instead of the KF. The SSKFs give results that are similar to the KF in such cases with the advantages of fewer computations, less complexity and avoiding inversion problems. However, they give poor estimation performance during the transient period and are not optimal if the noise is non-stationary, [(Simon, 2006), (Bar-Shalom, Li, & Kirubarajan, 2001) and (Grewal & Andrews, 2001)].

The SSKFs use a constant gain that can be obtained by using several methods such as the following, [(Simon, 2006), (Bar-Shalom, Li, & Kirubarajan, 2001), (Painter, Kerstetter, & Jowers, 1990), (Ogle & Blair, 2002), (Yo & Kim, 2003) and (Ogle & Blair, 2004)]:

- Numerical simulation: The final value that the gain approaches in time varying KF is determined. This value is then used as the KF steady state gain, (Simon, 2006).

- **Alpha-Beta Filter:** This is a simple filter that uses a pre-computed fixed gain for estimation. This filter is used in tracking problems involving one measurement and a second order Newtonian system defined as follow, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

$$\begin{aligned}\mathbf{x}_k &= \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \begin{bmatrix} T_s^2/2 \\ T_s \end{bmatrix} w_{k-1} \\ \mathbf{z}_k &= [1 \quad 0] \mathbf{x}_k + v_k\end{aligned}\tag{2.14}$$

Where T_s is the sampling time. The a priori and a posteriori error covariance matrices, and the KF gain, $\mathbf{K}_{\alpha\beta}$, are assumed to be constant and computed as functions of two parameters; α and β , as follows, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

$$\mathbf{K}_{\alpha\beta} = [\alpha \quad \beta/T]^T\tag{2.15}$$

$$\mathbf{P}_{\alpha\beta,k|k-1} = \begin{bmatrix} \frac{\alpha Q}{1-\alpha} & \frac{\beta Q}{T_s(1-\alpha)} \\ \frac{\beta Q}{T_s(1-\alpha)} & \left(\frac{\alpha}{T_s} + \frac{\beta}{2T_s}\right) \frac{\beta Q}{T_s(1-\alpha)} \end{bmatrix}\tag{2.16}$$

$$\mathbf{P}_{\alpha\beta,k|k} = \begin{bmatrix} \alpha R & \beta R/T_s \\ \beta R/T_s & \left(\frac{\alpha}{T_s} - \frac{\beta}{2T_s}\right) \frac{\beta Q}{T_s(1-\alpha)} \end{bmatrix}\tag{2.17}$$

Where $\mathbf{P}_{\alpha\beta,k|k-1}$ and $\mathbf{P}_{\alpha\beta,k|k}$ are the a priori and the a posteriori covariance matrices obtained by the Alpha-Beta KF. $Q = E(w_k w_k^T)$ and $R = E(v_k v_k^T)$ are the system and measurement noise covariance matrices, respectively. α and β are functions of the ratio between Q and R , $\rho = QT_s^2/R$, as follows, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

$$\alpha = \frac{-1}{8}(\rho^2 + 8\rho - (\rho + 4)\sqrt{\rho^2 + 8\rho})\tag{2.18}$$

$$\beta = \frac{1}{4}(\rho^2 + 4\rho - \rho\sqrt{\rho^2 + 8\rho})\tag{2.19}$$

- **Alpha-Beta-Gamma Filter:** This filter is similar to the Alpha-Beta filter. However, it is used in tracking problems involving one measurement and for third order Newtonian systems as follows, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

$$\mathbf{x}_k = \begin{bmatrix} 1 & T_s & T_s^2/2 \\ 0 & 1 & T_s \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \begin{bmatrix} T_s^2/2 \\ T_s \\ 1 \end{bmatrix} w_{k-1} \quad 2.20$$

$$z_k = [1 \quad 0 \quad 0] \mathbf{x}_k + v_k$$

The a posteriori error covariance matrices, $\mathbf{P}_{\alpha\beta,k|k}$, and the KF gain, $\mathbf{K}_{\alpha\beta\gamma}$, are assumed to be constant and functions of these parameters: α , β and γ as follows, [(Simon, 2006) and (Bar-Shalom, Li, & Kirubarajan, 2001)]:

$$\mathbf{K}_{\alpha\beta\gamma} = [\alpha \quad \beta/T_s \quad \gamma/2T_s^2]^T \quad 2.21$$

$$\mathbf{P}_{\alpha\beta,k|k} = \begin{bmatrix} \alpha R & \beta R/T_s & \gamma R/2T_s^2 \\ \beta R/T_s & \frac{8\alpha\beta + \gamma(\beta - 2\alpha - 4)}{8T_s^2(1-\alpha)} R & \frac{\beta(2\beta - \gamma)}{4T_s^3(1-\alpha)} R \\ \gamma R/2T_s^2 & \frac{\beta(2\beta - \gamma)}{4T_s^3(1-\alpha)} R & \frac{\gamma(2\beta - \gamma)}{4T_s^4(1-\alpha)} R \end{bmatrix} \quad 2.22$$

α , β and γ are functions of the ratio between the process and the measurement noise covariance matrices, $\rho = QT_s^2/R$, as follows, (Simon, 2006):

$$\alpha = 1 - \left(a - \frac{b}{3a} - \frac{c}{3} \right)^2 \quad 2.23$$

$$\beta = 2 \left(1 - \left(a - \frac{b}{3a} - \frac{c}{3} \right) \right) \quad 2.24$$

$$\gamma = 2\rho \left(a - \frac{b}{3a} - \frac{c}{3} \right) \quad 2.25$$

Where $a = \left[\frac{-d + \sqrt{d^2 + 4b^3/27}}{2} \right]^{1/3}$, $b = e - \frac{c^2}{3}$, $d = \frac{2c^3}{27} - \frac{ce}{3} - 1$, $c = \frac{\rho}{2} - 3$
and $e = \rho/2 + 3$.

The Alpha-Beta and Alpha-Beta-Gamma KF have advantages in terms of computational time, complexity and avoiding inversion problems. However, they are only applicable to certain system models, and they give poor performance in non-stationary application, (Simon, 2006).

2.3.4 Modifying the KF to make it applicable to non-linear systems.

Several implementations have been applied to the KF to make it applicable to non-linear systems defined as follows, [(Simon, 2006) and (Kreyszig, 2000)]:

$$\begin{aligned}\mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}, u_{k-1}) + \mathbf{w}_{k-1} \\ \mathbf{z}_k &= \mathbf{g}(\mathbf{x}_k) + \mathbf{v}_k\end{aligned}\quad 2.26$$

These implementations vary according to their objectives, i.e. increasing accuracy, stability and/or robustness, reducing computational time and/or complexity, and handling non-Gaussian noise distributions. In this subsection, a selected number of implementations for discrete non-linear systems are discussed.

2.3.4.1 The Perturbation Kalman Filter

In the *Perturbation KF* (PKF), also known as the *Linearized KF*, the nonlinear model of the system is replaced by its linearized form about a nominal state trajectory and input by using the first order Taylor Series Approximation, as follows, [(Barker, Brown, & Martin, 1995), (Ormsby, Raquet, & Maybeck, 2006) and (Negenborn, 2003)]:

$$\begin{aligned}\Delta \mathbf{x}_k &= \boldsymbol{\Phi}_{k-1} \Delta \mathbf{x}_{k-1} + \boldsymbol{\eta}_{k-1} \Delta u_{k-1} + \mathbf{w}_{k-1} \\ \Delta \mathbf{z}_k &= \mathbf{H}_k \Delta \mathbf{x}_k + \mathbf{v}_k,\end{aligned}\quad 2.27$$

Where $\boldsymbol{\Phi}_{k-1} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_{k-1,nom}, u=u_{k-1,nom}}$, $\mathbf{H}_k = \left. \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\mathbf{x}_{k,nom}}$, $\boldsymbol{\eta}_{k-1} = \left. \frac{\partial \mathbf{f}}{\partial u} \right|_{\mathbf{x}=\mathbf{x}_{k-1,nom}, u=u_{k-1,nom}}$, $\Delta \mathbf{x}_k = \mathbf{x}_k - \mathbf{x}_{k,nom}$ and $\Delta u_{k-1} = u_{k-1} - u_{k-1,nom}$.

The PKF estimates the derivatives with respect to the nominal values as shown in Fig 2.6. This filter is not suitable for systems that are highly non-linear where the first Taylor Series Approximation is not sufficiently accurate. Moreover, this filter needs the nominal state trajectory for linearization which may be unavailable or may differ from the actual trajectory due to modeling errors and noise.

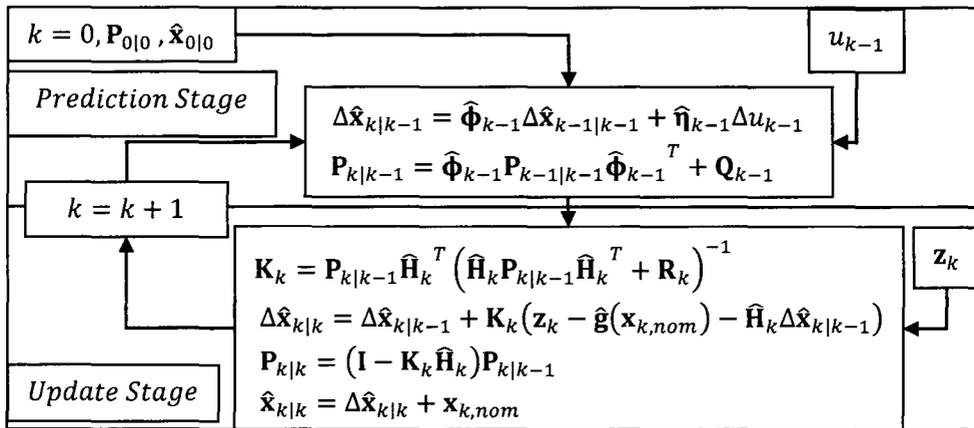


Fig 2.6: The Perturbation Kalman Filter, (Negenborn, 2003).

2.3.4.2 The Extended Kalman Filter

A revised version of the LKF was later proposed and was referred to as the *Kalman-Schmidt Filter* or the *Extended Kalman Filter* (EKF), [(Simon, 2006), (Grewal & Andrews, 2001), (Lary & Mussa, 2004) and (Leu & Baratti, 2000)]. This filter uses the non-linear model of the system to predict the a priori estimates. The system model is then linearized around the a posteriori estimate from the previous time step, and the measurement equation is linearized around the a priori estimation in order to calculate the Kalman gain, and the a priori and the a posteriori covariance matrices. This gives better performance than the PKF, (Simon, 2006). The EKF process is summarized in Fig 2.7, where $\hat{\Phi}_{k-1} = \left. \frac{\partial \hat{f}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1|k-1}, u=u_{k-1}}$ and $\mathbf{H}_k = \left. \frac{\partial \hat{g}}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}}$.

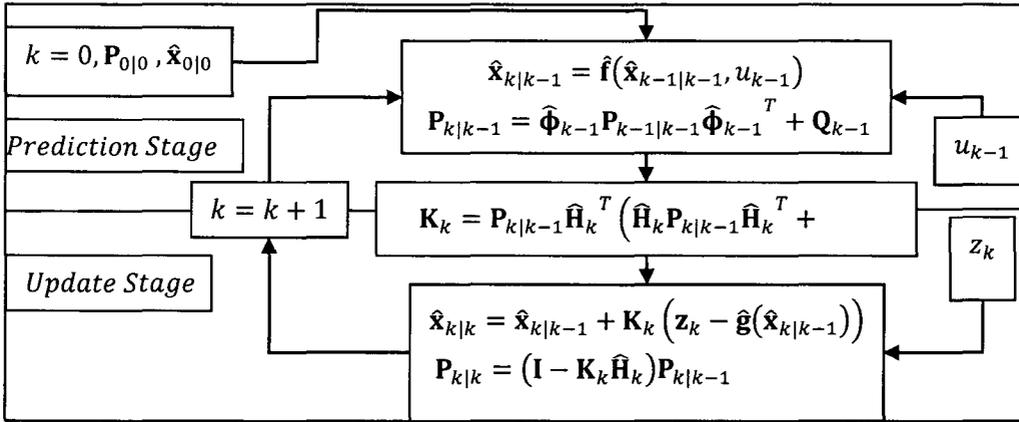


Fig 2.7: The Extended Kalman Filter, (Simon, 2006).

2.3.4.3 The Iterated Extended Kalman Filter

The EKF linearizes the measurement matrix using the Taylor Series Approximation around the a priori estimate, even though its value is not quite accurate. To improve the filter performance, the *Iterated Kalman Filter* (IKF) is used, [(Simon, 2006), (Bar-Shalom, Li, & Kirubarajan, 2001), (Shojaie, Ahmadi, & Shahri, 2007), (Zhang, Zhou, & Duan, 2006) and (Hyland, 2002)]. At the first iteration, the IKF procedure is similar to the EKF. Later on, the base of deriving the measurement matrix is replaced by the a posteriori estimate from the previous iteration, $\hat{\mathbf{x}}_{k|k}(r-1)$, as shown in Fig 2.8. The process is repeated for several iterations, N . This method reduces the error due to the linearization process where it uses a refined base of derivation which is more accurate than an a priori estimate. However, the filter performance becomes poor for

highly non-linear systems that cannot be represented by a first-order Taylor Series Approximation, (Simon, 2006).

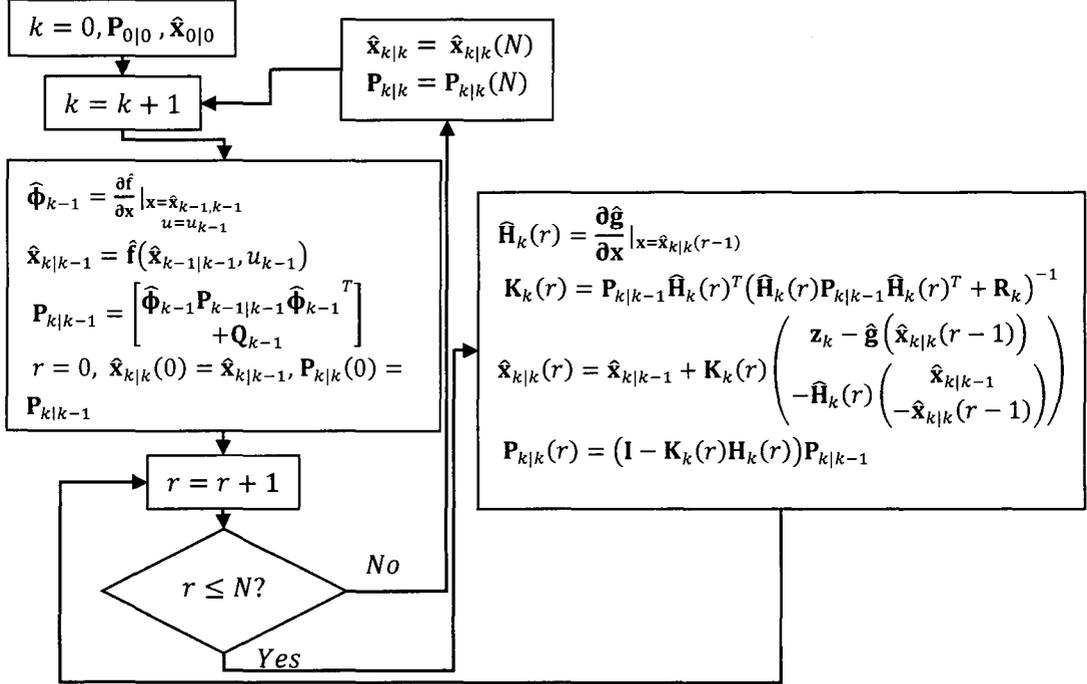


Fig 2.8: The Iterated Kalman Filter, (Simon, 2006).

2.3.4.4 Higher Order Extended Kalman Filter

The *Higher order EKF* is used to reduce the error due to choosing the first order Taylor Series Approximation as a linearizing process. When the system is highly non-linear, higher order terms of Taylor Series Approximation are used to reduce the truncation error due to the neglected terms. However, adding these terms makes the filter more complicated and requires more computational time. For example, the *second-order EKF* in [(Simon, 2006), (Dungate, Theobald, & Nurse, 1999), (Bayard & Kang, 2003) and (Athans, Wishner, & Bertolini, 1968)] needs to calculate the second derivative of the measurement and the system matrices, and to calculate additional correctives terms, $\boldsymbol{\pi}_{1,k}$ and $\boldsymbol{\pi}_{2,k}$, that are added to the a priori and the a posteriori estimate equations due to the additional Taylor Series Approximation terms as follows, (Simon, 2006):

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \hat{f}(\hat{\mathbf{x}}_{k|k-1}, u_{k-1}) + \boldsymbol{\pi}_{1,k} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k \left(\mathbf{z}_k - \hat{\mathbf{g}}(\hat{\mathbf{x}}_{k|k-1}) \right) + \boldsymbol{\pi}_{2,k} \end{aligned} \quad 2.28$$

Where $\boldsymbol{\varphi}_i = \begin{bmatrix} 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & & \downarrow & & & \\ & & & \text{ith column} & & & \end{bmatrix}^T$, $\boldsymbol{\pi}_{1,k} = \frac{1}{2} \sum_{i=1}^n \boldsymbol{\varphi}_i \text{Tr} \left[\frac{\partial^2 \hat{f}}{\partial \mathbf{x}_i^2} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k-1|k-1}} \mathbf{P}_{k-1|k-1} \right]$, and $\boldsymbol{\pi}_{2,k} = \frac{1}{2} \mathbf{K}_k \sum_{i=1}^m \boldsymbol{\varphi}_i \text{Tr} \left[\frac{\partial^2 \hat{g}}{\partial \mathbf{x}_i^2} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k-1}} \mathbf{P}_{k|k-1} \right]$.

As the system non-linearity increases, this method becomes unfeasible in terms of computational time.

2.3.4.5 The Sigma-point Kalman Filter

The *Sigma-point Kalman Filter* (SPKF) draws a certain number of points, called sigma points, from the probability distribution function projected for the states. Then the SPKF projects these points by using the system's nonlinear model to obtain an a posteriori estimate for the state's probability distribution hence avoiding the requirement for linearization. The SPKF is based on the weighted statistical linear regression method which linearizes the non-linear model statistically, [(Simon, 2006), (Van Der Merwe & Wan, 2004), (Ambadan & Tang, 2009), (Wang, Wang, Liao, & Liu, 2009), (Sadhu, Mondal, Srinivasan, & Ghoshal, 2006), (Schenkendorf, Kremling, & Mangold, 2009) and (Ali, Deriche, & Landolsi, 2009)]. This leads to a more accurate linearization technique than the Taylor Series Approximation with the advantages of eliminating the need to calculate the Jacobian matrices and the accommodation of noise distributions that are not Gaussian, (Tang, Zhao, & Zhang, 2008). The SPKFs family includes:

- The *Unscented KF* (UKF) described in [(Simon, 2006), (Van Der Merwe & Wan, 2004), (Ambadan & Tang, 2009), (Wang, Wang, Liao, & Liu, 2009), (Sadhu, Mondal, Srinivasan, & Ghoshal, 2006), (Schenkendorf, Kremling, & Mangold, 2009) and (Ali, Deriche, & Landolsi, 2009)]. The UKF obtains a minimal set of sigma points around the mean, as shown in Fig 2.9 and table 2.1, using one of the

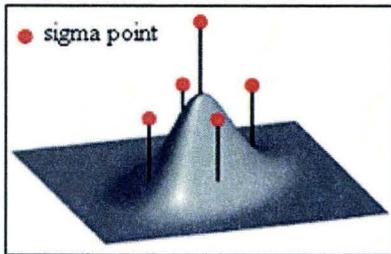


Fig 2.9: Sigma-Points for $n = 2$, (Van Der Merwe & Wan, 2004).

many unscented transformations that have been proposed, i.e. *unscented* (Simon, 2006), *general unscented* [(Simon, 2006) and (Ambadan & Tang, 2009)], *simplex unscented* [(Simon, 2006), (Tang, Zhao, & Zhang, 2008), (Kim & Shin, 2005) and (Julier, 2003)] or *spherical unscented* [(Simon, 2006), (Tang, Zhao, & Zhang, 2008), (Kim & Shin, 2005) and (Julier, 2003)] transformations.

By propagating these points through the non-linear system model and by using an associated weight factor, the mean and the covariance of the system are approximated; e.g. Fig 2.10 shows that the estimated mean and covariance obtained by using the sigma points (for a system that consists of two states and a nonlinear output matrix) approximate the actual system mean and covariance, where $\hat{\mathbf{X}}_i$ and $\hat{\mathbf{Z}}_i$ are the estimated state and measurement vectors of the i^{th} sigma point.

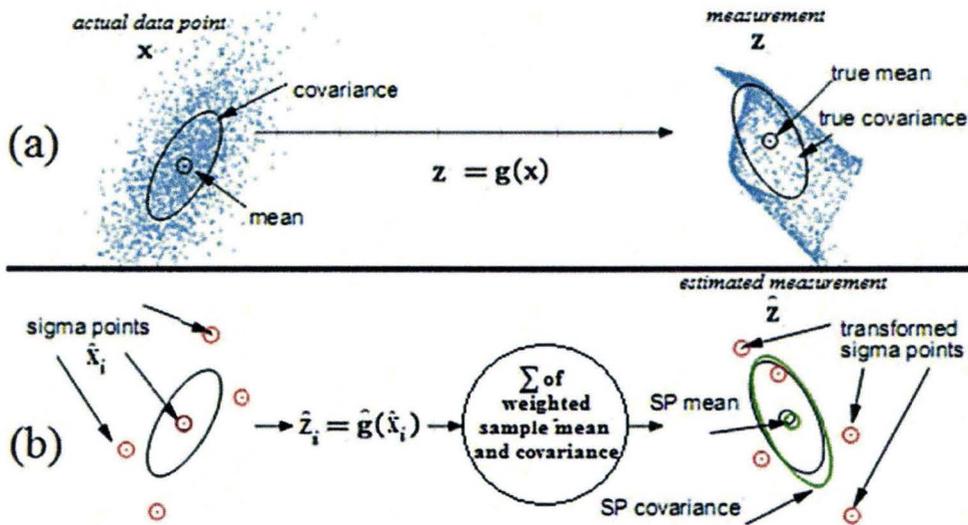


Fig 2.10: (a) The actual system states and their nonlinear measurement (b) The Sigma-Point KF's estimates, (Van Der Merwe & Wan, 2004).

The unscented transformations share similar principles although they differ in terms of the number of sigma points, how to choose these points, and how to calculate their weights. The unscented transformation gives accurate results up to a third order Taylor Series Approximation for Gaussian distributions, (Simon, 2006), and second order Taylor Series Approximation for non-Gaussian distributions, (Van Der Merwe & Wan, 2004). Conversely, the general unscented transformation gives a chance to improve the results by using a designing factor, λ , that improves the approximation accuracy to a higher order Taylor Series Approximation. The simplex unscented transformation is used to reduce the computational time and uses almost half of the sigma points. However, it has stability limitations when systems are of high order, [(Simon, 2006) and (Julier, 2003)]. The spherical unscented

transformation is used to overcome the simplex unscented transformation stability problem by rearranging the sigma points and their weights. Fig 2.11 summarizes these unscented methods and table 2.1 summarizes their parameters (W_i , \mathbf{q}_1 and \mathbf{q}_2). Note that q is the number of the sigma points, $\mathbf{P}_1 = \mathbf{P}_{k-1|k-1}$, $\mathbf{P}_2 = \mathbf{P}_{k|k-1}$, $(\mathbf{a})_i$ is the i row of \mathbf{a} , W_i , \mathbf{X}_i and \mathbf{Z}_i are the assigned weight, the estimate and its measurement for the i^{th} sigma point, respectively, and \mathbf{q}_1 and \mathbf{q}_2 are the selection parameters of the sigma points for the a priori and a posteriori estimates, respectively.

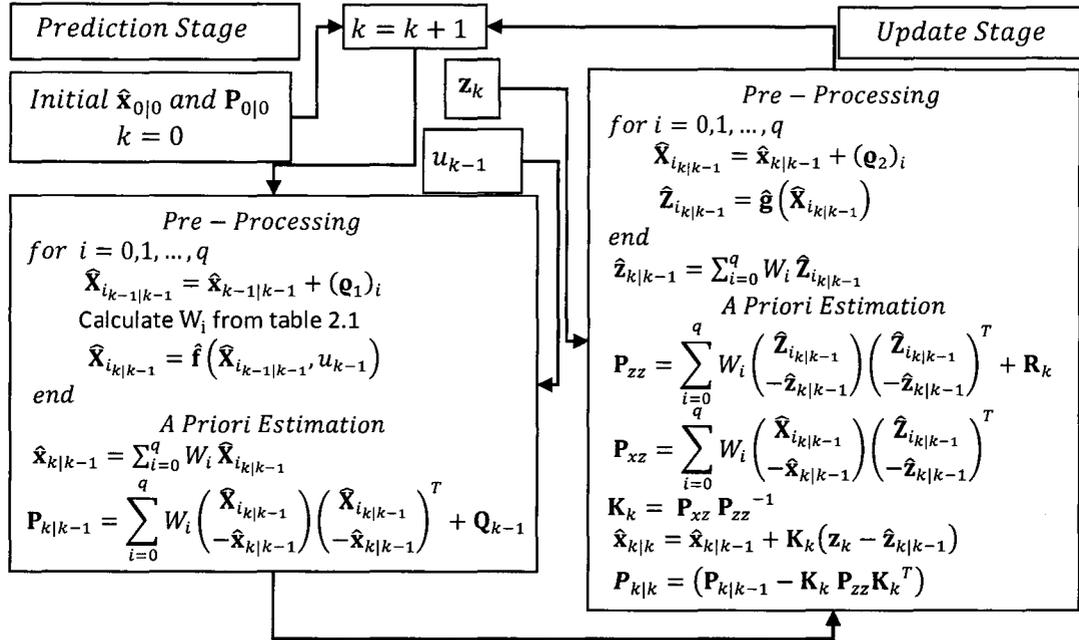


Fig 2.11: The Unscented Kalman Filter, (Simon, 2006).

Method	$(\mathbf{e}_j)_i^l, \quad \begin{matrix} j = 1, 2 \\ i = 1, 2, \dots, q \end{matrix}$	$W_i, i = 1, 2, \dots, q$	q
UKF	$(\mathbf{e}_j)_i = \begin{cases} 0 & i = 0 \\ \left(\sqrt{n\mathbf{P}_j} \right)_i^T & 1 \leq i \leq n \\ -\left(\sqrt{n\mathbf{P}_j} \right)_i^T & n+1 \leq i \leq 2n \end{cases}$	$W_i = \begin{cases} 0 & i = 0 \\ \frac{1}{2n} & i \neq 0 \end{cases}$	$2n + 1$
General UKF	$(\mathbf{e}_j)_i = \begin{cases} 0 & i = 0 \\ \left(\sqrt{(n+\lambda)\mathbf{P}_j} \right)_i^T & 1 \leq i \leq n \\ \left(\sqrt{(n+\lambda)\mathbf{P}_j} \right)_i^T & n+1 \leq i \leq 2n \end{cases}$	$W_i = \begin{cases} \frac{\lambda}{n+\lambda} & i = 0 \\ \frac{1}{2(n+\lambda)} & i \neq 0 \end{cases}$	$2n + 1$
Simplex UKF	$(\mathbf{e}_j)_i = \sqrt{\mathbf{P}_j} \boldsymbol{\rho}_i^n$ As $\boldsymbol{\rho}_i^n$ is obtained recursively as follows: $\boldsymbol{\rho}_0^1 = 0$ and $\boldsymbol{\rho}_1^1 = \boldsymbol{\rho}_2^1 = \frac{-1}{\sqrt{2W_1}}$, (the superscript is the recursive index) for $l = 2, \dots, n$ (number of the states) $\boldsymbol{\rho}_i^l = \begin{cases} \begin{bmatrix} \boldsymbol{\rho}_0^{l-1} \\ 0 \end{bmatrix} & i = 0 \\ \begin{bmatrix} \boldsymbol{\rho}_i^{l-1} \\ -1 \end{bmatrix} & 1 \leq i \leq l \\ \begin{bmatrix} \mathbf{0}_{l-1 \times 1} \\ l \end{bmatrix} & i = l+1 \end{cases}$ end	W_0 is chosen as $W_0 \in [0, 1)$ $W_i = \begin{cases} 2^{-n}(1 - W_0) & 1 \leq i \leq 2 \\ 2^{i-n}(W_1) & i > 2 \end{cases}$	$n + 2$
Spherical UKF	Similar to the simplex UKF except that $\boldsymbol{\rho}_i^l = \begin{cases} \begin{bmatrix} \boldsymbol{\rho}_0^{l-1} \\ 0 \end{bmatrix} & i = 0 \\ \begin{bmatrix} \boldsymbol{\rho}_i^{l-1} \\ -1 \end{bmatrix} & 1 \leq i \leq l \\ \begin{bmatrix} \mathbf{0}_{l-1 \times 1} \\ l \end{bmatrix} & i = l+1 \end{cases}$	W_0 is chosen as $W_0 \in [0, 1)$ $W_i = \frac{1 - W_0}{n + 1}$	$n + 2$

Table 2.1: The differences between the UKF methods, (Simon, 2006).

- The ***Sigma-Point Central Difference KF*** (CDKF) described in [(Nrgaard, Poulsen, & Ravn, 2000), (Zhang, Gao, & Tian, 2008) and (Jihua, Nanning, Zejian, & Qiang, 2009)]. This filter shares a similar principle to the UKF as it uses $2n + 1$ sigma points, (Jihua, Nanning, Zejian, & Qiang, 2009), propagates these point through the nonlinear system, and uses the linear weighted regression to calculate $\hat{\mathbf{x}}_{k|k-1}$, $\hat{\mathbf{z}}_{k|k-1}$, \mathbf{P}_{xx} and \mathbf{P}_{zz} . However, the idea of CDKF is to approximate the system model by linearizing the nonlinear matrices using the Taylor Series Approximation (as done in the EKF) and then replacing the derivatives with their numerical Stirling's polynomial interpolation forms (NSPI), (Sadati & Ghaffarkhah, 2007), that represent the derivative as central divided differences as follows, (Henrici, 1964):

$$\partial f^{(n)}(x) = \frac{1}{2} \left(f^{(n-1)} \left(x + \frac{T_s}{2} \right) - f^{(n-1)} \left(x - \frac{T_s}{2} \right) \right) \quad 2.29$$

As a consequence of using the NSPI, the CDKF method differs from the UKF method in terms of selecting the sigma points and their weights, and calculating the covariance matrices, as shown in Fig 2.12. The CDKF has several advantages over the EKF and the UKF:

- Accuracy: it has been found that the CDKF method gives superior performance over the UKF and EKF, [(Chen, 2003), (Simon, 2006) and (Van Der Merwe, 2004)].
- Ease of tuning: the CDKF uses one control parameter, T_{cd} , which is easy to be obtained as shown in (Van Der Merwe, 2004), e.g. it is set to be $T_{cd} = \sqrt{3}$ for Gaussian distributions. On the other hand, the UKF has at least one control parameter that needs to be tuned, i.e. W_0 for the simplex and the spherical UKF, and λ for the general UKF. Moreover, the tuning process varies according to the application which makes the UKF tuning process not unique.

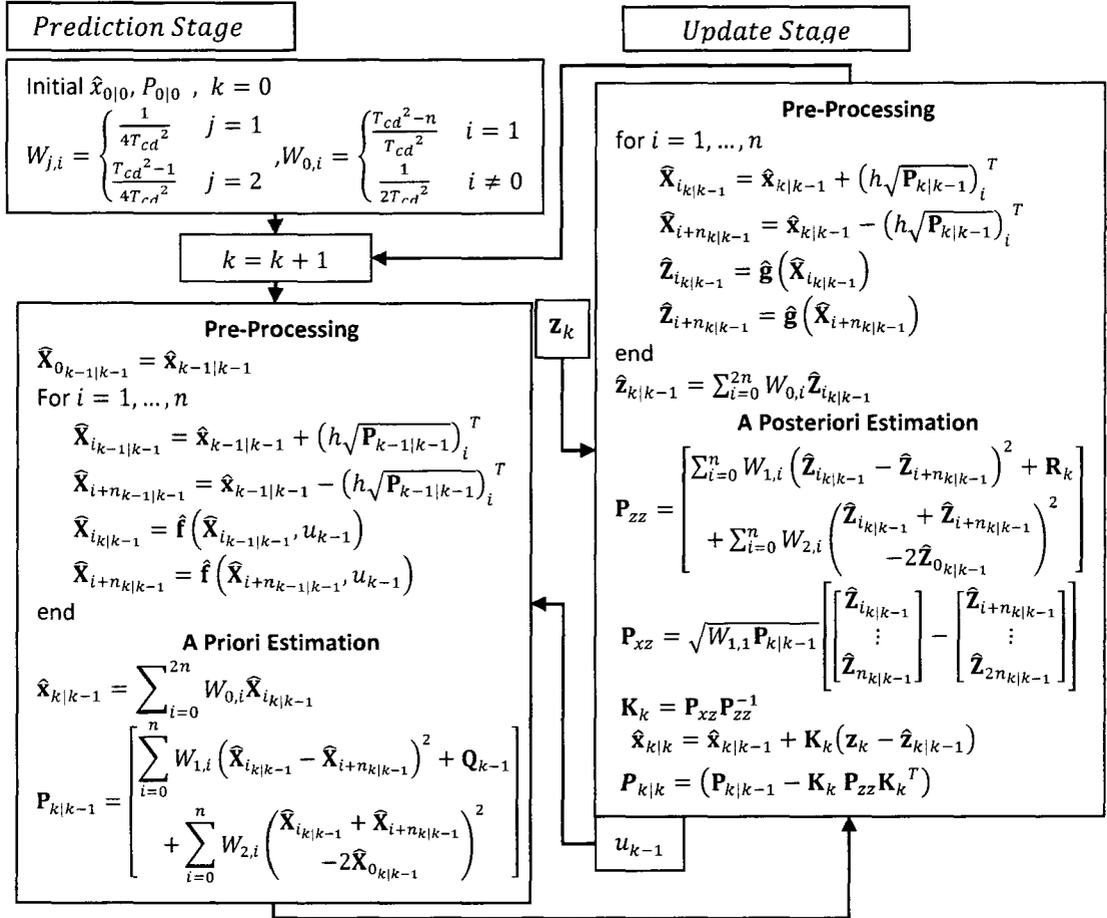


Fig 2.12: The Sigma-Point Central Difference Kalman Filter, [(Van Der Merwe, 2004) and (Van Der Merwe, 2004)].

2.3.5 Combination of the KF with intelligent techniques.

The KF has been combined with intelligent techniques, i.e. *fuzzy logic* [(Nguyen & Walker) and (Yager & Zadeh, 1992)], to improve its performance in term of robustness and stability. Examples of these algorithms are discussed in this subsection.

2.3.5.1 Carrasco's Fuzzy Observer/EKF method

A fuzzy observer was implemented by R. Carrasco et al. in (Carrasco, Cipriano, & Carelli, 2005). It was used for estimating a simple mobile robot posture defined by two coordinates (x, y) and a heading angle, φ . The nonlinear model was approximated using Takagi-Sugeno fuzzification with the heading angle as an input, as shown in Fig 2.13. Five linearized systems were obtained, L_1, L_2, \dots, L_5 , where each L_i is obtained by

linearizing the system around one of five values of the heading angle, which were (in degrees): ± 0 , ± 90 , ± 90 , ± 180 and ± 180 . Each value of the angle represents the peak of the corresponding fuzzification membership as shown in Fig 2.14. The shapes of the fuzzy membership functions were optimized using a genetic algorithm. For each linearized model, a KF was used to

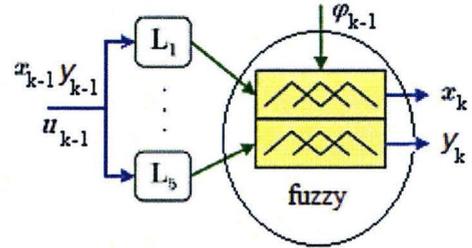


Fig 2.13: The Carrasco's Fuzzy nonlinear model, (Carrasco, Cipriano, & Carelli, 2005).

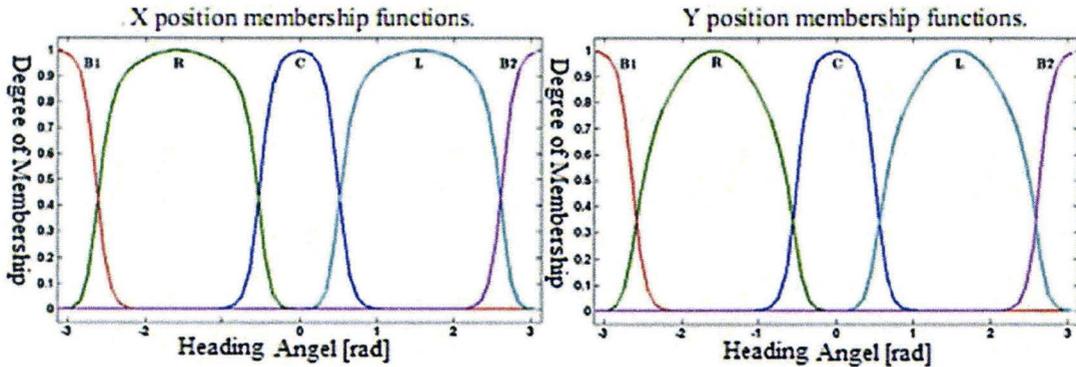


Fig 2.14: The Fuzzy membership function, (Carrasco, Cipriano, & Carelli, 2005).

estimate the states (including the heading angle). At each time step, the estimated heading angle, which is the average of the filters' heading angles at that time step, is used as an input to the fuzzy system to obtain the other estimated states (x, y) as shown in Fig 2.15. The resultant error covariance matrix is a linear combination of the filters' error covariance matrices. The

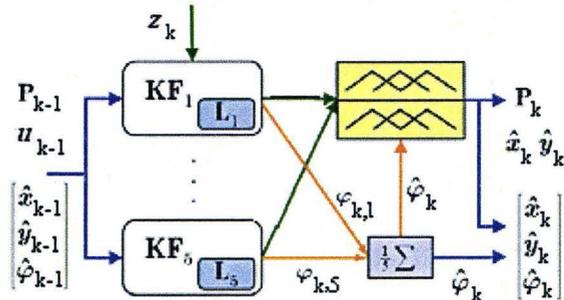


Fig 2.15: The Carrasco's Fuzzy Kalman Filter, (Carrasco, Cipriano, & Carelli, 2005).

resultant observer is more accurate, robust and easier to tune compared to the EKF.

2.3.5.2 Simon's Fuzzy /KF method

D. Simon in (Simon, 2003) represented the nonlinear system with Takagi-Sugeno fuzzy models. By breaking the nonlinear model into approximated fuzzy piecewise linear models, the estimation problem becomes simpler and fits with the KF approach. Each of

these models is then estimated using a KF with steady state gain. By combining the filters' outputs, the estimates are then retrieved. The input for the fuzzy system is a function of the states at the previous time step, while the resultant fuzzification weights are used as part of the estimation process for each fuzzy model.

In this method, the time-varying KF is represented by combining the KFs with steady state gains. The advantage of this method is that it reduces the computational time and complexity. The steady state gain is obtained before the experiment starts; thereby there is no need to calculate or update the error covariance matrices. Another advantage is that the fuzzy system approximates the nonlinearity in the system without the need to linearize it by using the Taylor Series Approximation.

2.3.5.3 Matia's Fuzzy /KF method

The Matia's *Fuzzy-Kalman Filter* was implemented by F. Matia et al. in (Matia, Jimenez, Rodriguez-Losada, & Al-Hadithi, 2004). This filter uses the possibility distribution (obtained from the fuzzy membership functions) instead of Gaussian distribution to accommodate unsymmetrical noise distributions. The basic idea is to represent the uncertainties by fuzzy logic, i.e. asymmetric trapezoidal membership functions for the error in states and for the noise, to overcome the EKF limitations in terms of non-Gaussian and asymmetric uncertainties. The resulting membership functions that represent the states, noise and measurements are then propagated through the nonlinear model. At each time step, the a posteriori estimates are considered as the center of gravity of the resultant compound membership function. According to (Matia, Jimenez, Rodriguez-Losada, & Al-Hadithi, 2004), this method gives a better estimation for the states of non-linear systems than the EKF. Using the fuzzy approach has benefits over the Gaussian assumptions in terms of approximating the noise, handling larger modeling errors, rejection of outliers, and having less sensitivity to errors in initial conditions.

2.3.6 Using adaptive techniques.

2.3.6.1 Noise level adjustment

The *noise level adjustment method* (proposed in (Bar-Shalom, Li, & Kirubarajan, 2001)) is an adaptive technique that changes the process noise covariance matrix settings for the filter. This method starts by assuming a low level of process noise. The efficiency of the current noise level is tested using one of the innovation tests described in (Bar-Shalom, Li, & Kirubarajan, 2001) such as the following:

Test 1: *Normalized innovation squared* (NIS) presented in [(Bar-Shalom, Li, & Kirubarajan, 2001), (Jilkovand & Li, 2002) and (chan, Hu, & Plant, 1979)].

This test uses the ratio between the covariance of the innovation, ($E(\mathbf{v}_k^T \mathbf{v}_k)$), and the calculated covariance matrix of the output's estimation error, \mathbf{S}_k , as follows, (Bar-Shalom, Li, & Kirubarajan, 2001):

$$\mathbf{NIS}_k = \mathbf{v}_k^T \mathbf{S}_k^{-1} \mathbf{v}_k \quad 2.30$$

Where \mathbf{v}_k is the innovation vector and it is defined as follows:

$$\mathbf{v}_k = \mathbf{z}_k - \hat{\mathbf{H}}_k \hat{\mathbf{x}}_{k|k-1} \quad 2.31$$

and \mathbf{S}_k is defined as follows:

$$\mathbf{S}_k = \hat{\mathbf{H}}_k \mathbf{P}_{k|k-1} \hat{\mathbf{H}}_k^T + \mathbf{R}_k \quad 2.32$$

The NIS value is tested to be less than a threshold, ε .

Test 2: ***Moving average*** presented in (Bar-Shalom, Li, & Kirubarajan, 2001). This test takes a window that contains l successive state points from $k - l + 1$ to k , and slides forward in time. The NIS is calculated for each point and then the mean of the resultant NIS values is tested against the threshold ε .

If the test fails; i.e. the value of NIS exceeds the threshold ε , then the process covariance matrix is scaled up using a scaling factor, known as a fudge factor. The process is repeated until the test method condition is satisfied.

2.3.6.2 Noise level switching

The ***noise level switching*** method described in (Bar-Shalom, Li, & Kirubarajan, 2001) is similar to the noise level adjustment method, except that there are two known levels of process noise to choose from. The choice is obtained according to the mentioned innovation testing methods in section (2.3.6.1), where two threshold values are used; upcrossing and downcrossing thresholds. If the process starts with a low process noise level (has small covariance matrix), and the innovation test fails; i.e. its value exceeds the upcrossing threshold, then a higher process noise level (has large covariance matrix) is used. The new level holds until the test value falls below the downcrossing threshold, then the process noise level is changed back to the lower level.

2.3.6.3 Multiple Models (hybrid system)

The Multiple Models approach (MM) (presented in [(Bar-Shalom, Li, & Kirubarajan, 2001), (Li & Bar-Shalom, 1996), (Li X. , 2000), (Li, Zwi, & Zwang, 1999), (Li, Zwang, & Zwi, 1999), (Li & Zwang, 2000), (Li, Jikov, & Ru, 2005) and (Ainsleigh, 2007)]) assumes a finite number of models, r , where the system can be any one of them.

These models, $M_1 \dots M_r$, include all possible system structures and/or parameters. The system model is assumed to be fixed during the estimation. Each model at time k has the following:

- 1- **The likelihood function** ($\Lambda_{i,k}$) is defined as the Gaussian probability density function (PDF) of the model i innovation ($\mathbf{v}_{i,k}$) assuming a zero mean and a covariance matrix of $\mathbf{S}_{i,k}$ (defined in equation (2.32)), and it has the following form:

$$\Lambda_{i,k} = \wp(\mathbf{v}_{i,k}; \mathbf{0}; \mathbf{S}_{i,k}) \quad 2.33$$

Where $\wp(x; \bar{x}_j, P_j)$ is the Gaussian PDF with average of \bar{x}_j and covariance of P_j .

- 2- A prior probability, $\boldsymbol{\mu}_{i,k-1}$, is used to update the a posteriori probability, $\boldsymbol{\mu}_{i,k}$, of the model i (M_i) as follows, (Bar-Shalom, Li, & Kirubarajan, 2001):

$$\boldsymbol{\mu}_{i,k} = \frac{\Lambda_{i,k} \boldsymbol{\mu}_{i,k-1}}{\sum_{j=1}^r \Lambda_{j,k} \boldsymbol{\mu}_{j,k-1}} \quad 2.34$$

The posterior estimate of M_i , $\hat{\mathbf{x}}_{i,k|k}$, and its error covariance matrix, $\mathbf{P}_{i,k|k}$, are calculated for each model using a bank of r Kalman Filters, and then they are combined as follows [(Bar-Shalom, Li, & Kirubarajan, 2001) and (Ainsleigh, 2007)]:

$$\hat{\mathbf{x}}_{k|k} = \sum_{j=1}^r \boldsymbol{\mu}_{j,k} \hat{\mathbf{x}}_{j,k|k} \quad 2.35$$

$$\mathbf{P}_{k|k} = \sum_{j=1}^r \boldsymbol{\mu}_{j,k} \left[\mathbf{P}_{j,k|k} + (\hat{\mathbf{x}}_{j,k|k} - \hat{\mathbf{x}}_{k|k})(\hat{\mathbf{x}}_{j,k|k} - \hat{\mathbf{x}}_{k|k})^T \right] \quad 2.36$$

The multiple model method is summarized in Fig 2.16. This method is limited to include a small number of models that are close to the system model [(Li & Bar-Shalom, 1996), (Li X. , 2000) and (Li X. , 2000)]. If the number of models is increased, the computational time increases and the performance may not improve because some unwanted models would corrupt the results.

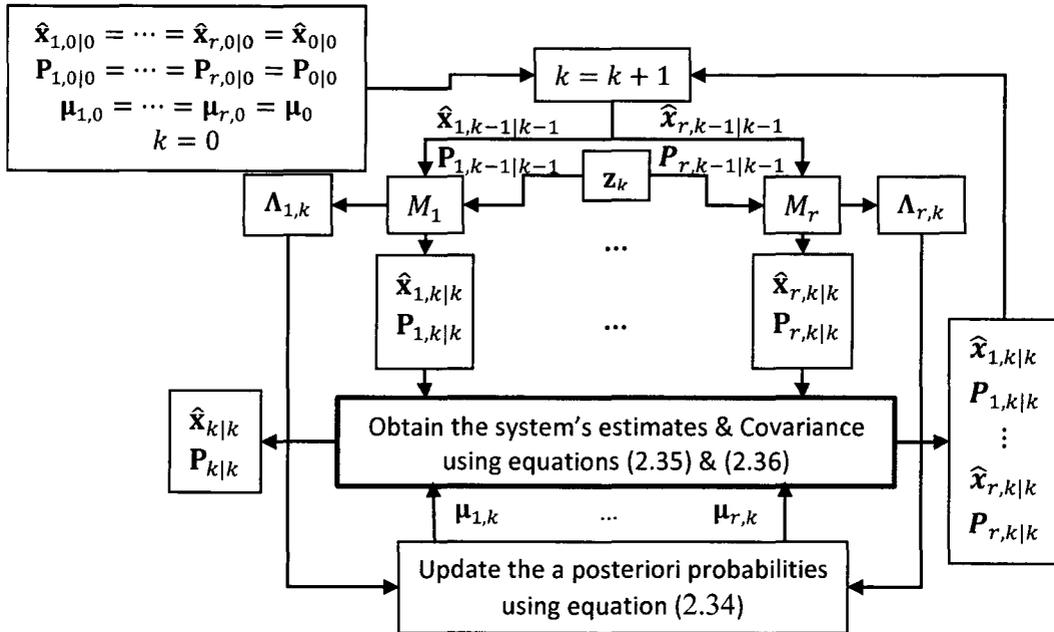


Fig 2.16: The Multiple Model algorithm, (Bar-Shalom, Li, & Kirubarajan, 2001).

2.3.6.4 Model Switching

The *Model Switching* (MS) approach (proposed in [(Bar-Shalom, Li, & Kirubarajan, 2001), (Ho T. , 2008), (Li X. , 2000), (Li, Zwi, & Zwang, 1999), (Ainsleigh, 2007) and (Chang & Athans, 1978)]) assumes a finite number of models and the system's model can be anyone of them at time k . This means that there is a chance of the model changing during the operation. A procedure similar to the multiple models approach is used, while taking into account predefined changing sequences. Each possible sequence is considered as a path as shown in Fig 2.17

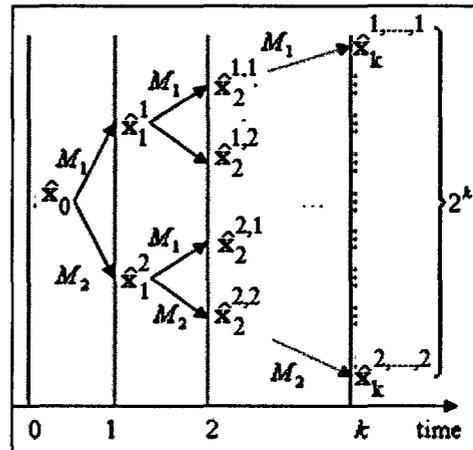


Fig 2.17: The Model Switching Approach for $r = 2$ at time k , (Chang & Athans, 1978).

(assuming $\hat{x}_i^{d_1 \dots d_i}$ is the estimate at time i using model sequence of $M_{d_1} \dots M_{d_i}$ as its path and $d_1 \dots d_i$ for this example have values of 1 or 2). The number of paths increases exponentially with respect to time. For example, if the system has a model from N

available models at every time step, then at the beginning (time 0) the system path has a probability space of N . At the next time step, the system path has a probability space of $N \times N$ possible sequences. The same principle holds during the next time step, which leads to N^k possible sequences at time k .

The states of each complete path are calculated using the Kalman Filter, and the resultant states are fused together to obtain the system states at any given time. This algorithm is an optimal algorithm. However, it is obvious that this method can be computationally very demanding.

2.3.6.5 Interacting Multiple Models

The *Interacting Multiple Models* (IMMs) method (described in [(Bar-Shalom, Li, & Kirubarajan, 2001), (Johnston & Krishnamurthy, 2001), (Mihaylova, Lefebvre, Stafetti, Btuyninckx, & De Schutter, 2002) and (Li & Bar-Shalom, 1994)]) calculates a mixture with different weights, referred to as mixing probabilities, for each possible model and calculates the corresponding previous a posteriori estimate using the likelihood functions. The mixing probability is calculated at each time step as follows, (Bar-Shalom, Li, & Kirubarajan, 2001):

$$\boldsymbol{\mu}_{i|j,k-1|k-1} = \frac{P_{rij} \boldsymbol{\mu}_{i,k-1}}{\sum_{i=1}^r P_{rij} \boldsymbol{\mu}_{i,k-1}} \quad 2.37$$

Where $\boldsymbol{\mu}_{i|j,k-1|k-1}$ is the mixing probability of the path sequence M_i at time $k-1$ to M_j at time k , P_{rij} is the probability of switching to the model M_j from the current model M_i , and $\boldsymbol{\mu}_{i,k-1}$ is the probability that the model M_i is the actual system model. By interacting these estimates and covariance matrices with the mixing probabilities, new modified values of the initial estimate and covariance matrix for the model M_j (i.e. $\hat{\mathbf{x}}_{0,j,k-1|k-1}$ and $\mathbf{P}_{0,j,k-1|k-1}$, respectively) are obtained for each corresponding possible model as follows, (Bar-Shalom, Li, & Kirubarajan, 2001):

$$\hat{\mathbf{x}}_{0,j,k-1|k-1} = \sum_{i=1}^r \hat{\mathbf{x}}_{0,i,k-1|k-1} \boldsymbol{\mu}_{i|j,k-1|k-1} \quad 2.38$$

$$\mathbf{P}_{0,j,k-1|k-1} = \sum_{i=1}^r \boldsymbol{\mu}_{i|j,k-1|k-1} \left[\mathbf{P}_{i,k-1|k-1} + \begin{pmatrix} \hat{\mathbf{x}}_{i,k-1|k-1} \\ -\hat{\mathbf{x}}_{0,j,k-1|k-1} \end{pmatrix} \begin{pmatrix} \hat{\mathbf{x}}_{i,k-1|k-1} \\ -\hat{\mathbf{x}}_{0,j,k-1|k-1} \end{pmatrix}^T \right] \quad 2.39$$

The values of $\hat{\mathbf{x}}_{0,j,k-1|k-1}$ and $\mathbf{P}_{0,j,k-1|k-1}$ summarize the old history of the model M_j up to the time $k-1$. The model probability, $\boldsymbol{\mu}_{j,k-1}$, is then updated as follows, (Bar-Shalom, Li, & Kirubarajan, 2001):

$$\mu_{j,k} = \frac{\Lambda_j \sum_{i=1}^r P_{r_{ij}} \mu_{i,k-1}}{\sum_{j=1}^r \Lambda_j \sum_{i=1}^r P_{r_{ij}} \mu_{i,k-1}} \quad 2.40$$

Using the a posteriori value of each filter; $\hat{\mathbf{x}}_{j,k-1|k-1}$ and $\mathbf{P}_{j,k-1|k-1}$, and the corresponding model probability; $\mu_{j,k}$, then the system a posteriori estimate and covariance matrix; $\hat{\mathbf{x}}_{k-1|k-1}$ and $\mathbf{P}_{k-1|k-1}$, can be calculated using equations (2.35) and (2.36). This method is summarized in Fig 2.18.

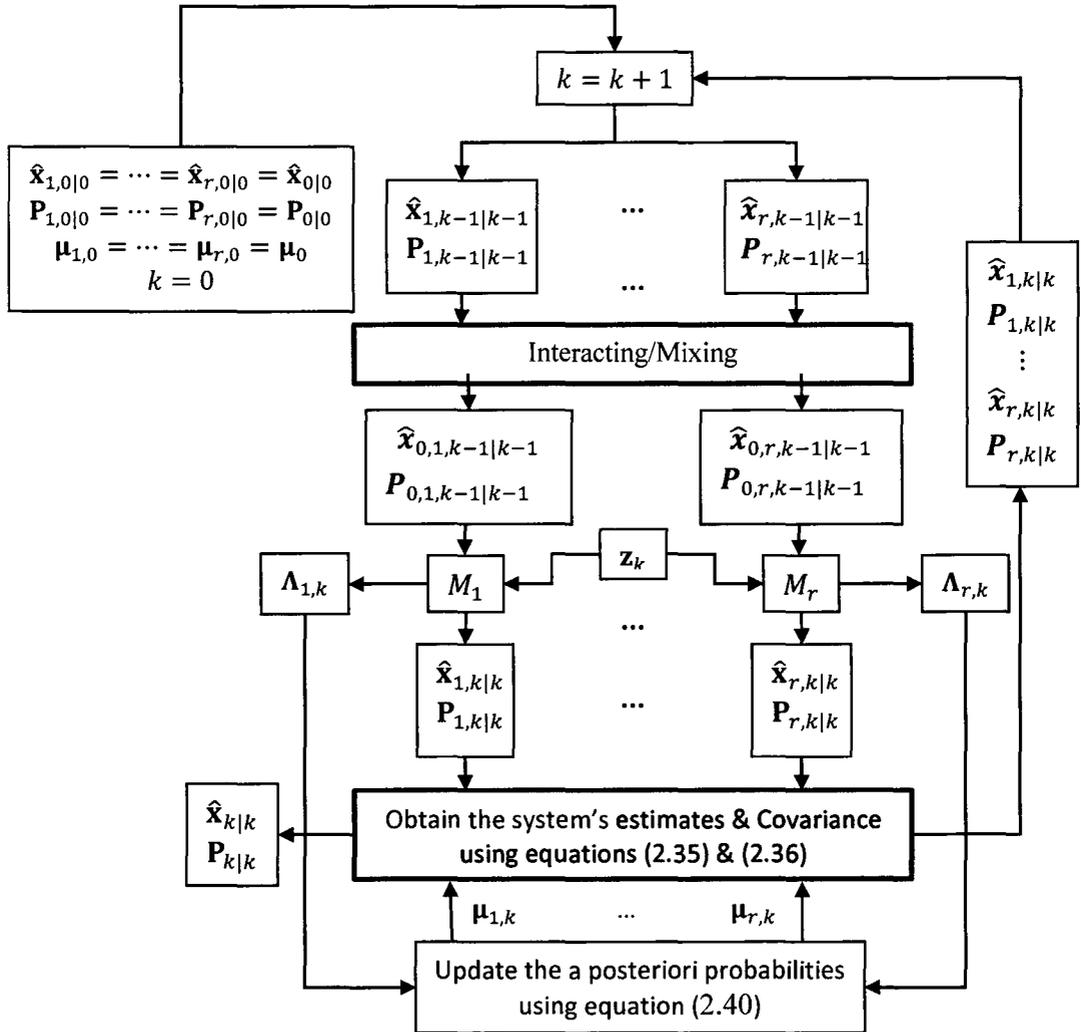


Fig 2.18: The Interacting Multiple Model Algorithm, (Li & Bar-Shalom, 1994).

Chapter Three:

Literature Review Part II - Sliding Mode Observers

3.1. Introduction to Variable Structure and Sliding Mode Control and Systems

In the 1940's, variable structure theory became popular in Soviet literature. They were based on systems that contained discontinuities in the differential equations. The systems make use of a discontinuity hyperplane that divides the state space into regions where the dynamic equations are continuous. These systems are referred to as variable structure systems, since their state space is segmented and the system dynamics changes as the state trajectories cross the discontinuity hyperplanes. *Variable structure control* (VSC) is a method that switches the control gain according to state space segment in which the state trajectory resides; thus resulting in a discontinuous control input, [(Utkin, Guldner, & Shi, 1999), (Utkin, 1977), (Utkin, 1983) and (Monsees, 2002)].

A mechanical system with Coulomb Friction as shown in Fig 3.1-a is a good example of a *variable structure system*, where the discontinuity part is shown in Fig 3.1-b and is presented as follows, (Utkin, Guldner, & Shi, 1999):

$$m\ddot{x} + kx + c \times \text{sgn}(\dot{x}) = 0 \quad 3.1$$

Where m , k and c are the mass, the spring stiffness and the damping constant, respectively, and x is the system displacement.

A popular method referred to as the *sliding mode control* (SMC) is a special form of the VSC. It makes use of a discontinuous control input in a manner that forces the state trajectory towards the switching hyperplane. A switching gain forces the state trajectory to remain close to the hyperplane, and slide along it. SMC has been demonstrated to be very robust to uncertainties and external disturbances. Prior to the 1980's, VSC and SMC methodologies were only considered in the continuous time domain. It wasn't until 1985 that a discrete sliding mode approach was presented, [(Milosavljevic, 1985) and (Monsees, 2002)]. Later, a stability condition for a discrete SMC was clearly formulated,

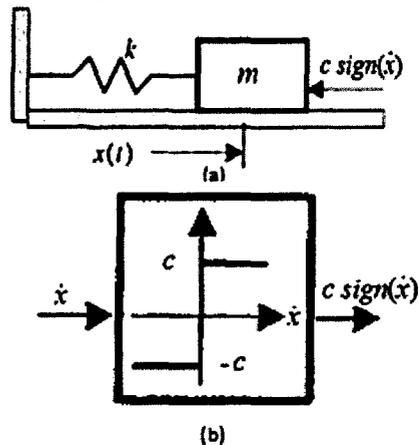


Fig 3.1: (a) Mechanical System with Coulomb Friction, (b) Discontinuously Coulomb Friction, (Utkin, Guldner, & Shi, 1999).

and is now commonly used in the design of discrete controllers, [(Utkin, Guldner, & Shi, 1999), (Monsees, 2002) and (Utkin, 1977)].

An early example of sliding mode control (SMC) was presented by Nikolski in the 1930's, where a relay controller is applied to manage a ship's course, (Utkin, Guldner, & Shi, 1999). The proposed ship model is a second order system that consists of the course and Rudder angles as shown in Fig 3.2-a. Using the SMC, the trajectories are forced to follow a pre-prescribed pattern, which is represented by the sliding surface $S: \{\delta + c_1\phi = 0\}$ as shown in Fig 3.2-b.

3.2. The Sliding Mode Control

The SMC is used to derive the state trajectory of a plant to reach a sliding surface and thereafter to slide along it. While on the sliding surface and subjected to bounded disturbances, the system can be shown to be stable and robust in its performance. While on the sliding surface, the system is less sensitive to uncertainties and perturbation under ideal conditions, (Utkin, Guldner, & Shi, 1999). The sliding surface for a system is defined usually in terms of the demanded and actual state trajectories. While in sliding mode, the error dynamics are governed by the sliding surface properties. For example, Fig 3.3-b shows how a sliding surface can be used as part of an SMC to stabilize an unstable second order subsystem (in phase-variable canonical form). Once the state trajectory reaches the sliding

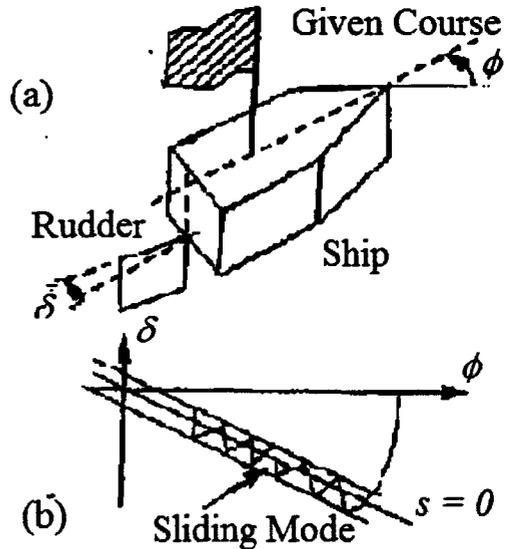


Fig 3.2: (a) Ship Model and its states, (b) Sliding Mode Control of Ship's Course, (Utkin, Guldner, & Shi, 1999).

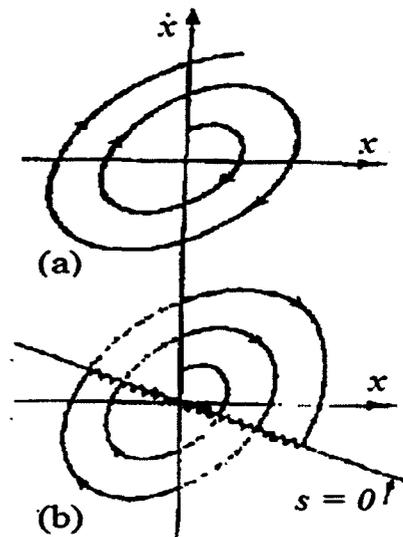


Fig 3.3:(a) Unstable 2nd order Subsystem, (b) SMC effectiveness, (Utkin, Guldner, & Shi, 1999).

surface, the system changes its dynamic behaviour to that defined by the sliding surface and the state trajectory slides along the surface to reach the destination. In sliding mode control, a discontinuous input signal that switches between two values (e.g. c , $-c$) is used to direct the state trajectory towards the switching surface and then to retain them on the surface. If the switching frequency is not infinite, or if there are delays in switching, then the state trajectory chatters in the vicinity of the sliding surface in the region φ defined in Fig 3.4. This condition is referred to as the “real” sliding mode. Other factors contributing to chattering and the real-sliding mode condition are: dead zone, hysteresis and delays, and the bandwidth of the actuator system. The chattering is considered to be the main drawback of the SMC and it results in reducing accuracy which affects the system’s stability, as it excites the un-modelled system dynamics, and wear or higher losses due to its high-frequency switching, [(Utkin, Guldner, & Shi, 1999) and (Perruquetti & Barbot , 2002)].

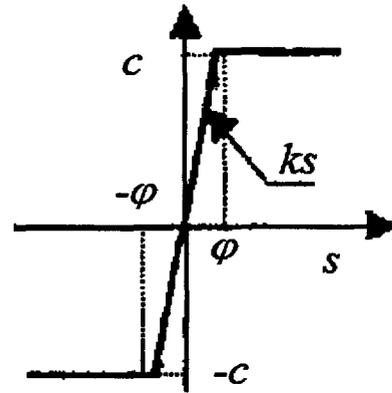


Fig 3.4: Real Discontinuous Control Signal – Chattering, (Utkin, Guldner, & Shi, 1999).

The chattering effect in SMC applications can be reduced by using a smoothing boundary layer where the control signal is interpolated according to the distance of the state trajectory with respect to the sliding surface when within specified boundary layer, (Utkin, Guldner, & Shi, 1999). This action and the smoothing boundary layer properties are further discussed later in chapter four.

3.3. Sliding Mode Observer

3.3.1. Introduction

Sliding Mode Observers (SMOs) were developed in the 1980s as robust estimation techniques against parametric uncertainties and systematic nonlinearity. These types of observers were developed as a natural extension to the Variable Structure (VSC) and Sliding mode (SMC) control, [(Spurgeon, 2008) and (Rolink, Boukhobza, & Sauter, 2006)]. Since the properties of the SMC are necessary for certain applications (i.e. fault construction and detection) and due to the duality between the control and the observer theories, a series of investigations have been reported on the SMO based on the SMC, (Hakiki, Mazari, Liazid, & Djaber, 2006).

An early example of the SMO was done by Slotine in 1986 and 1987. In (Slotine, Hedrick, & Misawa, 1987) and (Slotine, Hedrick, & Misawa, 1986), Slotine et al. modified the Luenberger Observer by adding a discontinuous element that tolerates the nonlinearity in the system. They described their observer using the SMC mold and explored the effectiveness of its gains mathematically. At the same period, Walcott et al. published a landmark paper, (Walcott, Corless, & Zak, 1987), on the linear SMO algorithm and its design methodologies. In that paper, Walcott et al. divided the system into linear and nonlinear subsystems, and then developed the SMO foundations. They specified the conditions that the linear and nonlinear parts of the system should satisfy in order to make the SMO applicable and stable. The basic idea of their work was to solve the constrained Lyapunov problem using algebraic tools. Later in 1994, Edwards et al. expanded the design algorithm to a more general form using symbolic manipulation and defined an explicit design algorithm, [(Edwards, Spurgeon, & Patton, 2000) and (Edwards & Spurgeon, 1994)]. At the beginning of the 21st century, Tan and Edwards presented their observer as an extended version of the Walcott and Zak observer. The proposed observer has less constraint and involves a simpler design method with a different sliding surface, (Tan & Edwards, 2003). These observers have been further modified lately in terms of their gains, the applicable constraints, the sliding hyperplane definition, and the complexity of their target application.

The main limitation of the above mentioned observers was that they were formulated to deal with either nonlinearities or uncertainties in the system. This raised several issues that have led to different forms and implementations based on the Walcott and Zak, and Tan and Edwards observers. Edwards et al. used Linear Matrix Inequality to optimize their design, (Edwards, 2004). Jing et al. manipulated the observer's gains to investigate the separation process between nonlinearity and uncertainties in the system, (Jing, Jing, Changfan, & Cheng, 2008).

3.3.2. SMO principles

In general, the SMO defines a hyperplane, which is referred to as the **sliding surface**, and applies a discontinuous force on the estimate to make it reach the hyperplane. The estimates then remain within a subspace surrounding the sliding hyperplane, referred to as the Existence Subspace. This leads to a motion that consists of three phases; **Reachability**, **Injection** and **Sliding**, (Qaiser, Bhatti, Samar, Iqbal, & Qadir, 2008). The former phase takes place first; as its aim is to bring the estimates to the hyperplane from their initial conditions in finite time. Once the first phase is complete, the other phases are activated simultaneously. The injection phase prevents the estimate

from leaving the existence subspace, while the sliding phase forces the estimated errors to slide along the hyperplane towards the origin. Fig 3.5 illustrates these phases for a second order system considering the error in the first and the second states are e_1 and e_2 , respectively.

The reachability phase is considered the most important phase and it should be designed carefully as it impacts the other phases. If during reachability, the SMO is designed to have a small gain, the observer will have a slow response and may not converge to the sliding hyperplane. This may result in poor performance and may cause instability. Otherwise, with a large gain, the observer will have a fast response and its robustness will be enhanced. However, the observer's performance becomes more sensitive to

measurement noise. During the injection phase large gains contribute to the creation of an artificial noise referred to as *Chattering*. Under this circumstance, when this method is used as a feedback to the SMC, chattering adds a high-frequency element to the control input that is highly undesirable and may excite high frequency dynamics, (Edwards, Spurgeon, & Patel, 2007).

The action of the injection phase under ideal conditions gives the observer the necessary robustness against uncertainties, modeling errors, and/or nonlinearities in the system, [(Edwards, Spurgeon, & Patel, 2007) and (Spurgeon, 2008)]. The average value of the discontinuous action, which is referred to as the *equivalent injection signal*, contains information that can be used to extract the magnitude of uncertainties, modeling errors and/or nonlinearities in the system, (Spurgeon, 2008). The equivalent injection signal is obtained by filtering the discontinuous switching action signal using a digital filter, (Hashimoto, Utkin, Xu, Suzuki, & Harashima, 1990). The various forms of sliding mode observers, the equivalent injection signal, and information extraction as mentioned above are reviewed in the following sections.

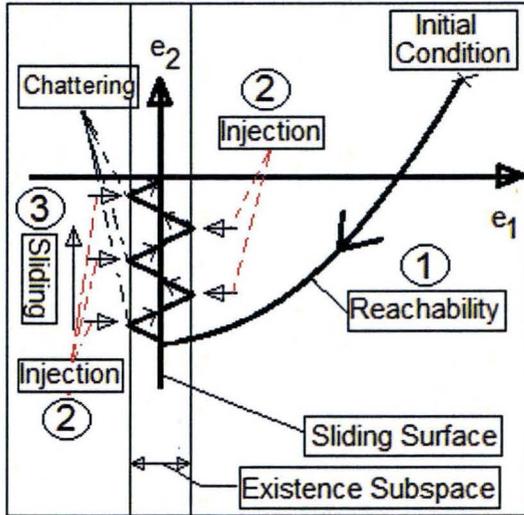


Fig 3.5: Sliding Mode Observer's Phases.

3.4. Discontinuous observer

The *discontinuous observer* is an observer that feeds back the output error between the observer and the system as a discontinuous signal, [(Edwards, Spurgeon, & Patel, 2007), (Hernandez & Barbot, 1996) and (Drakunov & Utkin, 1995)]. It is considered as an *equivalent output error injection* SMO type that is applied on a continuous linear system with the following structure:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad 3.2$$

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) \quad 3.3$$

Where the system, input and output matrices are time invariant and the output matrix should have the following form, (Edwards, Spurgeon, & Patel, 2007):

$$\mathbf{H} = [0_{m \times (n-m)} \quad \mathbf{I}_{m \times m}] \quad 3.4$$

If the output matrix does not have the structure of equation (3.4), a coordinate transformation, \mathbf{T} , is applied on the states to convert the output matrix structure to equation (3.4), (Haskara, Ozguner, & Utkin, 1996). The resultant transformed system is defined as follow, (Edwards, Spurgeon, & Patel, 2007):

$$\dot{\mathbf{y}}(t) = \mathbf{A}'\mathbf{y}(t) + \mathbf{B}'\mathbf{u}(t) \quad 3.5$$

$$\mathbf{z}(t) = \mathbf{H}'\mathbf{y}(t) \quad 3.6$$

Where $\mathbf{y}(t) = \mathbf{T}^{-1}\mathbf{x}(t) = \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix}$, $\mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, $\mathbf{B}' = \mathbf{T}^{-1}\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}$,

$\mathbf{H}' = \mathbf{H}\mathbf{T} = [0_{m \times (n-m)} \quad \mathbf{I}_{m \times m}]$, $\mathbf{A}_{11} \in \mathbb{R}^{(n-m) \times (n-m)}$, $\mathbf{A}_{21} \in \mathbb{R}^{m \times (n-m)}$, $\mathbf{A}_{12} \in \mathbb{R}^{(n-m) \times m}$, $\mathbf{A}_{22} \in \mathbb{R}^{m \times m}$, $\mathbf{B}_1 \in \mathbb{R}^{(n-m) \times 1}$, $\mathbf{B}_2 \in \mathbb{R}^{m \times 1}$, $\mathbf{y}_1(t) \in \mathbb{R}^{(n-m) \times 1}$, $\mathbf{y}_2(t) \in \mathbb{R}^{m \times 1}$ and $\mathbf{z}(t) = \mathbf{y}_2(t)$.

The discontinuous observer has a structure that is similar to the system structure, which is described in equations (3.5) and (3.6). A discontinuous term, which consists of a gain, \mathbf{K}_s , and of a sign function of the output error, \mathbf{N} , is added to obtain the transformed state vector estimate as follows:

$$\dot{\hat{\mathbf{y}}}(t) = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{y}}_1(t) \\ \hat{\mathbf{y}}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix} \mathbf{u}(t) + \mathbf{K}_s \mathbf{N} \quad 3.7$$

Where $\mathbf{K}_s = \begin{bmatrix} \mathbf{K}_{s,1} \\ \mathbf{I}_{m \times m} \end{bmatrix}$, $\mathbf{K}_{s,1} \in \mathbb{R}^{(n-m) \times m}$, $\mathbf{N} = \rho \text{sign}(\mathbf{e}_{\mathbf{y}_2})$ and the output error is defined as $\mathbf{e}_{\mathbf{y}_2}(t) = \mathbf{y}_2(t) - \hat{\mathbf{y}}_2(t)$. From equations (3.6) and (3.7) the estimated error is defined as follows:

$$\begin{bmatrix} \dot{\mathbf{e}}_1(t) \\ \dot{\mathbf{e}}_{y_2}(t) \end{bmatrix} (t) = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{e}_1(t) \\ \mathbf{e}_{y_2}(t) \end{bmatrix} - \mathbf{K}_s \mathbf{N} \quad 3.8$$

Where $\mathbf{e}_1 = \mathbf{e}_{y_1}$. As discussed in section (3.3.2), the scalar ρ should be large enough to force the estimated trajectory to reach the sliding surface in finite time, where the sliding surface is defined to be equal to \mathbf{e}_{y_2} , (Haskara, Ozguner, & Utkin, 1996). Once the estimated trajectory reaches the sliding surface, it starts to slide along it (sliding phase) and never leave it (injection phase). The equivalent output error injection is obtained during the sliding motion, ($\mathbf{e}_{y_2} = \dot{\mathbf{e}}_{y_2} = 0$). It represents the average energy applied on the estimate to maintain it on the sliding surface and it is obtained by filtering the discontinuous gain, \mathbf{N} , by using a low pass filter. Using equation (3.8) and the condition $\mathbf{e}_{y_2} = \dot{\mathbf{e}}_{y_2} = 0$, the equivalent output error injection value is proposed as follows, (Edwards, Spurgeon, & Patel, 2007):

$$\mathbf{N}_{eq} = \mathbf{A}_{21} \mathbf{e}_1 \quad 3.9$$

By substituting equation (3.9) into equation (3.8), the vector $\dot{\mathbf{e}}_1$ can be obtained as follows, (Edwards, Spurgeon, & Patel, 2007):

$$\dot{\mathbf{e}}_1 = (\mathbf{A}_{11} - \mathbf{K}_{s,1} \mathbf{A}_{21}) \mathbf{e}_1 \quad 3.10$$

The term $(\mathbf{A}_{11} - \mathbf{K}_{s,1} \mathbf{A}_{21})$ must have stable eigenvalues to guarantee the SMO stability, [(Edwards, Spurgeon, & Patel, 2007) and (Drakunov & Utkin, 1995)]. In addition to stability, the gain $\mathbf{K}_{s,1}$ impacts the convergence speed and the dynamic response of the observer, [(Haskara, Ozguner, & Utkin, 1996) and (Edwards, Spurgeon, & Patel, 2007)].

Hahimoto et al. proposed an observer that is similar to the Discontinuous Observer. However, they ignored the condition of equation (3.4) and assumed that the output matrix could be time varying with fixed rank. They proposed a method that maps the output to the hidden states by taking the output for several time steps and then mapping them to the states using the output and the system matrices (i.e. by using the Observability matrix), (Hashimoto, Utkin, Xu, Suzuki, & Harashima, 1990).

Their observer design can be used for unknown input signal reconstruction and fault detection and isolation.

The discontinuous observers have several limitations that can be summarized as follows, [(Edwards, Hebden, & Spurgeon, 2005) and (Edwards, 2004)]:

- This type of observer assumes that the system and the measurement noise do not exist. Presence of noise impacts the efficiency and the performance of the observer.
- They are sensitive to the design parameter, ρ . If ρ has a small value, it will result in a slow convergence rate to the sliding surface, or lead to instability. On the other hand, a large value of ρ leads to chattering.
- This type of observer assumes that the nonlinear system consists of a linear part, which is considered to be the true system's structure, and a nonlinear part, which is considered to be the uncertainties. This assumption holds if the linear part is dominant. In other words, the ratio between the dynamic contribution of the nonlinear part over the linear part is small, (Madani & Benallegue, 2007). This assumption is not valid for all applications. For systems that are nonlinear, their model is linearized around the previous estimate (Bandyopadhyay, Gandhi, & Kurode, 2009). The linearization increases computational complexity and adds to estimation error.

3.5. Observer with linear and discontinuous injection

The type of observers with linear and discontinuous corrective terms entails a simple modification that is applied to the discontinuous observer. This modification ensures that the observer has fast convergence to the surface and reduces chattering, (Edwards, Spurgeon, & Patel, 2007). The sliding patch is defined as the area in the estimated error trajectories in which the sliding and injection phases occurs. From Fig 3.5, the sliding patch is the region on the sliding surface that is bounded by $|e_2| \leq K_{1,s}$ and $e_1 = 0$. Slotine et al. in (Slotine, Hedrick, & Misawa, 1987) and Walcott and Zak in (Walcott & Zak, 1987) provided the earliest published examples of this type of observers. Their observer designs consisted of a linear gain and the output error. Although they worked independently, their observer gains, \mathbf{K} , have the same structure which is defined as follows:

$$\mathbf{K} = \mathbf{K}_s \mathbf{N} + \mathbf{K}_l \mathbf{e}_z \quad 3.11$$

where $K_l \in \mathbb{R}^{n \times m}$, is the proportional gain and it is designed to satisfy requirements related to dynamic response as well as convergence rate and optimality (Slotine, Hedrick, & Misawa, 1987). The gain \mathbf{K}_l forces the term $(\mathbf{A}' - \mathbf{K}_l \mathbf{H}')$ to have negative eigenvalues and guarantees stability given bounded uncertainties, [(Aurora & Ferrara, 2007) and (Chaal, Jovanovic, & Busawon, 2009)].

Adding the gain K_l increases the complexity of the observer's design, nonetheless it improves the observer performance in terms of robustness and stability (Spurgeon, 2008). In the following sub-sections, three specific forms of this type of observer are discussed in detail, namely the Slotine et al. observer, the Walcott and Zak observer, and the Convex Parameterization.

3.5.1. Slotine et al. Observer

Slotine et al. presented a combined observer with a gain that is composed of a linear and a discontinuous term, (Slotine, Hedrick, & Misawa, 1987). This observer can be applied to nonlinear systems that have the following canonical form:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) + \mathbf{B}\boldsymbol{\mu}(t, \mathbf{z}, u) \quad 3.12$$

$$z(t) = \mathbf{H}\mathbf{x}(t) \quad 3.13$$

where $\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \vdots & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & 0 & \ddots & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ and it has dimensions of $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$ with

dimensions of $\mathbf{B} \in \mathbb{R}^{n \times 1}$, $\mathbf{H} = [1 \ 0 \ \dots \ 0]$ with dimensions of $\mathbf{H} \in \mathbb{R}^{1 \times n}$, and $\boldsymbol{\mu}(t, \mathbf{z}, u)$ is an unknown nonlinear function with dimensions of $\boldsymbol{\mu}(t, \mathbf{z}, u) \in \mathbb{R}^{1 \times 1}$.

Slotine et al. explored the mathematical background behind using the gains \mathbf{K}_l and \mathbf{K}_s , and tested their effectiveness on the observer's performance. They did their investigations on a second order system with the same structure of equations (3.12) and (3.13). The observer has the following form:

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{B}u(t) + \mathbf{B}\hat{\boldsymbol{\mu}}(t, \mathbf{z}, u) + \mathbf{K}_l e_z + \mathbf{K}_s N \quad 3.14$$

$$\hat{z}(t) = \mathbf{H}\hat{\mathbf{x}}(t) \quad 3.15$$

where $\hat{\boldsymbol{\mu}}(t, \mathbf{z}, u)$ is a nonlinear estimated function for $\boldsymbol{\mu}(t, \mathbf{z}, u)$, $e_z = e_1 = z - \hat{z}$, $N = \text{sgn}(e_z)$ with $\mathbf{N} \in \mathbb{R}^{1 \times 1}$, $\mathbf{K}_l = \begin{bmatrix} K_{l,1} \\ K_{l,2} \end{bmatrix}$, $\mathbf{K}_s = \begin{bmatrix} K_{s,1} \\ K_{s,2} \end{bmatrix}$, and the error in $\hat{\boldsymbol{\mu}}(t, \mathbf{z}, u)$, $\Delta\boldsymbol{\mu}$, is assumed to be bounded by $K_{s,2}$. Slotine et al. reported the following.

- If the observer's gains $K_{l,1}$, $K_{l,2}$ and $K_{s,1}$ are designed to have zero values, and the gain $K_{s,2}$ is set to have a value greater than $|\Delta\boldsymbol{\mu}|$, then the SMO's behaviour is bounded-input bounded-output stable and oscillated with a magnitude that depended on the initial conditions and the value of $K_{s,2}$ as shown in Fig 3.6.

- The addition of the gain $K_{s,1}$ results in two types of motions; one before reaching the sliding patch and the other in the sliding mode, as shown in Fig 3.7. The first type of motion forces the trajectory towards the sliding patch which represents the reachability phase. Once the trajectory reaches that region, which is on the e_2 -axis and bordered by the interval $(-K_{s,1}, K_{s,1})$, the sliding and injection phases occur. The gain $K_{s,1}$ affects the reachability phase, as the convergence to the patch becomes faster when $K_{s,1}$ is increased, as shown in Fig 3.8.

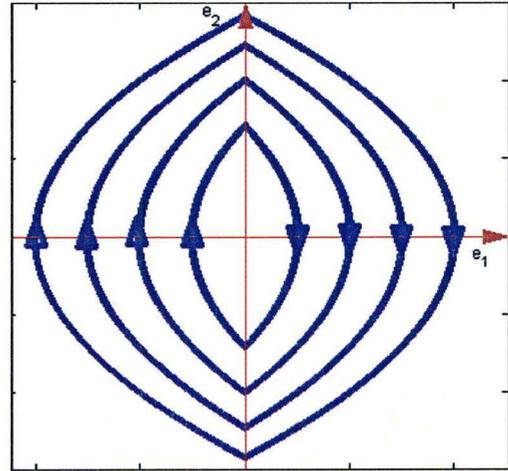


Fig 3.6: The Slotine et al. Observer behaviour for $K_{L,1} = K_{L,2} = K_{s,1} = 0$ and $K_{s,2} \geq \Delta\mu$.

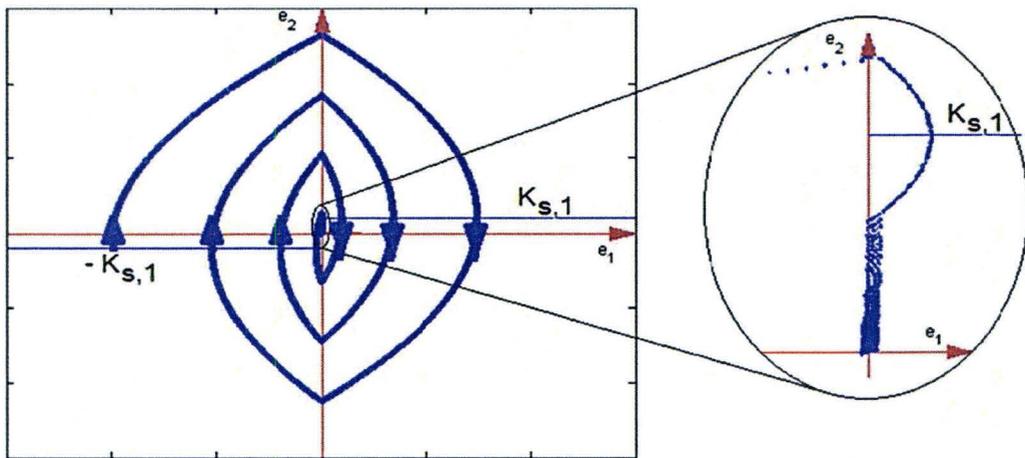


Fig 3.7: The Slotine et al. Observer behaviour for $K_{s,1} \neq 0$.

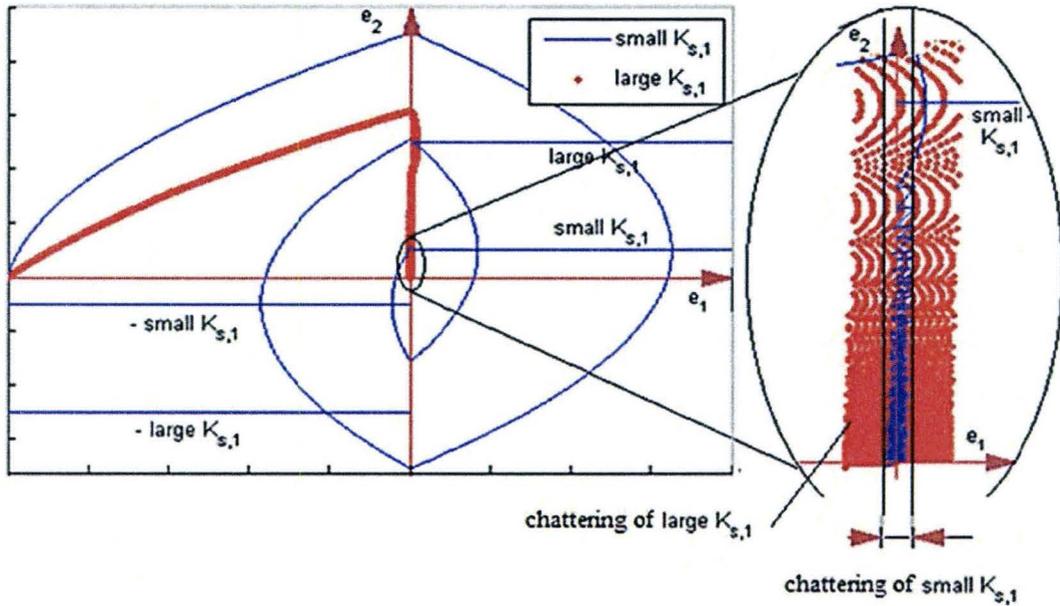


Fig 3.8: The effect of the gain $K_{s,1}$.

- The sliding motion inside the patch depends on the gains $K_{s,1}$ and $K_{s,2}$ as follows:
 - The dynamics on the patch decays exponentially with the ratio $K_{s,2}/K_{s,1}$.
 - The chattering amplitude depends only on the value $K_{s,1}$, and is proportional to $K_{s,1}$'s magnitude.
- The gain $K_{l,1}$ increases the size of the “Direct Attraction”, (Slotine, Hedrick, & Misawa, 1987). From Fig 3.7, the observer switches direction around the axis $K_{s,1}$, for positive e_1 , and $-K_{s,1}$, for negative e_1 . By adding the gain $K_{l,1}$, these axes start to have a slope and they are not horizontal anymore, as shown in Fig 3.9. In (Slotine, Hedrick, & Misawa, 1987), the slopes of these axes have the same values as the added gains as shown in Fig 3.7. This leads to a faster reachability without affecting the sliding patch and motion. Moreover, it does not affect the chattering amplitude.

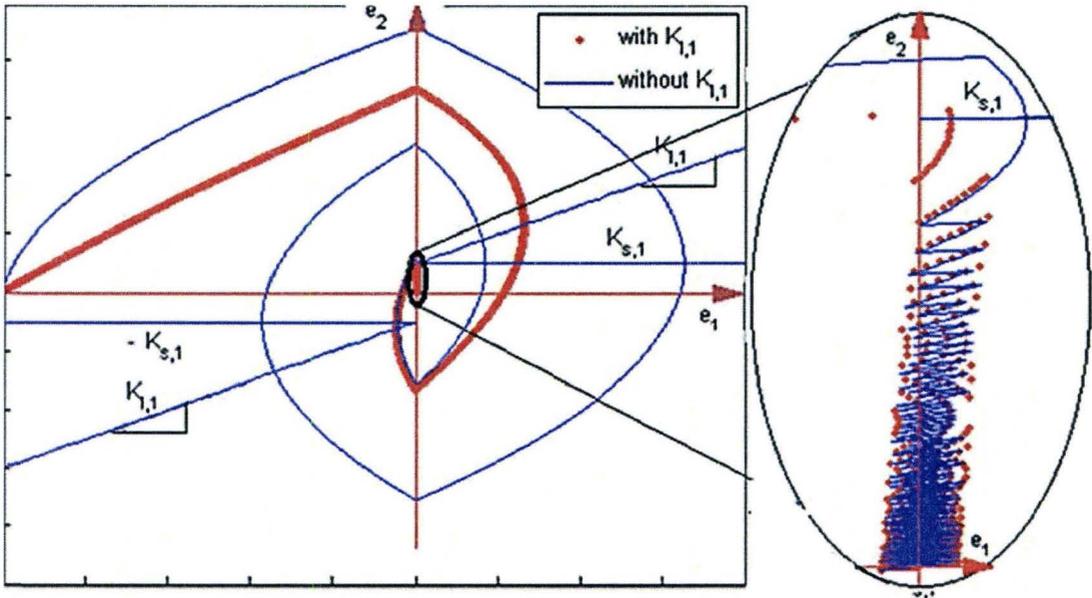


Fig 3.9: The effect of the gain $K_{l,1}$.

- The gain $K_{l,2}$ affects the reachability of the observer as shown in Fig 3.10. As the gain value is increased, the convergence of the observer becomes slower. However, this has a negligible impact on the size of the sliding patch.

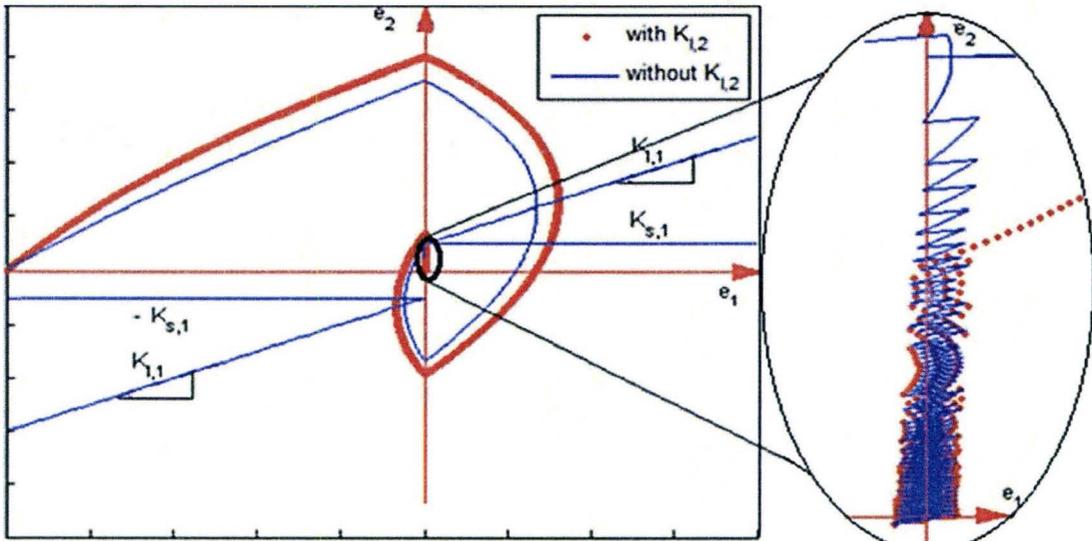


Fig 3.10: The effect of the gain $K_{l,2}$.

- Designing the gains $K_{l,1}$ and $K_{l,2}$ is dependent on the application. For example, Slotine et al. in (Slotine, Hedrick, & Misawa, 1987) used two methods; the Kalman Filter's gain and the Luenberger Observer's gain.

3.5.2. The Walcott and Zak Observer

Walcott and Zak presented their observer design as a modified form of the discontinuous observer to accommodate a multi-input signal as well as the nonlinearities/uncertainties in the system, [(Edwards, Spurgeon, & Patel, 2007), (Spurgeon, Edwards, & Foster, 1996) and (Spurgeon, 2008)]. In this method, the system is assumed to consist of a linear part and of a bounded unknown nonlinear part, $\mu(t, z, u)$, as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{B}\boldsymbol{\mu}(t, z, u) \quad 3.16$$

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) \quad 3.17$$

Where $\mathbf{u}(t)$ and $\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u})$ are a multi-input signal and a multi-dimensional nonlinear function, respectively, and $\mathbf{u}(t)$, $\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u}) \in \mathbb{R}^{p \times 1}$, and $\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u})$ is bounded by $\rho(t, \mathbf{z}, \mathbf{u})$ such that $\|\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u})\| \leq \rho(t, \mathbf{z}, \mathbf{u})$.

The measurement matrix is subjected to the condition of equation (3.4). Otherwise a transformation, \mathbf{T} , is applied and the resultant transformed system, \mathbf{y} , has the following structure:

$$\dot{\mathbf{y}}(t) = \mathbf{A}'\mathbf{y}(t) + \mathbf{B}'\mathbf{u}(t) + \mathbf{B}'\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u}) \quad 3.18$$

$$\mathbf{z}(t) = \mathbf{H}'\mathbf{y}(t) \quad 3.19$$

Where $\mathbf{H}' = \mathbf{H}\mathbf{T} = [\mathbf{0}_{m \times (n-m)} \quad \mathbf{I}_{m \times m}]$, $\mathbf{y}(t) = \mathbf{T}^{-1}\mathbf{x}(t) = \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix}$, $\mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$, $\mathbf{B}' = \mathbf{T}^{-1}\mathbf{B}$, $\mathbf{y}_1(t) \in \mathbb{R}^{(n-m) \times 1}$, $\mathbf{y}_2(t) \in \mathbb{R}^{m \times 1}$, and $\mathbf{B}' \in \mathbb{R}^{n \times p}$.

The Walcott and Zak observer has a structure that is similar to the Slotine observer in equation (3.11). However, the discontinuous part in that equation, $\mathbf{K}_s\mathbf{N}$, and the linear proportional gain, \mathbf{K}_l , have different forms. The matrices \mathbf{N} , \mathbf{K}_s and \mathbf{K}_l , are defined as follows:

$$\mathbf{N} = \boldsymbol{\rho}(t, \mathbf{z}, \mathbf{u})\text{sign}(\mathbf{C}\mathbf{e}_z) \quad 3.20$$

$$\mathbf{K}_s = \mathbf{P}^{-1}\mathbf{H}'\mathbf{T}\mathbf{C}^T \quad 3.21$$

$$\mathbf{K}_l = \begin{bmatrix} \mathbf{A}_{12} \\ \mathbf{A}_{22} - \mathbf{A}_{22}^s \end{bmatrix} \quad 3.22$$

Where:

\mathbf{P} is a square positive definite matrix with size of $\mathbf{P} \in \mathbb{R}^{n \times n}$ and is chosen to satisfy the following condition:

$$\mathbf{P}(\mathbf{A}' - \mathbf{K}_l \mathbf{H}') + (\mathbf{A}' - \mathbf{K}_l \mathbf{H}')^T \mathbf{P} < 0 \quad 3.23$$

\mathbf{C} is a design matrix with a size of $\mathbf{C} \in \mathbb{R}^{p \times m}$, and is used to define the sliding surface as $\mathbf{C}\mathbf{e}_y$, while subject to the following condition:

$$\mathbf{P}\mathbf{B}' = (\mathbf{C}\mathbf{H}')^T \quad 3.24$$

\mathbf{A}_{22}^s is a design matrix that has the same size as \mathbf{A}_{22} and with negative eigenvalues. The Walcott and Zak observer has several limitations as its applications are limited due to the numerous conditions that must be satisfied. These include the conditions of equations (3.23) and (3.7), in addition to the following, (Spurgeon, 2008) and (Edwards, Spurgeon, & Patel, 2007):

- $rank(\mathbf{H}\mathbf{B}) = p$, where p is the length of $\boldsymbol{\mu}$. If this condition is satisfied, then the outputs are linked to the source of uncertainties (observable).
- The matrix \mathbf{A}_{22} must have negative eigenvalues.
- The system transfer function defined as $\mathbf{G}(s) = \mathbf{H}(s\mathbf{I}_{n \times n} - \mathbf{A})^{-1}\mathbf{B}$, must not have any invariant positive zeros or poles.

3.5.3. Convex Parameterization/Tan and Edwards's Observer

In 2003, Tan and Edwards presented their observer as a modified form of the Walcott and Zak's observer, (Tan & Edwards, 2003). Their main aim was to make it applicable to more general problems with the added advantages of having extra degrees of freedom and requiring less design effort [(Edwards, Spurgeon, & Patel, 2007) and (Spurgeon, 2008)]. Later, the Tan and Edwards observer was referred to as the Convex Parameterization (Edwards, Spurgeon, & Patel, 2007). In their design, the uncertainties/nonlinearities' matrix has a more general form and it is not limited to multiplication by the input matrix. The system is assumed to have the following form:

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{D}\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u}) \quad 3.25$$

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) \quad 3.26$$

Where $\mathbf{D} = \begin{bmatrix} \mathbf{0}_{(n-q) \times q} \\ \mathbf{D}_2 \end{bmatrix}$ with dimensions of $\mathbf{D} \in \mathbb{R}^{n \times q}$, $\mathbf{H} = [\mathbf{0}_{m \times (n-m)} \quad \mathbf{T}_0]$, $\mathbf{u}(t) \in \mathbb{R}^{p \times 1}$ is the input and $\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u}) \in \mathbb{R}^{q \times 1}$ is a nonlinear function bounded by $\boldsymbol{\rho}(t, \mathbf{z}, \mathbf{u})$ such that $\|\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u})\| \leq \boldsymbol{\rho}(t, \mathbf{z}, \mathbf{u})$.

From equations (3.25) and (3.26), the measurement matrix, \mathbf{H} , and the uncertainties/nonlinearities input matrix, \mathbf{D} , must have specific forms. If the system does not have this canonical form, a coordinate transformation, \mathbf{T} , is applied to the system's equations, and the resultant system structure is redefined as follows:

$$\dot{\mathbf{y}}(t) = \mathbf{A}'\mathbf{y}(t) + \mathbf{B}'\mathbf{u}(t) + \mathbf{D}'\boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u}) \quad 3.27$$

$$\mathbf{z}(t) = \mathbf{H}'\mathbf{y}(t) \quad 3.28$$

Where $\mathbf{y}(t) = \mathbf{T}^{-1}\mathbf{x}(t) = \begin{bmatrix} \mathbf{y}_1(t) \\ \mathbf{y}_2(t) \end{bmatrix}$, $\mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \left\{ \mathbf{A}_{21} = \begin{bmatrix} \mathbf{A}_{211} \\ \mathbf{A}_{212} \end{bmatrix} \right\} & \mathbf{A}_{22} \end{bmatrix}$, $\mathbf{B}' = \mathbf{T}^{-1}\mathbf{B}$, $\mathbf{H}' = \mathbf{H}\mathbf{T} = [\mathbf{0}_{m \times (n-m)} \quad \mathbf{T}_0]$, $\mathbf{D}' = \mathbf{T}^{-1}\mathbf{D} = \begin{bmatrix} \mathbf{0}_{(n-q) \times q} \\ \mathbf{D}_2 \end{bmatrix}$, $\mathbf{T}_0 \in \mathbb{R}^{m \times m}$, $\mathbf{D}_2 \in \mathbb{R}^{q \times q}$, $\mathbf{A}_{11} \in \mathbb{R}^{(n-m) \times (n-m)}$, $\mathbf{A}_{21} \in \mathbb{R}^{m \times (n-m)}$, $\mathbf{A}_{12} \in \mathbb{R}^{(n-m) \times m}$, $\mathbf{A}_{22} \in \mathbb{R}^{m \times m}$, $\mathbf{A}_{211} \in \mathbb{R}^{(m-q) \times (n-m)}$, $\mathbf{A}_{212} \in \mathbb{R}^{q \times (n-m)}$, $\mathbf{y}_1(t) \in \mathbb{R}^{(n-m) \times 1}$, and $\mathbf{y}_2(t) \in \mathbb{R}^{m \times 1}$.

According to Walcott and Zak, the system must satisfy the following conditions in order to have a solution:

- $\text{rank}(\mathbf{H}\mathbf{D}) = q$, i.e. the uncertainties/nonlinearities are explicitly linked to the measurement hence they are Observable.
- The system transfer function, defined as $\mathbf{G}(s) = \mathbf{H}(s\mathbf{I}_{n \times n} - \mathbf{A})^{-1}\mathbf{B}$, must not have any invariant positive poles.

The Convex Parameterization observer has a structure that is similar to the Slotine's, and the Walcott and Zak's observers. However, its discontinuous part, $\mathbf{K}_s\mathbf{N}$, has a different form with \mathbf{N} defined as:

$$\mathbf{N} = \boldsymbol{\rho}(t, \mathbf{z}, \mathbf{u}) \left| \mathbf{P}_0 \mathbf{T}_0 \begin{bmatrix} \mathbf{0}_{(m-q) \times q} \\ \mathbf{D}_2 \end{bmatrix} \right| \text{sign}(\mathbf{e}_z) \quad 3.29$$

Where the matrix \mathbf{P}_0 is a symmetric positive definite matrix defined as a function of another positive definite matrix $\mathbf{P} \in \mathbb{R}^{n \times n}$ defined as:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & -\mathbf{P}_1[\mathbf{K}_{s,1} \quad \mathbf{0}_{(n-m) \times q}] \\ -[\mathbf{K}_{s,1} \quad \mathbf{0}_{(n-m) \times q}]^T \mathbf{P}_1 & \mathbf{P}_2 + [\mathbf{K}_{s,1} \quad \mathbf{0}_{(n-m) \times q}]^T \mathbf{P}_1 [\mathbf{K}_{s,1} \quad \mathbf{0}_{(n-m) \times q}] \end{bmatrix} \quad 3.30$$

\mathbf{P} is chosen to satisfy the condition of equation (3.23). The matrix \mathbf{P}_0 is defined as:

$$\mathbf{P}_0 = \mathbf{T}_0 \mathbf{P}_2 \mathbf{T}_0^T \quad 3.31$$

Note that \mathbf{T}_0 is the lower portion of the measurement matrix.

The sliding surface is defined in terms of the output error. The gain \mathbf{K}_s has a different structure from the other observers and is defined as follows:

$$\mathbf{K}_s = \begin{bmatrix} [\mathbf{K}_{s,1} & \mathbf{0}_{(n-m) \times q}] \mathbf{T}_0^T \\ \mathbf{T}_0^T \end{bmatrix} \mathbf{P}_0^{-1} \quad 3.32$$

where the gain $\mathbf{K}_{s,1}$ is chosen to force the term $(\mathbf{A}_{11} - \mathbf{K}_{s,1} \mathbf{A}_{21})$ to be stable.

To obtain the linear proportional gain, \mathbf{K}_1 , another transformation, \mathbf{T}_L , is applied to the states to transform the measurement matrix into an identity matrix as follows:

$$\mathbf{T}_L = \begin{bmatrix} \mathbf{I}_{(n-m) \times (n-m)} & [\mathbf{K}_{s,1} & \mathbf{0}_{(n-m) \times q}] \\ \mathbf{0}_{m \times (n-m)} & \mathbf{T}_0 \end{bmatrix}^{-1} \quad 3.33$$

\mathbf{T}_L converted the system's equations (3.22) and (3.23) to the following:

$$\dot{\boldsymbol{\zeta}}(t) = \mathbf{A}'' \boldsymbol{\zeta}(t) + \mathbf{B}'' \mathbf{u}(t) + \mathbf{D}'' \boldsymbol{\mu}(t, \mathbf{z}, \mathbf{u}) \quad 3.34$$

$$\mathbf{z}(t) = \mathbf{H}'' \boldsymbol{\zeta}(t) \quad 3.35$$

where $\boldsymbol{\zeta}(t) = \mathbf{T}_L^{-1} \mathbf{y}(t) = \begin{bmatrix} \boldsymbol{\zeta}_1(t) \\ \boldsymbol{\zeta}_2(t) \end{bmatrix}$, $\mathbf{A}'' = \mathbf{T}_L^{-1} \mathbf{A}' \mathbf{T}_L = \begin{bmatrix} \mathbf{A}'_{11} & \mathbf{A}'_{12} \\ \mathbf{A}'_{21} & \mathbf{A}'_{22} \end{bmatrix}$, $\mathbf{B}'' = \mathbf{T}_L^{-1} \mathbf{B}'$, $\mathbf{H}'' = \mathbf{H}' \mathbf{T}_L = [\mathbf{0}_{m \times (n-m)} \quad \mathbf{I}_{m \times m}]$, $\mathbf{D}'' = \mathbf{T}_L^{-1} \mathbf{D}' = \begin{bmatrix} \mathbf{0}_{(n-q) \times q} \\ \mathbf{T}_0 \mathbf{D}_2 \end{bmatrix}$, and \mathbf{A}'_{11} , \mathbf{A}'_{12} , \mathbf{A}'_{21} and \mathbf{A}'_{22} had the same dimensions as \mathbf{A}_{11} , \mathbf{A}_{12} , \mathbf{A}_{21} and \mathbf{A}_{22} , respectively. The gain \mathbf{K}_1 is defined as follows:

$$\mathbf{K}_1 = \begin{bmatrix} \mathbf{A}'_{12} \\ \mathbf{A}'_{22} - \mathbf{A}_{22}^s \end{bmatrix} \quad 3.36$$

where \mathbf{A}_{22}^s is a design matrix that has negative eigenvalues. The discontinuous gain, \mathbf{K}_s , has a simpler form and is defined as follows:

$$\mathbf{K}_s = \begin{bmatrix} \mathbf{0}_{(n-m) \times m} \\ \mathbf{P}_0^{-1} \end{bmatrix} \quad 3.37$$

This observer needs less design effort as compared to the Walcott and Zak observer. Its applications are however limited due to numerous conditions that must be satisfied, and due to the fact that there are no system and measurement noise. A revised and simpler version of this observer was proposed in 1996. The revised version uses the sliding surface of the Walcott and Zak observer. However, the matrix \mathbf{C} has a fixed value that depends on the matrices \mathbf{P}_2 , \mathbf{D}_2 and \mathbf{T}_0 . The gain \mathbf{K}_s and the matrix \mathbf{D} are chosen to be equal to the input matrix, \mathbf{B} , (Spurgeon, Edwards, & Foster, 1996).

3.6. SMO applications

The SMO is robust to uncertainties, nonlinearities, modeling errors, and noise as shown experimentally by Kim et al., (Kim, Kim, & Youn, 2006). Moreover, the discontinuous part in the SMO gain can provide additional information for estimation. Therefore, the SMO has been used in applications where robustness and signal reconstruction are of primary concern such as the applications in the following subsections.

3.6.1. Fault detection and isolation

The main aims of fault detection and isolation algorithms are to detect unusual changes in the system's condition (e.g. faults), and to identify their source or location. Using the sliding mode observer with the equivalent output error injection techniques, a fault can be detected and tracked. A common strategy for fault detection and diagnosis is to assume that the fault condition results in additive dynamic effects. For example, if a system fault results in the additive term $\mu(t, z, u)$ and a sensor fault results in $\mu_1(t, z, u)$, these additive terms can be extracted as, [(Hakiki, Mazari, Liazid, & Djaber, 2006) and (Tan & Edwards, 2004)]:

$$\mu(t, z, u) = (\mathbf{D}_2^T \mathbf{D}_2)^{-1} \mathbf{D}_2^T \mathbf{N}_{eq}(t, z, u) \quad 3.38$$

$$\mu_1(t, z, u) = (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12})^{-1} \mathbf{N}_{eq}(t, z, u) \quad 3.39$$

Where $\mathbf{N}_{eq}(t, z, u)$ is the equivalent output error injection obtained by filtering the discontinuous signal, \mathbf{A}_{11} , \mathbf{A}_{12} , \mathbf{A}_{21} and \mathbf{A}_{22} are sub-matrices from the system matrix, and \mathbf{D}_2 is a sub-matrix from the fault's input matrix (refer to section 3.5.3 for further details). A digital filter can be used for obtaining the signal \mathbf{N}_{eq} . This has however two disadvantages:

- The filtered signal will lose some of its information.
- The filtered signal will be phase shifted in real time applications.

The SMO for fault detection and isolation can be found in the following applications:

- Tan and Edwards used their observer to detect and identify faults caused by parametric uncertainties in a DC motor, (Tan & Edwards, 2004).

- In (Hakiki, Mazari, Liazid, & Djaber, 2006), an observer that is similar to the Tan and Edwards observer was used to reconstruct the system and sensor faults for an inverted pendulum attached to a cart.
- Drakunov et al. used the discontinuous SMO to reconstruct an electro-motive force of a car's electric circuit, (Drakunov, Utkin, Zarei, & Miller, 1996).
- Chen and Moskwa used the SMO to obtain a cylinder pressure in a nonlinear multi-cylinder engine based on measurements of the crankshaft dynamics, (Chen & Moskwa, 1997). Their observer relied on linearizing the system matrix for greater simplicity, and applied the EKF gain instead of the proportional gain in the SMO. The attempt of optimality impacted robustness and the observer reportedly diverged when subjected to large modeling errors.
- Zhao et al. used the SMO to obtain the road grade and bank angles as well as estimating the lateral and longitudinal velocities of a vehicle on a 'non-flat road', (Zhao, Liu, & Chen, 2009). The nonlinearities were linearized around the estimates and used in the functional definition of the discontinuous gain. Despite linearization, the observer provided robust performance.

3.6.2. Sensorless control scheme.

For sensorless control applications, an observer is needed to extract the hidden states from measurements, and then feed these to the control scheme. The observer must be robust; otherwise the stability and performance of the controller are compromised. It should be noted that for this kind of application, the observer's model must be close to that of the system's. Otherwise, the results are compromised. Examples of the SMO used in sensorless control are as follows.

- McCann et al. used the SMO in its discontinuous form to estimate the position and the velocity of a rotor and feed these back to control a switched reluctance motor in an automotive hydraulic brake system, [(McCann & Husain, 1997) and (McCann, Islam, & Husain, 2001)].
- Yan and Utkin used the SMO to estimate the stator current and the hidden rotor flux of an Induction machine using the equivalent output error injection method, (Yan & Utkin, 2002).

- Bandyopadhyay et al. used the SMO to estimate and feedback the states of a container's slosh-free motion (as shown in Fig 3.11, [(Bandyopadhyay, Gandhi, & Kurode, 2009) and (Kurode, Bandyopadhyay, & Gandhi, 2009)]. Due to the nonlinearity in the system matrix, the error equation was linearized around the current measurement error. The measurement signal was filtered using a digital filter to improve the observer performance. However, filtering leads to delay in time. The SMO showed fast response and convergence rate, and high robustness performance despite linearization truncation

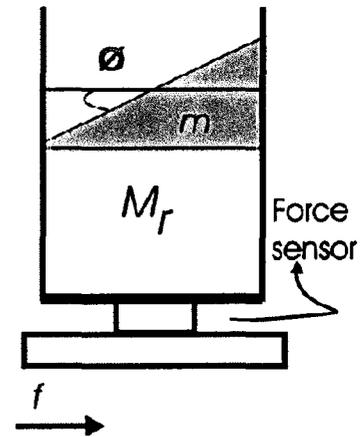


Fig 3.11: Slosh mass example, (Bandyopadhyay, Gandhi, & Kurode, 2009).

error, modelling errors, noise and filtering techniques.

- Chao and Shen used the SMO on a highly nonlinear system referred to as the “three axis four wire type pickup actuator” (this actuator is used in high data-density optical disks), (Chao & Shen, 2009). The Slotine’s observer (presented in section (3.5.1)) was used on the linearized system matrix. The observer demonstrated high resolution and performance around the required positional accuracy in the nano-meter range, (Chao & Shen, 2009).

3.6.3. Parameter estimation.

Observers can be used to estimate system parameters as well as states. This can be done by defining the investigated parameter as a dummy state resulting in a nonlinear model representation. The SMO for parameter estimation can be found in the following applications:

- ❖ Hasan and Husain used the SMO for parameter estimation for an induction motor, (Hasan & Husain, 2009). The observer that they used is similar to that of the Tan and Edwards (presented in section (3.5.3)), where the uncertainty function, $\mu(t, z, u)$, is defined as the error in the system matrix and the states, $\Delta A x(t)$. Their method was based on the equivalent output error injection and the Lyapunov’s second method for stability. The proposed observer had advantages over the EKF as it did not require linearization, it was simpler and it had a faster response.

- ❖ Aurora and Ferrara designed a more complex observer than Hasan and Husain's to accommodate higher modeling errors (with errors up to 150% of the nominal values), (Aurora & Ferrara, 2007). Their observer consisted of the following:
 - A discontinuous switching term that is a function of the estimates and measurements.
 - A proportional term that is linked to the output error.
 - An integral of the discontinuous term.

In this method, the choice of the Lyapunov function allows a functional definition of the parameters in a differential equation form. These equations are used in reformulating the system model to obtain an augmented state vector consisting of the original states and the parameters, [(Aurora & Ferrara, 2007) and (Aurora, Ferrara, & Levant, 2001)].

Chapter Four:

The Smooth Variable Structure Filter for Systems with Full Ranked Measurement Matrix

4.1 Historical and mathematical background

4.1.1 The Variable Structure Filter

In 2003, a new estimation method referred to as the Variable Structure Filter (VSF) was proposed for its applications to linear systems. This method is a recursive predictor-corrector filter that is based on the sliding mode concept. The VSF defines a hyperplane and then applies a corrective action that forces the estimate to go back and forth cross that plane. Although this method uses a discontinuous corrective action, it has a different structure than other Sliding Mode Observers (SMO) [(Habibi & Burton, 2002), (Habibi & Burton, 2003), (Wang, Burton, & Habibi, 2009), and (Habibi, Burton, & Chinniah, 2002)]. The VSF's advantages are its predictor-corrector form and its stronger physical meaning for the uncertainties and noise compared to the SMO [(Habibi & Burton, 2003), (Habibi & Burton, 2004) and (Habibi & Burton, 2007)]. The VSF has been developed for the systems described by the following equation:

$$\begin{aligned}\mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}u_{k-1} + \mathbf{w}_{k-1} \\ \mathbf{z}_k &= \mathbf{H}\mathbf{x}_k + \mathbf{v}_k\end{aligned}\quad 4.1$$

Where \mathbf{A} , \mathbf{B} and \mathbf{H} are the time invariant system, input and measurement matrices. The VSF is a predictor corrector method and can be summarized as follows, (Habibi & Burton, 2003):

1 - Prediction Stage:

The a priori state estimate is obtained by using an estimated model of the system as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \hat{\mathbf{A}}\hat{\mathbf{x}}_{k-1|k-1} + \hat{\mathbf{B}}u_{k-1} \\ \hat{\mathbf{z}}_{k|k-1} &= \hat{\mathbf{H}}\hat{\mathbf{x}}_{k|k-1}\end{aligned}\quad 4.2$$

2 - Correction Stage:

A corrective gain is calculated and used for refining the a priori estimate into its a posteriori form as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{VSF_k} \\ \hat{\mathbf{z}}_{k|k} &= \hat{\mathbf{H}}\hat{\mathbf{x}}_{k|k}\end{aligned}\quad 4.3$$

The term \mathbf{K}_{VSF_k} is the VSF's gain and is defined as follows, (Habibi & Burton, 2003):

$$\begin{aligned} \mathbf{K}_{VSF_k} = & \hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ \left[\left(|\hat{\mathbf{H}}\hat{\mathbf{A}}| \left\{ \boldsymbol{\Upsilon}|\hat{\mathbf{H}}^+| \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| + \left| \hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ (\mathbf{H}\mathbf{A}\mathbf{H}^+ - \hat{\mathbf{H}}\hat{\mathbf{A}}\hat{\mathbf{H}}^+)_{\max} \mathbf{z}_k \right| \right. \right. \right. \\ & + \left. \left. \left[|\hat{\mathbf{H}}^+| + |\hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+| \left(\begin{array}{c} (\mathbf{H}\mathbf{A}\mathbf{H}^+ - \hat{\mathbf{H}}\hat{\mathbf{A}}\hat{\mathbf{H}}^+)_{\max} \\ + \hat{\mathbf{H}}\hat{\mathbf{A}}\hat{\mathbf{H}}^+ + \mathbf{I}_{m \times m} \end{array} \right) \right] (\mathbf{v})_{\max} \right. \right. \\ & + \left. \left. \left| \hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ (\mathbf{H}\mathbf{B} - \hat{\mathbf{H}}\hat{\mathbf{B}})_{\max} u_k \right| \right. \right. \\ & \left. \left. + \left[|\hat{\mathbf{A}}^{-1}| + \left| \hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ (\mathbf{H} - \hat{\mathbf{H}})_{\max} \right| \right] (\mathbf{w})_{\max} \right\} \right] \circ \mathbf{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) \end{aligned} \quad 4.4$$

Where $\mathbf{e}_{\mathbf{z}_k|k-1}$ is the a priori estimation error and it is defined as $\mathbf{e}_{\mathbf{z}_k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_k|k-1$, $\boldsymbol{\Upsilon} \in \mathbb{R}^{n \times n}$ is a diagonal matrix with $\Upsilon_{ii} \geq 1$, the notation $(\cdot)_{\max}$ represents the upper bound of the element inside the bracket, and $\mathbf{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right)$ is a vector defined as follows:

$$\mathbf{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) = \begin{bmatrix} \mathit{sgn} \left(e_{z_{1k}|k-1} \right) \\ \vdots \\ \mathit{sgn} \left(e_{z_{mk}|k-1} \right) \end{bmatrix} \quad 4.5$$

Where $e_{z_{i_k}|k-1}$, $i = 1 \dots m$ is the i^{th} element in the vector $\mathbf{e}_{\mathbf{z}_k|k-1}$.

From equation (4.4), the terms $\left[|\hat{\mathbf{H}}^+| + |\hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+| \left(\begin{array}{c} (\mathbf{H}\mathbf{A}\mathbf{H}^+ - \hat{\mathbf{H}}\hat{\mathbf{A}}\hat{\mathbf{H}}^+)_{\max} \\ + \hat{\mathbf{H}}\hat{\mathbf{A}}\hat{\mathbf{H}}^+ + \mathbf{I}_{m \times m} \end{array} \right) \right] (\mathbf{v})_{\max}$ and $\left[|\hat{\mathbf{A}}^{-1}| + \left| \hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ (\mathbf{H} - \hat{\mathbf{H}})_{\max} \right| \right] (\mathbf{w})_{\max}$ are related to the measurement and the system noise, respectively. The terms $\left| \hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ (\mathbf{H}\mathbf{A}\mathbf{H}^+ - \hat{\mathbf{H}}\hat{\mathbf{A}}\hat{\mathbf{H}}^+)_{\max} \mathbf{z}_k \right|$ and $\left| \hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ (\mathbf{H}\mathbf{B} - \hat{\mathbf{H}}\hat{\mathbf{B}})_{\max} u_k \right|$ are functions of the uncertainties in the system, input and measurement matrices. The gain of the VSF has some similarities with previously considered SMO methods. The first term of the VSF gain, $\hat{\mathbf{A}}^{-1}\hat{\mathbf{H}}^+ \left[\left(|\hat{\mathbf{H}}\hat{\mathbf{A}}| \boldsymbol{\Upsilon}|\hat{\mathbf{H}}^+| \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \right) \right]$, is the proportional gain, while the remaining terms result in a discontinuous gain which is large enough to overcome the effect of the uncertainties. The advantages of using the gain of equation (4.4) can be summarized as follows:

- The VSF gain is explicitly linked to upper bounds of uncertainties and noise levels.
- The VSF design process is simpler compared to the SMO strategies previously outlined. Note that once the upper bounds of the uncertainties and noise level are established, the gain can be readily calculated.

The disadvantages of the VSF can be summarized as follows:

- The estimate is not optimal
- It is only applicable to linear systems with time invariant models.
- Chattering presents.

The VSF gain is discontinuous and can lead to chattering. Chattering can be eliminated by using a smoothing boundary layer, Ψ . The boundary layer is designed to have a larger width than the upper bound of uncertainties. The VSF's smoothing boundary layer has a different form than the SMO and has a constant width as follows, (Habibi & Burton, 2003):

$$\begin{aligned} \Psi = & \left| \hat{\mathbf{A}}^{-1} \hat{\mathbf{H}}^+ \right| \left| \hat{\mathbf{H}} \hat{\mathbf{A}} \right| \left| \hat{\mathbf{A}}^{-1} \hat{\mathbf{H}}^+ (\mathbf{H} \hat{\mathbf{A}} \mathbf{H}^+ - \hat{\mathbf{H}} \hat{\mathbf{A}} \hat{\mathbf{H}}^+)_{\max} (\mathbf{z})_{\max} \right| \\ & + \left[\left| \hat{\mathbf{H}}^+ \right| + \left| \hat{\mathbf{A}}^{-1} \hat{\mathbf{H}}^+ \right| \left(\left(\mathbf{H} \hat{\mathbf{A}} \mathbf{H}^+ - \hat{\mathbf{H}} \hat{\mathbf{A}} \hat{\mathbf{H}}^+ \right)_{\max} \right. \right. \\ & \left. \left. + \hat{\mathbf{H}} \hat{\mathbf{A}} \hat{\mathbf{H}}^+ + \mathbf{I}_{m \times m} \right) \right] (\mathbf{v})_{\max} \\ & + \left| \hat{\mathbf{A}}^{-1} \hat{\mathbf{H}}^+ (\mathbf{H} \mathbf{B} - \hat{\mathbf{H}} \hat{\mathbf{B}})_{\max} (\mathbf{u})_{\max} \right| \\ & + \left[\left| \hat{\mathbf{A}}^{-1} \right| + \left| \hat{\mathbf{A}}^{-1} \hat{\mathbf{H}}^+ (\mathbf{H} - \hat{\mathbf{H}})_{\max} \right| \right] (\mathbf{w})_{\max} \end{aligned} \quad 4.6$$

Using the boundary layer reduces the accuracy of the estimate. The relationship between the estimated error and the boundary layer width has been described in (Habibi & Burton, 2003). They are directly related and the estimation error increases as the boundary layer width is increased.

4.1.2 The Extended Variable Structure Filter

In 2006, a revised version of the VSF, referred to as the Extended Variable structure Filter (EVSF), was proposed for nonlinear systems with linear measurement matrices. Similarly to the EKF, the EVSF linearizes the system model using the previous a posteriori estimates, and uses a nonlinear system model to obtain the a priori estimates. It then calculates the EVSF gain using the linearized system matrix to obtain the a posteriori values, (Habibi, 2006).

The EVSF assumes that the nonlinear system and measurement equations are described in their discretized form by the following equations:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}, u_{k-1}) + \mathbf{w}_{k-1} \\ \mathbf{z}_k &= \mathbf{H} \mathbf{x}_k + \mathbf{v}_k \end{aligned} \quad 4.7$$

Note that the measurement matrix is a time invariant linear matrix. The EVSF process is similar to the VSF's, and it is summarized as follows, (Habibi, 2006).

1 - Prediction Stage:

The a priori state estimate is obtained by using the nonlinear model of the system as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \hat{\mathbf{f}}(\hat{\mathbf{x}}_{k-1|k-1}, u_{k-1}) \\ \hat{\mathbf{z}}_{k|k-1} &= \hat{\mathbf{H}}\hat{\mathbf{x}}_{k|k-1}\end{aligned}\quad 4.8$$

2 - Correction Stage:

A corrective gain is calculated and used for refining the a priori estimate into its a posteriori form as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{EVSF_k} \\ \hat{\mathbf{z}}_{k|k} &= \hat{\mathbf{H}}\hat{\mathbf{x}}_{k|k}\end{aligned}\quad 4.9$$

The \mathbf{K}_{EVSF_k} is the EVSF's gain and it is defined as follows, (Habibi, 2006):

$$\begin{aligned}\mathbf{K}_{EVSF_k} &= \hat{\Phi}_{k-1}^{-1} \hat{\mathbf{H}}^+ \left[\left\{ \left| (\hat{\mathbf{H}})_{\max} \right| \left| (\hat{\Phi})_{\max} \right| \left[\mathbf{Y} \left| (\hat{\mathbf{H}}^+)_{\max} \right| \mathbf{e}_{z_{k|k-1}} \right] \right. \right. \\ &\quad \left. \left. + \left| (\hat{\Phi}^{-1})_{\max} \right| \left| \left(\mathbf{f}(\mathbf{x}_{k-1}, u_{k-1}) - \hat{\mathbf{f}}(\hat{\mathbf{x}}_{k-1|k-1}, u_{k-1}) \right)_{\max} \right| \right. \right. \\ &\quad \left. \left. + \left[\left| (\hat{\Phi}^{-1})_{\max} \right| \left| (\hat{\mathbf{H}}^+)_{\max} \right| + (\hat{\mathbf{H}}^+)_{\max} \right] (\mathbf{v})_{\max} \right. \right. \\ &\quad \left. \left. + \left| (\hat{\Phi}^{-1})_{\max} \right| (\mathbf{w})_{\max} \right\} \circ \text{sgn} \left(\mathbf{e}_{z_{k|k-1}} \right) \right]\end{aligned}\quad 4.10$$

Where $\hat{\Phi}_k$ is the linearized system matrix around the a posteriori estimate and is obtained by using the Taylor Series approximation, $\hat{\Phi}_k = \frac{\partial \hat{\mathbf{f}}}{\partial \mathbf{x}} \Big|_{\mathbf{x}=\hat{\mathbf{x}}_{k|k}}$. The EVSF has the same advantages and disadvantages as the VSF. A further disadvantage of the EVSF is that it uses a linearized system model at each time step to calculate the corrective gain.

4.1.3 The Smooth Variable Structure Filter

In 2007, a revised version of the VSF, referred to as the Smooth Variable Structure Filter (SVSF), was proposed, (Habibi, 2007). The SVSF is a predictor corrector filter that is based on the SMC principles and can be applied to both linear and nonlinear systems. A requirement of this filter is that the system is differentiable and hence the word "smooth" is used to name this filter. The filter also requires that the system under consideration be observable, (Habibi, 2007).

The SVSF's derivation depends on the rank of the measurement matrix (number of independent measurements compared to the number of states). If the measurement matrix has partial rank (number of independent measurements is fewer than the number of states), the SVSF's gain is calculated by using Luenberger's reduced order technique as discussed later in chapter five, (Habibi, 2007). This chapter considers the case of systems that have full rank measurement matrix (number of independent measurements is equal to the number of states). For these systems, the SVSF process can be summarized as follows:

- The filter's model (linear or nonlinear) is used to obtain the a priori estimate. The filter's structure for linear systems is given in equation (4.2) and for nonlinear systems in equation (4.8). Similarly to the EVSF, the SVSF requires that the measurement matrix be linear.
- The SVSF then refines the a priori estimate into an a posteriori form by applying the SVSF's gain as follows:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{SVSF_k} \quad 4.11$$

The \mathbf{K}_{SVSF_k} for a measurement matrix **with full rank** is defined as follows:

$$\mathbf{K}_{SVSF_k} = \mathbf{H}_k^+ \left(\boldsymbol{\gamma} \left| \mathbf{e}_{\mathbf{z}_{k-1|k-1}} \right| + \left| \mathbf{e}_{\mathbf{z}_{k|k-1}} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_{k|k-1}} \right) \quad 4.12$$

Where $\boldsymbol{\gamma} \in \mathbb{R}^{n \times n}$ is a diagonal matrix with $\gamma_{ii} \leq 1$, and $\mathbf{e}_{\mathbf{z}_{k|k-1}}$ and $\mathbf{e}_{\mathbf{z}_{k-1|k-1}}$ are the a priori and the a posteriori output estimation error vectors and are defined as follows:

$$\begin{aligned} \mathbf{e}_{\mathbf{z}_{k|k-1}} &= \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} \\ \mathbf{e}_{\mathbf{z}_{k-1|k-1}} &= \mathbf{z}_{k-1} - \hat{\mathbf{z}}_{k-1|k-1} \end{aligned} \quad 4.13$$

- The a posteriori estimated output is obtained as follows:

$$\hat{\mathbf{z}}_{k|k} = \hat{\mathbf{H}}_k \hat{\mathbf{x}}_{k|k} \quad 4.14$$

The SVSF has two sets of indicators of performance associated with each state. The **primary indicators** of performance are the estimated errors, and the **secondary indicators** of performance are chattering signals resulting from the application of the discontinuous gains. This gives the SVSF the ability to explicitly point out and extract information on modeling uncertainties.

4.2 The SVSF's stability and gain derivations

The SVSF is a robust recursive predictor-corrector estimation method that can effectively deal with uncertainties associated with initial conditions and modeling errors. It guarantees bounded-input bounded-output stability and the convergence of the estimation process by using the Lyapunov stability condition. The derivation of SVSF's gain and its stability conditions can be found in (Habibi, 2007) and are summarized in the following subsections.

4.2.1 The Lyapunov stability theorem

Let \mathbf{M}_k be a Lyapunov function defined in terms of the a posteriori estimation error, such that:

$$\mathbf{M}_k = \mathbf{e}_{\mathbf{z}_{k|k}}^T \mathbf{e}_{\mathbf{z}_{k|k}} > 0 \quad 4.15$$

The estimation process is stable if:

$$\Delta \mathbf{M}_k < 0 \quad 4.16$$

where $\Delta \mathbf{M}_k$ represents the change in the Lyapunov function and in this case is defined as follows:

$$\Delta \mathbf{M}_k = \mathbf{e}_{\mathbf{z}_{k|k}}^T \mathbf{e}_{\mathbf{z}_{k|k}} - \mathbf{e}_{\mathbf{z}_{k-1|k-1}}^T \mathbf{e}_{\mathbf{z}_{k-1|k-1}} \quad 4.17$$

By substituting equation (4.17) into equation (4.16) and then rearranging, the following equation is obtained:

$$\mathbf{e}_{\mathbf{z}_{k|k}}^T \mathbf{e}_{\mathbf{z}_{k|k}} < \mathbf{e}_{\mathbf{z}_{k-1|k-1}}^T \mathbf{e}_{\mathbf{z}_{k-1|k-1}} \quad 4.18$$

Equation (4.18) is satisfied using the following equation:

$$\left| \mathbf{e}_{\mathbf{z}_{k|k}} \right| < \left| \mathbf{e}_{\mathbf{z}_{k-1|k-1}} \right| \quad 4.19$$

To remove the absolute operator, $| \cdot |$, both sides are expressed in the form of diagonal matrices and then they are multiplied with their transpose as follows:

$$\text{diag}(\mathbf{e}_{\mathbf{z}_{k|k}}) \text{diag}(\mathbf{e}_{\mathbf{z}_{k|k}})^T < \text{diag}(\mathbf{e}_{\mathbf{z}_{k-1|k-1}}) \text{diag}(\mathbf{e}_{\mathbf{z}_{k-1|k-1}})^T \quad 4.20$$

Where $\text{diag}(\mathbf{e}_{\mathbf{z}_{k|k}})$ is the diagonal matrix of $\mathbf{e}_{\mathbf{z}_{k|k}}$.

From equation (4.1), the a posteriori output estimation error is obtained as follows (assuming the output matrix is well known):

$$\mathbf{e}_{\mathbf{z}_{k|k}} = \mathbf{H}_k \mathbf{e}_{\mathbf{x}_{k|k}} + \mathbf{v}_k \quad 4.21$$

By substituting equation (4.21) into equation (4.20), equation (4.22) is obtained as:

$$\begin{pmatrix} \text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \\ + \text{diag}(\mathbf{v}_k) \text{diag}(\mathbf{v}_k) \\ + \text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \text{diag}(\mathbf{v}_k) \\ + \text{diag}(\mathbf{v}_k) \text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \end{pmatrix} < \begin{pmatrix} \text{diag}(\mathbf{H}_{k-1} \mathbf{e}_{x_{k-1|k-1}}) \text{diag}(\mathbf{H}_{k-1} \mathbf{e}_{x_{k-1|k-1}}) \\ + \text{diag}(\mathbf{v}_{k-1}) \text{diag}(\mathbf{v}_{k-1}) \\ + \text{diag}(\mathbf{H}_{k-1} \mathbf{e}_{x_{k-1|k-1}}) \text{diag}(\mathbf{v}_{k-1}) \\ + \text{diag}(\mathbf{v}_{k-1}) \text{diag}(\mathbf{H}_{k-1} \mathbf{e}_{x_{k-1|k-1}}) \end{pmatrix} \quad 4.22$$

If the measurement noise is stationary white, then by taking the expectation of both sides, equation (4.22) is transformed to equation (4.23).

$$E \left[\begin{pmatrix} \text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \\ + \text{diag}(\mathbf{v}_k) \text{diag}(\mathbf{v}_k) \end{pmatrix} \right] < E \left[\begin{pmatrix} \text{diag}(\mathbf{H}_{k-1} \mathbf{e}_{x_{k-1|k-1}}) \text{diag}(\mathbf{H}_{k-1} \mathbf{e}_{x_{k-1|k-1}}) \\ + \text{diag}(\mathbf{v}_{k-1}) \text{diag}(\mathbf{v}_{k-1}) \end{pmatrix} \right] \quad 4.23$$

Where $E \left[\text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \text{diag}(\mathbf{v}_k) \right]$ and $E \left[\text{diag}(\mathbf{v}_k) \text{diag}(\mathbf{H}_k \mathbf{e}_{x_{k|k}}) \right]$ vanish due to the white noise assumption. For a diagonal, positive and time-invariant measurement matrix, equation (4.23) is reduced to equation (4.24).

$$E \left[\text{diag}(\mathbf{e}_{x_{k|k}}) \text{diag}(\mathbf{e}_{x_{k|k}}) \right] < E \left[\text{diag}(\mathbf{e}_{x_{k-1|k-1}}) \text{diag}(\mathbf{e}_{x_{k-1|k-1}}) \right] \quad 4.24$$

Note that the assumptions pertaining to the measurement matrix are realistic since most applications use linear sensors as feedback in their operations. Moreover, these sensors are well calibrated and their structures are well known, (Habibi, 2007). Equation (4.24) is equivalent to the following:

$$E \left(\left| \mathbf{e}_{x_{k|k}} \right| \right) < E \left(\left| \mathbf{e}_{x_{k-1|k-1}} \right| \right) \quad 4.25$$

From equation (4.25), the expectation of the a posteriori estimation error is reduced in time (it converges towards the origin) which means that the filter is stable.

4.2.2 The derivation of the SVSF's gain

The SVSF's gain is derived to guarantee the stability condition of equation (4.19). Moreover, the gain must be larger than the uncertain dynamics of the estimation process, yet it should be bounded for bounded-input bounded-output stability.

Let $\boldsymbol{\gamma}$ be a diagonal positive matrix with dimensions $\boldsymbol{\gamma} \in \mathbb{R}^{n \times n}$ and with elements less than unity, i.e. $0 < \gamma_{ii} < 1$, then:

$$\boldsymbol{\gamma} \left| \mathbf{e}_{z_{k-1|k-1}} \right| < \left| \mathbf{e}_{z_{k-1|k-1}} \right| \quad 4.26$$

Adding the term $\left| \mathbf{e}_{z_{k|k-1}} \right|$ to both sides leads to the following:

$$\boldsymbol{\gamma} \left| \mathbf{e}_{z_{k-1|k-1}} \right| + \left| \mathbf{e}_{z_{k|k-1}} \right| < \left| \mathbf{e}_{z_{k-1|k-1}} \right| + \left| \mathbf{e}_{z_{k|k-1}} \right| \quad 4.27$$

The absolute value of the SVSF gain multiplied by the measurement matrix is set to be equal to the left hand side of equation (4.27) as follows:

$$|\widehat{\mathbf{H}}_k \mathbf{K}_{SVSF_k}| = \gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| + \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \quad 4.28$$

The sign of the gain is made equal to the sign of the a priori estimation error, $\mathbf{e}_{\mathbf{z}_k|k-1}$. This leads to:

$$\mathbf{K}_{SVSF_k} = \widehat{\mathbf{H}}_k^+ \left(\gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| + \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) \quad 4.29$$

Note that the proposed gain satisfies the conditions of being larger than the a priori estimation error. By applying the gain to the a priori estimate, and by substituting equations (4.29) and (4.12) in equation (4.11), the a posteriori estimated measurement is obtained as follows:

$$\widehat{\mathbf{z}}_k = \widehat{\mathbf{z}}_k|k-1 + \left(\gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| + \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) \quad 4.30$$

Subtracting equation (4.30) from the measurement \mathbf{z}_k leads to the following:

$$\mathbf{e}_{\mathbf{z}_k|k} = \mathbf{e}_{\mathbf{z}_k|k-1} - \left(\gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| + \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) \quad 4.31$$

Equation (4.31) can be rewritten in a simpler form by substituting $\left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right)$ by $\mathbf{e}_{\mathbf{z}_k|k-1}$ as follows:

$$\begin{aligned} \mathbf{e}_{\mathbf{z}_k|k} &= \mathbf{e}_{\mathbf{z}_k|k-1} - \gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) - \mathbf{e}_{\mathbf{z}_k|k-1} \\ &= -\gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) \end{aligned} \quad 4.32$$

By taking the absolute of both sides of equation (4.32), equation (4.33) is obtained as follows:

$$\left\{ \left| \mathbf{e}_{\mathbf{z}_k|k} \right| = \gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| \right\} < \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| \quad 4.33$$

Equation (4.33) proves that the error decays in time, which means that equations (4.19) and (4.25) are satisfied and the filter is stable.

4.3 Application of the SVSF to systems with a measurement matrix of full rank

In this section, the application of the SVSF to linear systems that have a measurement matrix with full rank is considered in further detail. The novel contributions of research reported in this section are the following:

- It is noted that there are two existence subspaces. One as previously reported for the a priori estimate and one newly determined for the a posteriori estimate. This research describes these subspaces and develops mathematical formulas for them. The mathematical formula of the a priori existence subspace is different than the previously reported one. This work is listed in subsection (4.3.1).
- The characteristics of the smoothing boundary layer are investigated, and a time-varying smoothing boundary layer is proposed in section (4.3.2) to improve the SVSF's performance.
- Two explicit equations are obtained for chattering in the a priori and the a posteriori estimates using the SVSF sense. These equations link the chattering amplitude to the source and level of model uncertainties.
- The chattering equations can be used for calculating and obtaining modeling errors and uncertainties, thereby establishing a method for model refinement. This method is developed in subsection (4.3.3).

4.3.1 The SVSF's existence subspaces

The SVSF forces the estimate towards the true state trajectory and then retains it within a subspace, referred to as the *existence subspace*. This occurs both for the a priori and the a posteriori estimates as shown in Fig 4.1. The widths of the existence subspaces are functions of the uncertainties, errors in the initial conditions, and/or modeling errors. This means that the widths are unknown and time varying as shown in Fig 4.2 for the a priori existence subspace. The a posteriori existence subspace defines a region that surrounds and encloses the true trajectory in which the estimate may exist in its a posteriori form. Its width is equal to the difference between the width of the a priori existence subspace and the amplitude of the corrective gain, as shown in Fig 4.1. In the following subsections, these two subspaces are considered in detail.

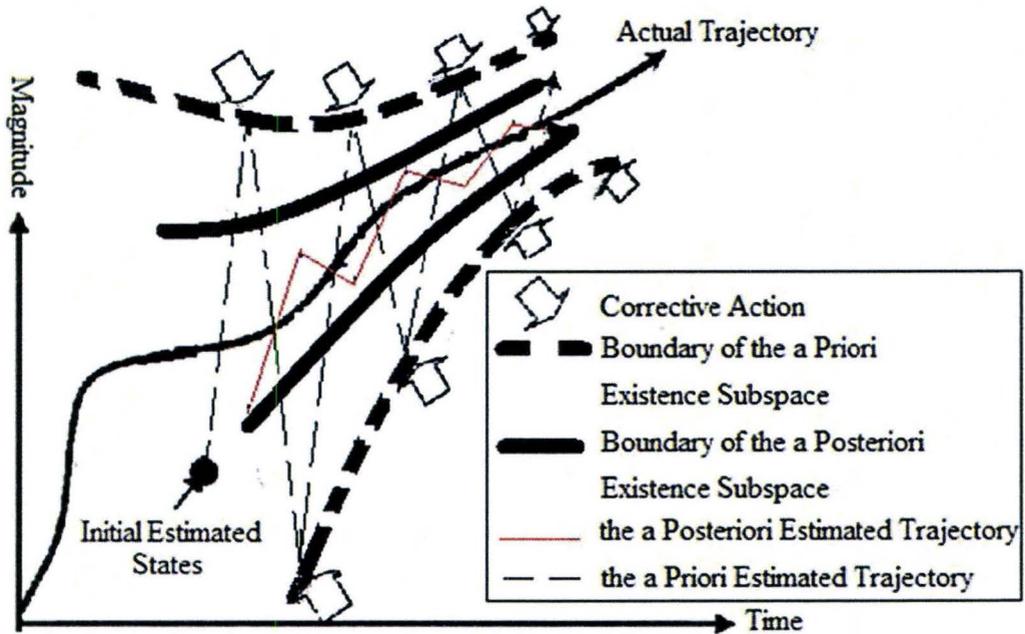


Fig 4.1: The Definition of the Existence Subspace, (Habibi I, 2008).

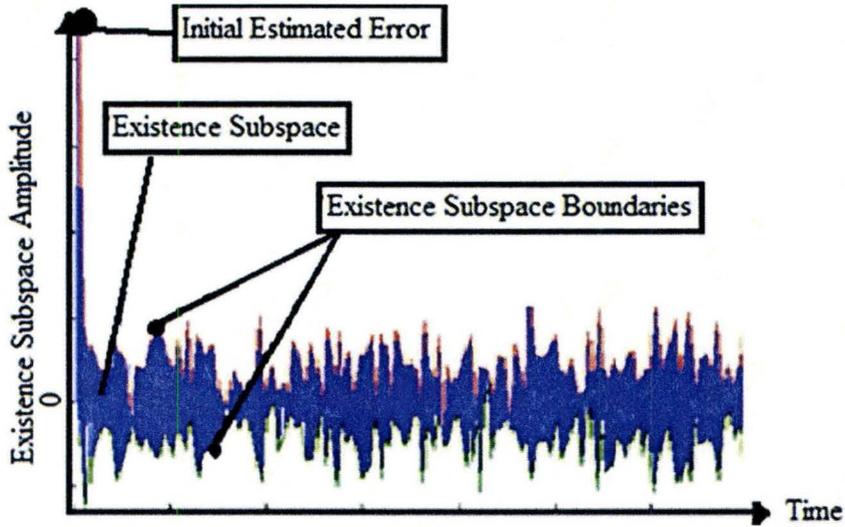


Fig 4.2: The Existence Subspace versus time.

4.3.1.1. The a priori existence subspace

In this section, an equation is derived to quantify the width of the a priori existence subspace.

The a priori estimation error is defined as follows:

$$\mathbf{e}_{z_k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k - \hat{\mathbf{H}}_k \hat{\mathbf{x}}_{k|k-1} \quad 4.34$$

Substituting equation (4.2) into equation (4.34) results in the following:

$$\begin{aligned} \mathbf{e}_{z_k|k-1} &= \mathbf{H}_k (\mathbf{A}_{k-1} \mathbf{x}_{k-1} + \mathbf{B}_{k-1} u_{k-1} + \mathbf{w}_{k-1}) \\ &\quad - \hat{\mathbf{H}}_k (\hat{\mathbf{A}}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \hat{\mathbf{B}}_{k-1} u_{k-1}) + \mathbf{v}_k \end{aligned} \quad 4.35$$

Substituting equations (4.1) and (4.2) into equation (4.35) gives the following:

$$\begin{aligned} \mathbf{e}_{z_k|k-1} &= \mathbf{H}_k (\mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{z}_{k-1} - \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} + \mathbf{B}_{k-1} u_{k-1} + \mathbf{w}_{k-1}) \\ &\quad - \hat{\mathbf{H}}_k (\hat{\mathbf{A}}_{k-1} \hat{\mathbf{H}}_k^{-1} \hat{\mathbf{z}}_{k-1|k-1} + \hat{\mathbf{B}}_{k-1} u_{k-1}) + \mathbf{v}_k \end{aligned} \quad 4.36$$

From equations (4.30) and (4.34), the estimated measurement is obtained as follows:

$$\hat{\mathbf{z}}_{k-1|k-1} = \mathbf{z}_{k-1} + \gamma \left| \mathbf{e}_{z_{k-2}|k-2} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \quad 4.37$$

Substituting equation (4.37) into equation (4.36) yields the following:

$$\begin{aligned} \mathbf{e}_{z_k|k-1} &= \mathbf{H}_k (\mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{z}_{k-1} + \mathbf{B}_{k-1} u_{k-1} + \mathbf{w}_{k-1} - \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1}) \\ &\quad - \hat{\mathbf{H}}_k (\hat{\mathbf{A}}_{k-1} \hat{\mathbf{H}}_k^{-1} \mathbf{z}_{k-1} + \hat{\mathbf{B}}_{k-1} u_{k-1}) + \mathbf{v}_k \\ &\quad - \hat{\mathbf{H}}_k \hat{\mathbf{A}}_{k-1} \hat{\mathbf{H}}_k^{-1} \gamma \left| \mathbf{e}_{z_{k-2}|k-2} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \end{aligned} \quad 4.38$$

Rearranging equation (4.38) and substituting equation (4.33) results in the following:

$$\begin{aligned} \mathbf{e}_{z_k|k-1} &= \left(\mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} - \hat{\mathbf{H}}_k \hat{\mathbf{A}}_{k-1} \hat{\mathbf{H}}_k^{-1} \right) \mathbf{z}_{k-1} + \left(\mathbf{H}_k \mathbf{B}_{k-1} - \hat{\mathbf{H}}_k \hat{\mathbf{B}}_{k-1} \right) u_{k-1} \\ &\quad + \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_k - \mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} \\ &\quad - \hat{\mathbf{H}}_k \hat{\mathbf{A}}_{k-1} \hat{\mathbf{H}}_k^{-1} \gamma \left| \mathbf{e}_{z_{k-2}|k-2} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \end{aligned} \quad 4.39$$

Further to equation (4.32) the a posteriori estimation error is related to the initial condition as follows:

$$\begin{aligned} \mathbf{e}_{z_{k-1}|k-1} &= -\gamma \left| \mathbf{e}_{z_{k-2}|k-2} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \\ &= -\gamma \left| -\gamma \left| \mathbf{e}_{z_{k-3}|k-3} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-2}|k-3} \right) \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \\ &= -\gamma^2 \left| \mathbf{e}_{z_{k-3}|k-3} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \\ &= -\gamma^2 \left| -\gamma \left| \mathbf{e}_{z_{k-4}|k-4} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-3}|k-4} \right) \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \\ &= -\gamma^3 \left| \mathbf{e}_{z_{k-4}|k-4} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \\ &\quad \vdots \\ \mathbf{e}_{z_{k-1}|k-1} &= -\gamma^{k-1} \left| \mathbf{e}_{z_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k-1}|k-2} \right) \end{aligned} \quad 4.40$$

Substituting equation (4.40) into equation (4.39) yields the following:

$$\begin{aligned} \mathbf{e}_{\mathbf{z}_k|k-1} = & \left(\mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} - \widehat{\mathbf{H}}_k \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \right) \mathbf{z}_{k-1} + \left(\mathbf{H}_k \mathbf{B}_{k-1} - \widehat{\mathbf{H}}_k \widehat{\mathbf{B}}_{k-1} \right) u_{k-1} \\ & + \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_k - \mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} \\ & - \widehat{\mathbf{H}}_k \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{\mathbf{z}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_{k-1}|k-2} \right) \end{aligned} \quad 4.41$$

The existence subspace is then defined as follows:

$$\begin{aligned} \mathbf{e}_{\mathbf{x}_k|k-1} = & \left(\mathbf{A}_{k-1} \mathbf{H}_k^{-1} - \mathbf{H}_k^{-1} \widehat{\mathbf{H}}_k \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \right) \mathbf{z}_{k-1} \\ & + \left(\mathbf{B}_{k-1} - \mathbf{H}_k^{-1} \widehat{\mathbf{H}}_k \widehat{\mathbf{B}}_{k-1} \right) u_{k-1} + \mathbf{w}_{k-1} - \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} \\ & - \mathbf{H}_k^{-1} \widehat{\mathbf{H}}_k \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{\mathbf{z}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_{k-1}|k-2} \right) \end{aligned} \quad 4.42$$

If $\mathbf{H}_k = \widehat{\mathbf{H}}_k$, then equation (4.42) is reduced to the following:

$$\begin{aligned} \mathbf{e}_{\mathbf{x}_k|k-1} = & \left(\mathbf{A}_{k-1} \mathbf{H}_k^{-1} - \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \right) \mathbf{z}_{k-1} + \left(\mathbf{B}_{k-1} - \widehat{\mathbf{B}}_{k-1} \right) u_{k-1} + \mathbf{w}_{k-1} \\ & - \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} - \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{\mathbf{z}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_{k-1}|k-2} \right) \end{aligned} \quad 4.43$$

Note that equation (4.43) is defined in terms of the state vector and takes into account the switching action. By the inclusion of this term, it differs from the existence subspace reported in (Habibi, 2007) that is defined in terms of the measurement vector. From equation (4.42), the terms $\left(\left(\mathbf{A}_{k-1} \mathbf{H}_k^{-1} - \mathbf{H}_k^{-1} \widehat{\mathbf{H}}_k \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \right) \mathbf{z}_{k-1} \right)$ and $\left(\left(\mathbf{B}_{k-1} - \mathbf{H}_k^{-1} \widehat{\mathbf{H}}_k \widehat{\mathbf{B}}_{k-1} \right) u_{k-1} \right)$ capture the influence of the modeling errors. The term $\left(\mathbf{w}_{k-1} - \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} \right)$ quantifies the impact of the system and measurement noise. The last term in equation (4.42), $\left(\mathbf{H}_k^{-1} \widehat{\mathbf{H}}_k \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{\mathbf{z}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_{k-1}|k-2} \right) \right)$, describes the effects of the uncertainty in initial conditions and its impact on the a priori existence subspace. According to the latter term, the effect of the error in initial conditions decays in time at a rate of $\mathbf{H}_k^{-1} \widehat{\mathbf{H}}_k \widehat{\mathbf{A}}_{k-1} \widehat{\mathbf{H}}_k^{-1} \boldsymbol{\gamma}^{k-1}$, and becomes negligible as $k \rightarrow \infty$. Then, the width of the a priori existence subspace becomes a function of the uncertainties, noise, and modeling errors.

The a priori existence subspace represents the error in the a priori estimate. In other words, it describes the chattering of the a priori estimate around the true trajectory. In this research, the magnitude of the resultant chattering is referred to as the *a priori chattering*. Because of the predictor-corrector nature of the SVSF and its gain, the a priori chattering is different from the chattering observed in other SMOs. The differences are summarized in table 4.1.

The SMO's chattering	The SVSF's a priori chattering
Chattering is caused by applying a large corrective action on the estimate.	Chattering is caused by the magnitude of uncertainties and modeling errors.
Chattering amplitude is equal to the difference between the amplitude of the corrective action and the filter's uncertainties and modeling errors	Chattering amplitude is linked to the width of the a priori existence subspace
Chattering affects the estimate as it adds an artificial noise to it.	The SVSF's gain removes the a priori chattering as discussed later in section (4.3.1.2). Therefore, it does not affect the estimate.
Some uncertainties and modeling errors may be reconstructed from chattering by filtering the discontinuous corrective action using a low pass filter (referred to as the equivalent injection signal).	Chattering contains the uncertainties and modeling errors, and can be extracted for model refinement as will be discussed in section (4.3.3).

Table 4.1: The differences between chattering of SMO and the SVSF a priori chattering.

4.3.1.2. The a posteriori existence subspace

As discussed earlier, the corrective action must have a magnitude that is larger than the uncertain dynamics of the estimation process in order to guarantee the filter stability. The difference between the width of the a priori existence subspace and the SVSF's gain is equal to the width of the a posteriori existence subspace, and has a different sign from the a priori existence subspace, as shown in Fig 4.1. The a posteriori existence subspace is derived in this section.

Further to equations (4.40) and (4.32), the a posteriori estimation error at time k is defined as follows:

$$\begin{aligned}
 \mathbf{e}_{\mathbf{z}_k|k} &= \mathbf{e}_{\mathbf{z}_k|k-1} - \left\{ \mathbf{H}_k \mathbf{K}_{SVSF_k} = \left(\gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| + \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) \right\} \\
 &= -\gamma \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) = -\gamma^k \left| \mathbf{e}_{\mathbf{z}_0|0} \right| \circ \text{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right)
 \end{aligned} \tag{4.44}$$

Note that equation (4.44) shows that the a priori chattering does not affect the a posteriori estimate as it is removed by the corrective gain. The a posteriori existence subspace for the state vector is defined as follows:

$$\mathbf{e}_{x_{k|k}} = -\boldsymbol{\gamma}^k \left| \mathbf{e}_{z_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{k|k-1}} \right) - \mathbf{v}_k \quad 4.45$$

Equation (4.44) shows that the a posteriori estimate chatters around the measurement. This chattering decays with time until the estimated output converges to the measured output. Equation (4.44) is further proof of the filter's stability. Note that the estimated output converges to the measurement which contains the true trajectory blurred with measurement noise. Therefore, the estimation is sensitive to the measurement noise as shown in equation (4.45). In this research, the resultant chattering is referred to as the *a posteriori chattering*. The differences and similarities between this chattering and that of the SMOs' are summarized in table 4.2.

The SMO's chattering	The SVSF's a posteriori chattering
Chattering is caused by applying the corrective action on the previous estimate.	Chattering is caused by applying the corrective action on the a priori estimate.
Chattering's amplitude is equivalent to the difference between the corrective action and the filter's uncertainties and modeling errors.	Chattering amplitude is linked to the difference between the corrective action and the a priori existence subspace.
Chattering affects the estimate as it adds an artificial noise to it.	Chattering affects the estimate as it adds an artificial noise to it. However, it decays with time.
Chattering is sensitive to system and measurement noise.	Chattering is sensitive to measurement noise.

Table 4.2: The differences and similarity between chattering of SMO and the SVSF's a posteriori chattering.

The a posteriori chattering decays with time and becomes a function of measurement noise after a few time steps depending on the coefficient $\boldsymbol{\gamma}$. Henceforth, the a priori chattering will be considered for information extraction, and the a posteriori chattering will be used for the design of the smoothing boundary layer width.

4.3.2 The smoothing boundary layer

As discussed earlier, the estimate has two levels of chattering; the a priori and the a posteriori chattering. Although the latter decays with time, it causes the a posteriori estimate to become more sensitive to measurement noise. In order to eliminate the a priori

and the a posteriori chattering, and to reduce the sensitivity to noise, the sign function in equation (4.12) is replaced by a smoothing function with a known boundary layer referred to as the *smoothing boundary layer*. Inside the smoothing boundary layer, the corrective action is interpolated based on the ratio between the amplitude of the output's a priori estimation error and the smoothing boundary layer's width. Outside the smoothing boundary layer, the discontinuous corrective action with its full amplitude is applied. The SVSF assigns and requires one smoothing boundary layer per estimate. The following equation defines the SVSF's gain with the smoothing boundary layer:

$$\mathbf{K}_{SVSF_k} = \hat{\mathbf{H}}_k^{-1} \left(\boldsymbol{\gamma} \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| + \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \right) \circ \mathbf{sat} \left(\mathbf{e}_{\mathbf{z}_k|k-1}, \boldsymbol{\Psi}_k \right) \quad 4.46$$

Where $\boldsymbol{\Psi}_k$ is a vector consisting of n –smoothing boundary layer widths at time k (one for each state), and \mathbf{sat} is a vector of the saturation functions and it is defined as follows:

$$\mathbf{sat} \left(\mathbf{e}_{\mathbf{z}_k|k-1}, \boldsymbol{\Psi}_k \right) = \left[\text{sat} \left(e_{z_{1k}|k-1}, \psi_{1k} \right) \quad \dots \quad \text{sat} \left(e_{z_{nk}|k-1}, \psi_{nk} \right) \right]^T \quad 4.47$$

Where $\text{sat} \left(e_{z_{ik}|k-1}, \psi_{ik} \right)$, $i = 1, \dots, n$ is the saturation function and is defined as follows:

$$\text{sat} \left(e_{z_{ik}|k-1}, \psi_{ik} \right) = \begin{cases} e_{z_{ik}|k-1} / \psi_{ik} & e_{z_{ik}|k-1} \leq \psi_{ik} \\ \text{sgn} \left(e_{z_{ik}|k-1} \right) & e_{z_{ik}|k-1} > \psi_{ik} \end{cases} \quad 4.48$$

The smoothing boundary layer must be larger than the uncertain dynamics associated with each estimate to remove the a priori and the a posteriori chattering, and smooth the a posteriori estimate. In order to understand the effectiveness of the smoothing boundary layer, the widths of the existence subspaces are mathematically examined for different smoothing boundary layer widths as follows.

The gain in equation (4.46) is rewritten as follows:

$$\mathbf{K}_{SVSF_k} = \hat{\mathbf{H}}_k^{-1} \mathbf{S}_{at} \left(\mathbf{e}_{\mathbf{z}_k|k-1}, \boldsymbol{\Psi}_k \right) \left(\left(\boldsymbol{\gamma} \left| \mathbf{e}_{\mathbf{z}_{k-1}|k-1} \right| + \left| \mathbf{e}_{\mathbf{z}_k|k-1} \right| \right) \circ \mathbf{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right) \right) \quad 4.49$$

Where $\mathbf{S}_{at} \left(\mathbf{e}_{\mathbf{z}_k|k-1}, \boldsymbol{\Psi}_k \right)$ is a positive diagonal matrix and it is defined as follows:

$$\mathbf{S}_{at} \left(\mathbf{e}_{\mathbf{z}_k|k-1}, \boldsymbol{\Psi}_k \right) = \begin{bmatrix} \left| \text{sat} \left(e_{z_{1k}|k-1}, \psi_{1k} \right) \right| & & 0 \\ & \ddots & \\ 0 & & \left| \text{sat} \left(e_{z_{nk}|k-1}, \psi_{nk} \right) \right| \end{bmatrix} \quad 4.50$$

Note that $\mathbf{sat} \left(\mathbf{e}_{\mathbf{z}_k|k-1}, \boldsymbol{\Psi}_k \right) = \mathbf{S}_{at} \left(\mathbf{e}_{\mathbf{z}_k|k-1}, \boldsymbol{\Psi}_k \right) \mathbf{sgn} \left(\mathbf{e}_{\mathbf{z}_k|k-1} \right)$. Further to the a posteriori estimation error of equation (4.31), and by using the gain \mathbf{K}_{SVSF_k} defined in equation (4.46), the resultant a posteriori estimated error is obtained as follows:

$$\begin{aligned}
\mathbf{e}_{z_k|k} &= \mathbf{e}_{z_k|k-1} - \mathbf{S}_{at}(\mathbf{e}_{z_k|k-1}, \Psi_k) \left((\gamma |\mathbf{e}_{z_{k-1}|k-1}| + |\mathbf{e}_{z_k|k-1}|) \circ \text{sgn}(\mathbf{e}_{z_k|k-1}) \right) \\
&= \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{z_k|k-1}, \Psi_k) \right) \mathbf{e}_{z_k|k-1} \\
&\quad - \gamma \mathbf{S}_{at}(\mathbf{e}_{z_k|k-1}, \Psi_k) \left(|\mathbf{e}_{z_{k-1}|k-1}| \circ \text{sgn}(\mathbf{e}_{z_k|k-1}) \right)
\end{aligned} \tag{4.51}$$

Note that equation (4.44) is not valid anymore because of the use of a smoothing boundary layer and the \mathbf{S}_{at} function.

The mathematical formula of equation (4.51) specifies the smoothing boundary layer properties. If the width of the smoothing boundary layer is smaller than the amplitude of the a priori estimation error, then the matrix \mathbf{S}_{at} becomes as follows:

$$\mathbf{S}_{at}(\mathbf{e}_{z_k|k-1}, \Psi_k) = \mathbf{I}_{n \times n} \tag{4.52}$$

and equation (4.51) is reduced to equation (4.44). The existence subspaces and the chattering amplitudes remain unaffected. Conversely, if the smoothing boundary layer has a width that is more than ten times that of the a priori output estimation error's amplitude, then the filter's a priori and a posteriori estimates will have similar values as the corrective action becomes insignificant and negligible as follows:

$$\lim_{\mathbf{S}_{at}(\mathbf{e}_{z_k|k-1}, \Psi_k) \rightarrow 0} \mathbf{K}_{SVSF_k} = 0 \tag{4.53}$$

Then, the a priori existence subspace will have the same width as that of the a posteriori existence subspace as follows:

$$\lim_{\mathbf{S}_{at}(\mathbf{e}_{z_k|k-1}, \Psi_k) \rightarrow 0} \mathbf{e}_{x_k|k} = \lim_{\mathbf{S}_{at}(\mathbf{e}_{z_k|k-1}, \Psi_k) \rightarrow 0} \mathbf{e}_{x_k|k-1} \tag{4.54}$$

Both subspaces would expand in time if the uncertainties and/or modeling errors are increased. Note that the a posteriori existence subspace affects the a priori existence subspace. Without the boundary layer, the a posteriori estimate would contain noise and an error term that decays in time (refer to equation (4.45)). By propagating this estimate, the a priori existence subspace is obtained, which is a function of uncertainties and modeling errors (refer to equation (4.42)). The corrective action overcomes uncertainties and will retrain the stability of the filter. If the error in the a posteriori error increases by applying a large smoothing function, then the error in the a priori also increases. If the corrective action is not large enough to overcome uncertainties, then the error increases with time and the filter becomes unstable. This can be proven by modifying equation (4.43) by using equations (4.34) and (4.51) to include the boundary layer effects as follows:

$$\mathbf{e}_{\mathbf{x}_{k|k}} = \begin{bmatrix} (\mathbf{A}_{k-1}\mathbf{H}_k^{-1} - \mathbf{H}_k^{-1}\hat{\mathbf{H}}_k\hat{\mathbf{A}}_{k-1}\hat{\mathbf{H}}_k^{-1})\mathbf{z}_{k-1} \\ + (\mathbf{B}_{k-1} - \mathbf{H}_k^{-1}\hat{\mathbf{H}}_k\hat{\mathbf{B}}_{k-1})\mathbf{u}_{k-1} \\ + \mathbf{w}_{k-1} - \mathbf{A}_{k-1}\mathbf{H}_k^{-1}\mathbf{v}_{k-1} + \mathbf{H}_k^{-1}\hat{\mathbf{H}}_k\hat{\mathbf{A}}_{k-1}\hat{\mathbf{H}}_k^{-1}\mathbf{e}_{\mathbf{z}_{k-1|k-1}} - \mathbf{K}_{SVSF_k} \end{bmatrix} \quad 4.55$$

Where \mathbf{K}_{SVSF_k} approaches zero as the function \mathbf{S}_{at} approaches zero. Note that the a posteriori existence subspace accumulates with time; thereby it may become unstable.

A larger width of the boundary layer causes a slower convergence rate and degrades filter performance as shown in Fig 4.3. These occur if the corrective action is not large enough to force the estimated trajectory to cross the actual trajectory as follows.

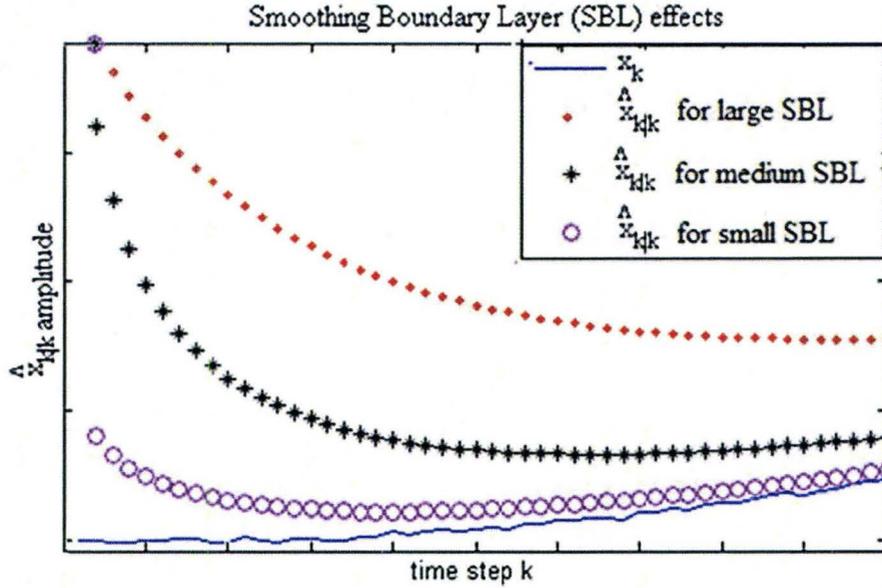


Fig 4.3: The Smoothing Boundary Layer Effects on the Convergence Rate.

If two SVSFs are applied to a system, each with a boundary layer; i.e. Ψ_1 and Ψ_2 , and assuming $\Psi_1 > \Psi_2$ and $\gamma_1 = \gamma_2 = \mathbf{0}_{n \times n}$, then the a posteriori errors for these two filters are obtained as follows:

$$\mathbf{e}_{\mathbf{z},1k|k} = \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{\mathbf{z},1k|k-1}, \Psi_1) \right) \mathbf{e}_{\mathbf{z},1k|k-1} \quad 4.56$$

$$\mathbf{e}_{\mathbf{z},2k|k} = \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{\mathbf{z},2k|k-1}, \Psi_2) \right) \mathbf{e}_{\mathbf{z},2k|k-1} \quad 4.57$$

Where $\mathbf{e}_{\mathbf{z},i_k|k}$ and $\mathbf{e}_{\mathbf{z},i_k|k-1}$, $i = 1,2$ are the a posteriori and the a priori errors pertaining to the estimation output at time k , respectively, and the subscript i signifies the choice of the filter.

For comparison purposes, the filters are assumed to have the same initial condition at time k_0 which indicates that they have the same a priori estimate at time $k_0 + 1$. Further to this assumption, the ratio between equations (4.56) and (4.57) is obtained as follows:

$$\begin{aligned} & \text{diag} \left(\left| \mathbf{e}_{z,2k_0+1|k_0+1} \right| \right) \text{diag} \left(\left| \mathbf{e}_{z,1k_0+1|k_0+1} \right| \right)^{-1} \\ &= \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_2 \right) \right) \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_1 \right) \right)^{-1} \end{aligned} \quad 4.58$$

If $\Psi_1 > \Psi_2$, then $\Psi_1^{-1} < \Psi_2^{-1}$. Multiplying both sides with the absolute value of the output estimations error, $\left| \mathbf{e}_{z,1k_0+1|k_0} \right|$, leads to the following:

$$\begin{aligned} & \left\{ \begin{array}{l} \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_1 \right) \\ = \text{diag} \left(\left| \mathbf{e}_{z,1k_0+1|k_0} \right| \circ \Psi_1^{-1} \right) \end{array} \right\} < \left\{ \begin{array}{l} \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_2 \right) \\ = \text{diag} \left(\left| \mathbf{e}_{z,1k_0+1|k_0} \right| \circ \Psi_2^{-1} \right) \end{array} \right\} \text{ and} \\ & \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_1 \right) \right) > \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_2 \right) \right) \end{aligned} \quad 4.59$$

Rearranging equation (4.59) gives the following:

$$\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_2 \right) \right) \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_1 \right) \right)^{-1} < \mathbf{I}_{n \times n} \quad 4.60$$

Further to equations (4.60) and (4.58), the following is obtained:

$$\left| \mathbf{e}_{z,2k|k} \right| < \left| \mathbf{e}_{z,1k|k} \right| \quad 4.61$$

Equation (4.61) proves that the error increases by increasing the smoothing boundary layer resulting in a slower convergence rate.

The width of the smoothing boundary layer affects the amplitude of the a posteriori chattering. If the width of the smoothing boundary layer increases then the chattering amplitude decreases, and vice versa. By increasing the width of the smoothing boundary layer, the amplitude of the discontinuous gain represented by the term $\mathbf{S}_{at} \left(\mathbf{e}_{z,k|k-1}, \Psi \right) \gamma \left| \mathbf{e}_{z,k-1|k-1} \right| \circ \text{sgn} \left(\mathbf{e}_{z,k|k-1} \right)$ decreases around the sliding surface; hence the a posteriori chattering is reduced. If the same filters (1 and 2) are used with the following assumptions:

- $\Psi_1 > \Psi_2$
- $\gamma_1 = \gamma_2$
- $\mathbf{e}_{z,1k_0|k_0} = \mathbf{e}_{z,2k_0|k_0}$

then the discontinuous gain of filter 1 has a smaller energy than the corresponding discontinuous gain of filter 2 because $S_{at}(\mathbf{e}_{z,1k_0+1|k_0}, \Psi_1) < S_{at}(\mathbf{e}_{z,2k_0+1|k_0}, \Psi_2)$. Therefore, the chattering amplitude decreases as the boundary layer width increases although the estimation error increases.

If the width of the smoothing boundary layer is chosen to be larger than the width of the a priori existence subspace and the difference between them is made to be small, then chattering is removed and the error in the output estimation is limited. This implies that the smoothing boundary layer must be time varying and dependent on the width of the a priori existence subspace. In this sense, the width of the smoothing boundary layer can be designed to obtain a better performance in terms of the a posteriori estimation error as will be shown in section (4.3.2.1).

When a smoothing boundary layer is used, the a priori chattering becomes equal to the difference between the width of the smoothing boundary layer and the amplitude of the output's a priori estimation error. Therefore, the width of the smoothing boundary layer determines the presence and the level of the a priori chattering. If the smoothing boundary layer is overestimated then chattering is removed. However, if due to changes in the system, additional uncertainties are added such that the amplitude of the output's a priori estimation error grows larger than the width of the smoothing boundary layer, then chattering will be observed, (Habibi, 2007). For example, the a priori chattering signal has been tracked for a second order system that is made to have parametric changes at time steps $t_1 = 4000$ and $t_2 = 7000$. These changes last for 1000 and 2500 time steps, respectively. The smoothing boundary layer was designed to enclose the existence subspace for the system before the parametric changes ($t < 4000$ time steps). Fig 4.4 shows the a priori chattering when uncertainties are injected into the model at time $t_1 = 4000$ and $t_2 = 7000$ time steps. Moreover, the figure shows the lasting period of each uncertainty injection. The SVSF is very sensitive to added uncertainties and exhibits chattering that can be used for detecting the inception of a change in the system. This capability is very useful for certain applications such as requiring early fault detection.

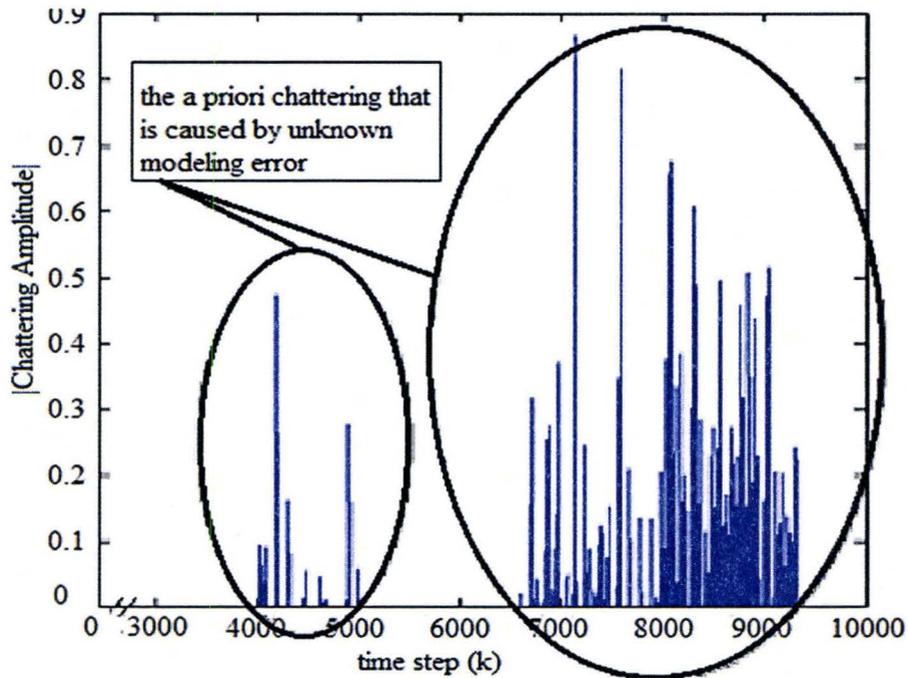


Fig 4.4: The a Priori Chattering as an indicator of parametric changes.

4.3.2.1 The time varying smoothing boundary layer

The width of the boundary layer has a large impact on the filter's performance. If it is too small it will result in chattering. If it is too large, it reduces estimation accuracy. The smoothing boundary layer can be designed to have time varying width while being linked to the a priori output estimation error, (Gadsden & Habibi, 2011). However, using the smoothing boundary layer implies an unknown increment in the width of the a priori existence subspace as shown in equation (4.55). In order to improve the filter's performance without affecting the existence subspace, a novel strategy is presented here for a time varying smoothing boundary layer by using two SVSFs. These are executed in parallel as shown in Fig 4.5. The first filter (the primary filter) uses a smoothing boundary layer with zero width, which means that the a priori existence subspace is described by equation (4.42). Further to section (4.3.1.2) and equation (4.44), the a posteriori chattering decays and becomes negligible as $k \rightarrow \infty$. As such, chattering is mainly present in the a priori estimate. This provides an exact indication of the existence subspaces' widths for the second filter (the secondary filter), and maintains the estimation stability. The secondary filter is then used to obtain another estimation vector with information derived from filter 1 as follows:

- The primary filter's a posteriori estimate, which was obtained at its previous time step, is propagated through to the secondary filter's model to obtain its estimate in the a priori form.
- The secondary filter's a priori estimate is then refined using a corrective action with a time-varying smoothing boundary layer that is linked to the a priori existence subspace derived from filter 1.

Note that the a posteriori estimate from the secondary filter is not used further in calculation of the a priori estimates for the next time interval, which ensures numerical stability of the estimation process. Moreover, the secondary filter and the primary filter share the same a priori existence subspace.

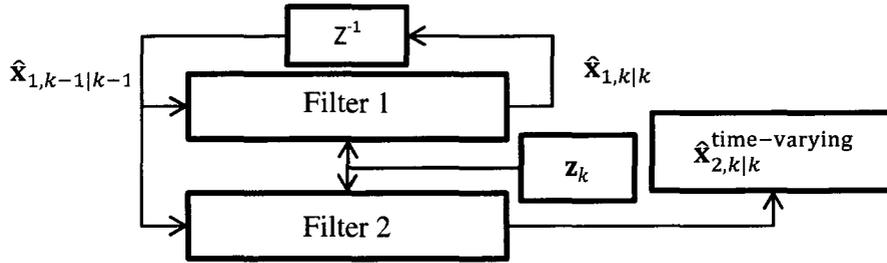


Fig 4.5: The time varying SVSF Structure.

The time varying smoothing boundary layer, Ψ_{tv_k} , in this thesis is designed to minimize the expected value of the quadratic cost function J_k with probabilities conditional on the measurement defined as follows, (Grewal & Andrews, 2001):

$$E[J_k | \mathbf{z}_{0,\dots,k}] = E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T \mathbf{M}(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) | \mathbf{z}_{0,\dots,k}] \quad 4.62$$

Where \mathbf{M} is a weighting matrix that is a symmetric positive definite matrix with size of $\mathbf{M} \in \mathbb{R}^{n \times n}$. This condition implies that the a posteriori estimation error is unbiased as follows:

$$\begin{aligned} \frac{\partial E[J_k | \mathbf{z}_{0,\dots,k}]}{\partial \hat{\mathbf{x}}_{k|k}} &= -2\mathbf{M}E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) | \mathbf{z}_{0,\dots,k}] = \mathbf{0}_{n \times 1} \\ &\rightarrow E[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k}) | \mathbf{z}_{0,\dots,k}] = \mathbf{0}_{n \times 1} \end{aligned} \quad 4.63$$

If no modeling errors are present, then the a posteriori estimation error for the secondary filter is defined by substituting the output's a priori estimation error of the primary filter (equation (4.41)) into the a posteriori estimation error of the secondary filter (equation (4.57)) as follows:

$$\begin{aligned}
\mathbf{e}_{x,2k|k} &= \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,2k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) \mathbf{e}_{z,2k|k-1} - \mathbf{v}_k \right) \Big|_{\mathbf{e}_{z,2k|k-1} = \mathbf{e}_{z,1k|k-1}} \\
&= \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) \mathbf{e}_{z,1k|k-1} - \mathbf{v}_k \\
&= \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\begin{pmatrix} \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_k \\ -\mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} \end{pmatrix}, \boldsymbol{\Psi}_{tv_k} \right) \right) \begin{pmatrix} \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_k \\ -\mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} \end{pmatrix} - \mathbf{v}_k
\end{aligned} \tag{4.64}$$

The expectation of equation (4.64) is zero assuming that the system and measurement noise are white. Therefore, the a posteriori estimation error is unbiased. The gain \mathbf{S}_{at} can be obtained by minimizing the expectation of the following quadratic cost function:

$$\mathbf{J}_{2k} = \left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T \right] \circ \mathbf{I}_{n \times n} \tag{4.65}$$

Note that the expectation of the cost function \mathbf{J}_{2k} represents the diagonal error covariance matrix. Therefore, the diagonal gain \mathbf{S}_{at} can only minimize the diagonal error covariance matrix. To derive this gain the following procedure is used.

Taking the expectation for equation (4.65), gives the following:

$$E[\mathbf{J}_{2k}] = E \left[(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k})^T \right] \circ \mathbf{I}_{n \times n} \tag{4.66}$$

Substituting equation (4.64) into equation (4.66) gives the following:

$$\begin{aligned}
E[\mathbf{J}_{2k}] &= E \left[\begin{aligned} &\left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) \left(\mathbf{e}_{z,1k|k-1} - \mathbf{v}_k \right) \right) \\ &\times \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) \left(\mathbf{e}_{z,1k|k-1} - \mathbf{v}_k \right) \right)^T \end{aligned} \right] \circ \mathbf{I}_{n \times n}
\end{aligned} \tag{4.67}$$

Taking the derivative of equation (4.67) with respect to \mathbf{S}_{at} (for the diagonal elements only) and setting it to zero leads to the following:

$$\begin{aligned}
E \left[\frac{\partial \mathbf{J}_{2k}}{\partial \mathbf{S}_{at}} \right] &= E \left[\begin{aligned} &\mathbf{e}_{z,1k|k-1} \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) \mathbf{e}_{z,1k|k-1} - \mathbf{v}_k \right)^T \\ &+ \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) \mathbf{e}_{z,1k|k-1} - \mathbf{v}_k \right) \mathbf{e}_{z,1k|k-1}^T \end{aligned} \right] \circ \mathbf{I}_{n \times n} = \mathbf{0}_{n \times n}
\end{aligned} \tag{4.68}$$

Rearranging equation (4.68) gives the following:

$$\begin{aligned}
E \left[\begin{aligned} &-\mathbf{e}_{z,1k|k-1} \mathbf{e}_{z,1k|k-1}^T \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) + \mathbf{e}_{z,1k|k-1} \mathbf{v}_k^T \\ &-\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1k|k-1}, \boldsymbol{\Psi}_{tv_k} \right) \right) \mathbf{e}_{z,1k|k-1} \mathbf{e}_{z,1k|k-1}^T + \mathbf{v}_k \mathbf{e}_{z,1k|k-1}^T \end{aligned} \right] \circ \mathbf{I}_{n \times n} = \mathbf{0}_{n \times n}
\end{aligned} \tag{4.69}$$

Further to equation (4.41) and using the white noise assumption, the term $E \left[\mathbf{e}_{z,1_{k|k-1}} \mathbf{v}_k^T \right]$ is obtained as follows:

$$E \left[\mathbf{e}_{z,1_{k|k-1}} \mathbf{v}_k^T \right] = E \left[\begin{pmatrix} \left(\mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} - \hat{\mathbf{H}}_k \hat{\mathbf{A}}_{k-1} \hat{\mathbf{H}}_k^{-1} \right) \mathbf{z}_{k-1} \\ + \left(\mathbf{H}_k \mathbf{B}_{k-1} - \hat{\mathbf{H}}_k \hat{\mathbf{B}}_{k-1} \right) \mathbf{u}_{k-1} \\ + \mathbf{H}_k \mathbf{w}_{k-1} + \mathbf{v}_k - \mathbf{H}_k \mathbf{A}_{k-1} \mathbf{H}_k^{-1} \mathbf{v}_{k-1} \end{pmatrix} \mathbf{v}_k^T \right] = E \left[\mathbf{v}_k \mathbf{v}_k^T \right] = \mathbf{R}_k \quad 4.70$$

Note that equation (4.70) is independent of the modeling errors. Substituting equation (4.70) into equation (4.69) gives the following:

$$\begin{aligned} & \left(-\mathbf{P}_{zz,1_{k|k-1}} \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1_{k|k-1}}, \boldsymbol{\Psi}_{tv_k} \right) \right)^T + 2\mathbf{R}_k \right. \\ & \quad \left. - \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1_{k|k-1}}, \boldsymbol{\Psi}_{tv_k} \right) \right) \mathbf{P}_{zz,1_{k|k-1}}^T \right) \circ \mathbf{I}_{n \times n} = \mathbf{0}_{n \times n} \end{aligned} \quad 4.71$$

Where $\mathbf{P}_{zz,1_{k|k-1}}$ is the output's a priori error covariance matrix and is defined as $\mathbf{P}_{zz,1_{k|k-1}} = E \left[\mathbf{e}_{z,1_{k|k-1}} \mathbf{e}_{z,1_{k|k-1}}^T \right]$. Since \mathbf{S}_{at} is diagonal, equation (4.71) can be reduced to the following:

$$\left(-\mathbf{P}_{zz,1_{k|k-1}} \left(\mathbf{I}_{n \times n} - \mathbf{S}_{at} \left(\mathbf{e}_{z,1_{k|k-1}}, \boldsymbol{\Psi}_{tv_k} \right) \right) + \mathbf{R}_k \right) \circ \mathbf{I}_{n \times n} = \mathbf{0}_{n \times n} \quad 4.72$$

Rearranging equation (4.72) gives the following:

$$\mathbf{S}_{at} \left(\mathbf{e}_{z,1_{k|k-1}}, \boldsymbol{\Psi}_{tv_k} \right) = \left(\left(\mathbf{P}_{zz,1_{k|k-1}} - \mathbf{R}_k \right) \circ \mathbf{I}_{n \times n} \right) \left(\mathbf{P}_{zz,1_{k|k-1}} \circ \mathbf{I}_{n \times n} \right)^{-1} \quad 4.73$$

The time-varying smoothing boundary layer is obtained by substituting equation (4.50) into equation (4.73) as follows:

$$\text{diag}(\boldsymbol{\Psi}_{tv_k}) = \text{diag} \left(\left| \mathbf{e}_{z,1_{k|k-1}} \right| \right) \left(\mathbf{P}_{zz,1_{k|k-1}} \circ \mathbf{I}_{n \times n} \right) \times \left(\begin{matrix} \mathbf{P}_{zz,1_{k|k-1}} \circ \mathbf{I}_{n \times n} \\ -\mathbf{R}_k \circ \mathbf{I}_{n \times n} \end{matrix} \right)^{-1} \quad 4.74$$

If no modeling errors are present and the measurement matrix is the identity matrix, then equation (4.74) can be rewritten using equation (4.41) as follows:

$$\begin{aligned} \text{diag}(\boldsymbol{\Psi}_{tv_k}) &= \text{diag} \left(\left| \mathbf{e}_{z,1_{k|k-1}} \right| \right) \left(\left(\mathbf{Q}_{k-1} + \mathbf{R}_k + \mathbf{A}_{k-1} \mathbf{R}_{k-1} \mathbf{A}_{k-1}^T \right) \circ \mathbf{I}_{n \times n} \right) \\ & \quad \times \left(\left(\mathbf{Q}_{k-1} + \mathbf{A}_{k-1} \mathbf{R}_{k-1} \mathbf{A}_{k-1}^T \right) \circ \mathbf{I}_{n \times n} \right)^{-1} \end{aligned} \quad 4.75$$

Substituting equation (4.75) into equation (4.73) yields the following:

$$\begin{aligned} \mathbf{S}_{at}(\mathbf{e}_{z,1_{k|k-1}}, \Psi_{tv_k}) \\ = \left(\left(\begin{array}{c} \mathbf{Q}_{k-1} \\ +\mathbf{A}_{k-1}\mathbf{R}_{k-1}\mathbf{A}_{k-1}^T \end{array} \right) \circ \mathbf{I}_{n \times n} \right) \left(\left(\begin{array}{c} \mathbf{Q}_{k-1} + \mathbf{R}_k \\ +\mathbf{A}_{k-1}\mathbf{R}_{k-1}\mathbf{A}_{k-1}^T \end{array} \right) \circ \mathbf{I}_{n \times n} \right)^{-1} \end{aligned} \quad 4.76$$

The time-varying smoothing boundary layer of equations (4.74) and (4.75) provide the SVSF with significant advantages over strategies such as the KF as follows:

- Stability is achieved, despite modeling uncertainties.
- The inversion operator of equations (4.73) and (4.74) is involved with diagonal positive matrices, which simplifies calculations and overcomes issues of the KF regarding the inversion operator.
- Unlike the KF, there is no need to predict and correct the error covariance matrix at each time step for the proposed gain as it depends only on \mathbf{Q}_{k-1} and \mathbf{R}_{k-1} . This reduces the filter complexity.
- If the system is linear and is stationary, the measurement matrix is the identity matrix, and its model is accurately estimated, then the term $\mathbf{S}_{at}(\mathbf{e}_{z,1_{k|k-1}}, \Psi_{tv_k})$ becomes time invariant, and it can be calculated offline as follows:

$$\mathbf{S}_{at}(\mathbf{e}_{z,1_{k|k-1}}, \Psi_{tv_k}) = ((\mathbf{Q} + \mathbf{A}\mathbf{R}\mathbf{A}^T) \circ \mathbf{I}_{n \times n}) \times ((\mathbf{Q} + \mathbf{R} + \mathbf{A}\mathbf{R}\mathbf{A}^T) \circ \mathbf{I}_{n \times n})^{-1} \quad 4.77$$

This improves the filter performance in terms of its complexity and computation time.

- The proposed algorithm does not depend on the error due to initial conditions as it becomes negligible with increasing time steps (depending on the matrix $\boldsymbol{\gamma}$). If $\boldsymbol{\gamma}$ is set to be zero, then the error due to initial conditions is eliminated in one time step. Moreover, the gain depends on the output's initial conditions which are available. On the other hand, the KF depends on the initial estimate and covariance matrices as they should be accurately estimated; otherwise its convergence is affected.

If modeling errors are present, the time-varying smoothing boundary layer can be obtained by taking an a priori chattering segment with large width at a time and assuming the following:

1. The measurement and system noise are white and are stationary.
2. The matrix \mathbf{S}_{at} is constant during that segment.

The output's a priori and a posteriori variance matrices of the segment; $\sigma_{\mathbf{e}_{z_d|d-1}}^2$ and $\sigma_{\mathbf{e}_{z_d|d}}^2$, respectively, are obtained by taking the variance of the output's a priori and a posteriori estimation error signals, respectively, as follows:

$$\begin{aligned}\sigma_{\mathbf{e}_{z_d|d-1}}^2 &= \frac{1}{d-1} \sum_{i=1}^d \left(\left(\mathbf{e}_{z_{i|i-1}} - \frac{1}{d} \sum_{j=1}^d \mathbf{e}_{z_{j|j-1}} \right) \left(\mathbf{e}_{z_{i|i-1}} - \frac{1}{d} \sum_{j=1}^d \mathbf{e}_{z_{j|j-1}} \right)^T \right) \\ \sigma_{\mathbf{e}_{z_d|d}}^2 &= \frac{1}{d-1} \sum_{i=1}^d \left(\left(\mathbf{e}_{z_{i|i}} - \frac{1}{d} \sum_{j=1}^d \mathbf{e}_{z_{j|j}} \right) \left(\mathbf{e}_{z_{i|i}} - \frac{1}{d} \sum_{j=1}^d \mathbf{e}_{z_{j|j}} \right)^T \right)\end{aligned}\quad 4.78$$

And the noise mean and variance are obtained from its expectation and covariance, using the *Law of Large Numbers*.

Lemma 4.1 - The Law of Large Numbers

The Law of Large Numbers states that: if d stationary uncorrelated random variables, such as v_i , $i = 1 \dots d$, share the same expectation value, then their average tends towards its expected value as $d \rightarrow \infty$ as follows, (Bar-Shalom & Li, 1993):

$$\text{If } E[v_1] = E[v_2] = \dots = E[v_d] = \bar{v} \text{ then } \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=1}^d v_i \right) = \bar{v} \quad 4.79$$

Where $\frac{1}{d} \sum_{i=1}^d v_i$ is the average of the random variables. The previous lemma can be extended to include the variance as follows: if d stationary uncorrelated random variables have zero means and share the same covariance value, then the variance of the random variables together tends towards the covariance matrix as $d \rightarrow \infty$. This statement implies the following:

$$\begin{aligned}\text{If } E[v_1 v_1^T] = \dots = E[v_d v_d^T] = \sigma^2 \text{ then} \\ \lim_{d \rightarrow \infty} \frac{1}{d-1} \sum_{i=1}^d \left(v_i - \frac{1}{d} \sum_{j=1}^d v_j \right) \left(v_i - \frac{1}{d} \sum_{j=1}^d v_j \right)^T = \sigma^2\end{aligned}\quad 4.80$$

With a segment, the time-varying function \mathbf{S}_{at} is assumed stationary and is obtained by minimizing the variance of the segment's a posteriori estimation error, $\sigma_{\mathbf{e}_{x_d|d}}^2$, which is defined as follows:

$$\sigma_{\mathbf{e}_{x_d|d}}^2 = \frac{1}{d-1} \sum_{i=1}^d \left[\left(\mathbf{e}_{x_{i|i}} - \frac{1}{d} \sum_{j=1}^d \mathbf{e}_{x_{j|j}} \right) \left(\mathbf{e}_{x_{i|i}} - \frac{1}{d} \sum_{j=1}^d \mathbf{e}_{x_{j|j}} \right)^T \right] \quad 4.81$$

Further to the law of large numbers, any segment from the measurement noise has variance that converges to the measurement noise covariance matrix when the segment's length becomes large. Moreover, if both the noise and the input are white signals, then the following are obtained:

$$\begin{aligned}\lim_{d \rightarrow \infty} \frac{1}{d} \sum_{i=1}^d \mathbf{w}_i &= \lim_{d \rightarrow \infty} \frac{1}{d} \sum_{i=1}^d \mathbf{v}_i = \mathbf{0}_{n \times 1} \\ \lim_{d \rightarrow \infty} \boldsymbol{\sigma}_{\mathbf{v}_k, \mathbf{w}_k} &= \lim_{d \rightarrow \infty} \boldsymbol{\sigma}_{\mathbf{z}_k, \mathbf{w}_k} = \mathbf{0}_{n \times n} \\ \lim_{d \rightarrow \infty} \boldsymbol{\sigma}_{\mathbf{u}_k, \mathbf{v}_k} &= \lim_{d \rightarrow \infty} \boldsymbol{\sigma}_{\mathbf{u}_k, \mathbf{w}_k} = \lim_{d \rightarrow \infty} \boldsymbol{\sigma}_{\mathbf{u}_k, \mathbf{z}_k} = \mathbf{0}_{1 \times n}\end{aligned}\quad 4.82$$

Where $\boldsymbol{\sigma}_{\mathbf{a}, \mathbf{b}}$ is the covariance between \mathbf{a} and \mathbf{b} . Substituting equations (4.64) and (4.82) into equation (4.81) and rearranging give the following:

$$\boldsymbol{\sigma}_{\mathbf{e}_{x_{d|d}}}^2 = \frac{1}{d-1} \sum_{i=1}^d \left[\left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{z,1_{|i-1}}, \boldsymbol{\Psi}_{tv_i}) \right) (\mathbf{e}_{z,1_{|i-1}} - \bar{\mathbf{e}}) - \mathbf{v}_k \right) \times \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{z,1_{|i-1}}, \boldsymbol{\Psi}_{tv_i}) \right) (\mathbf{e}_{z,1_{|i-1}} - \bar{\mathbf{e}}) - \mathbf{v}_k \right)^T \right] \quad 4.83$$

Where $\bar{\mathbf{e}}$ is the average of the output's a priori estimation error segment.

Taking the derivative of equation (4.83) with respect to \mathbf{S}_{at} (for the diagonal elements only) and setting it to zero lead to the following:

$$\begin{aligned}\frac{\partial \boldsymbol{\sigma}_{\mathbf{e}_{x_{d|d}}}^2 \circ \mathbf{I}_{n \times n}}{\partial \mathbf{S}_{at}} &= \frac{1}{d-1} \sum_{i=1}^d \left[\left(\mathbf{e}_{z,1_{|i-1}} - \bar{\mathbf{e}} \right) \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{z,1_{|i-1}}, \boldsymbol{\Psi}_{tv_i}) \right) (\mathbf{e}_{z,1_{|i-1}} - \bar{\mathbf{e}}) - \mathbf{v}_i \right)^T \right. \\ &\quad \left. + \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{z,1_{|i-1}}, \boldsymbol{\Psi}_{tv_i}) \right) (\mathbf{e}_{z,1_{|i-1}} - \bar{\mathbf{e}}) - \mathbf{v}_i \right) (\mathbf{e}_{z,1_{|i-1}} - \bar{\mathbf{e}})^T \right] \circ \mathbf{I}_{n \times n} \\ &= \mathbf{0}_{n \times n}\end{aligned}\quad 4.84$$

Rearranging equation (4.84) and using the diagonal principles and equation (4.82) give the following:

$$\begin{aligned}\lim_{d \rightarrow \infty} \left(\frac{\partial \boldsymbol{\sigma}_{\mathbf{e}_{x_{d|d}}}^2 \circ \mathbf{I}_{n \times n}}{\partial \mathbf{S}_{at}} \right) &= \left(\left(\mathbf{I}_{n \times n} - \mathbf{S}_{at}(\mathbf{e}_{z,1_{|i-1}}, \boldsymbol{\Psi}_{tv_i}) \right) \boldsymbol{\sigma}_{\mathbf{e}_{z,1_{d|d-1}}}^2 - \mathbf{R} \right) \circ \mathbf{I}_{n \times n} \\ &= \mathbf{0}_{n \times n}\end{aligned}\quad 4.85$$

Where $\boldsymbol{\sigma}_{\mathbf{e}_{z,1_{d|d-1}}}^2$ is the variance of the output's a priori estimation error (or the a priori chattering) of the first filter. Rearranging equation (4.85) gives the following:

$$S_{at}(\mathbf{e}_{z,1_{i|i-1}}, \Psi_{tv_i}) = \left((\sigma_{\mathbf{e}_{z,1_{d|d-1}}}^2 - \mathbf{R}) \circ \mathbf{I}_{n \times n} \right) (\sigma_{\mathbf{e}_{z,1_{d|d-1}}}^2 \circ \mathbf{I}_{n \times n})^{-1} \quad 4.86$$

The time-varying smoothing boundary layer is obtained by substituting equation (4.50) into equation (4.73) as follows:

$$diag(\Psi_{tv_k}) = diag(|\mathbf{e}_{z,1_{k|k-1}}|) (\sigma_{\mathbf{e}_{z,1_{d|d-1}}}^2 \circ \mathbf{I}_{n \times n}) \left((\sigma_{\mathbf{e}_{z,1_{d|d-1}}}^2 - \mathbf{R}) \circ \mathbf{I}_{n \times n} \right)^{-1} \quad 4.87$$

The boundary layer width in equation (4.87) is time varying although the function S_{at} is stationary. Note when there are modeling errors, the function S_{at} starts to approach the identity matrix as \mathbf{R} becomes significantly smaller than $\sigma_{\mathbf{e}_{z,1_{d|d-1}}}^2$.

The algorithm in Fig 4.5 can be extended to accommodate modeling errors when they present as shown in Fig 4.6. A threshold is created using the maximum absolute value of the output's a priori estimation error in equation (4.41) (when no modeling errors are present) as follows:

$$\delta = |\hat{\mathbf{A}}_{k-1}| |(\mathbf{v})_{max}| + |(\mathbf{w})_{max}| + |(\mathbf{v})_{max}| \quad 4.88$$

The output's a priori estimation error obtained from filter 1 is tested using the threshold of equation (4.88). If it is smaller than the threshold then the secondary filter uses the boundary layer of equation (4.75) (as no modeling errors are present). If the a priori estimation error is larger than the threshold, then the secondary filter takes a segment of that error and obtains the smoothing boundary layer by using equation (4.87) (as modeling errors are present).

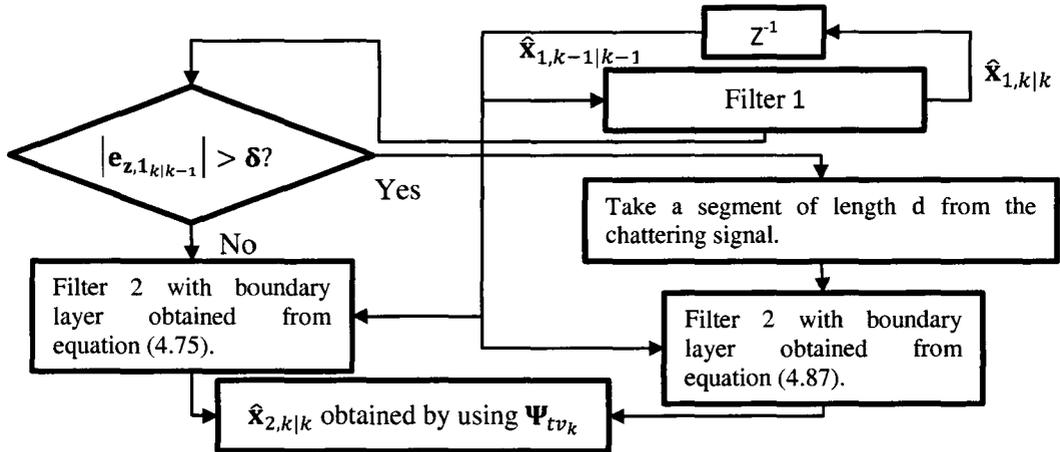


Fig 4.6: The SVSFs with time-varying smoothing boundary layer including the effects of modeling errors.

4.3.3 The chattering amplitudes and its information content

As discussed earlier, the a priori and the a posteriori estimates chatter within their respective existence subspaces. In order to eliminate chattering, a smoothing boundary layer is used. If the smoothing boundary layer has a width that is quite larger than the a priori existence subspace's width, then the a priori chattering, which is a function of the discrepancy between the smoothing boundary layer's width in comparison to the a priori existence subspace's width, is eliminated. Moreover, the corrective action is refined and its energy is reduced. By refining the corrective action, the estimate is forced towards the sliding hyper-plane without the necessity to cross it, or to cross it with smaller amplitude. This behavior reduces the a posteriori chattering amplitude and frequency. By refining the a posteriori chattering, the filter avoids and reduces the propagation of the chattering term through the two stages of the filter.

Conversely, if the smoothing boundary layer has a width that is smaller than the a priori existence subspace's width, then the a priori chattering occurs with an amplitude that is equivalent to the difference between the output's a priori estimation error and the smoothing boundary layer's width, (Habibi, 2007). The a posteriori chattering remains equal to equation (4.44) as the corrective action is applied with its full energy (the function **sat** becomes **sgn**).

The a priori chattering can be used to point out the source and amplitude of modeling errors. Chattering gives an indication that the current filter's model is uncertain and it needs to be re-estimated or tuned. Chattering thus provides an opportunity for combining the SVSF with adaptive techniques for model refinement as discussed later in chapter six.

According to equation (4.39), the SVSF has n -a priori chattering signals, one associated with each estimate. Thus, (assuming that the measurement matrix has full rank and is equal to the identity matrix) each chattering signal points out the modeling errors in the corresponding row of the system and input matrices. As such, the a priori chattering contents provide a means for extracting the modeling error explicitly. In this study, two approaches are presented depending on the input and measurement signal properties as follows:

- The first method uses the expectation operator of the a priori chattering and it is used for systems with non-stationary input and measurement.

- The second method is based on using the covariance between the a priori chattering signal and the measurement and the covariance between the a priori chattering and the input.

In order for these two methods to work, the measurement and system noise signals must be stationary (expectation and covariance matrix are time invariant). Since the expectation and the covariance matrix are not directly available, a segment of the a priori chattering is taken and its statistics are calculated; i.e. mean and variance. Using the law of large numbers, the means and variance matrices of the measurement and system noise in the a priori chattering segment approach their expectations and covariance matrices, respectively, and hence they are known.

4.3.3.1 Obtaining the chattering contents for non-stationary input using the average

The first method involves obtaining the modeling errors of a time invariant system from the a priori chattering signals by using an averaging technique. This can be done by taking a segment of length d (starting from the time step l , when the term \mathbf{y}^l becomes negligible, to time step $l + d - 1$) from the output's a priori estimation error, the input and the measurement signals, and then calculating their averages. From equation (4.39), the average of the output's a priori estimation error segment is obtained as follows (assuming a time invariant model):

$$\begin{aligned} \frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{e}_{z_{i|i-1}} &= \frac{1}{d} (\mathbf{H}\mathbf{A}\mathbf{H}^+ - \hat{\mathbf{H}}\hat{\mathbf{A}}\hat{\mathbf{H}}^+) \sum_{i=l}^{l+d-1} \mathbf{z}_{i-1} + \frac{1}{d} (\mathbf{H}\mathbf{B} - \hat{\mathbf{H}}\hat{\mathbf{B}}) \sum_{i=l}^{l+d-1} u_{i-1} \\ &\quad + \frac{1}{d} \sum_{i=l}^{l+d-1} (\mathbf{H}\mathbf{w}_{i-1} + \mathbf{v}_i - \mathbf{H}\mathbf{A}\mathbf{H}^+\mathbf{v}_{i-1}) \end{aligned} \quad 4.89$$

If d is chosen to be large enough to satisfy equations (4.79) and (4.80), then the averages of the measurement and system noise converge to their expectations which are zero (zero-mean assumption) as follows:

$$\lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} (\mathbf{H}\mathbf{w}_{i-1} + \mathbf{v}_i - \mathbf{H}\mathbf{A}\mathbf{H}^+\mathbf{v}_{i-1}) \right) = E[\mathbf{H}\mathbf{w}_{k-1} + \mathbf{v}_k - \mathbf{H}\mathbf{A}\mathbf{H}^+\mathbf{v}_{k-1}] = \mathbf{0}_{n \times 1} \quad 4.90$$

The a priori chattering at time i , $\mathbf{Ch}_{i|i-1}$, is then equal (when present) to the difference between the output's a priori estimation error and the smoothing boundary layer width as follows:

$$\mathbf{Ch}_{i|i-1} = \begin{bmatrix} Ch_{1|i|i-1} \\ \vdots \\ Ch_{n|i|i-1} \end{bmatrix} \text{ where } Ch_{j|i|i-1} = \begin{cases} 0 & |e_{z_{j|i|i-1}}| < \Psi_{j_i} \\ e_{z_{j|i|i-1}} - \Psi_{j_i} & |e_{z_{j|i|i-1}}| \geq \Psi_{j_i} \end{cases} \quad 4.91$$

If the measurement matrix is the identity matrix, system and measurement noise are white, and the smoothing boundary layer is set to have zero width, then equation (4.89) is simplified to the following:

$$\begin{aligned} \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{Ch}_{i|i-1} \right) &= \lim_{d \rightarrow \infty} \left(\frac{1}{d} (\mathbf{A} - \hat{\mathbf{A}}) \sum_{i=l}^{l+d-1} \mathbf{z}_{i-1} + \frac{1}{d} (\mathbf{B} - \hat{\mathbf{B}}) \sum_{i=l}^{l+d-1} u_{i-1} \right) \\ &= \lim_{d \rightarrow \infty} \left(\frac{1}{d} \Delta \mathbf{A} \sum_{i=l}^{l+d-1} \mathbf{z}_{i-1} + \frac{1}{d} \Delta \mathbf{B} \sum_{i=l}^{l+d-1} u_{i-1} \right) \end{aligned} \quad 4.92$$

Equation (4.92) (for systems with single input) has $(n+1) \times n$ unknowns which are the following:

- $n \times n$ modeling errors in the system matrix, $\Delta \mathbf{A}$.
- $n \times 1$ modeling errors in the input matrix, $\Delta \mathbf{B}$.

To solve these unknowns, $(n+1) \times n$ equations are needed. Each segment provides $(n+1)$ data points that represent the average values of n –measurement and one-input in that segment. Therefore using n segments provides n data points of vector $\begin{bmatrix} \sum_{i=l}^{l+d-1} \mathbf{z}_{i-1} \\ \sum_{i=l}^{l+d-1} \mathbf{u}_{i-1} \end{bmatrix}$, which gives $(n+1) \times n$ equations. This allows the problem to be solved. Using these vectors, equation (4.92) is then revised as follows:

$$\lim_{d \rightarrow \infty} \left(\frac{1}{d} \begin{bmatrix} \sum_{i=l_1}^{l_1+d-1} \mathbf{Ch}_{i|i-1} \\ \vdots \\ \sum_{i=l_n}^{l_n+d-1} \mathbf{Ch}_{i|i-1} \end{bmatrix}^T \right) = [\Delta \mathbf{A} \quad \Delta \mathbf{B}] \lim_{d \rightarrow \infty} \left(\frac{1}{d} \begin{bmatrix} \sum_{i=l_1}^{l_1+d-1} \mathbf{z}_{i-1} & \dots & \sum_{i=l_n}^{l_n+d-1} \mathbf{z}_{i-1} \\ \sum_{i=l_1}^{l_1+d-1} u_{i-1} & \dots & \sum_{i=l_n}^{l_n+d-1} u_{i-1} \end{bmatrix} \right) \quad 4.93$$

Where $l_i, i = 1, \dots, n$ represents the beginning of the i^{th} segment. Equation (4.93) is expressed in a simpler form as follows:

$$\mathbf{CH} = [\Delta \mathbf{A} \quad \Delta \mathbf{B}] \begin{bmatrix} \bar{\mathbf{Z}} \\ \bar{\mathbf{U}} \end{bmatrix} \quad 4.94$$

Where $\mathbf{CH} = \lim_{d \rightarrow \infty} \left(\frac{1}{d} \begin{bmatrix} \sum_{i=l_1}^{l_1+d-1} \mathbf{Ch}_{i|i-1} \\ \vdots \\ \sum_{i=l_n}^{l_n+d-1} \mathbf{Ch}_{i|i-1} \end{bmatrix}^T \right)$, $\bar{\mathbf{Z}} = \lim_{d \rightarrow \infty} \left(\frac{1}{d} \begin{bmatrix} \sum_{i=l_1}^{l_1+d-1} \mathbf{z}_{i-1} \\ \vdots \\ \sum_{i=l_n}^{l_n+d-1} \mathbf{z}_{i-1} \end{bmatrix}^T \right)$ and $\bar{\mathbf{U}} = \lim_{d \rightarrow \infty} \left(\frac{1}{d} \begin{bmatrix} \sum_{i=l_1}^{l_1+d-1} u_{i-1} \\ \vdots \\ \sum_{i=l_n}^{l_n+d-1} u_{i-1} \end{bmatrix}^T \right)$. Note that the input must not be stationary. From equation (4.94) the modeling errors can be obtained as follows:

$$[\Delta\mathbf{A} \quad \Delta\mathbf{B}] = \mathbf{CH} \begin{bmatrix} \bar{\mathbf{Z}} \\ \bar{\mathbf{U}} \end{bmatrix}^{-1} \quad 4.95$$

Equation (4.95) can be used to adapt the filter's model. Due to the expectation operator, the extracted parameters may differ from their real values. If the refined model is close to the system model, then the a priori chattering due to modeling error becomes negligible. Therefore, the a priori chattering becomes a function of the measurement and system noise, and its upper bound is obtained by equation (4.88). This upper bound is then used as the SVSF's new smoothing boundary layer. If the refined model differs from the system model due to the expectation operator, then the output's a priori estimation error exceeds the smoothing boundary layer and thus a priori chattering appears and indicates that the filter model needs to be re-tuned. The refinement's procedure is then repeated. Due to the choice of the segment length and the usage of the expectation operator, the refinement procedure may be repeated several times. Once the filter is tuned, the SVSF's smoothing boundary layer width changes according to equation (4.88). If modeling errors are present again, the smoothing boundary layer width is set to be zero in order to enable the refinement's process. After the refinement procedure is completed, the boundary layer is changed back to equation (4.88) using the new model parameters.

4.3.3.2 Obtaining the chattering contents for stationary random input with zero mean using the covariance

The drawback of the first method is that if the input is white (in the sense of $E(u_k) = 0$ and $E(u_k u_{k-1}) = 0$), then the inversion matrix in equation (4.95) becomes invalid and the modeling errors cannot be obtained. To overcome this issue, another method is proposed that involves obtaining the modeling errors of a time invariant system from the a priori chattering signals by using the variance technique. This can be done by taking a segment of length d time steps (starting from the time step l , when the term $\boldsymbol{\gamma}^l$ becomes negligible, to the time step $l + d - 1$) and then calculating the covariance between the a priori chattering, $\mathbf{Ch}_{i|i-1}$, and the measurement \mathbf{z}_{i-1} for this segment. If the

measurement matrix is set to the identity matrix, and the smoothing boundary layer is set to have zero width, then the covariance between $\mathbf{Ch}_{i|i-1}$ and \mathbf{z}_{i-1} , $(\boldsymbol{\sigma}_{\mathbf{Ch}_{d|d-1}, \mathbf{z}_{d-1}})$, is defined as follows:

$$\begin{aligned} \boldsymbol{\sigma}_{\mathbf{Ch}_{d|d-1}, \mathbf{z}_{d-1}} &= \frac{1}{d-1} \sum_{i=l}^{l+d-1} \left((\mathbf{Ch}_{i|i-1} - \overline{\mathbf{Ch}}_{k|k-1}) (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T \right) \\ \boldsymbol{\sigma}_{\mathbf{Ch}_{d|d-1}, \mathbf{z}_{d-1}} &= \frac{1}{d-1} \sum_{i=l}^{l+d-1} \begin{pmatrix} \Delta \mathbf{A} (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1}) (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T \\ + \Delta \mathbf{B} (u_{i-1} - \bar{u}_{k-1}) (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T \\ + (\mathbf{w}_{i-1} - \bar{\mathbf{w}}_{k-1}) (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T \\ + (\mathbf{v}_i - \bar{\mathbf{v}}_k) (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T \\ - \mathbf{A} (\mathbf{v}_{i-1} - \bar{\mathbf{v}}_{k-1}) (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T \end{pmatrix} \end{aligned} \quad 4.96$$

Where the notation \bar{a}_i represents the average of the segment around time i of the parameter a . If the segment length is consistent with the law of large numbers, and assuming that the noise and the input have zero means, then the following relationships are obtained:

$$\begin{aligned} \lim_{d \rightarrow \infty} (\bar{u}_{k-1}) &= 0, \quad \lim_{d \rightarrow \infty} (\bar{\mathbf{v}}_{k-1}) = \lim_{d \rightarrow \infty} (\bar{\mathbf{w}}_{k-1}) = \mathbf{0}_{n \times 1}, \\ \lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} (\mathbf{w}_{i-1} (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T) \right) &= \lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} (\mathbf{v}_i (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T) \right) = \mathbf{0}_{n \times n}, \\ \lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} (u_{i-1} (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T) \right) &= \mathbf{0}_{1 \times n} \text{ and} \\ \lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} (\mathbf{v}_{i-1} (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T) \right) &= \mathbf{R} \end{aligned} \quad 4.97$$

Substituting equation (4.97) into equation (4.96) gives the following:

$$\lim_{d \rightarrow \infty} (\boldsymbol{\sigma}_{\mathbf{Ch}_{d|d-1}, \mathbf{z}_{d-1}}) = \Delta \mathbf{A} \lim_{d \rightarrow \infty} (\boldsymbol{\sigma}_{\mathbf{z}_{d-1}}^2) - \mathbf{A} \mathbf{R} \quad 4.98$$

Where $\boldsymbol{\sigma}_{\mathbf{z}_{d-1}}^2 = \frac{1}{d-1} \sum_{i=l}^{l+d-1} ((\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1}) (\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1})^T)$. The modeling errors in the system matrix are then obtained by rearranging equation (4.98) as follows:

$$\Delta \mathbf{A} = \left(\lim_{d \rightarrow \infty} (\boldsymbol{\sigma}_{\mathbf{Ch}_{d|d-1}, \mathbf{z}_{d-1}}) + \hat{\mathbf{A}} \mathbf{R} \right) \left(\lim_{d \rightarrow \infty} (\boldsymbol{\sigma}_{\mathbf{z}_{d-1}}^2) - \mathbf{R} \right)^{-1} \quad 4.99$$

To estimate the input matrix from the chattering equation, the same segment is used and the cross-variance between the a priori chattering and the input signals for that segment is obtained. Further to the a priori chattering equation (equation (4.91)), and assuming the noise and the input are white signals with zero means, then the following is obtained:

$$\begin{aligned}
\sigma_{\mathbf{Ch}_{d|d-1}, u_{d-1}} &= \frac{1}{d-1} \sum_{i=l}^{l+d-1} \left((\mathbf{Ch}_{i|i-1} - \overline{\mathbf{Ch}}_{k|k-1}) u_{i-1}^T \right) \\
&= \frac{1}{d-1} \sum_{i=l}^{l+d-1} \begin{pmatrix} (\mathbf{A} - \widehat{\mathbf{A}})(\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1}) u_{i-1}^T \\ + (\mathbf{B} - \widehat{\mathbf{B}}) u_{i-1} u_{i-1}^T \\ + \mathbf{w}_{i-1} u_{i-1}^T + \mathbf{v}_i u_{i-1}^T - \mathbf{A} \mathbf{v}_{i-1} u_{i-1}^T \end{pmatrix}
\end{aligned} \tag{4.100}$$

The law of large numbers implies that if the input and the noise are uncorrelated, then for $d \rightarrow \infty$ the following relationships are obtained:

$$\begin{aligned}
\lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} (\mathbf{w}_{k-1} u_{k-1}^T) \right) &= \lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} (\mathbf{v}_k u_{k-1}^T) \right) \\
&= \lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} ((\mathbf{z}_{i-1} - \bar{\mathbf{z}}_{k-1}) u_{i-1}^T) \right) = \mathbf{0}_{n \times 1}
\end{aligned} \tag{4.101}$$

Substituting equation (4.101) into equation (4.100) gives the following:

$$\lim_{d \rightarrow \infty} (\sigma_{\mathbf{Ch}_{d|d-1}, u_{d-1}}) = \lim_{d \rightarrow \infty} \left(\frac{1}{d-1} \sum_{i=l}^{l+d-1} (\Delta \mathbf{B} u_{i-1} u_{i-1}^T) \right) = \Delta \mathbf{B} \lim_{d \rightarrow \infty} (\sigma_{u_{d-1}}^2) \tag{4.102}$$

Where $\sigma_{u_{d-1}}^2$ is the input's variance value. The modeling errors in the input matrix is then obtained as follows:

$$\Delta \mathbf{B} = \lim_{d \rightarrow \infty} (\sigma_{\mathbf{Ch}_{d|d-1}, u_{d-1}}) \left(\lim_{d \rightarrow \infty} (\sigma_{u_{d-1}}^2) \right)^{-1} \tag{4.103}$$

This algebraic algorithm can be used to refine the filter's model. The benefit of this algorithm over the first algorithm is that only one segment of length d is needed, thus reducing complexity and computation time. However, better knowledge of the noise is needed and an input signal that is white in nature is required. If the noise covariance matrices are wrongly estimated, then the modeling errors are not accurately reconstructed. In other words, this method is more sensitive to the noise model as compared to the first outlined algorithm.

4.3.3.3 Obtaining the chattering contents for non-stationary biased input with zero mean over the segment using the cross-correlation

The first method needs the mean of the input to be non-zero and the second method applies to an input signal that is white. If the input is a non-stationary biased signal such that its expectation is not zero, and it is varying with time but its mean is zero over the segment, then the previous two methods are not applicable. For example, the multi-steps input in Fig 4.7 has two levels; 1 for the first half of the segment and -1 for the second half. The mean of the segment is zero, but its autocorrelation, $E(u_k u_{k-1})$, is not zero.

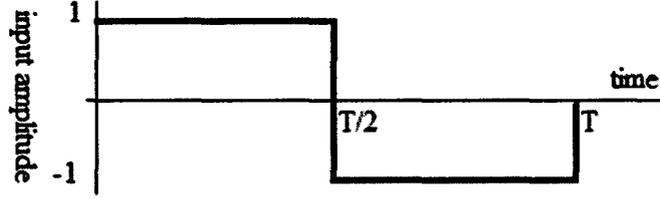


Fig 4.7: Multi-step input.

To solve this problem, another method is proposed that involves obtaining the modeling errors of a time invariant system from a segment of the a priori chattering signals by using the cross-correlation between chattering and the measurement signal that is delayed by two time steps, $C_{\text{Ch}_{d|d-1}, z_{d-2}}$, and between chattering and the input signal that is delayed by two time steps, $C_{\text{Ch}_{d|d-1}, u_{d-2}}$, as follows:

$$C_{\text{Ch}_{d|d-1}, z_{d-2}} = \frac{1}{d} \sum_{i=l}^{l+d-1} (\text{Ch}_{i|i-1} z_{i-2}^T) = \frac{1}{d} \sum_{i=l}^{l+d-1} \begin{pmatrix} \Delta \mathbf{A}(\mathbf{z}_{i-1} \mathbf{z}_{i-2}^T) + \Delta \mathbf{B}(u_{i-1} \mathbf{z}_{i-2}^T) \\ + (\mathbf{w}_{i-1} \mathbf{z}_{i-2}^T) \\ + (\mathbf{v}_i \mathbf{z}_{i-2}^T) - \mathbf{A}(\mathbf{v}_{i-1} \mathbf{z}_{i-2}^T) \end{pmatrix} \quad 4.104$$

$$C_{\text{Ch}_{d|d-1}, u_{d-2}} = \frac{1}{d} \sum_{i=l}^{l+d-1} (\text{Ch}_{i|i-1} u_{i-2}^T) = \frac{1}{d} \sum_{i=l}^{l+d-1} \begin{pmatrix} \Delta \mathbf{A}(\mathbf{z}_{i-1} u_{i-2}^T) + \Delta \mathbf{B}(u_{i-1} u_{i-2}^T) \\ + (\mathbf{w}_{i-1} u_{i-2}^T) \\ + (\mathbf{v}_i u_{i-2}^T) - \mathbf{A}(\mathbf{v}_{i-1} u_{i-2}^T) \end{pmatrix} \quad 4.105$$

Using the law of large numbers, the terms that contain the system and measurement noise vanish as $d \rightarrow \infty$. Therefore; equations (4.104) and (4.105) are simplified as follows:

$$\lim_{d \rightarrow \infty} (C_{\text{Ch}_{d|d-1}, z_{d-2}}) = \Delta \mathbf{A} \lim_{d \rightarrow \infty} (C_{z_{d-1}, z_{d-2}}) + \Delta \mathbf{B} \lim_{d \rightarrow \infty} (C_{u_{d-1}, z_{d-2}}) \quad 4.106$$

$$\lim_{d \rightarrow \infty} (C_{\text{Ch}_{d|d-1}, u_{d-2}}) = \Delta \mathbf{A} \lim_{d \rightarrow \infty} (C_{z_{d-1}, u_{d-2}}) + \Delta \mathbf{B} \lim_{d \rightarrow \infty} (C_{u_{d-1}, u_{d-2}}) \quad 4.107$$

Where $C_{z_{d-1}, z_{d-2}} = \frac{1}{d} \sum_{i=l}^{l+d-1} (\mathbf{z}_{i-1} \mathbf{z}_{i-2}^T)$, $C_{u_{d-1}, z_{d-2}} = \frac{1}{d} \sum_{i=l}^{l+d-1} (u_{i-1} \mathbf{z}_{i-2}^T)$, $C_{z_{d-1}, u_{d-2}} = \frac{1}{d} \sum_{i=l}^{l+d-1} (\mathbf{z}_{i-1} u_{i-2}^T)$ and $C_{u_{d-1}, u_{d-2}} = \frac{1}{d} \sum_{i=l}^{l+d-1} (u_{i-1} u_{i-2}^T)$.

The modeling errors can be reconstructed by rearranging equations (4.106) and (4.107), as follows:

$$\Delta \mathbf{A} = \left[\begin{array}{c} \left(\lim_{d \rightarrow \infty} (C_{\text{Ch}_{d|d-1}, u_{d-2}}) \left(\lim_{d \rightarrow \infty} (C_{u_{d-1}, u_{d-2}}) \right)^{-1} \lim_{d \rightarrow \infty} (C_{u_{d-1}, z_{d-2}}) \right) \\ - \lim_{d \rightarrow \infty} (C_{\text{Ch}_{d|d-1}, z_{d-2}}) \\ \times \left(\lim_{d \rightarrow \infty} (C_{z_{d-1}, u_{d-2}}) \left(\lim_{d \rightarrow \infty} (C_{u_{d-1}, u_{d-2}}) \right)^{-1} \lim_{d \rightarrow \infty} (C_{u_{d-1}, z_{d-2}}) \right) \\ - \lim_{d \rightarrow \infty} (C_{z_{d-1}, z_{d-2}}) \end{array} \right]^{-1} \quad 4.108$$

$$\Delta \mathbf{B} = \left[\begin{array}{c} \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, z_{d-2}}}) \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{z_{d-1, z_{d-2}}}) \right)^{-1} \lim_{d \rightarrow \infty} (\mathbf{C}_{z_{d-1, u_{d-2}}}) \right. \\ \times \left. \left(\lim_{d \rightarrow \infty} (C_{u_{d-1, u_{d-2}}}) \right)^{-1} - \lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, u_{d-2}}}) \left(\lim_{d \rightarrow \infty} (C_{u_{d-1, u_{d-2}}}) \right)^{-1} \right)^{-1} \\ \times \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{u_{d-1, z_{d-2}}}) \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{z_{d-1, z_{d-2}}}) \right)^{-1} \lim_{d \rightarrow \infty} (\mathbf{C}_{z_{d-1, u_{d-2}}}) \right) \\ \times \left(\lim_{d \rightarrow \infty} (C_{u_{d-1, u_{d-2}}}) \right)^{-1} - 1 \end{array} \right] \quad 4.109$$

Note that $C_{u_{d-1, u_{d-2}}}$ must not be zero.

This method shares the same advantages and disadvantages as the second method, as it needs only one segment to reconstruct the modeling errors while being more sensitive to the noise.

This chapter has explored the information contained in the chattering signal for a system that has a measurement matrix with full rank. It presents a method that can be used for extracting modeling uncertainties explicitly. In the next chapter, the information contained in the a priori chattering is investigated for a system with a measurement matrix that is not full rank.

4.3.4 Application of the SVSF to an electro-hydrostatic actuator that has a full-rank measurement matrix

The Electro-Hydrostatic Actuator (EHA) is a “pump controlled hydraulic system” that is used in the aerospace industry; i.e. airplanes aileron, (Wang S. , 2007). The EHA is an integrated unit that consists of an electrical motor, bi-directional pump, pressure and position sensors, and a linear actuator. Its hydraulic circuit is shown in Fig 4.8, [(Habibi & Burton, 2007) and (Habibi, Burton, & Sampson, 2006)]. In this thesis, the EHA developed in (Habibi & Burton, 2007) as shown in Fig 4.9 is used as a benchmark to test the performance and the stability of the SVSF with a time varying boundary layer, and to test the algorithms proposed for extraction of the information contained in the chattering signal.

The proposed EHA has been modeled and verified experimentally in (Habibi, Burton, & Sampson, 2006) and is characterized as a piece-wise linear system as follows:

$$G(s) = \frac{b'}{s(s^2 + 2\xi\omega_n s + \omega_n^2)} \quad 4.110$$

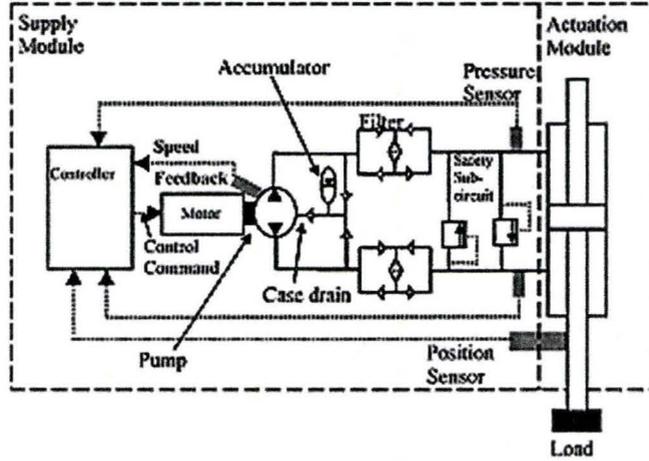


Fig 4.8: The components of the EHA proposed in (Habibi & Burton, 2007).

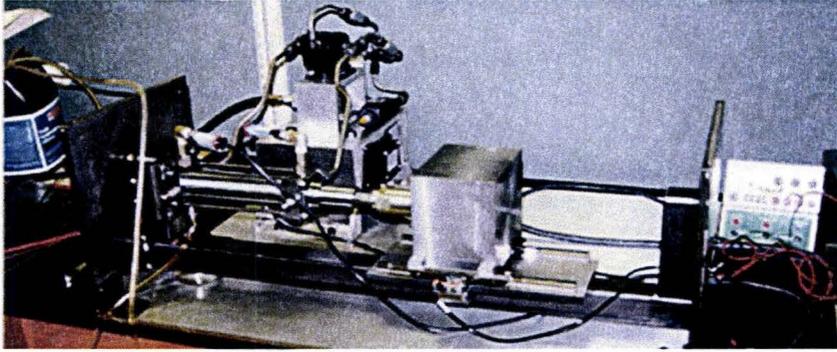


Fig 4.9: The prototype of the EHA proposed in (Habibi & Burton, 2007).

Where $b' = \frac{2D_p\beta_e A_e}{MV_0}$, $\omega_n = \sqrt{\frac{2\beta_e A_e^2}{MV_0}}$, $\xi = \frac{1(B_e V_0 + L\beta_e M)}{2\sqrt{2MV_0\beta_e A_e^2}}$ and their parameters are

defined in table 4.3. This table shows that the effective bulk modulus (which is hard to be measured) may vary with time. (Habibi & Burton, 2007) and (Habibi, 2007) used the VSF and the SVSF to estimate the effective bulk modulus of the proposed EHA, respectively. The EHA model of equation (4.111) can be represented in a discretized state space form (including the system noise) as follows:

$$\begin{bmatrix} x_{1k+1} \\ x_{2k+1} \\ x_{3k+1} \end{bmatrix} = \begin{bmatrix} 1 & T_s & 0 \\ 0 & 1 & T_s \\ 0 & -T_s\omega_n^2 & 1 - 2T_s\xi\omega_n \end{bmatrix} \begin{bmatrix} x_{1k} \\ x_{2k} \\ x_{3k} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ b'T_s \end{bmatrix} u_k + \begin{bmatrix} w_{1k} \\ w_{2k} \\ w_{3k} \end{bmatrix} \quad 4.111$$

The measurement equation (including measurement noise) is represented as follows:

$$\begin{bmatrix} z_{1k+1} \\ z_{2k+1} \\ z_{3k+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1k+1} \\ x_{2k+1} \\ x_{3k+1} \end{bmatrix} + \begin{bmatrix} v_{1k+1} \\ v_{2k+1} \\ v_{3k+1} \end{bmatrix} \quad 4.112$$

Where x_{1k} , x_{2k} and x_{3k} represent the piston's position, velocity and acceleration, respectively.

Symbol	Description	Value
A_e	Piston Area	$5.05 \times 10^{-4} \text{ m}^2$
B_e	Load Friction	760 Ns/m
D_p	Pump Displacement	$1.69 \times 10^{-7} \text{ m}^3/\text{rad}$
L	Leakage Coefficient	$2.5 \times 10^{-11} \text{ m}^3/(\text{Pa} \cdot \text{s})$
M	Load Mass	20 Kg
V_0	Chamber Volume	$6.85 \times 10^{-5} \text{ m}^3$
β_e	Effective Bulk Modulus	$1.5 \times 10^8 - 2 \times 10^8 \text{ Pa}$

Table 4.3: The parameters of the EHA proposed in (Habibi & Burton, 2007).

In this section, the SVSF is used to:

- Estimate the states of the EHA using the SVSF with a time varying smoothing boundary layer as described in section (4.3.2.1).
- Estimate the states and parameters of the EHA using the information extracted from the chattering signal as discussed in section (4.3.3).

4.3.4.1. The SVSF with time-varying smoothing boundary layer

4.3.4.1.1. Simulation setup

The SVSF with the time-varying smoothing boundary layer is applied to the EHA described in equations (4.111) and (4.112). It is assumed that the effective bulk modulus has a value of $1.82 \times 10^8 \text{ Pa}$, and that the other parameters have their original values as listed in table 4.3. The sampling time is $T_s = 0.001 \text{ sec}$. The system can be represented as follows:

$$\begin{bmatrix} x_{1k+1} \\ x_{2k+1} \\ x_{3k+1} \end{bmatrix} = \begin{bmatrix} 1 & T_s & 0 \\ 0 & 1 & T_s \\ 0 & -67.7585 & 0.8956 \end{bmatrix} \begin{bmatrix} x_{1k} \\ x_{2k} \\ x_{3k} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0.0227 \end{bmatrix} u_k + \begin{bmatrix} w_{1k} \\ w_{2k} \\ w_{3k} \end{bmatrix} \quad 4.113$$

The measurement and system noise are white with a noise-to-signal ratio of 10% of their corresponding states. Equal weight is placed on both the system and the measurement

noise with the following variances, $\sigma_w^2 = \sigma_v^2 = \begin{bmatrix} 8.2 \times 10^{-12} & 0 & 0 \\ 0 & 2.2 \times 10^{-10} & 0 \\ 0 & 0 & 3 \times 10^{-6} \end{bmatrix}$.

Further to section (4.3.2), two SVSFs are used, both using a coefficient matrix $\boldsymbol{\gamma}$ with a value of $\boldsymbol{\gamma} = 0.02 \times \mathbf{I}_{3 \times 3}$. The first SVSF has a smoothing boundary layer with zero width, while the second SVSF uses a smooth boundary layer that is designed to have a time varying width described in equation (4.75). The states' initial condition vector is set to $\hat{\mathbf{x}}_0^{\text{SVSF}} = \mathbf{z}_0$ for both filters. The input consists of a random signal superimposed on step changes as shown in Fig 4.10. For comparison purposes, the simulation is repeated using the KF. The initial covariance matrix is set to be $\mathbf{P}_0 = \mathbf{R}$, the states' initial condition vector is chosen to be $\hat{\mathbf{x}}_0^{\text{KF}} = \mathbf{z}_0$, and the system and measurement noise covariance matrices are defined as $\mathbf{Q} = \mathbf{R} = \sigma_w^2$.

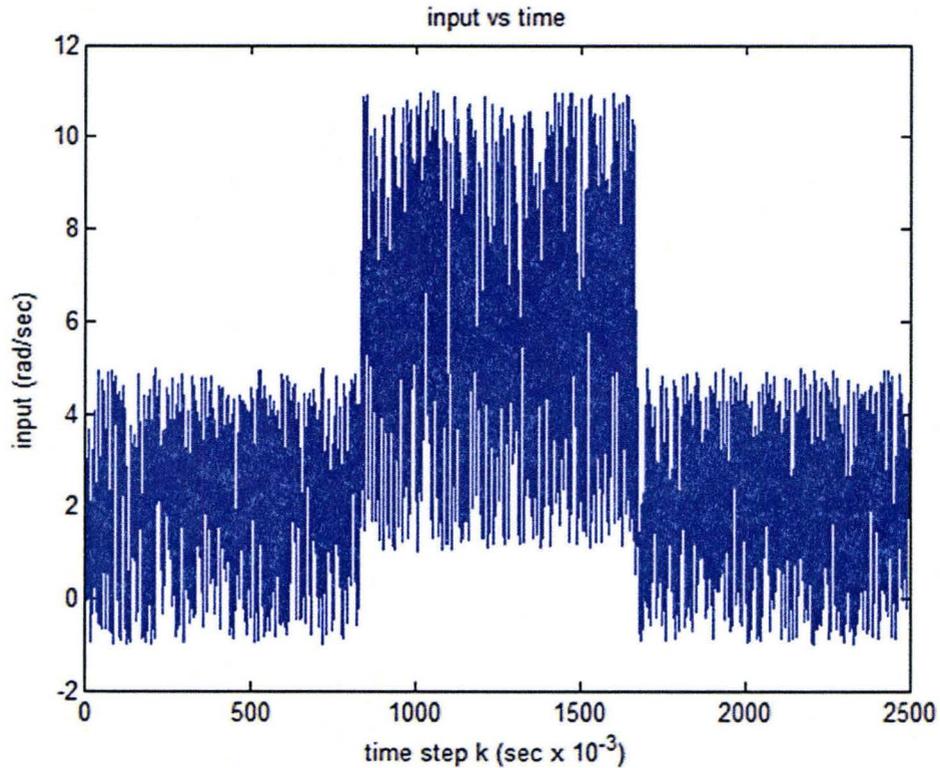


Fig 4.10: The input signal.

4.3.4.1.2. Simulation results

The results of the SVSF with the time-varying smoothing boundary layer compared to the KF are shown in Figs (4.11) to (4.13). The two methods, the SVSF and the KF, are compared in terms of the root mean square error ($RMSE_j$), which is defined as follows:

$$RMSE_j = \frac{\sqrt{\sum_{i=1}^{length(x)} (y_{j_i} - \hat{y}_{j_i})^2}}{length(x)} \text{ for } y = x_1, x_2, \text{ and } x_3 \quad 4.114$$

and the maximum absolute error ($MaxError_j$), which is equal to:

$$MaxError_j = \max(|y_{j_i} - \hat{y}_{j_i}|) \text{ for } y = x_1, x_2 \text{ and } x_3, \quad 4.115$$

The comparison is summarized in table 4.4. The results show that the proposed time varying smoothing boundary layer width is a successful choice as the SVSF has a performance that is quite similar to the KF's performance. The table shows that the KF is slightly better than the SVSF. Conversely, the SVSF has a superior performance when modeling errors are present, as discussed in section (4.3.2).

	$RMSE_j$	$MaxError_j$
x_1^{KF}	4.2×10^{-12}	8.8×10^{-06}
x_2^{KF}	1.2×10^{-10}	3.6×10^{-05}
x_3^{KF}	1.4×10^{-06}	4.3×10^{-03}
x_1^{SVSF}	6.2×10^{-12}	9.5×10^{-06}
x_2^{SVSF}	1.7×10^{-10}	4.7×10^{-05}
x_3^{SVSF}	1.8×10^{-06}	4.9×10^{-03}

Table 4.4: The root mean square error and the maximum absolute error of the estimated states.

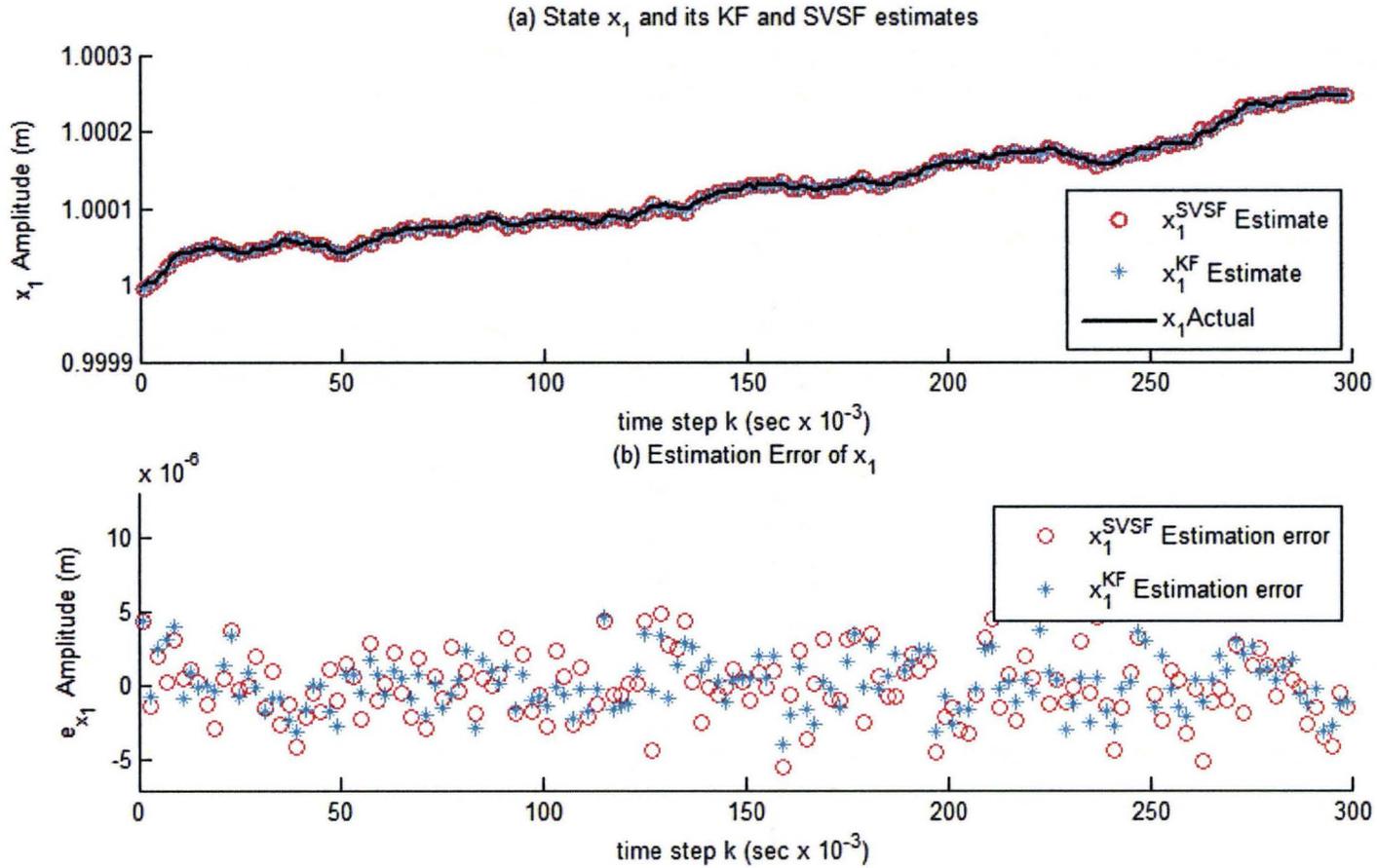


Fig 4.11: The first 300 points of (a) x_1 's actual and estimated values, (b) estimation error of x_1 , for the SVSF with time-varying smoothing boundary layer.

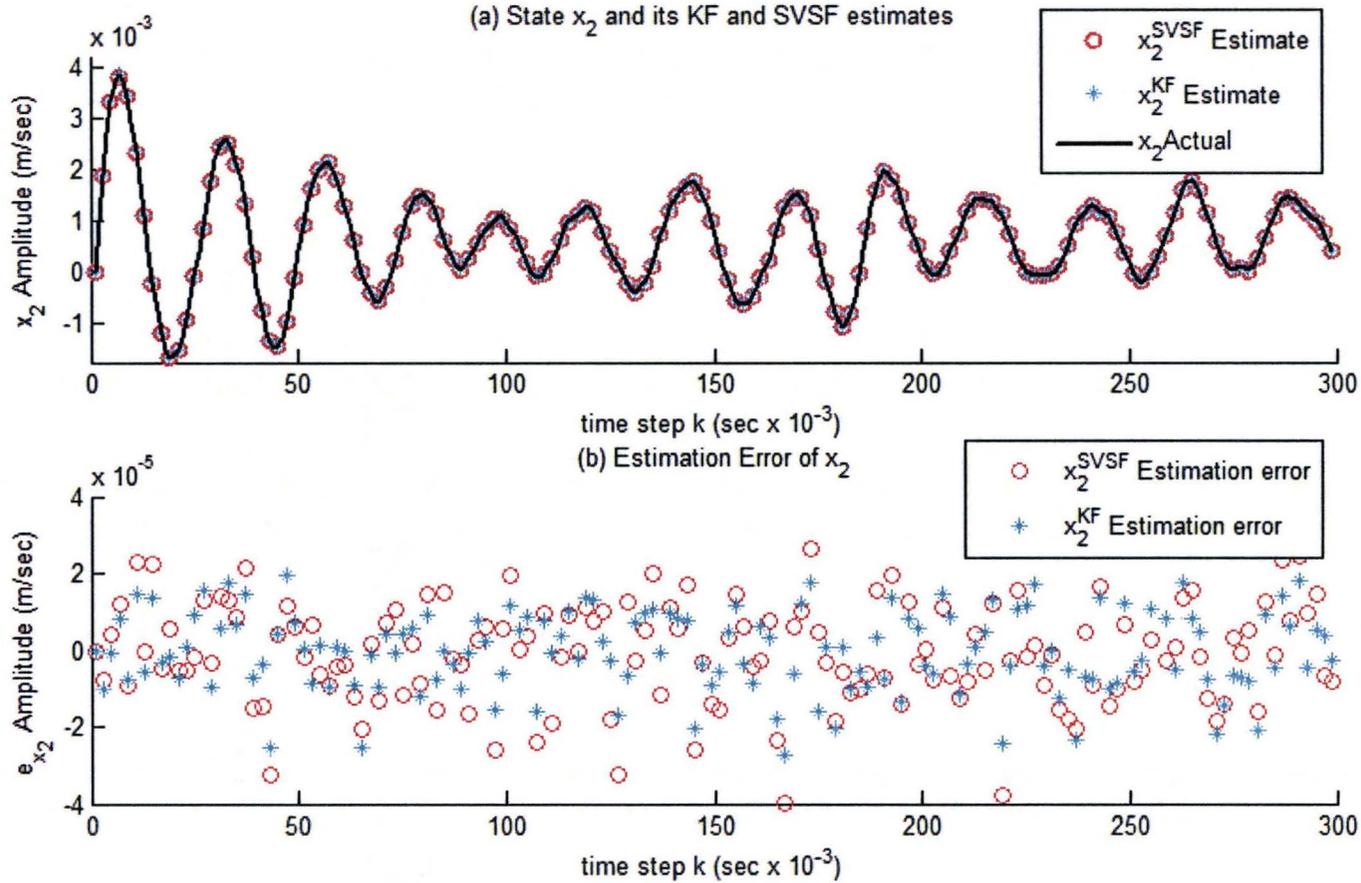


Fig 4.12: The first 300 points of (a) x_2 's actual and estimated values, (b) estimation error of x_2 , for the SVSF with time-varying smoothing boundary layer.

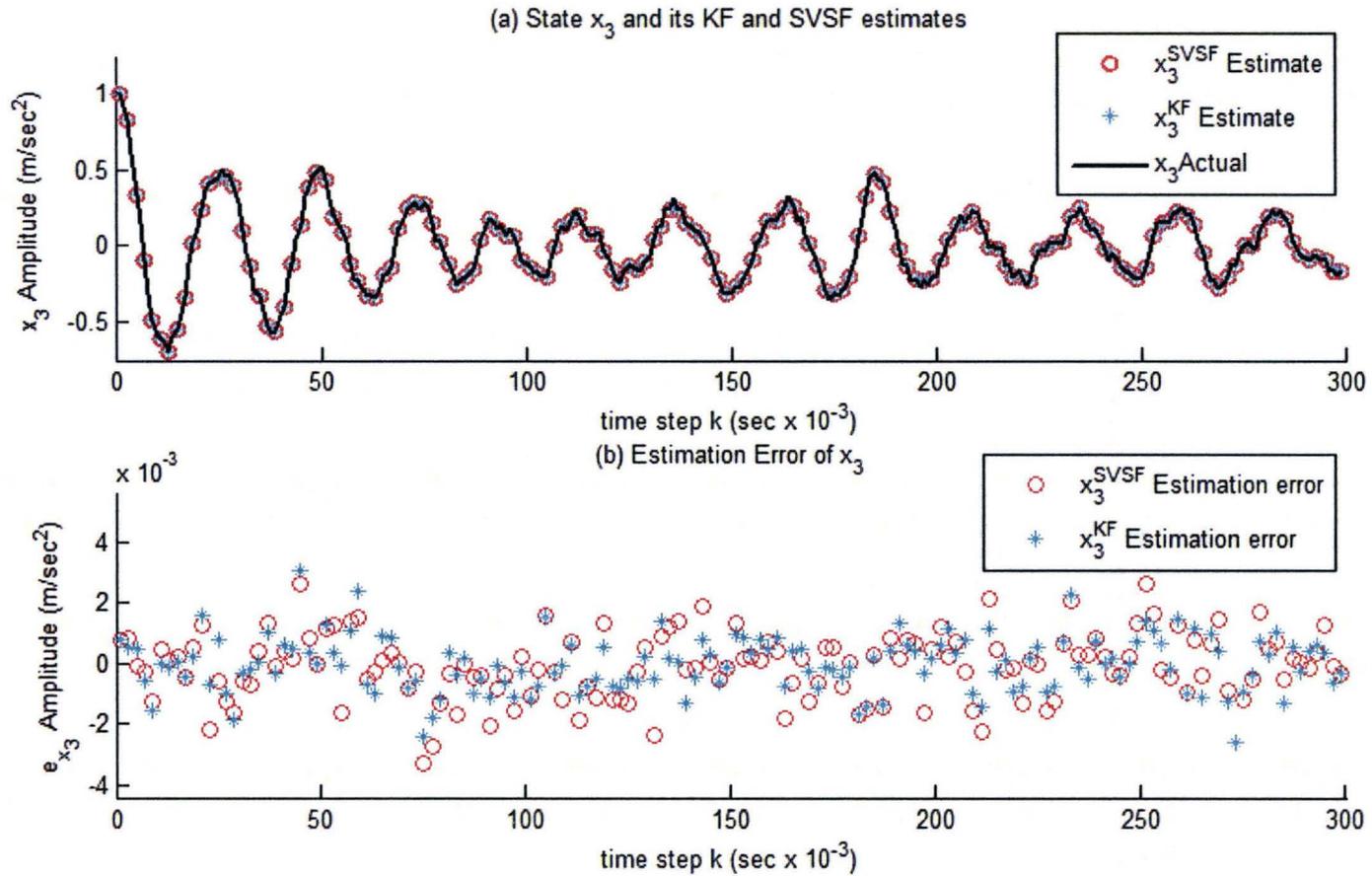


Fig 4.13: The first 300 points of (a) x_3 's actual and estimated values, (b) estimation error of x_3 , for the SVSF with time-varying smoothing boundary layer.

4.3.4.2. The SVSF with tuning process based on information contents in chattering

4.3.4.2.1. Simulation setup

The SVSF's chattering signal can be used to tune the filter model in order to reduce the modeling errors. This has been tested and verified by applying the SVSF to an EHA with an effective bulk modulus that is made to change its value from time to time where each change holds for a period of time. The other EHA parameters are assumed to have their original values from table 4.3. If the sampling time is $T_s = 0.001 \text{ sec}$ then the system has the following structure:

$$\mathbf{A} = \begin{bmatrix} 1 & 0.001 & 0 \\ 0 & 1 & 0.001 \\ 0 & -3.723\beta_e \times 10^{-7} & 1 - (0.038 + 3.65\beta_e \times 10^{-7}) \end{bmatrix} \begin{bmatrix} x_{1k} \\ x_{2k} \\ x_{3k} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1.25\beta_e \times 10^{-10} \end{bmatrix} u_k + \begin{bmatrix} w_{1k} \\ w_{2k} \\ w_{3k} \end{bmatrix} \quad 4.116$$

The effective bulk modulus, β_e , is made to change nine times and each change to last for at least 3000 time steps as shown in table 4.5. During the simulation time, the filter's model suffers from modeling errors that vary between 7 – 340% as shown in table 4.6.

Region #	Beginning time step	Ending time step	Region #	Beginning time step	Ending time step
1	1	4656	6	25329	30482
2	4657	9044	7	30483	34317
3	9045	14438	8	34318	39786
4	14439	19600	9	39787	43999
5	19601	25328			

Table 4.5: The beginning and ending of each region for the tuning process simulation.

The measurement and system noise are white noise with a noise-to-signal ratio of 10% of their corresponding states, and with the following variances, $\sigma_w^2 = \sigma_v^2 = \begin{bmatrix} 8.2 \times 10^{-12} & 0 & 0 \\ 0 & 2.2 \times 10^{-10} & 0 \\ 0 & 0 & 3 \times 10^{-6} \end{bmatrix}$. The SVSF's coefficient matrix $\boldsymbol{\gamma}$ has a value of

$\boldsymbol{\gamma} = 0.02 \times \mathbf{I}_{3 \times 3}$ and its smooth boundary layer is designed as $\boldsymbol{\Psi} = \begin{bmatrix} 2 \times 10^{-5} \\ 9 \times 10^{-5} \\ 1 \times 10^{-2} \end{bmatrix}$. The input

consists of a random signal superimposed on step changes as shown in Fig 4.14.

	region 1	region 2	region 3	region 4	region 5	region 6	region 7	region 8	region 9
	$ \text{error} /\%$								
β_e	340	17	9	100	54	8	100	54	7
a_{32}	340	17	9	100	54	8	100	54	7
a_{33}	25	3	1	5	6	1	5	6	1
b_3	340	17	9	100	54	8	100	54	7

Table 4.6: The % of the error between the system and the filter parameters for the tuning process simulation.

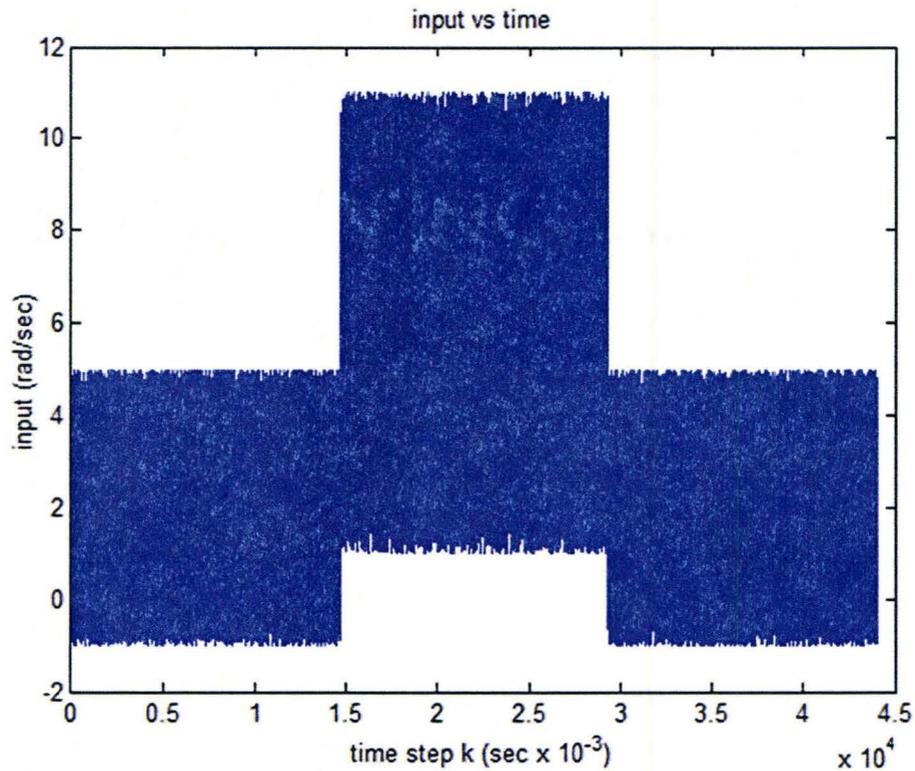


Fig 4.14: The input vs. time for the tuning process simulation.

4.3.4.2.2. Simulation results

The results of the SVSF that use the information contained in the chattering signal to tune the filter model are shown in Figs (4.15) to (4.20). These results show that the SVSF is a robust and stable filter that detects parametric changes once they occur (note that the latter property cannot be found in the KF or its other forms). The filter then uses a sample consisting of 500 points from the chattering signal to extract the modeling errors as discussed in section (4.3.3) (sample length satisfies the law of large numbers, and the noise vectors of the sample have errors in means and variances that are less than 5% compared to their expected values). Once the filter's model is tuned, the filter continues estimating the states. During the model reconstructing process, the results show that the stability is not affected by modeling errors.

The detection process in this section is based on the occurrence of chattering. If the width of the smoothing boundary layer is well-estimated, then chattering is removed. However, if modeling errors are present, then the amplitude of the output's a priori estimation error becomes larger than the width of the smoothing boundary layer and chattering is observed. Therefore, chattering can be used to detect modeling errors when they are present (as discussed in section (4.3.2)). The detection process is sensitive to the width of the smoothing boundary layer. If the smoothing boundary layer has a very small width (compared to the noise uncertainties and the system matrix), then many false alarms will be detected; e.g. false alarm detected in Figs (4.19) and (4.20) at time step 7489. A false alarm causes extra computation (as the tuning process is applied when it is not needed), and may increase the root mean square errors (as the variance of the estimated parameter increases and the reconstruction's sample at the false position may have parametric changes within it). As the width of the smoothing boundary layer is increased, the number of false alarms is reduced. However, this may cause a delay in detecting modeling errors. Fig 4.18 shows that it takes seven time steps to detect modeling errors in the third time interval that starts at time step 9,045. This causes the errors in these parameters to look like spikes at the beginning of the intervals. The delay factor depends on the differences between the smoothing boundary layer width and the a priori existence subspace width (which depends on the system matrix and the noise vectors, as shown in equation (4.42)). The SVSF detects parametric changes at time steps as summarized in table 4.7, which demonstrates the speed of the detection process. Table 4.8 summarizes the root mean square errors (defined in equation (4.117)) for the estimated states and parameters. This shows that the modeling errors are recoverable by using the chattering signal.

$$RMSE_j = \frac{\sqrt{\sum_{i=1}^{length(x)} (y_{j_i} - \hat{y}_{j_i})^2}}{length(x)} \text{ for } y = x_1, x_2, x_3, \xi, \omega_n \text{ and } b' \quad 4.117$$

Region #	Time step of fault being detected
1	1
2	4659
3	9052
4	14440
5	19602
6	25330
7	30484
8	34319
9	39790

Table 4.7: The estimated beginning of each region for the tuning process simulation.

	<i>RMSE</i>
x_1	5.3×10^{-12}
x_2	1.43×10^{-10}
x_3	2.14×10^{-06}
$a_{32} = -\omega_n^2 T_s$	0.2809
$a_{33} = 1 - 2\xi\omega_n T_s$	5.1×10^{-07}
$b'_3 = b' T_s$	3.2×10^{-08}

Table 4.8: The root mean square error of the tuning process simulation.

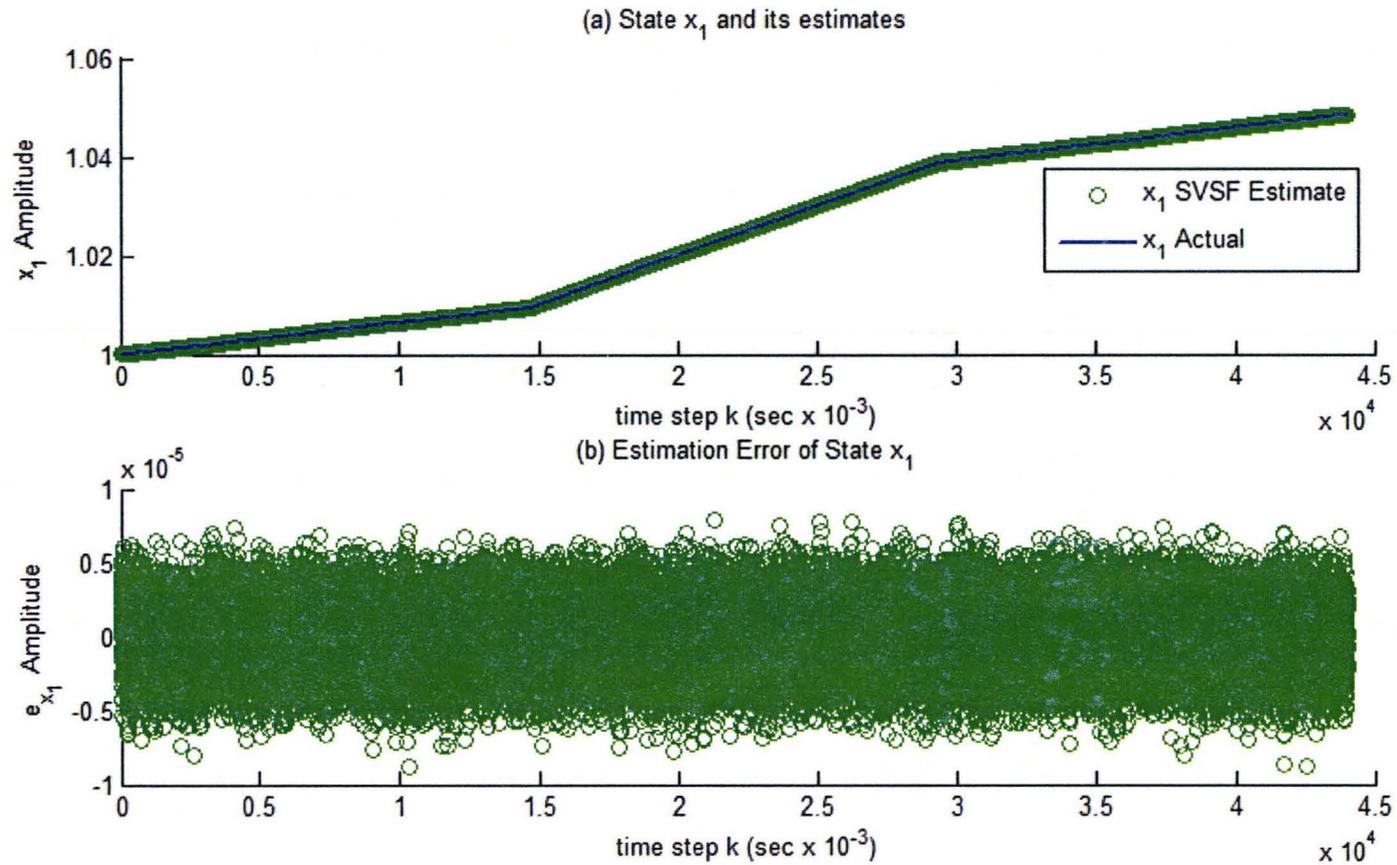


Fig 4.15: (a) x_1 's actual and estimated values, (b) estimation error of x_1 for the simulation of the tuning process obtained by using the information contained in the chattering signal.

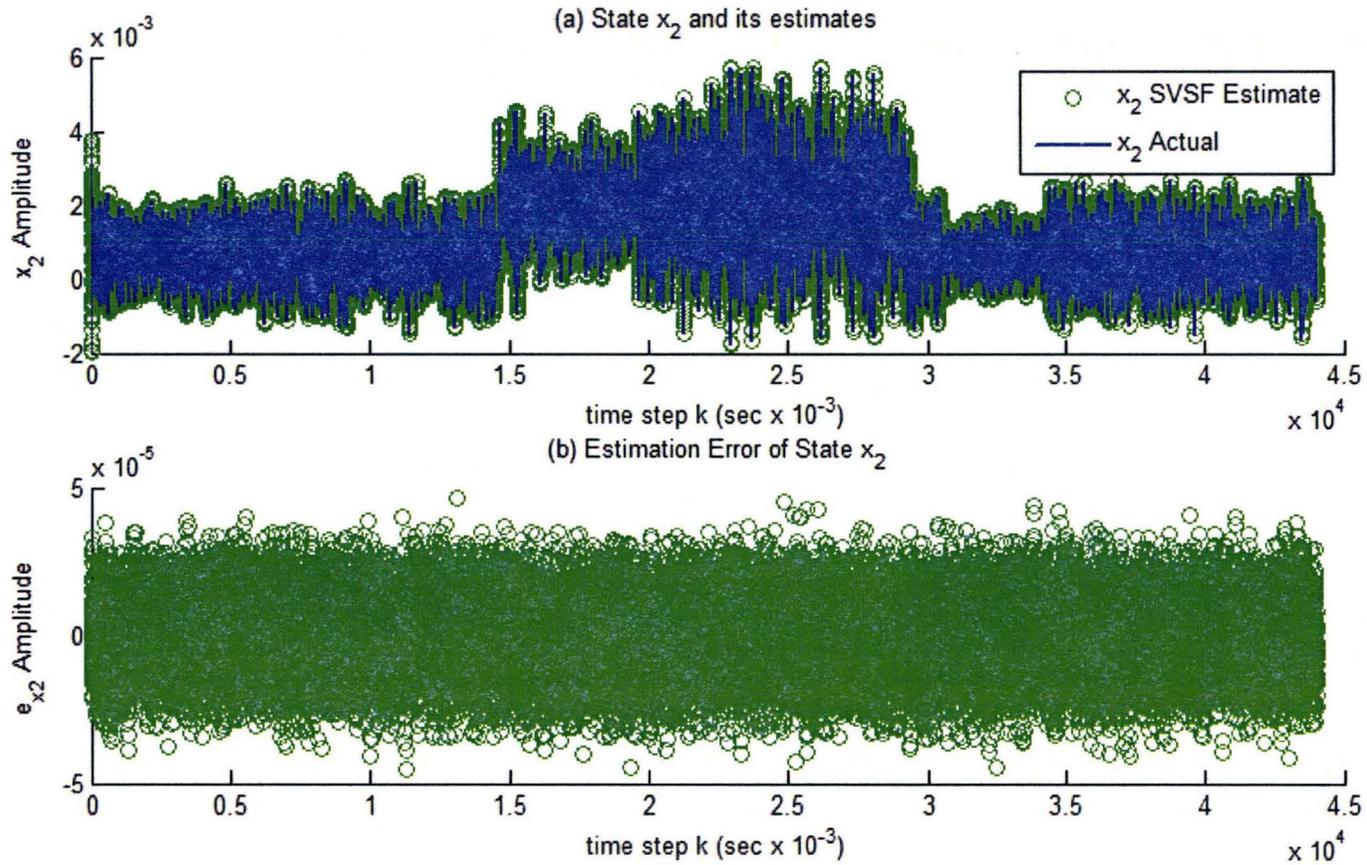


Fig 4.16: (a) x_2 's actual and estimated values, (b) estimation error of x_2 for the simulation of the tuning process obtained by using the information contained in the chattering signal.

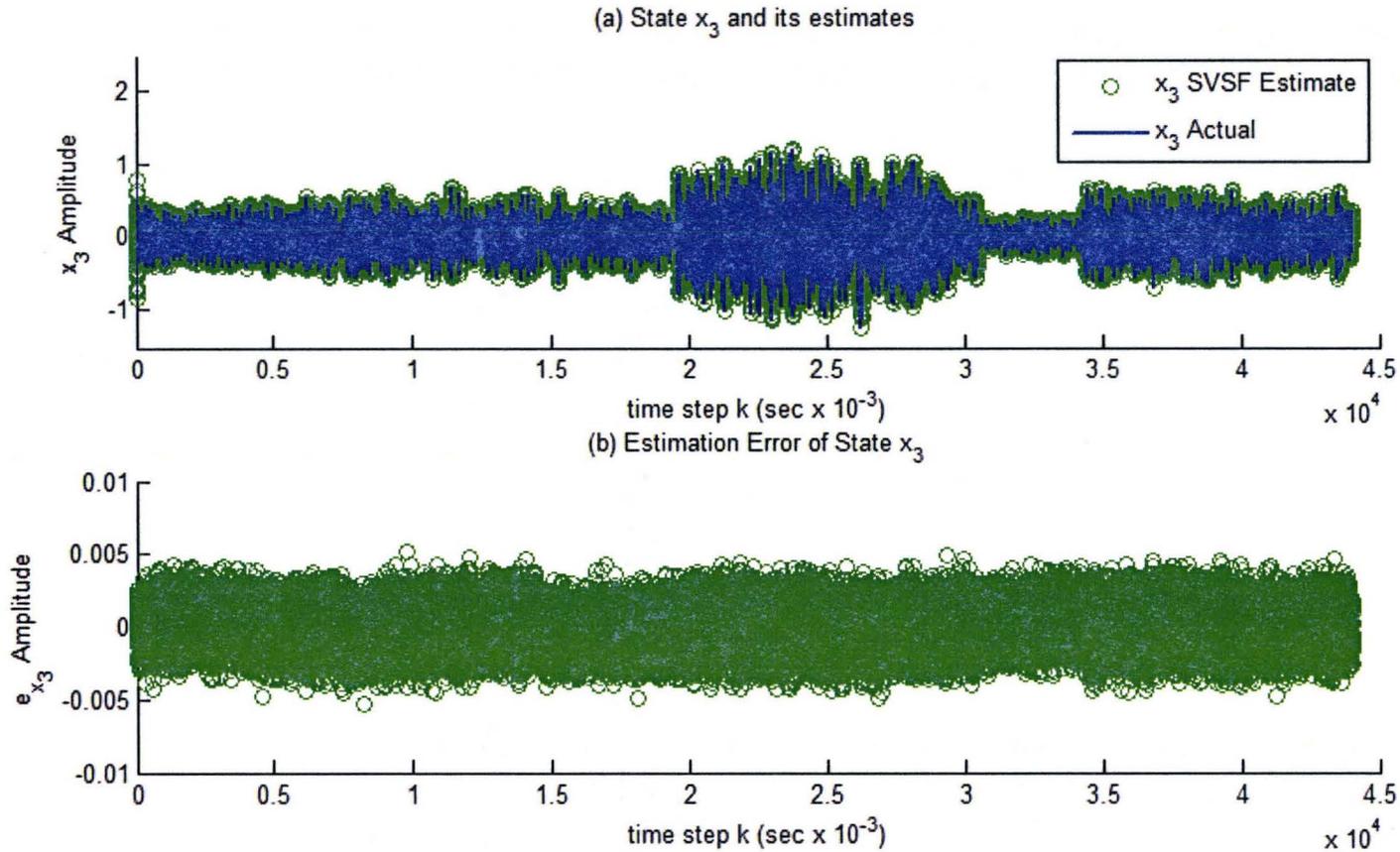


Fig 4.17: (a) x_3 's actual and estimated values, (b) estimation error of x_3 for the simulation of the tuning process obtained by using the information contained in the chattering signal.

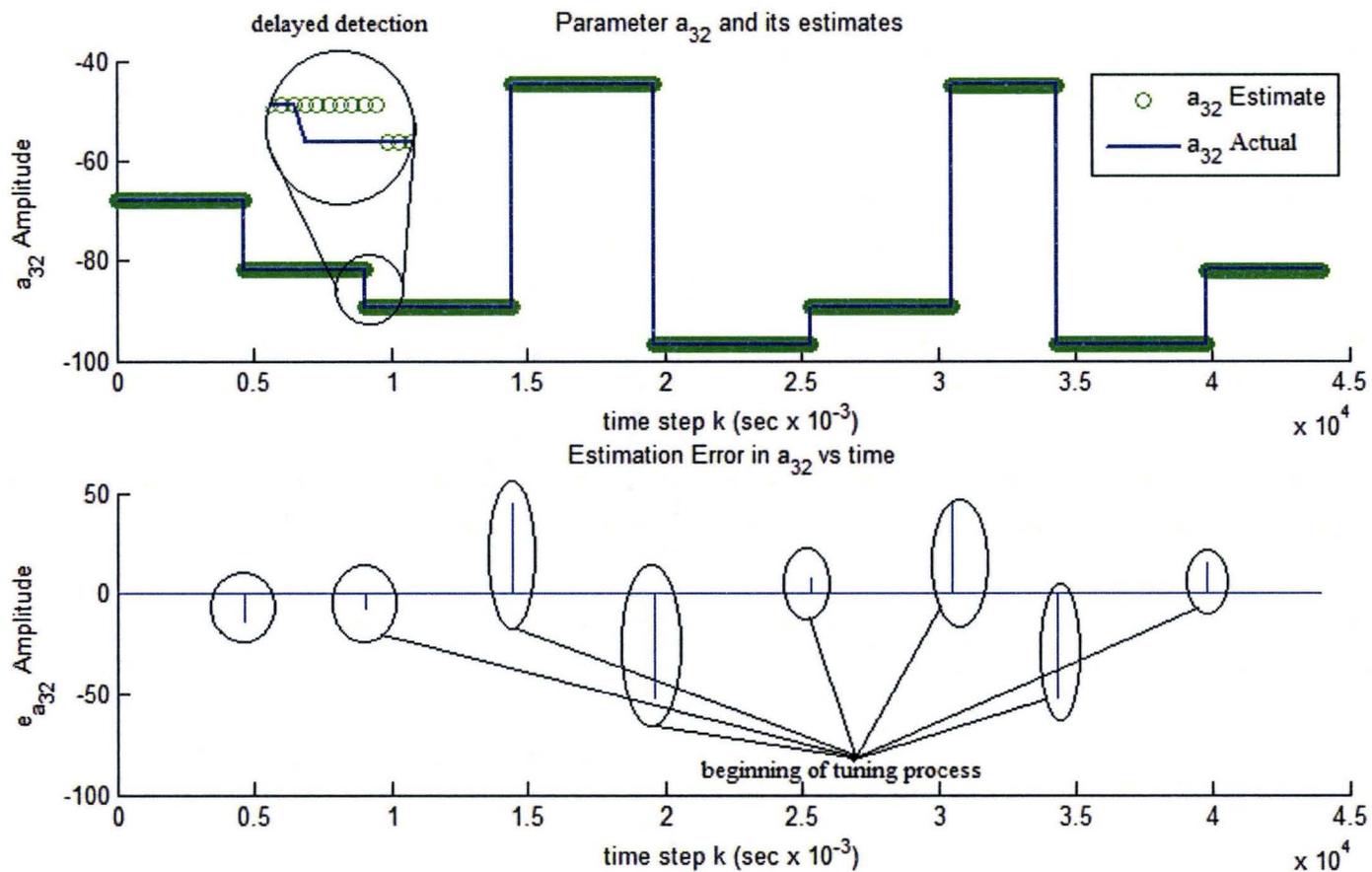


Fig 4.18: (a) a_{32} 's actual and estimated values, (b) estimation error of a_{32} for the simulation of the tuning process obtained by using the information contained in the chattering signal.

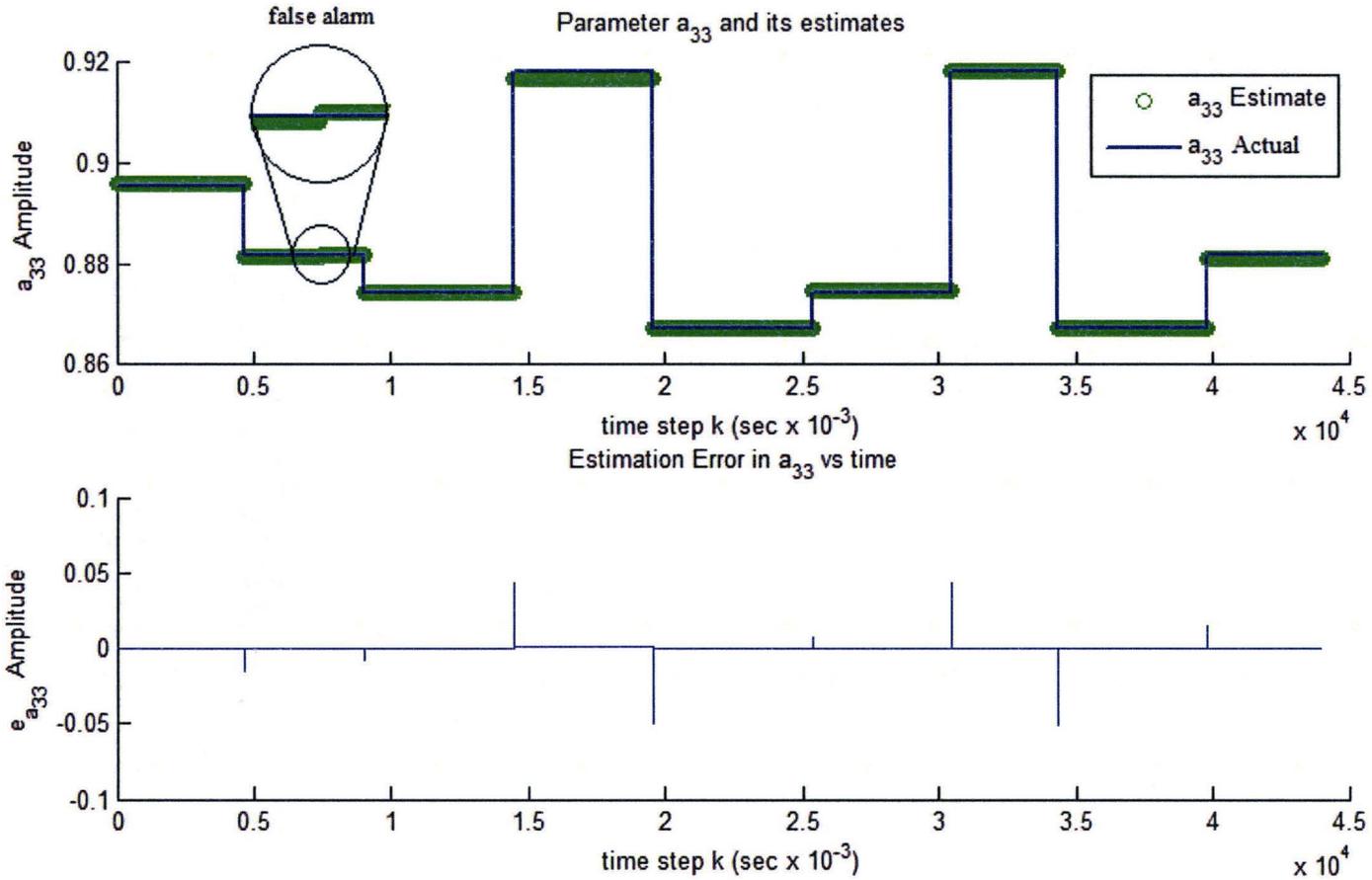


Fig 4.19: (a) a_{33} 's actual and estimated values, (b) estimation error of a_{33} for the simulation of the tuning process obtained by using the information contained in the chattering signal.

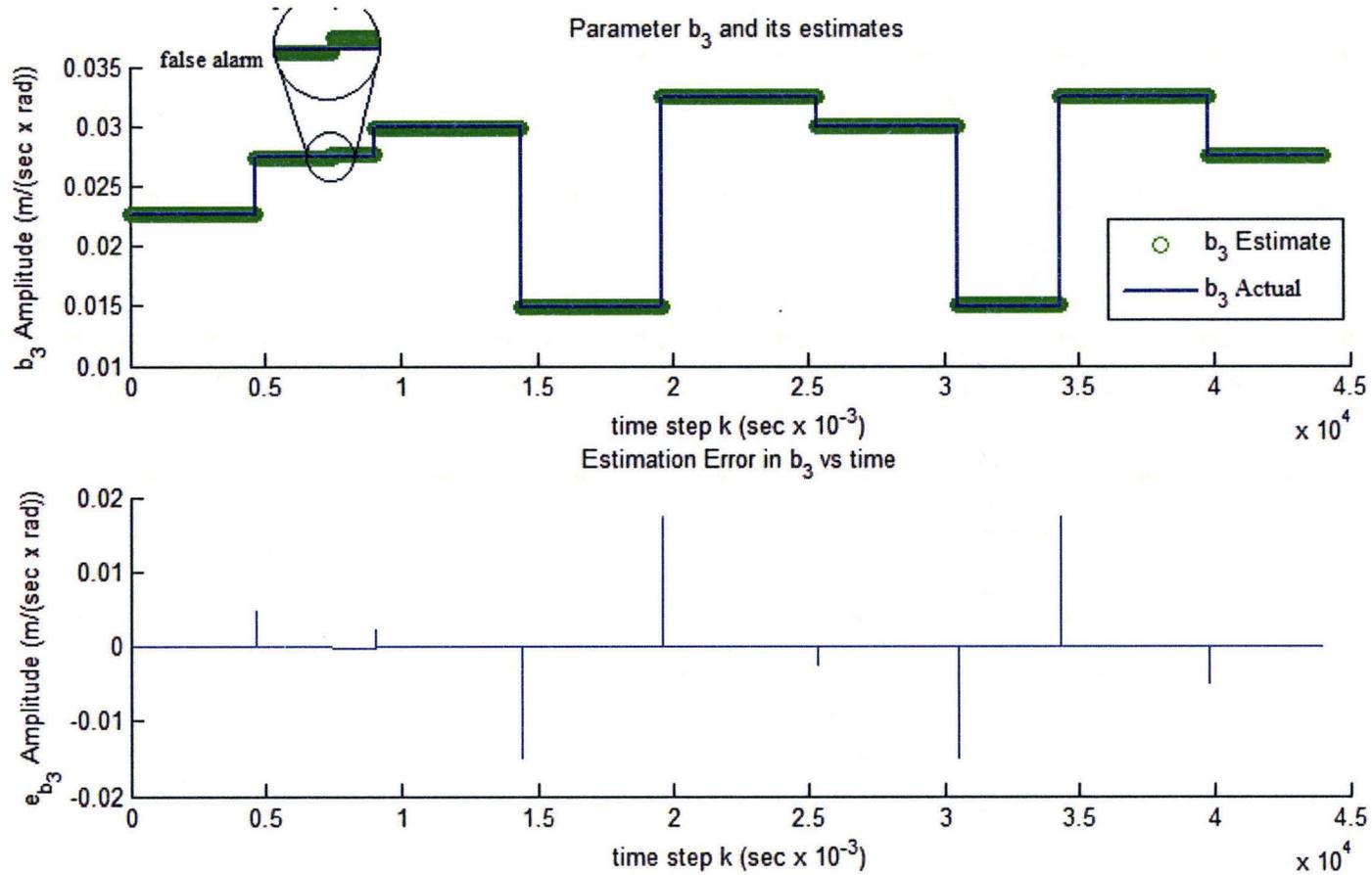


Fig 4.20: (a) b_3 's actual and estimated values, (b) estimation error of b_3 for the simulation of the tuning process obtained by using the information contained in the chattering signal.

Chapter Five:

The Toeplitz/Observability Smooth Variable Structure Filters for Systems with Measurement Matrix that is not Full Rank

5.1 Historical and mathematical background

Chapter four discussed the SVSF concepts and behaviour when it is applied to a system that has a measurement matrix of full-rank. However in most real applications, the measurement matrix does not have a full rank. Moreover, not all of the states are accessible and physically measurable, (Habibi, 2007). If the measurement matrix does not have a full rank, the SVSF is then combined with a reduced order technique that is similar to Luenberger's approach. The reduced order technique extracts the hidden states from the available measurements, and then uses them to compensate for the missing measurements. Thus, Luenberger's approach allows the application of the SVSF using the same algorithms that were described in chapter four. In order to apply Luenberger's reduced order form to the SVSF, the system must be observable and differentiable. The reduced order algorithm differs in approach for linear systems compared to nonlinear systems and it is discussed in the following subsections.

5.1.1 The SVSF algorithm for linear systems with partially ranked measurement matrices

The reduced order algorithm that is discussed in this subsection is based on (Habibi, 2007), and it is applicable for linear systems with the following time invariant measurement matrix:

$$\mathbf{H}_k = \mathbf{H} = [\mathbf{I}_{m \times m} \quad \mathbf{0}_{m \times (n-m)}] \quad 5.1$$

If the measurement (output) matrix does not have the structure of equation (5.1), a coordinate transformation, \mathbf{T} , is applied on the states to convert it to equation (5.1). The resultant transformed time-invariant system is defined as follow:

$$\begin{aligned} \mathbf{y}_k &= \mathbf{A}'\mathbf{y}_{k-1} + \mathbf{B}'\mathbf{u}_{k-1} + \mathbf{w}'_{k-1} \\ \mathbf{z}_k &= \mathbf{H}'\mathbf{y}_k + \mathbf{v}_k \end{aligned} \quad 5.2$$

Where $\mathbf{y}_k = \mathbf{T}^{-1}\mathbf{x}_k = \begin{bmatrix} \mathbf{y}_{1k} \\ \mathbf{y}_{2k} \end{bmatrix}$, $\mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, $\mathbf{B}' = \mathbf{T}^{-1}\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}$, $\mathbf{H}' = \mathbf{H}\mathbf{T} = [\mathbf{I}_{m \times m} \quad \mathbf{0}_{m \times (n-m)}]$, $\mathbf{A}_{11} \in \mathbb{R}^{m \times m}$, $\mathbf{A}_{21} \in \mathbb{R}^{(n-m) \times m}$, $\mathbf{A}_{12} \in \mathbb{R}^{m \times (n-m)}$, $\mathbf{A}_{22} \in \mathbb{R}^{(n-m) \times (n-m)}$, $\mathbf{B}_1 \in \mathbb{R}^{m \times 1}$, $\mathbf{B}_2 \in \mathbb{R}^{(n-m) \times 1}$, $\mathbf{y}_{1k} \in \mathbb{R}^{m \times 1}$, $\mathbf{y}_{2k} \in \mathbb{R}^{(n-m) \times 1}$, $\mathbf{w}'_k = \mathbf{T}^{-1}\mathbf{w}_k$, and $\mathbf{z}_k = \mathbf{y}_{1k} + \mathbf{v}_k$.

For the system that is described by equation (5.2), the SVSF's structure is as follows:

1 - Prediction Stage:

The a priori state estimate is obtained by using an estimated model of the system as follows:

$$\begin{aligned}\hat{\mathbf{y}}_{k|k-1} &= \hat{\mathbf{A}}' \mathbf{y}_{k-1|k-1} + \hat{\mathbf{B}}' \mathbf{u}_{k-1} \\ \hat{\mathbf{z}}_{k|k-1} &= \hat{\mathbf{H}}' \hat{\mathbf{y}}_{k|k-1}\end{aligned}\quad 5.3$$

2 - Corrective Stage:

A corrective gain is calculated and used for refining the a priori estimate into its a posteriori form as follows:

$$\begin{aligned}\hat{\mathbf{y}}_{k|k} &= \hat{\mathbf{y}}_{k|k-1} + \mathbf{K}_{SVSF-II_k} \\ \hat{\mathbf{z}}_{k|k} &= \hat{\mathbf{H}}' \hat{\mathbf{y}}_{k|k}\end{aligned}\quad 5.4$$

where $\mathbf{K}_{SVSF-II_k}$ is the SVSF's gain for a reduced order linear system and is defined as follows:

$$\mathbf{K}_{SVSF-II_k} = \begin{bmatrix} \mathbf{K}_{1_k} \\ \mathbf{K}_{2_k} \end{bmatrix}\quad 5.5$$

The upper partition of the gain, \mathbf{K}_{1_k} , is the SVSF's gain for the states that are directly linked to measurements and is defined as follows:

$$\begin{aligned}\mathbf{K}_{1_k} &= \hat{\mathbf{H}}'^+ \left(\left| \mathbf{e}_{z_k|k-1} \right| + \gamma_1 \left| \mathbf{e}_{z_{k-1}|k-1} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{z_k|k-1} \right) \\ &= \left(\left| \mathbf{e}_{z_k|k-1} \right| + \gamma_1 \left| \mathbf{e}_{z_{k-1}|k-1} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{z_k|k-1} \right)\end{aligned}\quad 5.6$$

The lower partition \mathbf{K}_{2_k} is the SVSF's gain for the lower partition of the states that are not directly related to the measurements. This gain is obtained using the Luenberger's method as follows:

Further to the Luenberger method, the system described by equation (5.2) is rewritten by using a revised state vector, \mathbf{y}'_k , as follows:

$$\begin{aligned}\mathbf{y}'_k &= \mathbf{A}' \mathbf{y}'_{k-1} + \mathbf{B}' \mathbf{u}_{k-1} + \mathbf{w}'_{k-1} - \mathbf{A}' \begin{bmatrix} \mathbf{v}_{k-1} \\ \mathbf{0}_{(n-m) \times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_k \\ \mathbf{0}_{(n-m) \times 1} \end{bmatrix} \\ \begin{bmatrix} \mathbf{z}_k \\ \mathbf{y}_{2_k} \end{bmatrix} &= \mathbf{A}' \begin{bmatrix} \mathbf{z}_{k-1} \\ \mathbf{y}_{2_{k-1}} \end{bmatrix} + \mathbf{B}' \mathbf{u}_{k-1} + \mathbf{w}'_{k-1} - \mathbf{A}' \begin{bmatrix} \mathbf{v}_{k-1} \\ \mathbf{0}_{(n-m) \times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_k \\ \mathbf{0}_{(n-m) \times 1} \end{bmatrix}\end{aligned}\quad 5.7$$

Both the a priori and the a posteriori estimates and their revised vectors ($\hat{\mathbf{y}}'_{k|k-1}$ and $\hat{\mathbf{y}}'_{k|k}$, respectively) are partitioned as $\hat{\mathbf{y}}_{k|k-1} = \begin{bmatrix} \hat{\mathbf{y}}_{1k|k-1} \\ \hat{\mathbf{y}}_{2k|k-1} \end{bmatrix}$, $\hat{\mathbf{y}}_{k|k} = \begin{bmatrix} \hat{\mathbf{y}}_{1k|k} \\ \hat{\mathbf{y}}_{2k|k} \end{bmatrix}$, $\hat{\mathbf{y}}'_{k|k-1} = \begin{bmatrix} \hat{\mathbf{z}}_{k|k-1} \\ \hat{\mathbf{y}}_{2k|k-1} \end{bmatrix}$ and $\hat{\mathbf{y}}'_{k|k} = \begin{bmatrix} \mathbf{z}_k \\ \hat{\mathbf{y}}_{2k|k} \end{bmatrix}$. Where $\hat{\mathbf{y}}_{1k|k-1}, \hat{\mathbf{z}}_{k|k-1}, \hat{\mathbf{y}}_{1k|k} \in \mathbb{R}^{m \times 1}$ and $\hat{\mathbf{y}}_{2k|k-1}, \hat{\mathbf{y}}_{2k|k} \in \mathbb{R}^{(n-m) \times 1}$. Further to the above discussion and by defining a revised measurement matrix as the identity matrix, equation (5.3) is rewritten as follows:

$$\hat{\mathbf{y}}'_{k|k-1} = \begin{bmatrix} \hat{\mathbf{z}}_{k|k-1} \\ \hat{\mathbf{y}}_{2k|k-1} \end{bmatrix} = \hat{\mathbf{A}}' \begin{bmatrix} \mathbf{z}_{k-1} \\ \hat{\mathbf{y}}_{2k-1|k-1} \end{bmatrix} + \hat{\mathbf{B}}' \mathbf{u}_{k-1} \quad 5.8$$

Subtracting equation (5.8) from (5.7) gives, (Habibi, 2007):

$$\begin{aligned} \begin{bmatrix} \mathbf{e}_{z_{k|k-1}} \\ \mathbf{e}_{y_{2k|k-1}} \end{bmatrix} &= \begin{bmatrix} \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} \\ \mathbf{y}_{2k} - \hat{\mathbf{y}}_{2k|k-1} \end{bmatrix} = \hat{\mathbf{A}}' \begin{bmatrix} 0 \\ \mathbf{e}_{y_{2k-1|k-1}} \end{bmatrix} + \mathbf{d}_{k-1} \\ &= \begin{bmatrix} \hat{\mathbf{A}}_{11} & \hat{\mathbf{A}}_{12} \\ \hat{\mathbf{A}}_{21} & \hat{\mathbf{A}}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{0}_{m \times 1} \\ \mathbf{e}_{y_{2k-1|k-1}} \end{bmatrix} + \begin{bmatrix} \mathbf{d}_{1k-1} \\ \mathbf{d}_{2k-1} \end{bmatrix} \end{aligned} \quad 5.9$$

Where $\hat{\mathbf{A}}_{11} \in \mathbb{R}^{m \times m}$, $\hat{\mathbf{A}}_{21} \in \mathbb{R}^{(n-m) \times m}$, $\hat{\mathbf{A}}_{12} \in \mathbb{R}^{m \times (n-m)}$, $\hat{\mathbf{A}}_{22} \in \mathbb{R}^{(n-m) \times (n-m)}$, $\mathbf{d}_{1k-1} \in \mathbb{R}^{m \times 1}$, $\mathbf{d}_{2k-1} \in \mathbb{R}^{(n-m) \times 1}$ and \mathbf{d}_{k-1} represents uncertainties and modeling errors and it is defined as follows:

$$\mathbf{d}_{k-1} = \begin{bmatrix} \mathbf{d}_{1k-1} \\ \mathbf{d}_{2k-1} \end{bmatrix} = \Delta \mathbf{A}' \begin{bmatrix} \mathbf{z}_{k-1} \\ \mathbf{y}_{2k-1} \end{bmatrix} + \Delta \mathbf{B}' \mathbf{u}_{k-1} - \mathbf{A}' \begin{bmatrix} \mathbf{v}_{k-1} \\ \mathbf{0}_{(n-m) \times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{w}'_{1k-1} + \mathbf{v}_k \\ \mathbf{w}'_{2k-1} \end{bmatrix} \quad 5.10$$

Further to equation (5.9) and using equation (5.8), the a posteriori and the a priori estimation errors are obtained as follows:

$$\begin{aligned} \mathbf{e}_{y_{2k-1|k-1}} &= \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} - \hat{\mathbf{A}}_{12}^{-1} \mathbf{d}_{1k-1} \\ \mathbf{e}_{y_{2k|k-1}} &= \hat{\mathbf{A}}_{22} \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} - \hat{\mathbf{A}}_{22} \hat{\mathbf{A}}_{12}^{-1} \mathbf{d}_{1k-1} + \mathbf{d}_{2k-1} \end{aligned} \quad 5.11$$

The measured component of equation (5.11) is used as the switching hyperplane for the lower partition, thus the lower gain is defined as follows, (Habibi, 2007):

$$\mathbf{K}_{2k} = \left(\left| \hat{\mathbf{A}}_{22} \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} \right| + \gamma_2 \left| \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} \right| \right) \circ \text{sgn} \left(\hat{\mathbf{A}}_{22} \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} \right) \quad 5.12$$

Substituting equations (5.12) and (5.6) into equation (5.5) gives the following:

$$\mathbf{K}_{SVSF-II_k} = \begin{bmatrix} \left| \mathbf{e}_{z_{k|k-1}} \right| + \gamma_1 \left| \mathbf{e}_{z_{k-1|k-1}} \right| \\ \left| \hat{\mathbf{A}}_{22} \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} \right| + \gamma_2 \left| \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} \right| \end{bmatrix} \circ \text{sgn} \left(\begin{bmatrix} \mathbf{e}_{z_{k|k-1}} \\ \hat{\mathbf{A}}_{22} \hat{\mathbf{A}}_{12}^{-1} \mathbf{e}_{z_{k|k-1}} \end{bmatrix} \right) \quad 5.13$$

Where $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ are diagonal matrices with elements less than one and size of $\boldsymbol{\gamma}_1 \in \mathbb{R}^{m \times m}$ and $\boldsymbol{\gamma}_2 \in \mathbb{R}^{(n-m) \times (n-m)}$, respectively. The mapping function ($\widehat{\mathbf{A}}_{12}$) needs to be invertible. Modelling errors in $\widehat{\mathbf{A}}_{12}$ and $\widehat{\mathbf{A}}_{22}$ impact the accuracy of the estimate. Even though the filter is stable (because the upper partition maintains the states close to the measurement), its performance is compromised. To improve the performance of this filter, a novel revised version of SVSF referred to as the Toeplitz/Observability SVSF is proposed in section (5.2).

5.1.2 The SVSF algorithm for non-linear systems with partially ranked measurement matrix

In (Habibi, 2007), an algorithm that is applicable to a nonlinear system with fewer measurements compared to the number of states was proposed. The proposed algorithm assumes the system has the structure of equation (4.7). If the measurement matrix does not have that structure, a coordinate transformation, \mathbf{T} , is applied on the states to convert the measurement matrix structure to equation (5.1). The resultant transformed system is defined as follow:

$$\begin{aligned} \mathbf{y}_k &= \mathbf{f}''(\mathbf{y}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}) = \begin{bmatrix} \mathbf{f}_{1k}(\mathbf{y}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}) \\ \mathbf{f}_{2k}(\mathbf{y}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}) \end{bmatrix} \\ \mathbf{z}_k &= \mathbf{H}'\mathbf{y}_k \end{aligned} \quad 5.14$$

Where $\mathbf{H}' = \mathbf{HT} = [\mathbf{I}_{m \times m} \quad \mathbf{0}_{m \times (n-m)}]$, $\mathbf{y}_k = \mathbf{T}^{-1}\mathbf{x}_k = \begin{bmatrix} \mathbf{y}_{1,k} \\ \mathbf{y}_{2,k} \end{bmatrix}$, $\mathbf{f}''(\mathbf{y}_{k-1}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1}) = \mathbf{T}^{-1}(\mathbf{f}(\mathbf{y}_{k-1}, \mathbf{u}_{k-1}, \mathbf{v}_{k-1}) + \mathbf{w}_{k-1})$, $\mathbf{y}_{1k}, \mathbf{f}_{1k} \in \mathbb{R}^{m \times 1}$, $\mathbf{y}_{2k}, \mathbf{f}_{2k} \in \mathbb{R}^{(n-m) \times 1}$, and $\mathbf{z}_k = \mathbf{y}_{1k}$.

To obtain the SVSF's gain, a methodology similar to the linear case is used. The system described by equation (5.14) is rewritten using the measurement vector as follows:

$$\begin{bmatrix} \mathbf{z}_k \\ \mathbf{y}_{2k} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{1k} \left(\begin{bmatrix} \mathbf{z}_{k-1} - \mathbf{v}_{k-1} \\ \mathbf{y}_{2k-1} \end{bmatrix}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1} \right) \\ \mathbf{f}_{2k} \left(\begin{bmatrix} \mathbf{z}_{k-1} - \mathbf{v}_{k-1} \\ \mathbf{y}_{2k-1} \end{bmatrix}, \mathbf{u}_{k-1}, \mathbf{w}_{k-1} \right) \end{bmatrix} + \begin{bmatrix} \mathbf{v}_k \\ \mathbf{0}_{(n-m) \times 1} \end{bmatrix} \quad 5.15$$

By rearranging equation (5.15), the following is obtained:

$$\begin{aligned} \mathbf{y}_{2k-1} &= \mathbf{f}_{1k}^{-1}(\mathbf{z}_{k-1}, \mathbf{z}_k, \mathbf{u}_{k-1}, \mathbf{v}_{k-1}, \mathbf{v}_k, \mathbf{w}_{k-1}) \\ \mathbf{y}_{2k-1} &= \mathbf{f}_{2k}^{-1}(\mathbf{z}_{k-1}, \mathbf{y}_{2k}, \mathbf{u}_{k-1}, \mathbf{v}_{k-1}, \mathbf{w}_{k-1}) \end{aligned} \quad 5.16$$

Where \mathbf{f}_{1k}^{-1} and \mathbf{f}_{2k}^{-1} are mapping functions that uniquely extract the states \mathbf{y}_{2k-1} from $\mathbf{z}_k, \mathbf{z}_{k-1}$ and \mathbf{u}_{k-1} . Using equation (5.16), a temporary estimate, $\bar{\boldsymbol{\Xi}}_{k-1}$, is obtained as follows:

$$\Xi_{k-1} = \begin{bmatrix} \Xi_{1k-1} \\ \Xi_{2k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{z}_{k-1} \\ \hat{\mathbf{f}}_1^{-1}(\mathbf{z}_k, \mathbf{z}_{k-1}, \mathbf{u}_{k-1}) \end{bmatrix} \quad 5.17$$

Where $\hat{\mathbf{f}}_1^{-1}$ is an estimated mapping function that uniquely extracts the hidden temporary states vector Ξ_{2k-1} from $\mathbf{z}_k, \mathbf{z}_{k-1}$ and \mathbf{u}_{k-1} . The intermediate Ξ variables are used as projected measurements for the reduced order nonlinear SVSF. In this research, the projected measurements are referred to as the *alternative measurements*. The SVSF structure is outlined as follows:

1 - Prediction Stage:

The a priori state estimate is obtained by using an estimated model of the system as follows:

$$\hat{\mathbf{y}}_{k|k-1} = \begin{bmatrix} \hat{\mathbf{y}}_{1k|k-1} \\ \hat{\mathbf{y}}_{2k|k-1} \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{f}}_{1k}(\Xi_{k-1}, \mathbf{u}_{k-1}) \\ \hat{\mathbf{f}}_{2k}(\Xi_{k-1}, \mathbf{u}_{k-1}) \end{bmatrix} \quad 5.18$$

$$\hat{\mathbf{y}}'_{k|k-1} = \begin{bmatrix} \hat{\mathbf{z}}_{k|k-1} \\ \hat{\mathbf{y}}_{2k|k-1} \end{bmatrix}$$

2 - Corrective Stage:

A corrective gain is calculated and used for refining the a priori estimate into its a posteriori form as follows:

$$\hat{\mathbf{y}}_{k|k} = \hat{\mathbf{y}}_{k|k-1} + \mathbf{K}_{SVSF-III_k}$$

$$\hat{\mathbf{y}}'_{k|k} = \begin{bmatrix} \hat{\mathbf{z}}_{k|k} \\ \hat{\mathbf{y}}_{2k|k} \end{bmatrix} \quad 5.19$$

where $\mathbf{K}_{SVSF-III_k}$ is the SVSF's gain for a reduced order nonlinear system and is defined as follows, (Habibi, 2007):

$$\mathbf{K}_{SVSF-III_k} = \left(\gamma \left| \mathbf{e}_{y_{k-1|k-1}} \right| + \left| \mathbf{e}_{y_{k|k-1}} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{y_{k|k-1}} \right) \quad 5.20$$

Where $\mathbf{e}_{y_{k|k}}$ and $\mathbf{e}_{y_{k|k-1}}$ are the a posteriori and the a priori estimation error, respectively, and they are defined as follows:

$$\mathbf{e}_{y_{k|k}} = \Xi_k - \hat{\mathbf{y}}'_{k|k} \quad 5.21$$

$$\mathbf{e}_{y_{k|k-1}} = \Xi_{k-1} - \hat{\mathbf{y}}'_{k|k-1}$$

In real applications, the measurement and system noise are unknown. Because of that and due to modeling errors, the alternative measurements vector, Ξ_k , contain a higher

level of noise due to differentiation. This means that the smoothing boundary layer will be large. The disadvantage of this algorithm is that it leads to a large existence subspace for the states that do not have an explicit measurement associated with them.

5.2 The Toeplitz/Observability SVSF

5.2.1. Introduction to the Toeplitz and Observability matrices

The *Observability matrix* is a mathematical tool that can be used to determine if the states can be uniquely extracted from a finite number of measurements. It is derived from the system and measurement matrices. The Observability matrix, \mathbf{O} , for a linear time invariant systems with n -states and **one-measurement** is defined as follows, (Kailath, 1980):

$$\mathbf{O} = [\mathbf{H}^T \quad \mathbf{A}^T \mathbf{H}^T \quad \dots \quad (\mathbf{A}^{n-1})^T \mathbf{H}^T]^T \quad 5.22$$

If the Observability matrix is of full rank, then in the absence of noise, the system is *observable*, and the states can be uniquely extracted from measurements using the following equation, (Kailath, 1980):

$$\mathbf{x}_k = \mathbf{O}^{-1}([z_k \quad z_{k+1} \quad \dots \quad z_{k+n-1}]^T - \mathbf{T}_o[u_k \quad u_{k+1} \quad \dots \quad u_{k+n-1}]^T) \quad 5.23$$

Where:

$$\mathbf{T}_o = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \mathbf{HB} & 0 & 0 & \dots & 0 \\ \mathbf{HAB} & \mathbf{HB} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{HA}^{n-2}\mathbf{B} & \mathbf{HA}^{n-3}\mathbf{B} & \dots & \mathbf{HB} & 0 \end{bmatrix} \quad 5.24$$

In the presence of system and measurement noise, equation (5.23) can be restated as:

$$\mathbf{x}_k = \mathbf{O}^{-1} \left(\begin{bmatrix} z_k \\ z_{k+1} \\ \vdots \\ z_{k+n-1} \end{bmatrix} - \mathbf{T}_o \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+n-1} \end{bmatrix} - \begin{bmatrix} v_k \\ v_{k+1} \\ \vdots \\ v_{k+n-1} \end{bmatrix} - \mathbf{T}_w \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \vdots \\ \mathbf{w}_{k+n-1} \end{bmatrix} \right) \quad 5.25$$

Where:

$$\mathbf{T}_w = \begin{bmatrix} \mathbf{0}_{1 \times n} & \mathbf{0}_{1 \times n} & \mathbf{0}_{1 \times n} & \dots & \mathbf{0}_{1 \times n} \\ \mathbf{H} & \mathbf{0}_{1 \times n} & \mathbf{0}_{1 \times n} & \dots & \mathbf{0}_{1 \times n} \\ \mathbf{HA} & \mathbf{H} & \mathbf{0}_{1 \times n} & \dots & \mathbf{0}_{1 \times n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{HA}^{n-2} & \mathbf{HA}^{n-3} & \dots & \mathbf{H} & \mathbf{0}_{1 \times n} \end{bmatrix} \quad 5.26$$

Proof:

The measurement equation defined in equation (5.2) can be written in recursive form as follows, (Hsu, 1995):

$$z_k = \mathbf{H} \left(\mathbf{A}^k \mathbf{x}_0 + \sum_{i=0}^{k-1} \left(\mathbf{A}^{k-1-i} (\mathbf{B}u_i + \mathbf{w}_i) \right) \right) + v_k \quad 5.27$$

Further to equation (5.27), the measurement at time step l is linked to the state at time step k as follows:

$$z_l = \mathbf{H} \left(\mathbf{A}^{l-k} \mathbf{x}_k + \sum_{i=k}^{l-1} \left(\mathbf{A}^{l-1-i} (\mathbf{B}u_i + \mathbf{w}_i) \right) \right) + v_l \quad 5.28$$

For a vector of measurements starting from time step k to time step $l = k + n - 1$, and by using equation (5.28), the following is obtained:

$$\begin{bmatrix} z_k \\ \vdots \\ z_{k+n-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}\mathbf{x}_k + v_k \\ \vdots \\ \mathbf{H} \left(\mathbf{A}^{n-1} \mathbf{x}_k + \sum_{i=k}^{k+n-2} \left(\mathbf{A}^{k+n-2-i} (\mathbf{B}u_i + \mathbf{w}_i) \right) \right) + v_{k+n-1} \end{bmatrix} \quad 5.29$$

Rearranging equation (5.29) gives the following:

$$\begin{bmatrix} z_k \\ \vdots \\ z_{k+n-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}\mathbf{O}_{n \times 1} \\ \vdots \\ \mathbf{H} \sum_{i=k}^{k+n-2} \left(\mathbf{A}^{k+n-2-i} \mathbf{B}u_i \right) \end{bmatrix} + \begin{bmatrix} \mathbf{H}\mathbf{O}_{n \times 1} \\ \vdots \\ \mathbf{H} \sum_{i=k}^{k+n-2} \left(\mathbf{A}^{k+n-2-i} \mathbf{w}_i \right) \end{bmatrix} + \begin{bmatrix} v_k \\ \vdots \\ v_{k+n-1} \end{bmatrix} + \begin{bmatrix} \mathbf{H} \\ \vdots \\ \mathbf{H}\mathbf{A}^{n-1} \end{bmatrix} \mathbf{x}_k \quad 5.30$$

By expanding the summation term in equation (5.30) and rewriting it in a matrix form, the following is obtained:

$$\begin{bmatrix} z_k \\ \vdots \\ z_{k+n-1} \end{bmatrix} = \mathbf{T}_o \begin{bmatrix} u_k \\ \vdots \\ u_{k+n-1} \end{bmatrix} + \begin{bmatrix} v_k \\ \vdots \\ v_{k+n-1} \end{bmatrix} + \mathbf{T}_w \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+n-1} \end{bmatrix} + \mathbf{O}\mathbf{x}_k \quad 5.31$$

Where \mathbf{T}_o and \mathbf{T}_w are lower triangular Toeplitz matrices that have the forms of equations (5.24) and (5.26), respectively. Equation (5.25) is then obtained by rearranging equation (5.31) and multiplying it with the inverse of the Observability matrix. Note that the Observability matrix should be invertible, otherwise equation (5.25) is invalid and the states or their expectation cannot be obtained.

From equations (5.23) and (5.25), the inputs are mapped to the states through the matrix, \mathbf{T}_o . In this research, the matrices \mathbf{T}_o and \mathbf{T}_w are referred to as the system Toeplitz

and the noise Toeplitz matrices, respectively. The estimated state vector of equation (5.25), $\hat{\mathbf{x}}_{TO_k}$, is then obtained as follows:

$$\hat{\mathbf{x}}_{TO_k} = \hat{\mathbf{O}}^{-1} \left(\begin{bmatrix} z_k \\ z_{k+1} \\ \vdots \\ z_{k+n-1} \end{bmatrix} - \hat{\mathbf{T}}_o \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+n-1} \end{bmatrix} \right) \quad 5.32$$

Where $\hat{\mathbf{O}}$ and $\hat{\mathbf{T}}_o$ are the estimated Observability and system Toeplitz matrices, respectively. Note that the error in estimation is equal to the following:

$$\begin{aligned} \mathbf{x}_k - \hat{\mathbf{x}}_{TO_k} = & \mathbf{O}^{-1} \left(\begin{bmatrix} z_k \\ z_{k+1} \\ \vdots \\ z_{k+n-1} \end{bmatrix} - \mathbf{T}_o \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+n-1} \end{bmatrix} - \begin{bmatrix} v_k \\ v_{k+1} \\ \vdots \\ v_{k+n-1} \end{bmatrix} - \mathbf{T}_w \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \\ \vdots \\ \mathbf{w}_{k+n-1} \end{bmatrix} \right) \\ & - \hat{\mathbf{O}}^{-1} \left(\begin{bmatrix} z_k \\ z_{k+1} \\ \vdots \\ z_{k+n-1} \end{bmatrix} - \hat{\mathbf{T}}_o \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+n-1} \end{bmatrix} \right) \end{aligned} \quad 5.33$$

If $\hat{\mathbf{O}}$ and $\hat{\mathbf{T}}_o$ are equal to \mathbf{O} and \mathbf{T}_o , respectively, then equation (5.33) is reduced to the following:

$$\mathbf{x}_k - \hat{\mathbf{x}}_{TO_k} = -\hat{\mathbf{O}}^{-1} ([v_k \ v_{k+1} \ \dots \ v_{k+n-1}]^T + \mathbf{T}_w [\mathbf{w}_k \ \mathbf{w}_{k+1} \ \dots \ \mathbf{w}_{k+n-1}]^T) \quad 5.34$$

5.2.2. The Toeplitz/Observability SVSF

The SVSF needs the measurement matrix to have full rank. If this condition is not valid, then the SVSF is combined with a method similar to the Luenberger algorithm as discussed in section (5.1). However, this method has some limitations due to the Luenberger algorithm as discussed in sections (1.3) and (5.1). Therefore, a new form of the SVSF is developed in this section to solve the problem associated with a lower number of measurements compared to the number of states.

From equation (5.34), if the system's **Observability and Toeplitz matrices are exactly known**, then the vector $\hat{\mathbf{x}}_{TO_k}$ is simply extracted from the measurement and the input, and it contains the vector \mathbf{x}_k 's information blurred with measurement and system noise, and their derivatives. Therefore, the vector $\hat{\mathbf{x}}_{TO_k}$ can be used to compensate the missing $(n - 1)$ measurements. For an observable system with only one measurement such that:

$$\mathbf{H} = [h_1 \ \mathbf{0}_{1 \times (n-1)}] \quad 5.35$$

then a new measurements vector, $\hat{\mathbf{z}}_{TO_k}$, can be defined as follows:

$$\hat{\mathbf{z}}_{TO_k} = \bar{\bar{\mathbf{H}}}_k \hat{\mathbf{x}}_{TO_k} \quad 5.36$$

Where $\hat{\mathbf{z}}_{TO_k}$ is the estimated measurement vector obtained by the Observability and Toeplitz matrices and has dimensions of $\hat{\mathbf{z}}_{TO_k} \in \mathbb{R}^{n \times 1}$. $\bar{\bar{\mathbf{H}}}_k$ is the new measurement matrix with dimensions of $\bar{\bar{\mathbf{H}}}_k \in \mathbb{R}^{n \times n}$ and has the following form:

$$\bar{\bar{\mathbf{H}}}_k = \begin{bmatrix} h_1 & 0 & 0 & 0 \\ 0 & h_1 & 0 & 0 \\ & & \ddots & \\ 0 & 0 & & h_1 & 0 \\ 0 & 0 & & 0 & h_1 \end{bmatrix} \quad 5.37$$

For the special case where $h_1 = 1$, then the new measurement matrix reduces to a unity matrix such that $\bar{\bar{\mathbf{H}}}_k = \mathbf{I}_{n \times n}$. In this research, the vectors $\hat{\mathbf{x}}_{TO_k}$ and $\hat{\mathbf{z}}_{TO_k}$ will be referred to as the *alternative states* and *alternative measurements*, respectively.

If the Observability and Toeplitz matrices are exactly known, this implies that the output matrix is also exactly known. Nonetheless modeling errors may be present in the system and the input matrices; e.g. if the system model is described in its Observability canonical form and the measurement represents the first state and its function is well known, then the Observability and the Toeplitz matrices are independent of the system parameters. In this research, the output matrix, \mathbf{H} , is assumed to be known as it is part of the SVSF assumption.

The vectors $\hat{\mathbf{x}}_{TO_k}$ and $\hat{\mathbf{z}}_{TO_k}$ are stochastic signals and require filtering strategies; i.e. the SVSF, for state and parameter estimation. In this section, the SVSF is applied by using the alternative measurements and their matrix, $\bar{\bar{\mathbf{H}}}_k$, as follows:

1 - Prediction Stage:

Further to equation (5.3), the a priori state estimate is obtained by using an estimated state space model of the system and replacing the measurement vector with the alternative measurement vector as follows:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \hat{\mathbf{A}}_{k-1} \hat{\mathbf{x}}_{k-1|k-1} + \hat{\mathbf{B}}_{k-1} u_{k-1} \\ \hat{\mathbf{z}}_{k|k-1} &= \bar{\bar{\mathbf{H}}}_k \hat{\mathbf{x}}_{k|k-1} \end{aligned} \quad 5.38$$

2 - Update Stage:

A corrective gain is calculated and used for refining the a priori estimate into its a posteriori form as follows:

$$\begin{aligned} \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{e_{TO_k}} \\ \hat{\mathbf{z}}_{k|k} &= \bar{\bar{\mathbf{H}}}_k \hat{\mathbf{x}}_{k|k} \end{aligned} \quad 5.39$$

where $\mathbf{K}_{e_{TO_k}}$ is the SVSF's gain for the Toeplitz/Observability SVSF and is defined as follows:

$$\mathbf{K}_{e_{TO_k}} = \bar{\mathbf{H}}_k^{-1} \left(\left| \mathbf{e}_{z_{TO_k}|k-1} \right| + \gamma \left| \mathbf{e}_{z_{TO_{k-1}}|k-1} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{z_{TO_k}|k-1} \right) \quad 5.40$$

Where $\mathbf{e}_{z_{TO_k}|k-1} = \hat{\mathbf{z}}_{TO_k} - \hat{\mathbf{z}}_{k|k-1}$ and $\mathbf{e}_{z_{TO_{k-1}}|k-1} = \hat{\mathbf{z}}_{TO_{k-1}} - \hat{\mathbf{z}}_{k-1|k-1}$.

5.2.3. Application of the Toeplitz/Observability SVSF into an electro-hydrostatic actuator

The Toeplitz/Observability SVSF has been tested on an electro-hydrostatic actuator described in chapter four, which has a third order system with the following system, input and measurement matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & T_s & 0 \\ 0 & 1 & T_s \\ 0 & -\omega_n^2 T_s & 1 - 2\xi\omega_n T_s \end{bmatrix} \quad 5.41$$

$$\mathbf{B} = [0 \quad 0 \quad BT_s]^T$$

$$\mathbf{H} = [1 \quad 0 \quad 0]$$

Where $\omega_n = 260.3 \text{ Hz}$, $B = 22.7 \frac{m}{\text{sec} \times \text{rad}}$ and $\xi = 0.2$. The sampling time is 0.001 sec .

The measurement noise is white noise with a noise-to-signal ratio of 5% with respect to the first state, and with a variance of $\sigma_v^2 = [1.2 \times 10^{-15}]$. The process noise is white with a noise-to-signal ratio of 5% of the state amplitudes, and with a variance of $\sigma_w^2 =$

$\begin{bmatrix} 2 \times 10^{-15} & 0 & 0 \\ 0 & 5 \times 10^{-10} & 0 \\ 0 & 0 & 1 \times 10^{-5} \end{bmatrix}$. For comparison purposes, the results of the

proposed algorithm are compared to corresponding results obtained from the KF. The system and measurement noise covariance matrices are defined as $\mathbf{Q} = \sigma_w^2$ and $\mathbf{R} = \sigma_v^2$, respectively, and the initial covariance matrix has a value of $\mathbf{P}_0 = \mathbf{I}_{3 \times 3}$. The SVSF's coefficient matrix γ has a value of $\gamma = 0.02 \times \mathbf{I}_{3 \times 3}$ and the SVSF's smoothing boundary

layer is designed as $\Psi = \begin{bmatrix} 5 \times 10^{-7} \\ 5 \times 10^{-4} \\ 5 \end{bmatrix}$. The input consists of a random signal superimposed

on step changes. The same simulation is performed again assuming that the model parameters are inaccurate and with values of ω_n, B and ξ are equal to 545.75 Hz , $10 \frac{m}{\text{sec} \times \text{rad}}$ and 0.302 , respectively. Note that the modeling errors vary between 50 – 110% of their actual values as shown in table 5.1.

	Actual	Estimated	Absolute % Error
B	22.7	10	56%
ω_n	260.3	545.75	110%
ξ	0.2	0.302	51%

Table 5.1: Actual and Estimated system parameters.

5.2.3.1. Simulation results

Two cases are considered in this section; the first one involves a known model without parametric uncertainties. The second involves an uncertain model with large parametric uncertainties as presented in table 5.1.

5.2.3.1.1. Results using an accurate model (No modeling errors)

The results of the Toeplitz/Observability SVSF compared to the KF for a known model are shown in Figs (5.1) to (5.3). These figures show that the KF gives slightly better estimated results for the states than the proposed algorithm when no modeling errors are presented (refer to table 5.2). The estimated states obtained by the Toeplitz/Observability SVSF are more sensitive to the noise and its derivatives than the ones obtained by the KF. Therefore, the proposed method is limited to small noise-to-signal ratios (as rule of thumbs, the noise-to-signal ratio of noise's derivatives to their corresponding states should be less than 10%). The sensitivity could be reduced by increasing the width of the smoothing boundary layer at the expense of accuracy as discussed in section (4.3.2). The two methods; the Toeplitz/Observability SVSF and the KF, are compared in terms of the root mean square error ($RMSE_j$) defined in equation (4.114). Table 5.2 summarizes the $RMSE_j$ results for both methods.

	KF	Toeplitz/Observability SVSF
Position $RMSE$	8.6×10^{-16}	3.2×10^{-15}
Velocity $RMSE$	1.5×10^{-9}	1.6×10^{-9}
Acceleration $RMSE$	5.6×10^{-5}	6.8×10^{-5}

Table 5.2: Comparison between the Toeplitz/Observability SVSF and the KF for a known estimated model.

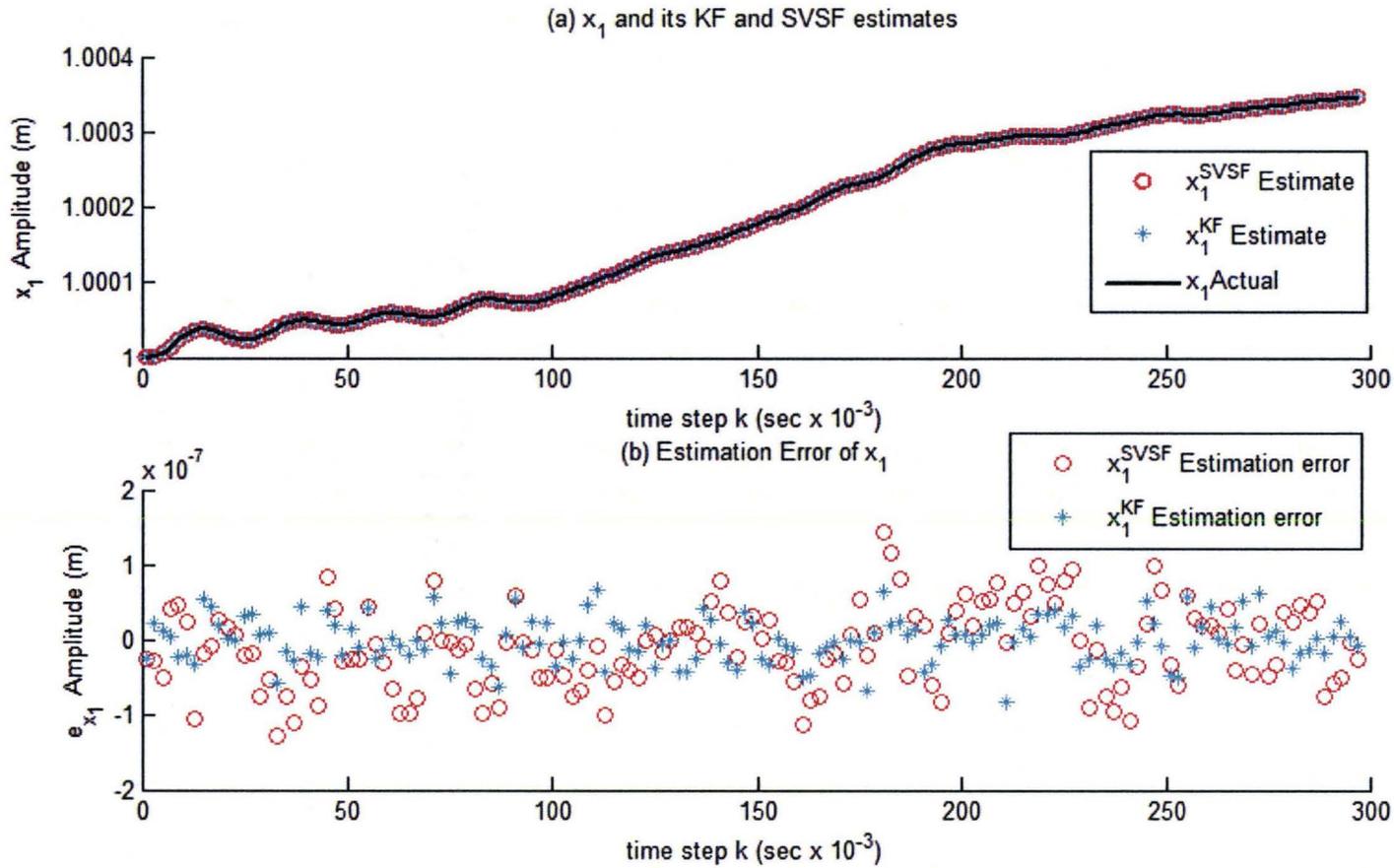


Fig 5.1: (a) x_1 's actual and estimated values, (b) estimation error of x_1 for the Toeplitz/Observability SVSF and the KF when there are no modeling errors.

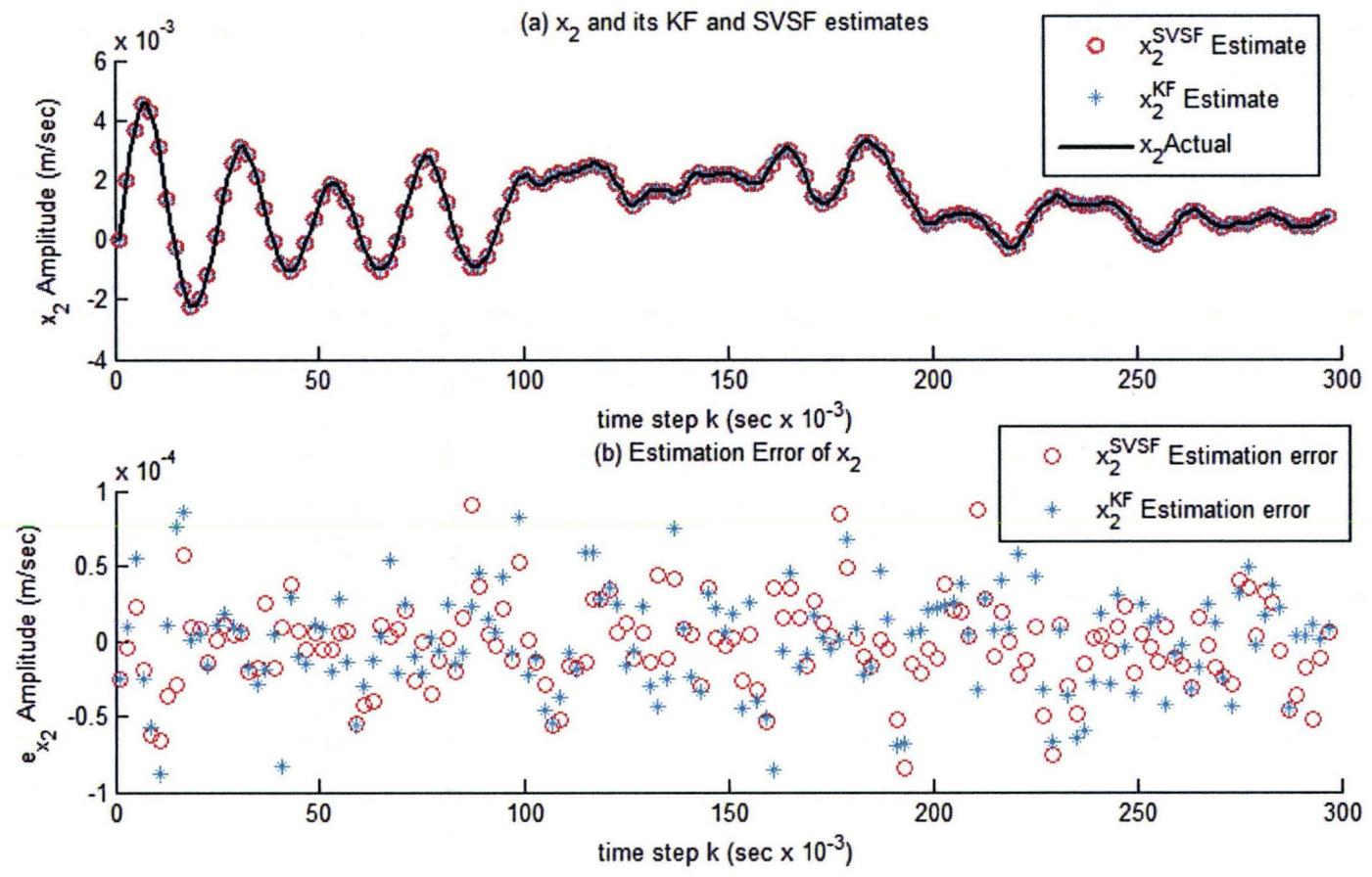


Fig 5.2: (a) x_2 's actual and estimated values, (b) estimation error of x_2 for the Toeplitz/Observability SVSF and the KF when there are no modeling errors.

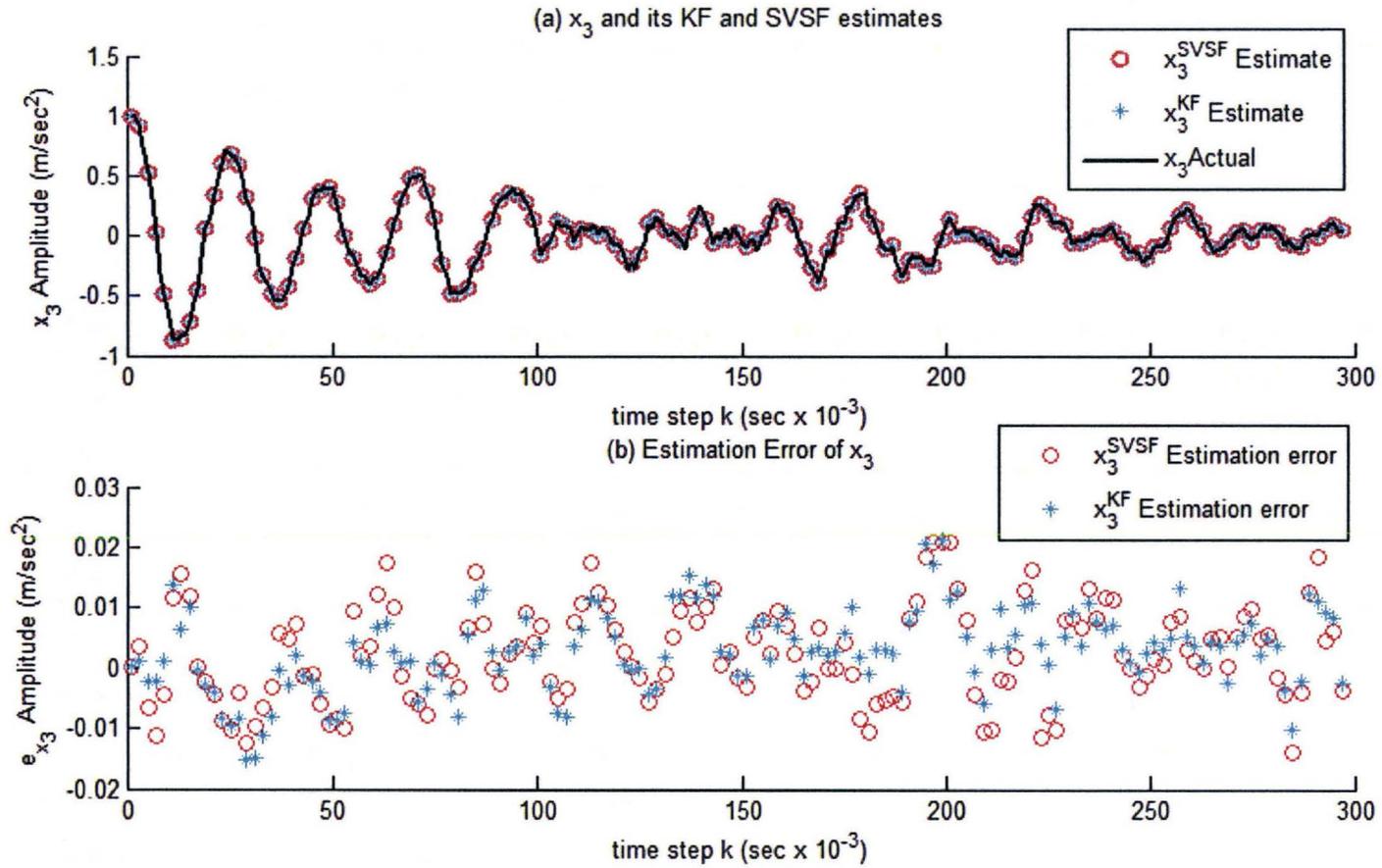


Fig 5.3: (a) x_3 's actual and estimated values, (b) estimation error of x_3 for the Toeplitz/Observability SVSF and the KF when there are no modeling errors.

5.2.3.1.2. Results using a non-accurate model

The results of the Toeplitz/Observability SVSF compared to the KF for a system model with large parametric uncertainties as presented in table 5.1 are shown in Figs (5.4) to (5.6). These figures show that the Toeplitz/Observability SVSF gives better results than the KF (the *RMSE* of x_1 , x_2 and x_3 obtained by the Toeplitz/Observability SVSF are smaller than their corresponding values obtained by the KF, by factors of 1/55, 1/400, and 1/2 times, respectively). The method is robust and it is not affected by the magnitude of the modeling errors. However, the estimates are sensitive to the noise amplitude and the width of the smoothing boundary layer. A confirmation of the conclusions is provided by a comparison of the root mean square errors as presented in table 5.3.

	KF	Toeplitz/Observability SVSF
Position <i>RMSE</i>	4.7×10^{-13}	8.5×10^{-15}
Velocity <i>RMSE</i>	3.3×10^{-6}	8.7×10^{-9}
Acceleration <i>RMSE</i>	0.3747	0.187

Table 5.3: Comparison between the Toeplitz/Observability SVSF and the KF for model with parameters uncertainties.

The proposed algorithm is limited to systems that:

- Have a single measurement.
- Are observable.
- Have Observability and Toeplitz matrices that are correctly estimated. As mentioned before, this statement does not imply that no modeling errors are present. There may be modeling errors in the system and the input matrices. The measurement matrix must however be known.

To remove these limitations, a general form is proposed in the following section that is referred to as the General Toeplitz/Observability SVSF.

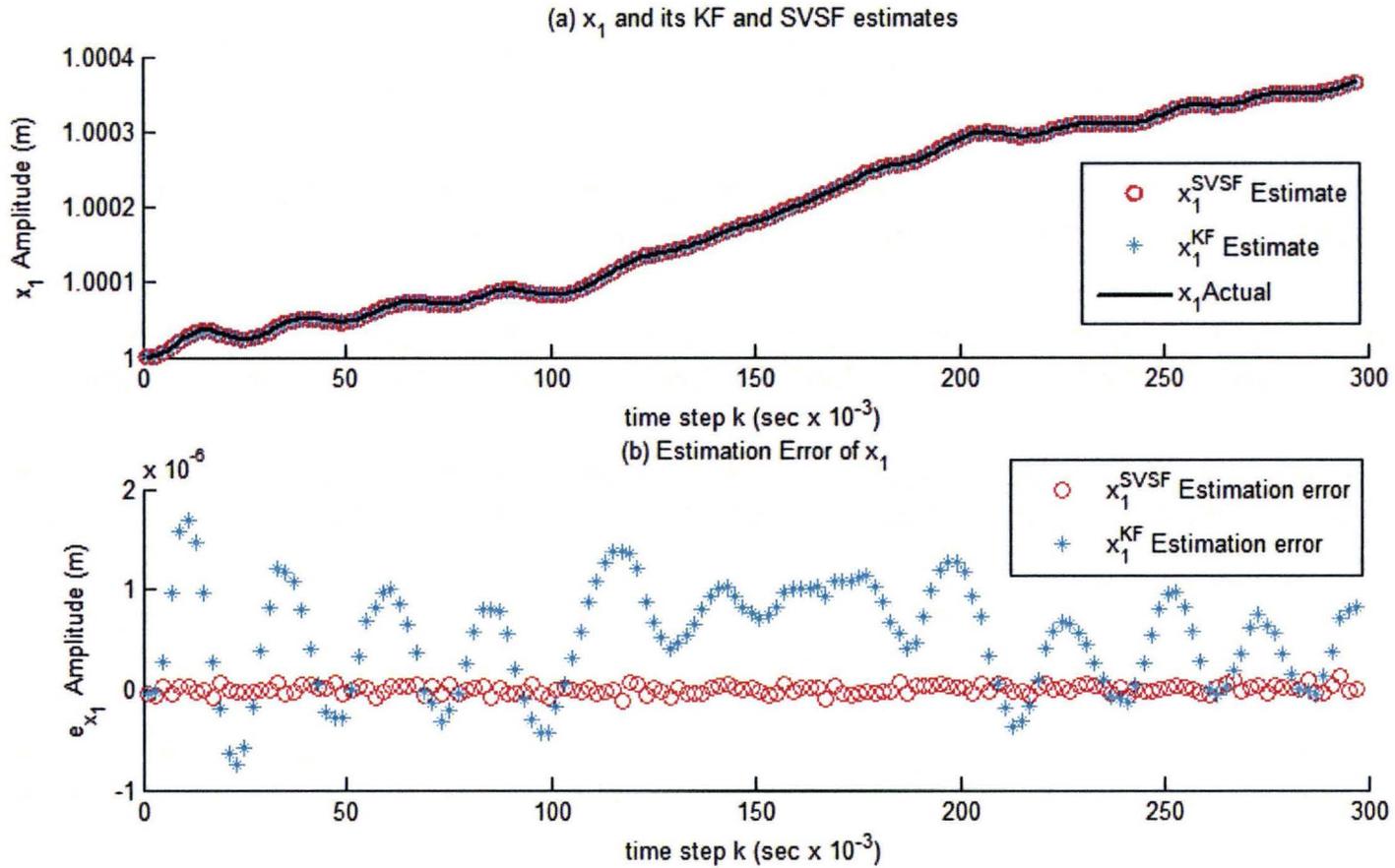


Fig 5.4: (a) x_1 's actual and estimated values, (b) estimation error of x_1 for the Toeplitz/Observability SVSF and the KF when modeling errors are present.

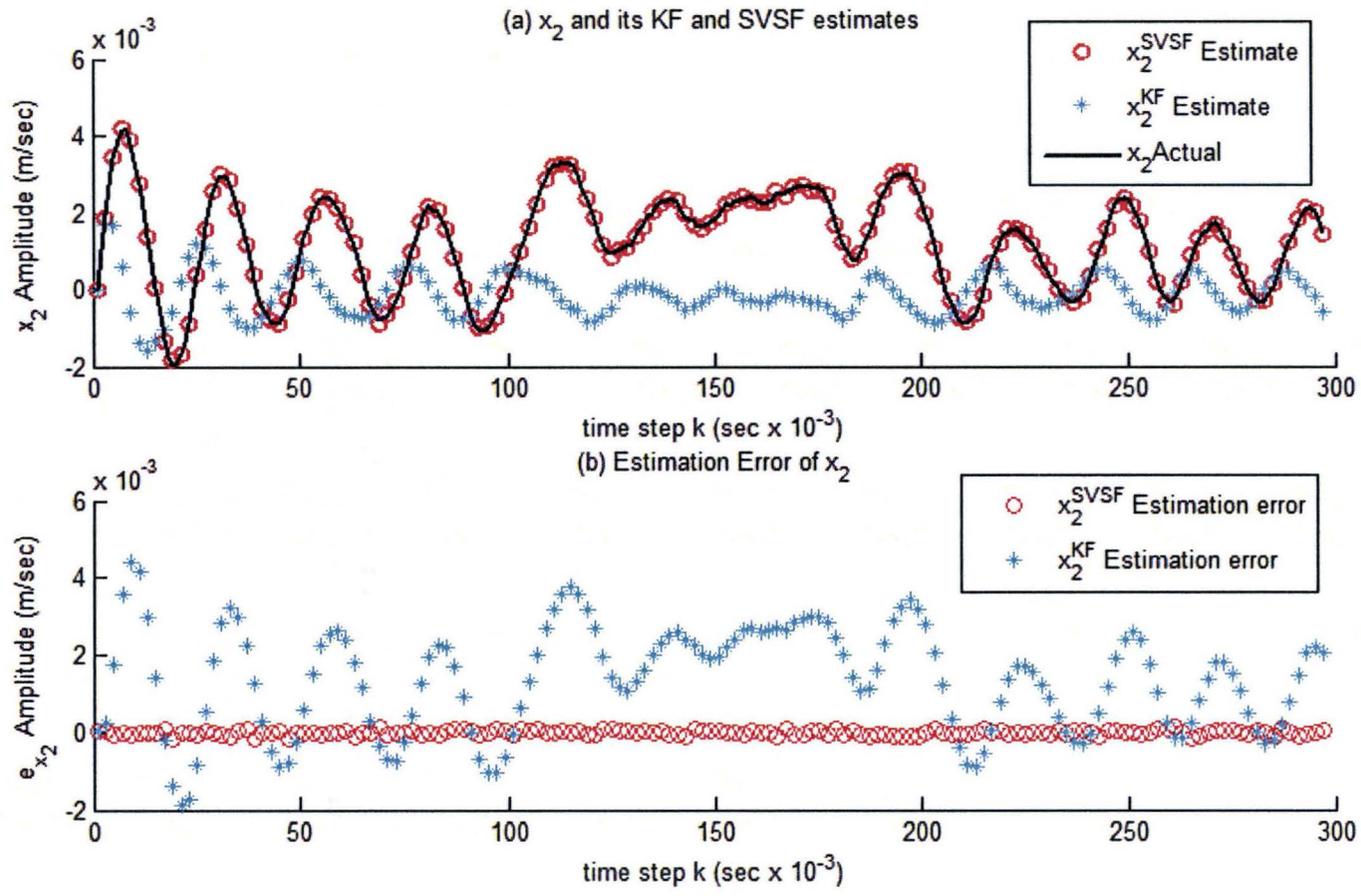


Fig 5.5: (a) x_2 's actual and estimated values, (b) estimation error of x_2 for the Toeplitz/Observability SVSF and the KF when modeling errors are present.

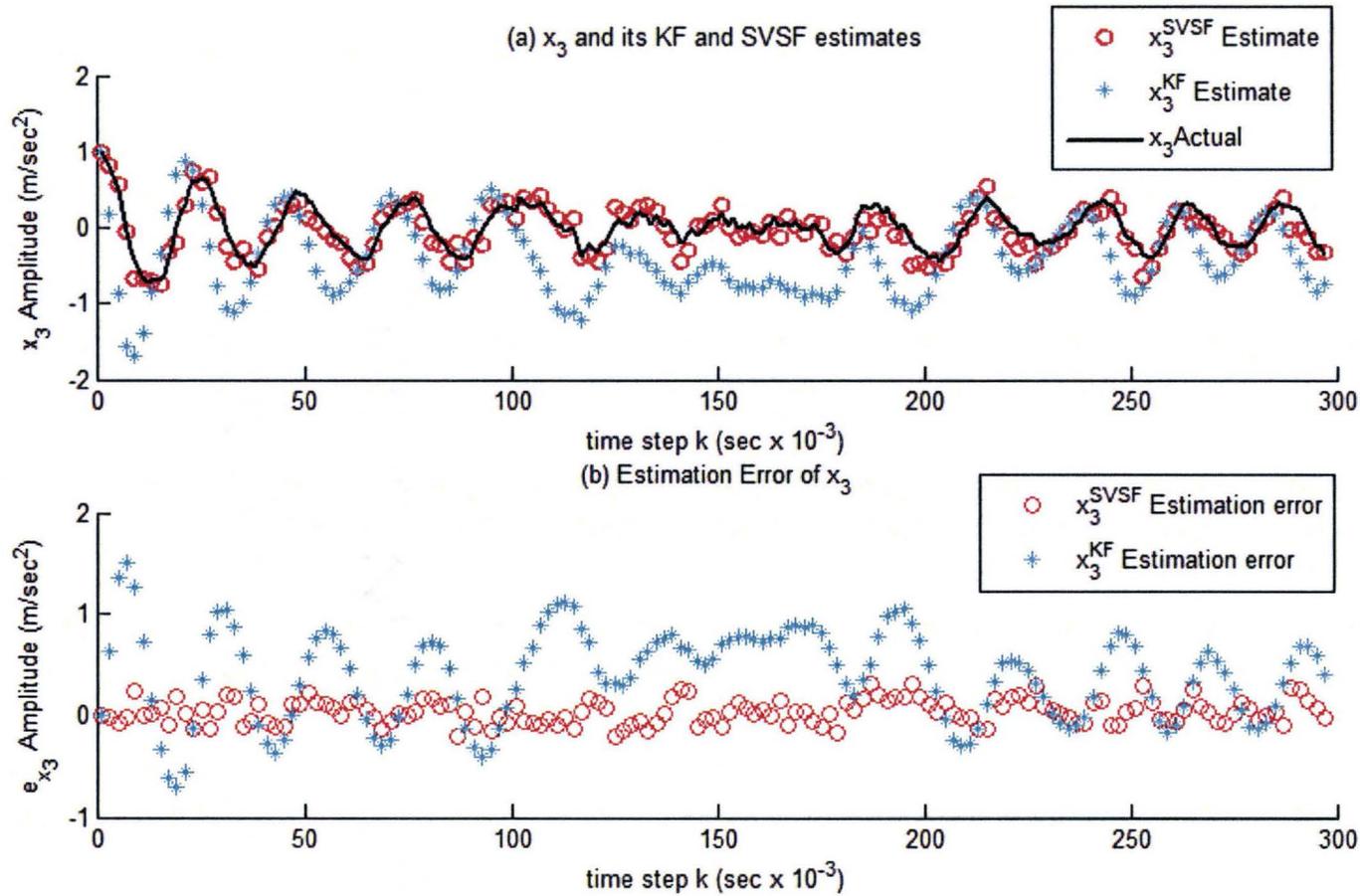


Fig 5.6: (a) x_3 's actual and estimated values, (b) estimation error of x_3 for the Toeplitz/Observability SVSF and the KF when modeling errors are present.

5.3 The General Toeplitz/Observability SVSF

5.3.1 Introduction to canonical forms representation

The state space representation of a transfer function is not unique and can have an infinite number of forms. In this section, four well-known canonical forms including the observer, controller, Observability and controllability canonicals forms are summarized. These canonical forms are important in mechanical and electrical applications as their states have strong physical meanings, [(Kailath, 1980) and (Nise, 2007)].

Consider the following system with a continuous transfer function as given in equation (5.42):

$$G(s) = \frac{b_{n-1}s^{n-1} + \dots + b_0}{s^n + a_{n-1}s^{n-1} + \dots + a_0} \quad 5.42$$

This system can also be represented in a state space form. The above mentioned state space canonical representations of the continuous system of equation (5.42) are given in table 5.4.

Canonical Form Name	System matrix	input matrix	measurement matrix
Observer	$\begin{bmatrix} -a_{n-1} & 1 & & & \\ -a_{n-2} & & 1 & & \\ & & & \ddots & \dots \\ -a_1 & & & & 1 \\ -a_0 & & & & \end{bmatrix}$	$\begin{bmatrix} b_{n-1} \\ b_{n-2} \\ \vdots \\ b_0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T$
Controller	$\begin{bmatrix} -a_{n-1} & 1 & & & \\ -a_{n-2} & & 1 & & \\ & & & \ddots & \dots \\ -a_1 & & & & 1 \\ -a_0 & & & & \end{bmatrix}^T$	$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\begin{bmatrix} b_{n-1} \\ b_{n-2} \\ \vdots \\ b_0 \end{bmatrix}^T$
Observability	$\begin{bmatrix} & & & -a_0 \\ & & & -a_1 \\ 1 & & & -a_2 \\ & 1 & \dots & \\ & & & \vdots \\ & & & 1 & -a_{n-1} \end{bmatrix}^T$	$\begin{bmatrix} b'_{n-1} \\ b'_{n-2} \\ \vdots \\ b'_0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T$
Controllability	$\begin{bmatrix} & & & -a_0 \\ & & & -a_1 \\ & & & -a_2 \\ 1 & & & \\ & 1 & \dots & \\ & & & \vdots \\ & & & 1 & -a_{n-1} \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\begin{bmatrix} b'_{n-1} \\ b'_{n-2} \\ \vdots \\ b'_0 \end{bmatrix}^T$

Table 5.4: The system, input and measurement matrices' of the continuous observer, controller, Observability and controllability canonical forms.

$$\text{Where } \begin{bmatrix} b'_{n-1} \\ b'_{n-2} \\ \vdots \\ b'_0 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ a_{n-1} & 1 & & \\ a_{n-2} & a_{n-1} & \dots & \\ \vdots & \vdots & & \\ a_0 & a_1 & & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_{n-1} \\ b_{n-2} \\ \vdots \\ b_0 \end{bmatrix}.$$

If the system's transfer function does not have zeros ($b_1, b_2 \dots b_{n-1} = 0$), then the system Toeplitz matrix becomes trivial. The corresponding Observability and system Toeplitz matrices are summarized in Table 5.5.

Canonical Form (CF)	Observability matrix	System Toeplitz matrix
Observer	$\begin{bmatrix} 1 & 0 & 0 \\ -a_{n-1} & 1 & 0 \\ \vdots & \vdots & \dots \\ \text{function of } (a_1 \dots a_{n-1}) & \text{function of } (a_2 \dots a_{n-1}) & 1 \end{bmatrix}$	$\mathbf{0}_{n \times n}$
Controller	$\begin{bmatrix} 0 & 0 & b_0 \\ \vdots & & \\ 0 & b_0 & \dots & 0 \\ b_0 & 0 & 0 & 0 \end{bmatrix}$	$\mathbf{0}_{n \times n}$
Observability	$\mathbf{I}_{n \times n}$	$\mathbf{0}_{n \times n}$
Controllability	$b_0 \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -a_{n-1} \\ \dots & \vdots & \\ 1 & \text{function of } (a_2 \dots a_{n-1}) & \text{function of } (a_1 \dots a_{n-1}) \end{bmatrix}$	$\mathbf{0}_{n \times n}$

Table 5.5: The Observability and Toeplitz matrices' structures of the observer, controller, Observability and controllability canonical forms.

According to table 5.5, if modeling errors are present, then the system should be presented in the Observability canonical form as the Observability and system Toeplitz matrices are independent of the system model parameters a_0, \dots, a_{n-1}, b_0 . As such, modeling errors would not impact the alternative states $\hat{\mathbf{x}}_{TO_k}$ as shown in equation (5.34).

5.3.2 Discrete Observability canonical form

In order to reduce the estimation error due to modeling uncertainties, the system model with a single input is converted to its Observability canonical form. This implies that for a system with the transfer function of equation (5.42) and assuming ($b_1, b_2 \dots b_{n-1} = 0$), the discretized system, measurement and input state space matrices have the following forms, (Nise, 2007):

$$\mathbf{A} = \begin{bmatrix} 1 & T_s & 0 & \dots & 0 \\ 0 & 1 & T_s & \ddots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & T_s \\ -a_0 T_s & -a_1 T_s & -a_2 T_s & \dots & 1 - a_{n-1} T_s \end{bmatrix} \quad 5.43$$

$$\mathbf{H} = [h_1 \quad 0 \quad \dots \quad 0]$$

$$\mathbf{B} = [0 \quad \dots \quad 0 \quad b_0 T_s]^T$$

Using the Observability representation leads to:

- An Observability matrix that is time invariant, with the dimension $n \times n$, and independent of the system's parameters, as follows:

$$\mathbf{O} = \begin{bmatrix} [1 \quad \mathbf{0}_{1 \times (n-1)}] \\ [1 \quad T_s \quad \mathbf{0}_{1 \times (n-2)}] \\ \vdots \\ \left[1 \quad (n-1)T_s \quad \sum_{j=3}^n (j-2)T_s^2 \quad \dots \quad (n-1)T_s^{n-2} \quad T_s^{n-1} \right] \end{bmatrix} \quad 5.44$$

- A system Toeplitz matrix with zero entries, as the terms $(\mathbf{H}\mathbf{B}, \dots, \mathbf{H}\mathbf{A}^{n-2}\mathbf{B})$ in equation (5.24) have values of zero.
- A noise Toeplitz matrix that is time invariant and is independent of the system parameters, as the terms $(\mathbf{H}, \dots, \mathbf{H}\mathbf{A}^{n-2})$ in equation (5.26) are constants.
- An estimation error that is a function of noise as shown in equation (5.34).

The above observations are illustrated by the following example:

Example 5.1:

If the system has the following transfer function:

$$G(s) = \frac{b_0}{s^3 + a_2 s^2 + a_1 s + a_0} = \frac{557 \times 10^3}{s^3 + 60s^2 + 28 \times 10^3 s} \quad 5.45$$

and assuming the time step is 0.001 sec, then the discretized system, measurement, and input state space matrices have the following forms:

$$\mathbf{A} = \begin{bmatrix} 1 & T_s & 0 \\ 0 & 1 & T_s \\ -a_0 T_s & -a_1 T_s & 1 - a_2 T_s \end{bmatrix} = \begin{bmatrix} 1 & 0.001 & 0 \\ 0 & 1 & 0.001 \\ 0 & -28 & 0.94 \end{bmatrix} \quad 5.46$$

$$\mathbf{H} = [h_1 \quad 0 \quad 0] = [1 \quad 0 \quad 0]$$

$$\mathbf{B} = [0 \quad \dots \quad 0 \quad b_0 T_s]^T = [0 \quad \dots \quad 0 \quad 557]^T$$

The observability, system Toeplitz and noise Toeplitz matrices are independent of the parameters of \mathbf{A} and \mathbf{B} , and are defined as follows:

$$\begin{aligned} \mathbf{O} &= \begin{bmatrix} \mathbf{H} \\ \mathbf{HA} \\ \mathbf{HA}^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & T_s & 0 \\ 1 & 2T_s & T_s^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0.001 & 0 \\ 1 & 0.002 & 10^{-6} \end{bmatrix} \\ \mathbf{T}_0 &= \begin{bmatrix} 0 & & \\ \mathbf{HB} & 0 & \\ \mathbf{HAB} & \mathbf{HB} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ \mathbf{T}_w &= \begin{bmatrix} \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} \\ \mathbf{H} & \mathbf{0}_{1 \times 3} & \mathbf{0}_{1 \times 3} \\ \mathbf{HA} & \mathbf{H} & \mathbf{0}_{1 \times 3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0.001 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \quad 5.47$$

5.3.3 General Observability Canonical Form

The Observability canonical form reduces the error in the estimated states due to modeling errors in matrices \mathbf{A} and \mathbf{B} . However, the system's model cannot be reduced to the form of equation (5.43) for systems that have multiple measurements and/or multiple degrees of freedom. In order to accommodate such systems, an extended form of the Observability canonical form referred to as the General Observability Canonical Form (GOCF) is presented in this study as follows:

The system is considered to consist of ($NSUB$) subsystems or $NSUB^{th}$ degrees of freedom. Assuming that each subsystem has an order of n_i and a first state (in its continuous time form) as $\mathcal{X}_i(t), i = 1 \dots NSUB$ then the overall system model is described by the following set of differential equations:

$$\begin{aligned} \mathcal{X}_1^{(n_1)}(t) &= - \sum_{l=1}^{n_1} a'_{1,l} \mathcal{X}_1^{(l-1)}(t) - \sum_{l=1}^{n_2} a'_{1,l+n_1} \mathcal{X}_2^{(l-1)}(t) - \dots \\ &\quad - \sum_{l=1}^{n_{NSUB}} a'_{1,l+\sum_{i=1}^{NSUB-1} n_{i_1}} \mathcal{X}_{NSUB}^{(l-1)}(t) + b'_1 u(t) \\ \mathcal{X}_2^{(n_2)}(t) &= - \sum_{l=1}^{n_1} a'_{2,l} \mathcal{X}_1^{(l-1)}(t) - \sum_{l=1}^{n_2} a'_{2,l+n_1} \mathcal{X}_2^{(l-1)}(t) - \dots \\ &\quad - \sum_{l=1}^{n_{NSUB}} a'_{2,l+\sum_{i=1}^{NSUB-1} n_{i_1}} \mathcal{X}_{NSUB}^{(l-1)}(t) + b'_2 u(t) \\ &\quad \vdots \\ \mathcal{X}_{NSUB}^{(n_{NSUB})}(t) &= - \sum_{l=1}^{n_1} a'_{NSUB,l} \mathcal{X}_1^{(l-1)}(t) - \sum_{l=1}^{n_2} a'_{NSUB,l+n_1} \mathcal{X}_2^{(l-1)}(t) - \dots \\ &\quad - \sum_{l=1}^{n_{NSUB}} a'_{NSUB,l+\sum_{i=1}^{NSUB-1} n_{i_1}} \mathcal{X}_{NSUB}^{(l-1)}(t) + b'_{NSUB} u(t) \end{aligned} \quad 5.48$$

The phase variables $x_1 \dots x_1^{(n_1-1)}, \dots, x_{NSUB} \dots x_{NSUB}^{(n_{NSUB}-1)}$ are assumed to be the states of the General Observability Canonical Form. They have a total number of $n = \sum_{i=1}^c n_i$ and are defined as follows:

$$\begin{aligned} x_1(t) &= X_1(t), x_2(t) = \dot{X}_1(t), \dots, x_{n_1}(t) = X_1^{(n_1-1)}(t), \\ x_{n_1+1}(t) &= X_2(t), x_{n_1+2}(t) = \dot{X}_2(t), \dots, x_{n_1+n_2}(t) = X_2^{(n_2-1)}(t) \\ &\vdots \\ x_{\sum_{l=1}^{NSUB-1} n_l+1}(t) &= X_{NSUB}(t), \dots, x_{\sum_{l=1}^{NSUB} n_l}(t) = X_{NSUB}^{(n_{NSUB}-1)}(t) \end{aligned} \quad 5.49$$

For simplicity, new indices that specify the locations (rows) of the last states of each subsystem according to the system matrix are defined as follows:

$$ENDSUB_i = \sum_{l=1}^i n_l, i = 1 \dots NSUB \quad 5.50$$

If m measurements are available for the system, then the measurement matrix is defined as follows:

$$\mathbf{H} = [\mathbf{H}_1 \quad \dots \quad \mathbf{H}_m]^T \quad 5.51$$

Where $\mathbf{H} \in \mathbb{R}^{m \times n}$ and $\mathbf{H}_1 \dots \mathbf{H}_m \in \mathbb{R}^{1 \times n}$. If each measurement is related to **only one of the states**, then each partition of the measurement matrix is defined as follows:

$$\mathbf{H}_i = [0 \quad \dots \quad 0 \quad 1 \quad 0 \quad \dots \quad 0], i = 1 \dots m \quad 5.52$$

\uparrow
 location LMS_i

Where LMS_i is the location (column) of the measured state i . These locations are sorted in the output matrix such that $LMS_1 < LMS_2 < \dots < LMS_m$. Further to equation (5.51) and using the definition of equation (5.52), the measurement matrix is defined as follows:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_m \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{1 \times (LMS_1-1)} & 1 & \mathbf{0}_{1 \times (n-LMS_1)} \\ \vdots & \vdots & \vdots \\ \mathbf{0}_{1 \times (LMS_m-1)} & 1 & \mathbf{0}_{1 \times (n-LMS_m)} \end{bmatrix} \quad 5.53$$

If the measurement matrix does not have the form of equation (5.53), then a special transformation can be applied to the measurements as discussed later in section (5.4).

Using the definition of the states in equation (5.49) and the indices in equation (5.50), a revised form of the model is defined as follows:

$\mathbf{x}_k = \mathbf{X}(t_k)$, $\mathbf{x}_k \in \mathbb{R}^{n \times 1}$, $\mathbf{B} \in \mathbb{R}^{n \times 1}$, $u_k = U(t_k)$, $u_k \in \mathbb{R}^{1 \times 1}$, $\mathbf{A} \in \mathbb{R}^{n \times n}$ and the parameters $a_{ENDSUB_{i,j}}$ and b_{ENDSUB_i} are defined as follows:

$$a_{ENDSUB_{i,j}} = \begin{cases} 1 - a'_{ENDSUB_{i,j}}T_s & ENDSUB_i = j \\ -a'_{ENDSUB_{i,j}}T_s & ENDSUB_i \neq j \end{cases}, i = 1 \dots NSUB \text{ and } j = 1 \dots n \quad 5.63$$

$$b_{ENDSUB_i} = b_i T_s, i = 1 \dots c \quad 5.64$$

The above mentioned representation is illustrated by the following example.

Example 5.2:

The system in Fig 5.7 has three degrees of freedom and can be expressed by the following differential equations:

$$M_1 \ddot{x}_1(t) + (c_1 + c_2 + c_4) \dot{x}_1(t) + (k_1 + k_2 + k_4)x_1(t) = c_2 \dot{x}_2(t) + c_4 \dot{x}_3(t) + k_2 x_2(t) + k_4 x_3(t) + b_1 u(t) \quad 5.65$$

$$M_2 \ddot{x}_2(t) + (c_2 + c_3) \dot{x}_2(t) + (k_2 + k_3)x_2(t) = c_2 \dot{x}_1(t) + c_3 \dot{x}_3(t) + k_2 x_1(t) + k_3 x_3(t) + b_2 u(t) \quad 5.66$$

$$M_3 \ddot{x}_3(t) + (c_3 + c_4) \dot{x}_3(t) + (k_3 + k_4)x_3(t) = c_3 \dot{x}_2(t) + c_4 \dot{x}_1(t) + k_4 x_1(t) + k_3 x_2(t) + b_3 u(t) \quad 5.67$$

If the parameters have the following values:

$$k_1 = k_2 = k_3 = k_4 = 10^3 \text{ N/m},$$

$$c_1 = c_2 = c_3 = c_4 = 10^3 \text{ N.sec/m},$$

$$b_1 = b_2 = b_3 = 10^3 \text{ and } M_1 = M_2 = M_3 = 1 \text{ kg}$$

and assuming the states are defined as follows:

$$x_{1_1} = x_1, x_{1_2} = \dot{x}_1, x_{2_1} = x_2, x_{2_2} = \dot{x}_2, x_{3_1} = x_3 \text{ and } x_{3_2} = \dot{x}_3$$

then the Observability canonical representation

of the discretized system is obtained as follows:

$$\begin{bmatrix} x_{1_1 k+1} \\ x_{1_2 k+1} \\ x_{2_1 k+1} \\ x_{2_2 k+1} \\ x_{3_1 k+1} \\ x_{3_2 k+1} \end{bmatrix} = \begin{bmatrix} 1 & T_s & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 1 & T_s & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & 0 & 0 & 1 & T_s \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} \begin{bmatrix} x_{1_1 k} \\ x_{1_2 k} \\ x_{2_1 k} \\ x_{2_2 k} \\ x_{3_1 k} \\ x_{3_2 k} \end{bmatrix} + \begin{bmatrix} 0 \\ b_1 T_s / m_1 \\ 0 \\ b_2 T_s / m_2 \\ 0 \\ b_3 T_s / m_3 \end{bmatrix} u_k \quad 5.68$$

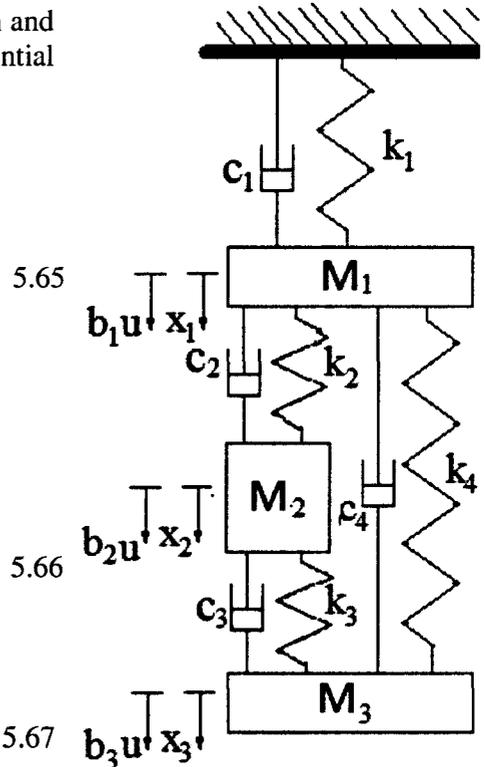


Fig 5.7: Example 2 – System with three degrees of freedom.

$$\text{Where } a_{21} = \frac{-(k_1+k_2+k_4)}{M_1} T_s, \quad a_{22} = 1 - \frac{(c_1+c_2+c_4)}{M_1} T_s, \quad a_{23} = \frac{k_2}{M_1} T_s, \quad a_{24} = \frac{c_2}{M_1} T_s, \quad a_{25} = \frac{k_4}{M_1} T_s, \\ a_{26} = \frac{c_4}{M_1} T_s, \quad a_{41} = \frac{k_2}{M_2} T_s, \quad a_{42} = \frac{c_2}{M_2} T_s, \quad a_{43} = \frac{-(k_2+k_3)}{M_2} T_s, \quad a_{44} = 1 - \frac{(c_2+c_3)}{M_2} T_s, \quad a_{45} = \\ \frac{k_3}{M_2} T_s, \quad a_{46} = \frac{c_3}{M_2} T_s, \quad a_{61} = \frac{k_4}{M_3} T_s, \quad a_{62} = \frac{c_4}{M_3} T_s, \quad a_{63} = \frac{k_3}{M_3} T_s, \quad a_{64} = \frac{c_3}{M_3} T_s, \quad a_{65} = \frac{-(k_4+k_3)}{M_3} T_s \quad \text{and} \\ a_{66} = 1 - \frac{(c_4+c_3)}{M_3} T_s.$$

If the sampling time has a value of 0.001 *sec*, then equation (5.68) is reduced to the following:

$$\begin{bmatrix} x_{1_{k+1}} \\ x_{1_{2k+1}} \\ x_{2_{k+1}} \\ x_{2_{2k+1}} \\ x_{3_{k+1}} \\ x_{3_{2k+1}} \end{bmatrix} = \begin{bmatrix} 1 & 0.001 & 0 & 0 & 0 & 0 \\ -3 & -2 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0.001 & 0 & 0 \\ 1 & 1 & -2 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0.001 \\ 1 & 1 & 1 & 1 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_{1_{1k}} \\ x_{1_{2k}} \\ x_{2_{1k}} \\ x_{2_{2k}} \\ x_{3_{1k}} \\ x_{3_{2k}} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} u_k \quad 5.69$$

If the displacements for all three masses are measured, then the measurement vector is defined as follows:

$$\begin{bmatrix} Z_{1k} \\ Z_{2k} \\ Z_{3k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{1_{1k}} & x_{1_{2k}} & x_{2_{1k}} & x_{2_{2k}} & x_{3_{1k}} & x_{3_{2k}} \end{bmatrix}^T \quad 5.70$$

Further to equations (5.61), (5.62) and (5.53), the indices m, LMS_1, LMS_2, LMS_3 , and $NSUB$ have values of 3, 1, 3, 5 and 3, respectively. The indices $ENDSUB_1, ENDSUB_2$ and $ENDSUB_3$ have values of 2, 4 and 6, respectively.

5.3.3.1. The General Observability matrix

As discussed in section (5.2.1), the full state vector at time k of an observable system can be extracted using successive data points of the system's measurement and input in conjunction with the Observability and the Toeplitz matrices. For an n^{th} order observable single input – single output (SISO) system, n successive data points of the measurement and the input are needed to obtain the n states. If the system has multiple measurements and/or has multiple degrees of freedom as discussed in the previous section, then equation (5.22) cannot be applied and the definition of the Observability and the Toeplitz matrices must be modified. For systems with multiple measurements, a proposed matrix referred to as the General Observability matrix that can be obtained to map the states of the observable system to measurement data segments as follows:

$$\begin{bmatrix} z_{1k} \\ \vdots \\ z_{1k+LM_1-1} \\ \vdots \\ z_{m_k} \\ \vdots \\ z_{m_k+LM_m-1} \end{bmatrix} = \mathbf{T}_{og} \begin{bmatrix} u_k \\ \vdots \\ u_{k+m_a-1} \end{bmatrix} + \begin{bmatrix} v_{1k} \\ \vdots \\ v_{1k+LM_1-1} \\ \vdots \\ v_{m_k} \\ \vdots \\ v_{m_k+LM_m-1} \end{bmatrix} + \mathbf{T}_{wg} \begin{bmatrix} w_k \\ \vdots \\ w_{k+m_a-1} \end{bmatrix} + \mathbf{O}_g \mathbf{x}_k \quad 5.71$$

Where \mathbf{T}_{og} and \mathbf{T}_{wg} are the General System and System Noise Toeplitz matrices. These are further discussed later in section (5.3.3.3). \mathbf{O}_g is the General Observability matrix and it is defined as follows:

$$\mathbf{O}_g = [\mathbf{H}_1 \ \mathbf{H}_1 \mathbf{A} \ \dots \ \mathbf{H}_1 \mathbf{A}^{LM_1-1} \ \mathbf{H}_2 \ \dots \ \mathbf{H}_2 \mathbf{A}^{LM_2-1} \ \dots \ \mathbf{H}_m \ \dots \ \mathbf{H}_m \mathbf{A}^{LM_m-1}]^T \quad 5.72$$

Where LM_i represents the segment's length that is needed from the measurement z_i (starting from time k to time $k + LM_i - 1$) in order to extract $LM_i - 1$ portions of the hidden states at time k (plus the first measured). For example, if at time k , LM_i has a value of five, then a segment consisting of five time steps from the measurement z_i , namely $z_{i_k}, z_{i_{k+1}}, \dots, z_{i_{k+4}}$, are needed to extract five of the states at time k from the measurement z_i .

Proof:

By taking LM_i successive data points from the measurement z_i starting from time step k to time step $k + LM_i - 1$, and mapping them to the system's state at time k using equations (5.27), (5.28), (5.29) and (5.30), the following is obtained:

$$\begin{bmatrix} z_{i_k} \\ \vdots \\ z_{i_{k+LM_i-1}} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_i \mathbf{x}_k + v_{i_k} \\ \vdots \\ \mathbf{H}_i \mathbf{A}^{LM_i-1} \mathbf{x}_k + \mathbf{H}_i \sum_{j=k}^{k+LM_i-2} (\mathbf{A}^{k+LM_i-2-j} (\mathbf{B} \mathbf{u}_j + \mathbf{w}_j)) + v_{i_{k+LM_i-1}} \end{bmatrix} \quad 5.73$$

Rearranging equation (5.73) gives the following:

$$\begin{bmatrix} z_{i_k} \\ \vdots \\ z_{i_{k+LM_i-1}} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_i \mathbf{0}_{n \times 1} \\ \vdots \\ \mathbf{H}_i \sum_{j=k}^{k+LM_i-2} (\mathbf{A}^{k+LM_i-2-j} \mathbf{B} \mathbf{u}_j) \end{bmatrix} + \begin{bmatrix} \mathbf{H}_i \mathbf{0}_{n \times 1} \\ \vdots \\ \mathbf{H}_i \sum_{j=k}^{k+LM_i-2} (\mathbf{A}^{k+LM_i-2-j} \mathbf{w}_j) \end{bmatrix} + \begin{bmatrix} v_{i_k} \\ \vdots \\ v_{i_{k+LM_i-1}} \end{bmatrix} + \begin{bmatrix} \mathbf{H}_i \\ \vdots \\ \mathbf{H}_i \mathbf{A}^{LM_i-1} \end{bmatrix} \mathbf{x}_k \quad 5.74$$

By expanding the summation term in equation (5.74) and rewriting it in a matrix form, the following is obtained:

$$\begin{bmatrix} z_{i_k} \\ \vdots \\ z_{i_{k+LM_i-1}} \end{bmatrix} = \mathbf{T}_{O_i} \begin{bmatrix} u_k \\ \vdots \\ u_{k+LM_i-1} \end{bmatrix} + \begin{bmatrix} v_{i_k} \\ \vdots \\ v_{i_{k+LM_i-1}} \end{bmatrix} + \mathbf{T}_{W_i} \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+LM_i-1} \end{bmatrix} + \mathbf{O}_i \mathbf{x}_k \quad 5.75$$

$i = 1 \dots m$

Where \mathbf{T}_{O_i} and \mathbf{T}_{W_i} are the system and system noise Toeplitz matrices for the measurement i , and they are similar to equations (5.24) and (5.26), respectively. \mathbf{O}_i is the Observability matrix of the measurement i and it has the following form:

$$\mathbf{O}_i = [\mathbf{H}_i \ \mathbf{H}_i \mathbf{A} \ \dots \ \mathbf{H}_i \mathbf{A}^{LM_i-1}]^T \quad 5.76$$

Further to equation (5.76), the measurement segments are linked to the state by combining their Observability and Toeplitz matrices as follows:

$$\begin{bmatrix} z_{1_k} \\ \vdots \\ z_{1_{k+LM_1-1}} \\ \vdots \\ z_{m_k} \\ \vdots \\ z_{m_{k+LM_m-1}} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{O_1} \\ \vdots \\ \mathbf{T}_{O_m} \end{bmatrix} \begin{bmatrix} u_k \\ \vdots \\ u_{k+LM_1-1} \\ \vdots \\ u_k \\ \vdots \\ u_{k+LM_m-1} \end{bmatrix} + \begin{bmatrix} v_{1_k} \\ \vdots \\ v_{1_{k+LM_1-1}} \\ \vdots \\ v_{m_k} \\ \vdots \\ v_{m_{k+LM_m-1}} \end{bmatrix} + \begin{bmatrix} \mathbf{T}_{W_1} \\ \vdots \\ \mathbf{T}_{W_m} \end{bmatrix} \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+LM_1-1} \\ \vdots \\ \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+LM_m-1} \end{bmatrix} + \begin{bmatrix} \mathbf{O}_1 \\ \vdots \\ \mathbf{O}_m \end{bmatrix} \mathbf{x}_k \quad 5.77$$

The matrices $\begin{bmatrix} \mathbf{T}_{O_1} \\ \vdots \\ \mathbf{T}_{O_m} \end{bmatrix}$ and $\begin{bmatrix} \mathbf{T}_{W_1} \\ \vdots \\ \mathbf{T}_{W_m} \end{bmatrix}$ have dimensions of $\begin{bmatrix} \mathbf{T}_{O_1} \\ \vdots \\ \mathbf{T}_{O_m} \end{bmatrix} \in \mathbb{R}^{(\sum_{i=1}^m LM_i) \times (\sum_{i=1}^m LM_i)}$ and

$\begin{bmatrix} \mathbf{T}_{W_1} \\ \vdots \\ \mathbf{T}_{W_m} \end{bmatrix} \in \mathbb{R}^{(\sum_{i=1}^m LM_i) \times (n \sum_{i=1}^m LM_i)}$, respectively. The size of these matrices can be

minimized by merging the resultant repeated input and system noise segments as follows:

$$\begin{bmatrix} \mathbf{T}_{O_1} \\ \vdots \\ \mathbf{T}_{O_m} \end{bmatrix} [u_k \ \dots \ u_{k+LM_1-1} \ \dots \ u_k \ \dots \ u_{k+LM_m-1}]^T \equiv \mathbf{T}_{O_g} \begin{bmatrix} u_k \\ \vdots \\ u_{k+m_a-1} \end{bmatrix} \text{ and} \quad 5.78$$

$$\begin{bmatrix} \mathbf{T}_{W_1} \\ \vdots \\ \mathbf{T}_{W_m} \end{bmatrix} [\mathbf{w}_k \ \dots \ \mathbf{w}_{k+LM_1-1} \ \dots \ \mathbf{w}_k \ \dots \ \mathbf{w}_{k+LM_m-1}]^T \equiv \mathbf{T}_{W_g} \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+m_a-1} \end{bmatrix}$$

Where m_a is the maximum value of LM_i . The resultant matrices; \mathbf{T}_o and \mathbf{T}_w , are referred to as the General System and General System Noise Toeplitz matrices and they have

dimensions of $\mathbf{T}_{O_g} \in \mathbb{R}^{(\sum_{i=1}^m LM_i) \times m_a}$ and $\mathbf{T}_{W_g} \in \mathbb{R}^{(\sum_{i=1}^m LM_i) \times (nm_a)}$, respectively. By combining the Observability matrices and using equation (5.76), the General Observability matrix is obtained as follows:

$$\mathbf{O}_g = \begin{bmatrix} \mathbf{O}_1 \\ \vdots \\ \mathbf{O}_m \end{bmatrix} = \left[[\mathbf{H}_1 \dots \mathbf{H}_1 \mathbf{A}^{LM_1-1}] \dots [\mathbf{H}_m \dots \mathbf{H}_m \mathbf{A}^{LM_m-1}] \right]^T \quad 5.79$$

Substituting equations (5.78) and (5.79) into equation (5.77) gives the following:

$$\begin{bmatrix} z_{1k} \\ \vdots \\ z_{1k+LM_1-1} \\ \vdots \\ z_{mk} \\ \vdots \\ z_{mk+LM_m-1} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{O_1} \\ \vdots \\ \mathbf{T}_{O_m} \end{bmatrix} \begin{bmatrix} u_k \\ \vdots \\ u_{k+LM_1-1} \\ \vdots \\ u_k \\ \vdots \\ u_{k+LM_m-1} \end{bmatrix} + \begin{bmatrix} v_{1k} \\ \vdots \\ v_{1k+LM_1-1} \\ \vdots \\ v_{mk} \\ \vdots \\ v_{mk+LM_m-1} \end{bmatrix} + \begin{bmatrix} \mathbf{T}_{W_1} \\ \vdots \\ \mathbf{T}_{W_m} \end{bmatrix} \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+LM_1-1} \\ \vdots \\ \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+LM_m-1} \end{bmatrix} + \begin{bmatrix} \mathbf{O}_1 \\ \vdots \\ \mathbf{O}_m \end{bmatrix} \mathbf{x}_k \quad 5.80$$

$$\rightarrow \begin{bmatrix} z_{1k} \\ \vdots \\ z_{1k+LM_1-1} \\ \vdots \\ z_{mk} \\ \vdots \\ z_{mk+LM_m-1} \end{bmatrix} = \mathbf{T}_{O_g} \begin{bmatrix} u_k \\ \vdots \\ u_{k+m_a-1} \end{bmatrix} + \begin{bmatrix} v_{1k} \\ \vdots \\ v_{1k+LM_1-1} \\ \vdots \\ v_{mk} \\ \vdots \\ v_{mk+LM_m-1} \end{bmatrix} + \mathbf{T}_{W_g} \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+m_a-1} \end{bmatrix} + \mathbf{O}_g \mathbf{x}_k$$

In order to obtain the full state vector, \mathbf{O}_g must have rank of n .

The General Observability matrix gives an opportunity to obtain the states of an observable system from data segments originating from more than one measurement signal. In this research, the mechanism that relates data segments associated with multiple measurement signals to system states using the new form of Observability and Toeplitz matrices is explored. Not only do these matrices accommodate multiple measurements but they also reduce the estimation error due to modeling errors.

5.3.3.1.1. Derivation of the General Observability matrix

From equation (5.22), the measurement's segment is mapped to the states through the Observability matrix. Therefore, the Observability matrix must have full rank for the system to be observable in order to obtain the full state vector from a segment of the measurement signal. For systems with multiple measurements, there are usually multiple segments available and the selection algorithm of these segments is more complicated

than the one presented in section (5.2.1) for the case of a single measurement. This is due to the overlap of information that can be extracted from the measurements.

A segment from each of the measurements may be related to some of the states, and the length of each segment represents the number of the states that can be extracted from it. The connection between each segment and its corresponding states is obtained through a matrix referred to as the sub-Observability matrix. As long as this matrix is not singular, the states can be determined. Therefore, the choice of the segment length is based on the rank of the sub-Observability matrix and can be determined by the following heuristic rule:

The segment length can be increased as long as the new data points provide new information, and the resultant sub-Observability matrix has full rank.

Further to the above rule, the maximum length of the segment may be determined iteratively by increasing the segment length one data point at a time, and then checking the rank of the resultant sub-Observability matrix. If the new data point does not increase the rank of the sub-Observability matrix then the process is stopped, and the maximum segment length is made equal to the rank of the sub-Observability matrix as follows:

$$\text{rank} \left(\begin{bmatrix} \mathbf{H}_i \\ \mathbf{H}_i \mathbf{A} \\ \vdots \\ \mathbf{H}_i \mathbf{A}^{LM_i-1} \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} \mathbf{H}_i \\ \mathbf{H}_i \mathbf{A} \\ \vdots \\ \mathbf{H}_i \mathbf{A}^{LM_i} \end{bmatrix} \right) = LM_i, \text{ for } i = 1 \dots m \quad 5.81$$

Note that the portion $\mathbf{H}_i \mathbf{A}^{LM_i}$ does not provide any new information, thus it will not be included. This condition implies that the value of LM_i represents the maximum segment length for the i^{th} -measurement signal needed for extracting LM_i number of states. Any additional length beyond this value will not provide any additional information. The following example illustrates the process for obtaining the segment maximum lengths.

Example 5.3

If the system is a third order with the following system and measurement matrices:

$$\mathbf{A} = \begin{bmatrix} 1 & T_s & 0 \\ 0 & 1 & T_s \\ 0 & -a_1 T_s & 1 - a_2 T_s \end{bmatrix} \quad 5.82$$

and

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad 5.83$$

respectively, then two data segments may be used to extract the full state vector. Their corresponding lengths; LM_1 and LM_2 , are obtained using equation (5.81) and as follows.

The maximum segment length of the first measurement, LM_1 , is obtained as follows:

- Iteration 0:

$$\begin{aligned} \mathbf{H}_1 &= [1 \quad 0 \quad 0] \\ \text{rank}([\mathbf{H}_1]) &= \text{rank}(\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}) = 1 \end{aligned} \quad 5.84$$

- Iteration 1:

$$\begin{aligned} \mathbf{H}_1 \mathbf{A} &= [1 \quad T_s \quad 0] \\ \text{rank} \left(\begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1 \mathbf{A} \end{bmatrix} \right) &= \text{rank} \left(\begin{bmatrix} [1 & 0 & 0] \\ [1 & T_s & 0] \end{bmatrix} \right) = 2 \end{aligned} \quad 5.85$$

- Iteration 2:

$$\begin{aligned} \mathbf{H}_1 \mathbf{A}^2 &= [1 \quad 2T_s \quad T_s^2] \\ \text{rank} \left(\begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1 \mathbf{A} \\ \mathbf{H}_1 \mathbf{A}^2 \end{bmatrix} \right) &= \text{rank} \left(\begin{bmatrix} [1 & 0 & 0] \\ [1 & T_s & 0] \\ [1 & 2T_s & T_s^2] \end{bmatrix} \right) = 3 \end{aligned} \quad 5.86$$

- Iteration 3:

$$\begin{aligned} \mathbf{H}_1 \mathbf{A}^3 &= [1 \quad 3T_s - a_1 T_s^3 \quad 3T_s^2 - a_2 T_s^3] \\ \text{rank} \left(\begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1 \mathbf{A} \\ \mathbf{H}_1 \mathbf{A}^2 \\ \mathbf{H}_1 \mathbf{A}^3 \end{bmatrix} \right) &= \text{rank} \left(\begin{bmatrix} [1 & 0 & 0] \\ [1 & T_s & 0] \\ [1 & 2T_s & T_s^2] \\ [1 & 3T_s - a_1 T_s^3 & 3T_s^2 - a_2 T_s^3] \end{bmatrix} \right) = 3 \end{aligned} \quad 5.87$$

Iteration is stopped and $LM_1 = 3$.

The maximum segment length of the second measurement, LM_2 , is obtained as follows.

- Iteration 0:

$$\begin{aligned} \mathbf{H}_2 &= [0 \quad 1 \quad 0] \\ \text{rank}([\mathbf{H}_2]) &= \text{rank}(\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}) = 1 \end{aligned} \quad 5.88$$

- Iteration 1:

$$\mathbf{H}_2\mathbf{A} = [0 \quad 1 \quad T_s]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H}_2 \\ \mathbf{H}_2\mathbf{A} \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [0 \quad 1 \quad 0] \\ [0 \quad 1 \quad T_s] \end{bmatrix} \right) = 2 \quad 5.89$$

- Iteration 2:

$$\mathbf{H}_2\mathbf{A}^2 = [0 \quad 1 - a_1T_s^2 \quad 2T_s - a_2T_s^2]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H}_2 \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \end{bmatrix} \right) \text{rank} \left(\begin{bmatrix} [1 \quad 0 \quad 0] \\ [0 \quad 1 \quad T_s] \\ [0 \quad 1 - a_1T_s^2 \quad 2T_s - a_2T_s^2] \end{bmatrix} \right) = 2 \quad 5.90$$

Iteration is stopped and $LM_2 = 2$.

Note that $LM_1 + LM_2 = 3 + 2 > 3$ which means that there is an overlap between the extracted information from the two segments. Therefore, more constraints are needed to reduce the information overlap.

The condition of equation (5.81) may lead to overlaps between the information extracted from the segments originating from multiple measurements with the measurements' information themselves. Example 5.3 demonstrates that information from $\mathbf{H}_1\mathbf{A}$ in equation (5.85) overlaps with information from equations (5.84) and (5.88). Hence, it does not provide any additional information and can be omitted. Therefore, the condition of equation (5.81) can be modified to reduce the segment maximum length of the measurement i when its information overlaps the information provided by another measurement as follows.

The maximum length of the segment is obtained iteratively by increasing the segment length by one successive data point at a time, and then checking the rank of a matrix that consists of combining the resulting sub-Observability and the measurement matrices. If the incremental data point at certain time does not increase the rank of the combined matrix then the process is stopped, and the maximum segment length is equal to the number of iterations as follows:

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_i\mathbf{A} \\ \vdots \\ \mathbf{H}_i\mathbf{A}^{LM_i-1} \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_i\mathbf{A} \\ \vdots \\ \mathbf{H}_i\mathbf{A}^{LM_i} \end{bmatrix} \right) = m + LM_i - 1, \text{ for } i = 1 \dots m \quad 5.91$$

Note that the rows in the measurement matrix contribute to obtaining the segment length of the measurement i . Therefore, the resulting length of this method is less than the length obtained by equation (5.81) as illustrated in the following example.

Example 5.4:

Further to Example 5.3 and using the condition of equation (5.91), LM_1 and LM_2 are obtained as follows.

The maximum segment length of the first measurement (position), LM_1 , is obtained as follows:

- Iteration 0:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad 5.92$$

$$\text{rank}([\mathbf{H}]) = \text{rank}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\right) = 2$$

- Iteration 1:

$$\mathbf{H}_1\mathbf{A} = [1 \quad T_s \quad 0]$$

$$\text{rank}\left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & T_s & 0 \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\right) = 2 \quad 5.93$$

Iteration is stopped and $LM_1 = 1$. Note that $\mathbf{H}_1\mathbf{A}$ overlaps with the measurement information, hence, it is omitted.

The maximum segment length of the second measurement (velocity), LM_2 , is obtained as follows:

- Iteration 0, recalling equation (5.92):

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad 5.92$$

$$\text{rank}([\mathbf{H}]) = \text{rank}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\right) = 2$$

- Iteration 1:

$$\mathbf{H}_2\mathbf{A} = [0 \quad 1 \quad T_s]$$

$$\text{rank}\left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_2\mathbf{A} \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & T_s \end{bmatrix}\right) = 3 \quad 5.94$$

- Iteration 2:

$$\mathbf{H}_2\mathbf{A}^2 = [0 \quad 1 - a_1T_s^2 \quad 2T_s - a_2T_s^2]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \end{bmatrix} \right) \text{rank} \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & T_s \\ [0 \quad 1 - a_1T_s^2 \quad 2T_s - a_2T_s^2] \end{bmatrix} \right) = 3 \quad 5.95$$

Iteration is stopped and $LM_2 = 2$. Note that $\mathbf{H}_2\mathbf{A}^2$ does not add to the rank and is omitted.

Note that the resulting segment lengths obtained in Example 5.4 are smaller than their corresponding lengths obtained in Example 5.3. Using the results of equations (5.93) and (5.95), the resultant rank is equal to the number of the states; therefore, the General Observability matrix is then defined using equation (5.72) as follows:

$$\mathbf{O}_g = [\mathbf{H}_1 \quad \mathbf{H}_2 \quad \mathbf{H}_2\mathbf{A}]^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & T_s \end{bmatrix} \quad 5.96$$

The relation between the measurement segment vector and the state vector is defined using equation (5.71) as follows:

$$\begin{bmatrix} z_{1k} \\ z_{2k} \\ z_{2k+1} \end{bmatrix} = \mathbf{T}_{og} \begin{bmatrix} u_k \\ u_{k+1} \end{bmatrix} + \begin{bmatrix} v_{1k} \\ v_{2k} \\ v_{2k+1} \end{bmatrix} + \mathbf{T}_{wg} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \end{bmatrix} + \mathbf{O}_g \mathbf{x}_k \quad 5.97$$

Rearranging equation (5.97) gives the following:

$$\mathbf{x}_k = \mathbf{O}_g^{-1} \left(\begin{bmatrix} z_{1k} \\ z_{2k} \\ z_{2k+1} \end{bmatrix} - \mathbf{T}_{og} \begin{bmatrix} u_k \\ u_{k+1} \end{bmatrix} - \begin{bmatrix} v_{1k} \\ v_{2k} \\ v_{2k+1} \end{bmatrix} - \mathbf{T}_{wg} \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k+1} \end{bmatrix} \right) \quad 5.98$$

The inverse of the General Observability matrix is obtained by using equation (5.96) as follows:

$$\mathbf{O}_g^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{T_s} & \frac{1}{T_s} \end{bmatrix} \quad 5.99$$

Equations (5.98) and (5.99) show that only the velocity segment is used to extract the third state (acceleration). This makes sense knowing that the acceleration is the second derivative of the position (and its uncertainties) and the first derivative of the velocity (and its uncertainties). In general, the first derivative of the velocity's

uncertainties has a lower bound than the second derivative of the position's uncertainties; therefore, the error in estimating the third state is reduced by using the velocity signal.

The General Observability matrix should have a rank of n to be observable. This implies that it has dimensions of $\mathbf{O} \in \mathbb{R}^{n \times n}$. However, choosing the segments' length using equation (5.91) alone may lead to overlaps between the segments' extracted information as shown in the following example.

Example 5.5

Further to Example 5.2, the segments' lengths of the three measurements are chosen as follows.

The segment length of the first measurement is obtained as follows:

- Iteration 0:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad 5.100$$

$$\text{rank}([\mathbf{H}]) = \text{rank}\left(\begin{bmatrix} [1 & 0 & 0 & 0 & 0 & 0] \\ [0 & 0 & 1 & 0 & 0 & 0] \\ [0 & 0 & 0 & 0 & 1 & 0] \end{bmatrix}\right) = 3$$

- Iteration 1:

$$\mathbf{H}_1\mathbf{A} = [1 \quad 0.001 \quad 0 \quad 0 \quad 0 \quad 0]$$

$$\text{rank}\left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} [1 & 0 & 0 & 0 & 0 & 0] \\ [0 & 0 & 1 & 0 & 0 & 0] \\ [0 & 0 & 0 & 0 & 1 & 0] \\ [1 & 0.001 & 0 & 0 & 0 & 0] \end{bmatrix}\right) = 4 \quad 5.101$$

- Iteration 2:

$$\mathbf{H}_1\mathbf{A}^2 = [0.997 \quad -0.001 \quad 0.001 \quad 0.001 \quad 0.001 \quad 0.001]$$

$$\text{rank}\left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_1\mathbf{A}^2 \end{bmatrix}\right) = \text{rank}\left(\begin{bmatrix} [1 & 0 & 0 & 0 & 0 & 0] \\ [0 & 0 & 1 & 0 & 0 & 0] \\ [0 & 0 & 0 & 0 & 1 & 0] \\ [1 & 0.001 & 0 & 0 & 0 & 0] \\ 10^{-3} \times [997 & -1 & 1 & 1 & 1 & 1] \end{bmatrix}\right) = 5 \quad 5.102$$

- Iteration 3:

$$\mathbf{H}_1\mathbf{A}^3 = [1.002 \quad 0.005 \quad -0.001 \quad -0.001 \quad -0.001 \quad -0.001]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_1\mathbf{A}^2 \\ \mathbf{H}_1\mathbf{A}^3 \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ [1 \quad 0.001 \quad 0 \quad 0 \quad 0 \quad 0] \\ 10^{-3} \times [997 \quad -1 \quad 1 \quad 1 \quad 1 \quad 1] \\ -10^{-3} \times [-1002 \quad -5 \quad 1 \quad 1 \quad 1 \quad 1] \end{bmatrix} \right) = 5 \quad 5.103$$

Iteration is stopped and $LM_1 = 3$.

The segment length of the second measurement is obtained as follows:

- Iteration 0, recall equation (5.100):

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\text{rank}([\mathbf{H}]) = \text{rank} \left(\begin{bmatrix} [1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0] \\ [0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0] \\ [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0] \end{bmatrix} \right) = 3 \quad 5.100$$

- Iteration 1:

$$\mathbf{H}_2\mathbf{A} = [0 \quad 0 \quad 1 \quad 0.001 \quad 0 \quad 0]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_2\mathbf{A} \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ [0 \quad 0 \quad 1 \quad 0.001 \quad 0 \quad 0] \end{bmatrix} \right) = 4 \quad 5.104$$

- Iteration 2:

$$\mathbf{H}_2\mathbf{A}^2 = [0.001 \quad 0.001 \quad 0.998 \quad 0 \quad 0.001 \quad 0.001]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ [0 \quad 0 \quad 1 \quad 0.001 \quad 0 \quad 0] \\ [0.001 \quad 0.001 \quad 0.998 \quad 0 \quad 0.001 \quad 0.001] \end{bmatrix} \right) = 5 \quad 5.105$$

- Iteration 3:

$$\mathbf{H}_2\mathbf{A}^3 = [-0.001 \quad -0.001 \quad 1 \quad 0.003 \quad 0 \quad 10^{-6}]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \\ \mathbf{H}_2\mathbf{A}^3 \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [1 \ 0 \ 0 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 1 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0] \\ [0 \ 0 \ 1 \ 0.001 \ 0 \ 0] \\ [0.001 \ 0.001 \ 0.998 \ 0 \ 0.001 \ 0.001] \\ [-0.001 \ -0.001 \ 1 \ 0.003 \ 0 \ 10^{-6}] \end{bmatrix} \right) = 6 \quad 5.106$$

- Iteration 4:

$$\mathbf{H}_2\mathbf{A}^4 = [0.005 \quad 0.005 \quad 0.993 \quad -0.003 \quad 0.002 \quad 0.002]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \\ \mathbf{H}_2\mathbf{A}^3 \\ \mathbf{H}_2\mathbf{A}^4 \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [1 \ 0 \ 0 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 1 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0] \\ [0 \ 0 \ 1 \ 0.001 \ 0 \ 0] \\ 10^{-3} \times [1 \ 1 \ 998 \ 0 \ 1 \ 1] \\ 10^{-3} \times [-1 \ -1 \ 1000 \ 3 \ 0 \ 10^{-3}] \\ 10^{-3} \times [5 \ 5 \ 993 \ -3 \ 2 \ 2] \end{bmatrix} \right) = 6 \quad 5.107$$

Iteration is stopped and $LM_2 = 4$.

The segment length of the third measurement is obtained as follows:

- Iteration 0:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\text{rank}([\mathbf{H}]) = \text{rank} \left(\begin{bmatrix} [1 \ 0 \ 0 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 1 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0] \end{bmatrix} \right) = 3 \quad 5.108$$

- Iteration 1:

$$\mathbf{H}_3\mathbf{A} = [0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0.001]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_3\mathbf{A} \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [1 \ 0 \ 0 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 1 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0.001] \end{bmatrix} \right) = 4 \quad 5.109$$

- Iteration 2:

$$\mathbf{H}_3\mathbf{A}^2 = [0.001 \quad 0.001 \quad 0.001 \quad 0.001 \quad 0.998 \quad 0]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_3\mathbf{A} \\ \mathbf{H}_3\mathbf{A}^2 \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [1 \ 0 \ 0 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 1 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0.001] \\ [0.001 \ 0.001 \ 0.001 \ 0.001 \ 0.998 \ 0] \end{bmatrix} \right) = 5 \quad 5.110$$

- Iteration 3:

$$\mathbf{H}_3\mathbf{A}^3 = [-0.001 \quad -0.001 \quad 0 \quad 10^{-6} \quad 1 \quad 0.003]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_3\mathbf{A} \\ \mathbf{H}_3\mathbf{A}^2 \\ \mathbf{H}_3\mathbf{A}^3 \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [1 \ 0 \ 0 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 1 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0.001] \\ [0.001 \ 0.001 \ 0.001 \ 0.001 \ 0.998 \ 0] \\ [-0.001 \ -0.001 \ 0 \ 10^{-6} \ 1 \ 0.003] \end{bmatrix} \right) = 6 \quad 5.111$$

- Iteration 4:

$$\mathbf{H}_3\mathbf{A}^4 = [0.005 \quad 0.005 \quad 0.002 \quad 0.002 \quad 0.993 \quad -0.003]$$

$$\text{rank} \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_3\mathbf{A} \\ \mathbf{H}_3\mathbf{A}^2 \\ \mathbf{H}_3\mathbf{A}^3 \\ \mathbf{H}_3\mathbf{A}^4 \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [1 \ 0 \ 0 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 1 \ 0 \ 0 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0] \\ [0 \ 0 \ 0 \ 0 \ 1 \ 0.001] \\ [0.001 \ 0.001 \ 0.001 \ 0.001 \ 0.998 \ 0] \\ [-0.001 \ -0.001 \ 0 \ 10^{-6} \ 1 \ 0.003] \\ [0.005 \ 0.005 \ 0.002 \ 0.002 \ 0.993 \ -0.003] \end{bmatrix} \right) = 6 \quad 5.112$$

Iteration is stopped and $LM_3 = 4$.

Note that $\sum_{i=1}^3 LM_i = 3 + 4 + 4 = 11 > 6$, thus there are overlaps in the extracted information. The General Observability matrix is obtained using equation (5.72) as follows:

$$\mathbf{O}_g = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 10^{-3} & 0 & 0 & 0 & 0 \\ 0.997 & -10^{-3} & 10^{-3} & 10^{-3} & 10^{-3} & 10^{-3} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 10^{-3} & 0 & 0 \\ 10^{-3} & 10^{-3} & 0.998 & 0 & 10^{-3} & 10^{-3} \\ -10^{-3} & -10^{-3} & 1 & 3 \times 10^{-3} & 0 & 10^{-6} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 10^{-3} \\ 10^{-3} & 10^{-3} & 10^{-3} & 10^{-3} & 0.998 & 0 \\ -10^{-3} & -10^{-3} & 0 & 10^{-6} & 1 & 3 \times 10^{-3} \end{bmatrix} \quad 5.113$$

To remove the overlapping information, the segments' lengths must be reduced. However, the reduced General Observability matrix is not unique. In this thesis, a novel method is used to reduce the segments' lengths to their minimum values (this reduces the uncertainties), and to obtain the General Observability matrix in a square matrix form.

Note that the General Observability matrix in equation (5.113) is not a square matrix. Mathematically, this means that the number of equations is larger than the number of unknowns and this may result in one of the following problems:

- Adding extra source of uncertainties without improving the information provided by the segments. As the segment length increases, the amplitude of the uncertainties becomes significantly large as higher derivatives of the noise are included. Moreover, larger General Observability matrices need larger amounts of memory and computational time especially when the inverse matrix is calculated.
- The overlapped information signals may contradict one another especially when modeling errors are present. This may lead to a system of equations with no solution as illustrated in Example 5.6.

Example 5.6:

Assume that two unknowns a and b are described by the following three equations:

$$2a + 2b = 1, 2a + b = 2 \text{ and } a + 2b = 1 \quad 5.114$$

Taking the first two equations, the solution of a and b are equal to 1.5 and -1, respectively. Note that the solution of the first two equations is not a solution for the third equation where $1.5 + 2 \times -1 = -0.5 \neq 1$. This means that there is no solution for this system of equations. Similarly, if the uncertainties are significantly large and/or modeling errors are present when obtaining equation (5.113), then the full state vector may not be extracted although the General Observability matrix has full rank. Therefore, the General

Observability matrix must be square matrix with rank equal to n and must consist of minimum segments' lengths in order to uniquely extract the full state vector from the measurement segments with minimum errors. In order to achieve these objectives, other conditions and constraints are needed. This matter is addressed in the following section.

5.3.3.2. The selection of the General Observability matrix for overlapped information

If the General Observability matrix leads to overlapping information, then the segment lengths ($LM_i, = 1 \dots m$) must be reduced such that a total length of n is obtained. This gives a possibility of several non-singular forms of the General Observability matrix depending on the combined choices of values for LM_i . In this subsection, a novel general selection procedure is developed to obtain a General Observability matrix that is square with full rank and independent of system parameters. The proposed algorithm is based on the following observations.

Further to examples (Example 5.2) and (Example 5.5) and using the same method of examples (Example 5.4) and (Example 5.5), the following observations can be made.

Observation 1: Further to Example 5.5, if the measurement matrix has a rank of 1 (only one measurement is available), then the General Observability matrix is singular unless the location of the measured state, LMS_1 , has a value of 3 or 5 as shown in table 5.6. These values change as the system parameters change.

Location of z_1 (LMS_1)	1	2	3	4	5	6
Rank of $[\mathbf{H} \ \mathbf{H}_1\mathbf{A} \ \mathbf{H}_1\mathbf{A}^2 \ \mathbf{H}_1\mathbf{A}^3 \ \mathbf{H}_1\mathbf{A}^4 \ \mathbf{H}_1\mathbf{A}^5]^T$	4	4	6	5	6	5

Table 5.6: The rank of the General Observability matrix of Example 5.2 given six states and one measurement; i.e. $\mathbf{H} = \mathbf{H}_1$.

In conclusion, if a system has a fewer number of measurements, m , compared to its degrees of freedom, $NSUB$; i.e. $m < NSUB$, then the Observability matrix has a high probability of being singular.

Observation 2: Further to Example 5.5, if the measurement matrix has a rank of 2 (two measurements are available), then selection of the LMS_1 and LMS_2 leads to the General Observability matrix having several forms, some of which are singular as shown in table 5.7. These forms are summarized as follows:

- If both measurements belong to the same sub-system, then the General Observability has a high probability of being singular.

- If the measurements belong to different sub-systems, then the General Observability's singularity depends on the system parameters and is obtained by trial and error (iteration). For example 2, if the first measurement belongs to the second sub-system and the second measurement belongs to the third sub-system then the Observability matrix has full rank as shown in table 5.7. If the parameters change, these locations change as well.

In conclusion, for a system with multiple degrees of freedom with $m < NSUB$, each possible length of the measurements' segments should be examined to obtain the full rank Observability matrix.

Location of z_1 (LMS_1)	Location of z_2 (LMS_2)	Rank of $\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_1\mathbf{A}^2 \\ \mathbf{H}_1\mathbf{A}^3 \\ \mathbf{H}_1\mathbf{A}^4 \end{bmatrix}$	Rank of $\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_1\mathbf{A}^2 \\ \mathbf{H}_1\mathbf{A}^3 \\ \mathbf{H}_2\mathbf{A} \end{bmatrix}$	Rank of $\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_1\mathbf{A}^2 \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \end{bmatrix}$	Rank of $\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \\ \mathbf{H}_2\mathbf{A}^3 \end{bmatrix}$	Rank of $\begin{bmatrix} \mathbf{H} \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_2\mathbf{A}^2 \\ \mathbf{H}_2\mathbf{A}^3 \\ \mathbf{H}_2\mathbf{A}^4 \end{bmatrix}$
1	2	4	4	4	4	4
1	3	5	6	5	6	6
1	4	5	6	6	6	6
1	5	5	6	5	6	6
1	6	5	6	6	6	6
2	3	5	6	6	6	6
2	4	5	6	5	6	6
2	5	5	6	6	6	6
2	6	5	6	5	6	6
3	4	5	4	4	5	6
3	5	6	6	5	6	6
3	6	6	6	6	6	6
4	5	6	6	6	6	6
4	6	6	6	5	6	6
5	6	5	4	4	5	6

Table 5.7: The rank of the General Observability matrix for example 2 given six states and two measurements i.e. $\mathbf{H} = [\mathbf{H}_1 \ \mathbf{H}_2]^T$.

Observation 3: if the measurement matrix has rank of 3 (equal to the number of the sub-systems, $NSUB$, in the multiple degrees of freedom system), then the rank of the General Observability matrix depends on the location of the measured states as shown in table 5.8 and must be outlined as follows:

- If at least two measurements belong to the same sub-system, then the General Observability matrix has a high probability of being singular.
- If the measurements belong to different sub-systems, then the General Observability matrix has a high probability of having full rank, depending on the location of the measured states and the system's parameters. From example 2, if the measured states are the first states of each corresponding sub-systems then the General Observability matrix is guaranteed to have full rank and is independent of the system's parameters.

Location of z_1 (LMS_1)	Location of z_2 (LMS_2)	Location of z_3 (LMS_3)	Rank of $\begin{bmatrix} H \\ H_3A \\ H_3A^2 \\ H_3A^3 \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_2A \\ H_2A^2 \\ H_2A^3 \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_1A \\ H_1A^2 \\ H_1A^3 \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_1A \\ H_2A \\ H_3A \\ H_3A^2 \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_2A \\ H_3A \\ H_3A^2 \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_2A \\ H_2A^2 \\ H_3A \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_1A \\ H_2A \\ H_2A^2 \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_1A \\ H_1A^2 \\ H_2A \end{bmatrix}$	Rank of $\begin{bmatrix} H \\ H_1A \\ H_1A^2 \\ H_3A \end{bmatrix}$	
1	2	3	6	5	5	5	5	5	6	5	4	5
1	2	4	6	5	5	5	6	5	6	5	4	5
1	2	5	6	5	5	5	5	5	6	5	4	5
1	2	6	6	5	5	5	6	5	6	5	4	5
1	3	4	6	5	6	5	5	6	4	5	5	5
1	3	5	6	6	5	6	6	6	6	6	6	6
1	3	6	6	6	6	6	6	6	6	6	6	6
1	4	5	6	6	6	6	6	6	6	6	6	6
1	4	6	6	6	5	6	6	5	6	5	6	6
1	5	6	6	5	6	5	5	6	4	5	5	5
2	3	4	6	5	6	5	5	6	4	5	5	6
2	3	5	6	6	5	6	6	6	6	6	6	6
2	3	6	6	6	6	6	6	6	6	6	6	6
2	4	5	6	6	6	6	6	6	6	6	6	6
2	4	6	6	6	5	6	6	6	6	6	6	6
2	5	6	6	5	6	5	5	6	4	5	5	6
3	4	5	6	6	5	5	5	5	6	5	4	5
3	4	6	6	6	5	5	6	5	6	5	4	5
3	5	6	6	5	6	5	5	6	4	5	5	5
4	5	6	6	5	6	5	5	6	4	5	5	6

Table 5.8: The rank of the General Observability matrix of example 2 given six states and three measurements i.e. $\mathbf{H} = [\mathbf{H}_1 \quad \mathbf{H}_2 \quad \mathbf{H}_3]^T$.

Example 5.7:

Further to the system in Example 5.1, the General Observability matrix for different single measurement locations is obtained in table 5.9. It can be observed from table 5.9 that the Observability matrix does not have full rank and is not independent of the parameters unless the measurement corresponds to the first state. The same conclusion is obtained from tables (5.7) and (5.8).

	$\mathbf{H} = [1 \ 0 \ 0]$	$\mathbf{H} = [0 \ 1 \ 0]$	$\mathbf{H} = [0 \ 0 \ 1]$
General Observability matrix	$\begin{bmatrix} 1 & 0 & 0 \\ 1 & T_s & 0 \\ 1 & 2T_s & T_s^2 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 - 28T_s \\ 0 & T_s & 1.94T_s \end{bmatrix}^T$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -28 & -54.32 \\ 1 & 0.94 & 0.8836 - 28T_s \end{bmatrix}^T$
Rank	3	2	2

Table 5.9: The General Observability matrix and its rank for example 1 given three states and one measurement.

In conclusion, the General Observability matrix of the General Observability Canonical Form is guaranteed to have a full rank that is independent of the parameters if at least each sub-system's first state is measured. Otherwise the General Observability matrix's rank depends on the system's parameters, measurement signal locations, and the selection of the data segments' lengths.

5.3.3.2.1. The General Observability matrix's conditions

In order to force the resultant General Observability matrix to be fully ranked, time invariant and independent of the system's parameters, the measurement matrix must qualify the following:

$$m \geq NSUB \quad 5.115$$

$$\begin{aligned} \text{For any } STSUB_i \in \{STSUB_1, \dots, STSUB_{NSUB}\} \text{ there is} \\ LMS_j \in \{LMS_1, \dots, LMS_m\} \text{ where } STSUB_i = LMS_j \end{aligned} \quad 5.116$$

Where LMS_j is the location of the measured state j and it is defined by equation (5.52), and $STSUB_i$ is the row that represents the beginning of the subsystem i . The conditions of equations (5.115) and (5.116) imply that if a multi-degree of freedom system exists with order of $NSUB$, then at least one measurement from each subsystem is needed and this measurement should represent the first state in that subsystem.

5.3.3.2.2. The General Observability matrix's – Selection procedure

If the system satisfies the conditions of equations (5.115) and (5.116), then the best selection of the General Observability matrix is obtained as follows.

- 1- The measurement matrix is divided into blocks; each block is related to one of the sub-systems.
- 2- To extract the hidden states, the measurements are taken in segments. The segment length of the measurement i , LM_i , is defined as a function of the measured states' locations of the measurement signals i and $i + 1$; i.e. LMS_i and LMS_{i+1} , respectively. This can be expressed as follows:

$$LM_i = \begin{cases} LMS_{i+1} - LMS_i & i < m \\ n - LMS_m + 1 & i = m \end{cases} \quad 5.117$$

Note that if the system qualifies equations (5.115) and (5.116), then the segment length can be obtained directly by equation (5.117) and there is no need to use the iterative process described in section (5.3.3.1.1). This reduces the time needed to prepare the segments. Moreover, the resultant General Observability matrix is a square matrix with rank of n .

- 3- The General Observability matrix is then divided into m –blocks in a block-diagonal form as shown in equation (5.118).

$$\mathbf{O}_g = \begin{bmatrix} \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1 \mathbf{A} \\ \vdots \\ \mathbf{H}_1 \mathbf{A}^{LM_1-1} \end{bmatrix} \\ \begin{bmatrix} \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_2 \mathbf{A}^{LM_2-1} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \mathbf{H}_m \\ \vdots \\ \mathbf{H}_m \mathbf{A}^{LM_m-1} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \mathbf{O}_{sub_1} & \mathbf{0}_{LM_1 \times (n-LM_1)} & & \\ \mathbf{0}_{LM_2 \times (LMS_2-1)} & \mathbf{O}_{sub_2} & & \\ & \vdots & \dots & \\ \mathbf{0}_{LM_m \times (LMS_m-1)} & & & \mathbf{O}_{sub_m} \end{bmatrix} \quad 5.118$$

Each block consists of the following:

- An Observability matrix that is similar to equation (5.44) and has a size of $LM_i \times LM_i$, where i is the block's order. This matrix is referred to as the sub-Observability matrix of block i , \mathbf{O}_{sub_i} . This matrix has full rank and its location in the General Observability matrix starts at row and column equal to LMS_i .

- Two zero matrices with sizes of $LM_i \times (LMS_i - 1)$ and $LM_i \times (n - LMS_i - LM_i + 1)$. The former matrix is located to the left side of the sub-Observability matrix, \mathbf{O}_{sub_i} , while the latter is located to its right side. Note that the sub-Observability and the zero matrices have the same number of rows, and the sum of their columns is equal to the number of states, n .
- 4- The inverse of the General Observability matrix has size of $\mathbf{O}_g^{-1} \in \mathbb{R}^{n \times n}$ and it can be obtained by replacing each sub-Observability matrix with its inverse.

$$\mathbf{O}_g^{-1} = \begin{bmatrix} \mathbf{O}_{sub_1}^{-1} & \mathbf{0}_{LM_1 \times (n-LM_1)} & & & \\ \mathbf{0}_{LM_2 \times (LMS_2-1)} & \mathbf{O}_{sub_2}^{-1} & \mathbf{0}_{LM_2 \times (n-LMS_2-LM_2+1)} & & \\ & \vdots & & & \\ & \mathbf{0}_{LM_m \times (LMS_m-1)} & & \mathbf{O}_{sub_m}^{-1} & \end{bmatrix} \quad 5.119$$

Example 5.8:

Further to Example 5.2, if the measurement matrix is defined as follows:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad 5.120$$

Then the locations of the measured states, LMS_i , are set to the values of 1, 3 and 5. The segments' lengths are defined using equation (5.117) as follows:

$$\begin{aligned} LM_1 &= LMS_2 - LMS_1 = 3 - 1 = 2 \\ LM_2 &= LMS_3 - LMS_2 = 5 - 3 = 2 \\ LM_3 &= n + 1 - LMS_3 = 6 + 1 - 5 = 2 \end{aligned} \quad 5.121$$

The General Observability matrix is then defined using equation (5.72) as follows:

$$\mathbf{O} = \begin{bmatrix} [\mathbf{H}_1] \\ [\mathbf{H}_1\mathbf{A}] \\ [\mathbf{H}_2] \\ [\mathbf{H}_2\mathbf{A}] \\ [\mathbf{H}_3] \\ [\mathbf{H}_3\mathbf{A}] \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} & 0 & 0 & 0 & 0 \\ \begin{bmatrix} 1 & T_s \end{bmatrix} & 0 & 0 & 0 & 0 \\ 0 & 0 & \begin{bmatrix} 1 & 0 \end{bmatrix} & 0 & 0 \\ 0 & 0 & \begin{bmatrix} 1 & T_s \end{bmatrix} & 0 & 0 \\ 0 & 0 & 0 & 0 & \begin{bmatrix} 1 & 0 \end{bmatrix} \\ 0 & 0 & 0 & 0 & \begin{bmatrix} 1 & T_s \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \mathbf{O}_{sub_1} & \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \mathbf{O}_{sub_2} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} & \mathbf{O}_{sub_3} \end{bmatrix} \quad 5.122$$

Where $\mathbf{O}_{sub_1} = \mathbf{O}_{sub_2} = \mathbf{O}_{sub_3} = \begin{bmatrix} 1 & 0 \\ 1 & T_s \end{bmatrix}$.

If the measurement matrix is defined as follows:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad 5.123$$

Then $LMS_i, i = 1 \dots 4$ have values of 1, 2, 3 and 5, and the segments' lengths are defined using equation (5.117) as follows:

$$\begin{aligned} LM_1 &= LMS_2 - LMS_1 = 2 - 1 = 1 \\ LM_2 &= LMS_3 - LMS_2 = 3 - 2 = 1 \\ LM_3 &= LMS_4 - LMS_3 = 5 - 3 = 2 \\ LM_4 &= n + 1 - LMS_4 = 6 + 1 - 5 = 2 \end{aligned} \quad 5.124$$

The General Observability matrix is then defined using equation (5.72) as follows:

$$O = \begin{bmatrix} [H_1] \\ [H_2] \\ [H_3] \\ [H_3A] \\ [H_4] \\ [H_4A] \end{bmatrix} = \begin{bmatrix} [1] & 0 & 0 & 0 & 0 & 0 \\ 0 & [1] & 0 & 0 & 0 & 0 \\ 0 & 0 & [1 & 0] & 0 & 0 \\ 0 & 0 & [1 & T_s] & 0 & 0 \\ 0 & 0 & 0 & 0 & [1 & 0] \\ 0 & 0 & 0 & 0 & [1 & T_s] \end{bmatrix} = \begin{bmatrix} O_{sub_1} & O_{1 \times 1} & O_{1 \times 2} & O_{1 \times 2} \\ O_{1 \times 1} & O_{sub_2} & O_{1 \times 2} & O_{1 \times 2} \\ O_{2 \times 1} & O_{2 \times 1} & O_{sub_3} & O_{2 \times 2} \\ O_{2 \times 1} & O_{2 \times 1} & O_{2 \times 2} & O_{sub_4} \end{bmatrix} \quad 5.125$$

Where $O_{sub_1} = O_{sub_2} = [1]$ and $O_{sub_3} = O_{sub_4} = \begin{bmatrix} 1 & 0 \\ 1 & T_s \end{bmatrix}$.

5.3.3.2.3. The General Observability matrix's – Modified selection procedure

If the system satisfies only the condition of equation (5.115); i.e. number of measurements is larger than the number of subsystems and at least one subsystem does not have a measurement related to its first state, then the best selection of the General Observability matrix is **iteratively** obtained as follows.

STEP 1: The system matrix is divided into blocks; each block is related to one of the subsystems that starts at row $SSUB_i$ and ends at row $ENDSUB_i, i = 1 \dots NSUB$ as shown in Fig 5.8.

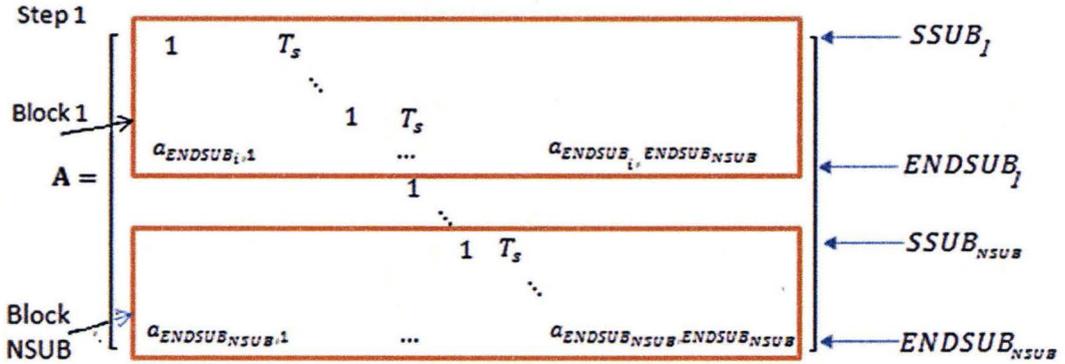


Fig 5.8: The modified selection procedure of the General Observability matrix - Step 1: dividing the system matrix into blocks

Example 5.9:

Further to Example 5.2, where the system matrix is defined as follows:

$$A = \begin{bmatrix} 1 & T_s & 0 & 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 1 & T_s & 0 & 0 \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & 0 & 0 & 0 & 1 & T_s \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} \quad 5.126$$

the system is divided into three blocks.

- Block 1 starts at row $SSUB_1 = 1$ and ends at rows $ENDSUB_1 = 2$
- Block 2 starts at row $SSUB_2 = 3$ and ends at rows $ENDSUB_2 = 4$
- Block 3 starts at row $SSUB_3 = 5$ and ends at rows $ENDSUB_3 = 6$

STEP 2: From Fig 5.9, each block is assigned to a portion of the state vector located between rows $SSUB_i$ and $ENDSUB_i$, $i = 1 \dots NSUB$, and it has length of n_i defined as follows:

$$n_i = \begin{cases} SSUB_{i+1} - SSUB_i & i < NSUB \\ n - SSUB_i + 1 & i = NSUB \end{cases}, i = 1 \dots NSUB \quad 5.127$$

Example 5.10:

Further to Example 5.9, the lengths of the state vector portions are obtained as follows:

$$\begin{aligned}
 n_1 &= SSUBS_2 - SSUBS_1 = 3 - 1 = 2 \\
 n_2 &= SSUBS_3 - SSUBS_2 = 5 - 3 = 2 \\
 n_3 &= n - SSUBS_3 + 1 = 6 - 5 + 1 = 2
 \end{aligned} \tag{5.128}$$

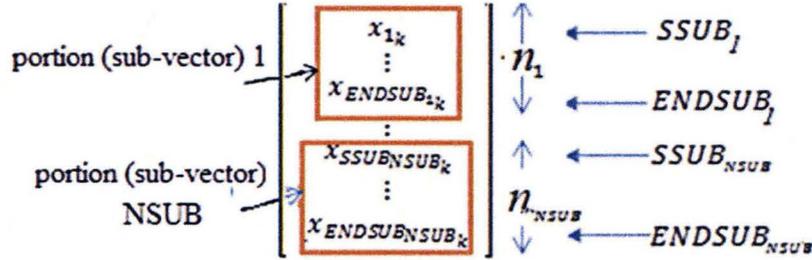


Fig 5.9: The modified selection procedure of the General Observability matrix - Step 2: dividing the state vector into sub-vectors according to step 1 blocks

STEP 3: Using the relation between the states and the measurements described by equation (5.52) and using the state portions (sub-vectors) of step 2, the measurement vector can be divided into groups. Each group is assigned to one of the blocks using the location of the measured state as follows:

$$z_j \in Blok_i \text{ If } SSUB_i \leq LMS_j \leq ENDSUB_i \tag{5.129}$$

Each block has a number of measurements m_i associated with it, where $i = 1 \dots NSUB$. The measurement matrix defined by equations (5.52) and (5.53) is sorted according to the location of the measured states. The measurement matrix is then divided into sub-matrices according to m_i . These sub-matrices are referred to as measurement blocks. Each measurement block is assigned to a portion of the measurement vector of length m_i located between the row indices $\sum_{l=1}^{i-1} m_l + 1$ and $\sum_{l=1}^i m_l$ of its corresponding measurement vector, \mathbf{z}_k as shown in Fig 5.10. The measurement sub-vector of block i is then defined as follows:

$$\mathbf{z}_{sub_{i_k}} = \left[Z_{(\sum_{l=1}^{i-1} m_l)+1_k} \quad Z_{(\sum_{l=1}^{i-1} m_l)+2_k} \quad \dots \quad Z_{(\sum_{l=1}^i m_l)_k} \right]^T \tag{5.130}$$

The locations of the measured states, which are represented by this sub-vector, are defined as follows:

$$LMS_{(\sum_{l=1}^{i-1} m_l)+1}, LMS_{(\sum_{l=1}^{i-1} m_l)+2}, \dots, LMS_{(\sum_{l=1}^i m_l)} \tag{5.131}$$

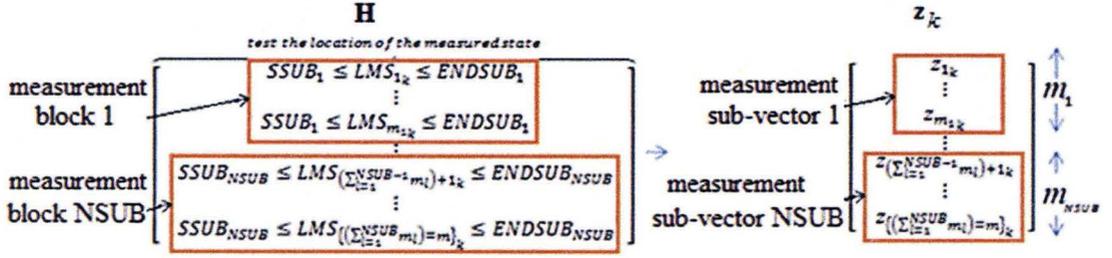


Fig 5.10: The modified selection procedure of the General Observability matrix - Step 3:
Assigning the measurements sub-vectors

For example, if the system has 2 measurements for the first sub-system, and one measurement for the second sub-system, then measurement sub-vector 1 has two measurements; i.e. ($m_1 = 2$) and measurement sub-vector 2 has one measurement; i.e. ($m_2 = 1$). According to equations (5.130) and (5.131), the measurements of sub-vector 1 are $z_{(\sum_{l=1}^0 m_l)+1k} = z_{0+1k} = z_{1k}$ and $z_{(\sum_{l=1}^1 m_l)_k} = z_{m_1k} = z_{2k}$ located at the rows $LMS_{(\sum_{l=1}^0 m_l)+1k} = LMS_{0+1k} = LMS_1$ and $LMS_{(\sum_{l=1}^1 m_l)_k} = LMS_{m_1k} = LMS_2$, respectively. The measurement of sub-vector 2 is $z_{(\sum_{l=1}^1 m_l)+1k} = z_{2+1k} = z_{3k}$ located at the row $LMS_{(\sum_{l=1}^1 m_l)+1k} = LMS_{2+1k} = LMS_3$.

Example 5.11:

Further to Example 5.9, if the measurement matrix is defined as follows:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \mathbf{H}_3 \end{bmatrix} \quad 5.132$$

Then the measurement vector is divided into blocks as follows:

$$z_1 \in \text{Blok}_1 \text{ because } SSUB_1 \leq LMS_{1k} \leq ENDSUB_1$$

$$z_2 \in \text{Blok}_2 \text{ because } SSUB_2 \leq LMS_{2k} \leq ENDSUB_2$$

$$z_3 \in \text{Blok}_3 \text{ because } SSUB_3 \leq LMS_{3k} \leq ENDSUB_3$$

$$\rightarrow m_1 = 1, m_2 = 1 \text{ and } m_3 = 1.$$

5.133

STEP 4: The initial segment length of each measurement is obtained by processing each block of the system matrix individually (later, this length may change). This can be done for block j by modifying equation (5.117) as follows:

$$LM_i = \begin{cases} LMS_{i+1} - LMS_i & i < \sum_{l=1}^j m_l \\ \sum_{l=1}^j n_l - LMS_{\sum_{q=1}^j m_q} + 1 & i = \sum_{l=1}^j m_l \end{cases} \quad 5.134$$

for $i = \sum_{l=1}^{j-1} m_l + 1, \dots, \sum_{l=1}^j m_l$ and $j = 1 \dots NSUB$

Equation (5.134) sets the initial block measurement segments' lengths to be similar to those obtained from the procedure in section 5.3.3.2.2; i.e. the initial segment length of the measurement i is the difference between the measured state locations LMS_i and at LMS_{i+1} as shown in Fig 5.11 (Note that the counter i starts in block j from the value $\sum_{l=1}^{j-1} m_l + 1$ and ends with the value $\sum_{l=1}^j m_l$). The last measurement in block j , $z_{\sum_{l=1}^j m_l}^j$, has an initial segment length equal to the difference between the row that represents the end of the block and the location of the measured state represented by the measurement $z_{\sum_{l=1}^j m_l}^j, LMS_{\sum_{l=1}^j m_l}$.

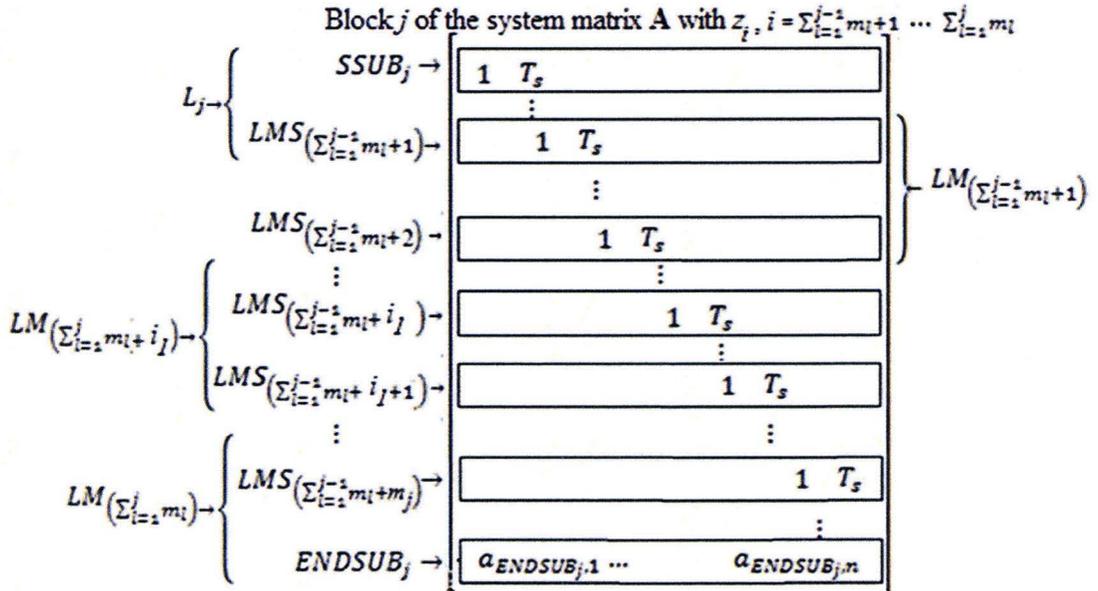


Fig 5.11: The modified selection procedure of the General Observability matrix - Step 4: Obtaining the initial measurement segments lengths.

Example 5.12:

Further to Example 5.9, the initial General Observability matrix is obtained as follows.

The initial segment lengths are obtained as follows:

$$\begin{aligned}
 LM_1 &= \sum_{l=1}^1 n_l - LMS_{\sum_{q=1}^1 m_q} + 1 = n_1 - LMS_1 + 1 = 2 - 1 + 1 = 2 \\
 LM_2 &= \sum_{l=1}^2 n_l - LMS_{\sum_{q=1}^2 m_q} + 1 = n_1 + n_2 - LMS_2 + 1 = 4 - 4 + 1 = 1 \\
 LM_3 &= \sum_{l=1}^3 n_l - LMS_{\sum_{q=1}^3 m_q} + 1 = n_1 + n_2 + n_3 - LMS_3 + 1 = 6 - 6 + 1 = 1
 \end{aligned} \tag{5.135}$$

And the initial General Observability matrix, \mathbf{O}_{g_0} , is defined as follows:

$$\mathbf{O}_{g_0} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1 \mathbf{A} \\ \mathbf{H}_2 \\ \mathbf{H}_3 \end{bmatrix} \tag{5.136}$$

STEP 5: The remaining states located in state's sub-vector j are between $SSUB_j$ and $LMS_{\sum_{l=1}^{j-1} m_{l+1}}$, which represents the difference between the beginning of this sub-vector and the location of the measured state obtained from its first measurement. These states have a length of L_j obtained as follows:

$$L_j = LMS_{\sum_{l=1}^{j-1} m_{l+1}} - SSUB_j, j = 1 \dots NSUB \tag{5.137}$$

The resultant initial General Observability matrix, \mathbf{O}_{g_0} , is then obtained from the initial segments and it has the following rank:

$$\text{rank}(\mathbf{O}_{g_0}) = n - \sum_{j=1}^{NSUB} L_j \tag{5.138}$$

Note that if the first measurement in each block represents its first state, then $\sum_{j=1}^{NSUB} L_j = 0$ and the General Observability matrix has full rank. Equation (5.138) shows that $\sum_{j=1}^{NSUB} L_j$ data points are needed to obtain the full state vector. To obtain these data points, the iterative procedure described in Fig 5.12 is used as follows.

- I. A list of the measurement segments is created. This list is referred to as the *expandable segment list* and it contains the measurement segments that can be

expanded to increase the rank of the initial General Observability matrix. Initially, all the segments are assumed to be expandable.

- II. The measurement with the shortest LM_i (segment length) from the expandable segment list is taken first, and its segment is expanded by one time step, $z_{i_{k+LM_i}}$. The rank of the combined matrix that consists of the initial General Observability matrix and the data of $\mathbf{H}_i \mathbf{A}^{LM_i}$ is tested. If the rank increases, then the initial General Observability matrix is expanded as follows:

$$\mathbf{O}_{g_{Iteration+1}} = \begin{bmatrix} \mathbf{O}_{g_{Iteration}} \\ \mathbf{H}_i \mathbf{A}^{LM_i} \end{bmatrix} \quad 5.139$$

Where $\mathbf{O}_{g_{Iteration+1}}$ is the new initial General Observability matrix for the next iteration, and $\mathbf{O}_{g_{Iteration}}$ is the old initial General Observability matrix.

- III. If the rank does not increase, then the selected measurement, z_i , has the maximum segment length and it cannot be expanded, hence it is removed from the list of step I.
- IV. The rank of the resultant initial General Observability matrix is examined. If the rank is equal to the number of states, then the algorithm stops. Otherwise, steps (II), (III) and (IV) are repeated.
- V. Once the iteration stops, the General Observability matrix is then considered to be the last update of the initial General Observability matrix.

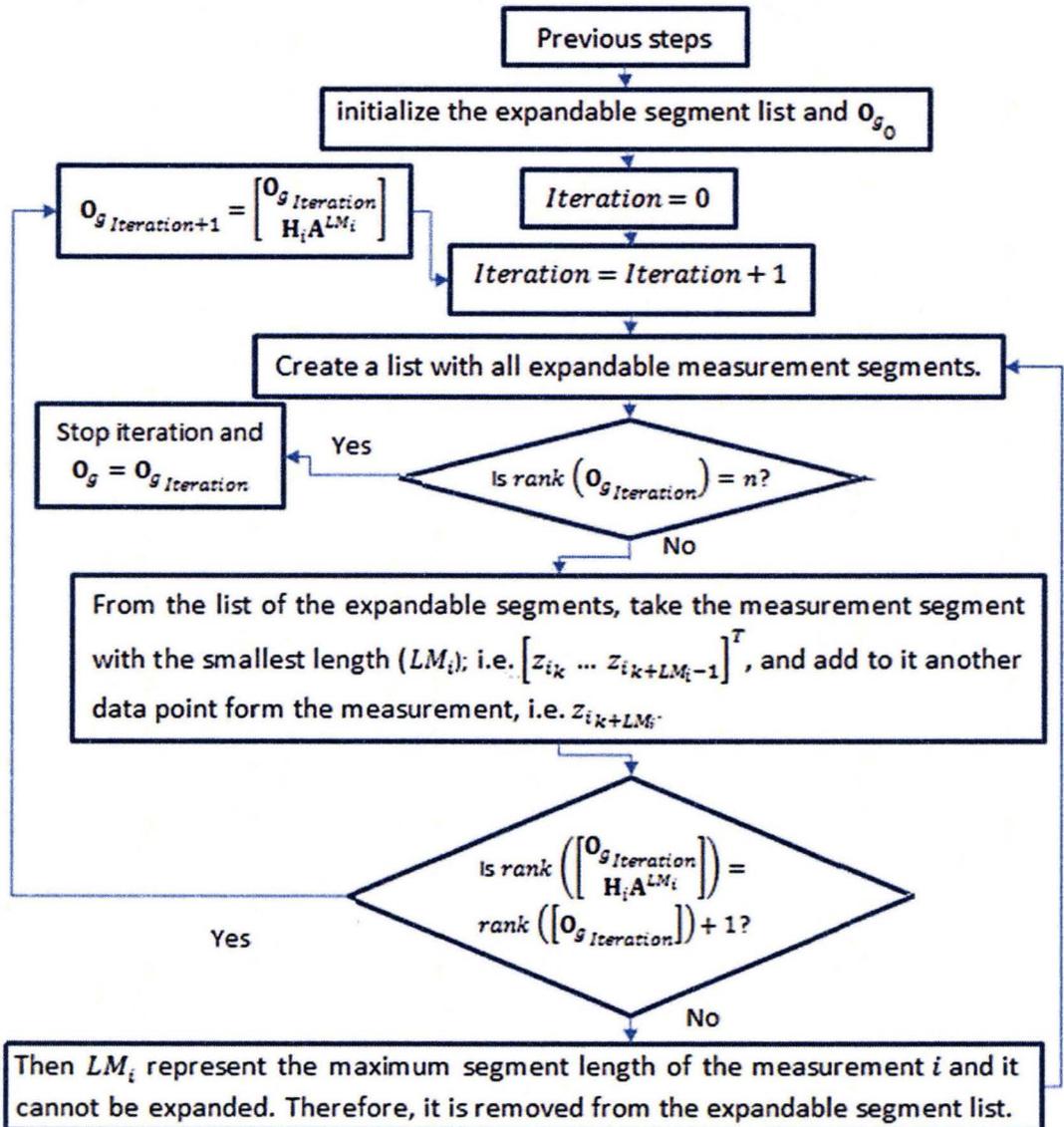


Fig 5.12: The iterative procedure of increasing the segment length in order to obtain the General Observability matrix with full rank

Example 5.13:

Further to Example 5.12, LMS_i has values of 1, 4 and 6, and the initial General Observability matrix has rank of 4 and is defined as follows:

$$\mathbf{O}_{g_0} = \begin{bmatrix} [\mathbf{H}_1] \\ [\mathbf{H}_1\mathbf{A}] \\ [\mathbf{H}_2] \\ [\mathbf{H}_3] \end{bmatrix} = \begin{bmatrix} [1 & 0] & 0 & 0 & 0 & 0 \\ [1 & T_s] & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad 5.140$$

Using equation (5.137), the remaining hidden states in each block, L_i , have lengths of:

$$\begin{aligned} L_1 &= \left(LMS_{\sum_{l=1}^{j-1} m_{l+1}} - SSUB_j \right) |_{j=1} = LMS_1 - SSUB_1 = 1 - 1 = 0 \\ L_2 &= \left(LMS_{\sum_{l=1}^{j-1} m_{l+1}} - SSUB_j \right) |_{j=2} = LMS_2 - SSUB_2 = 4 - 3 = 1 \\ L_3 &= \left(LMS_{\sum_{l=1}^{j-1} m_{l+1}} - SSUB_j \right) |_{j=3} = LMS_3 - SSUB_3 = 6 - 5 = 1 \end{aligned} \quad 5.141$$

The initial General Observability matrix needs two more independent rows to become a General Observability matrix with full rank. From step 5 in the modified procedure, the expandable segment list is created and initially it contains LM_1 , LM_2 and LM_3 .

Iteration 1:

Two segments have the shortest segment length (LM_i) in the list, which are z_2 and z_3 with $LM_2 = 1$ and $LM_3 = 1$, respectively. Each one can be used at this point. In this example, the measurement z_2 is used. Increasing the segment of z_2 by one data point, $\mathbf{H}_2\mathbf{A}$, increases the rank of the combined matrix (that consists of the initial General Observability matrix \mathbf{O}_{g_0} and the new data $\mathbf{H}_2\mathbf{A}$) by one as follows:

$$\text{rank} \left(\begin{bmatrix} \mathbf{O}_{g_0} \\ \mathbf{H}_2\mathbf{A} \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} [1 & 0] & 0 & 0 & 0 & 0 \\ [1 & T_s] & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ [1 & 1 & -2 & -1 & 1 & 1] \end{bmatrix} \right) = \text{rank}(\mathbf{O}_{g_0}) + 1 = 5 \quad 5.142$$

The segment length of the z_2 is then increased, and the new initial General Observability matrix becomes as follows:

$$\mathbf{O}_{g_1} = \begin{bmatrix} [\mathbf{H}_1] \\ [\mathbf{H}_1\mathbf{A}] \\ [\mathbf{H}_2] \\ [\mathbf{H}_2\mathbf{A}] \\ [\mathbf{H}_3] \end{bmatrix} = \begin{bmatrix} [1 & 0] & 0 & 0 & 0 & 0 \\ [1 & T_s] & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & -2 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad 5.143$$

The new initial General Observability matrix does not have full rank, so the procedure is repeated.

Iteration 2:

The expandable segment list is updated. The measurement with the shortest LM_i becomes z_3 with $LM_3 = 1$. Increasing the segment of z_3 by one data point, $\mathbf{H}_3\mathbf{A}$, increases the rank of the combined matrix (that consists of the initial General Observability matrix \mathbf{O}_{g_1} and the new data $\mathbf{H}_3\mathbf{A}$) by one as follows:

$$\text{rank} \left(\begin{bmatrix} \mathbf{O}_{g_1} \\ \mathbf{H}_3\mathbf{A} \end{bmatrix} \right) = \text{rank} \left(\begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & T_s \\ 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ -2 & -1 & 1 & 1 \end{bmatrix} \end{bmatrix} \right) = \text{rank}(\mathbf{O}_{g_1}) + 1 = 6 \quad 5.144$$

The segment length of the z_3 is then increased, and the new initial General Observability matrix becomes as follows:

$$\mathbf{O}_{g_2} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_2 \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_3 \\ \mathbf{H}_3\mathbf{A} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & T_s \\ 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ -2 & -1 & 1 & 1 \end{bmatrix} \end{bmatrix} \quad 5.145$$

The new initial General Observability matrix has full rank, so the iterative process is stopped and the General Observability matrix is defined as follows:

$$\mathbf{O}_g = \mathbf{O}_{g_2} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_2 \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_3 \\ \mathbf{H}_3\mathbf{A} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & T_s \\ 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ -2 & -1 & 1 & 1 \end{bmatrix} \end{bmatrix} \quad 5.146$$

The General Observability matrix of equation (5.146) depends on the system's parameters, thus the expanding algorithm may vary from system to system. If modeling errors are present, this will affect the states estimates. This research focuses on systems that follow the conditions of equations (5.115) and (5.116). Otherwise, the modeling errors cannot be constructed and the states will be wrongly estimated.

measurements. Therefore, by taking consecutive points of that measurement, these states can be extracted. Using the definition of the General Observability matrix in equation (5.72) and the General System and the General System Noise Toeplitz matrices in equations (5.147) and (5.148), the states can be obtained as follows:

$$\mathbf{x}_k = \mathbf{O}_g^{-1} \left(\begin{bmatrix} z_{1k} \\ \vdots \\ z_{1k+LM_1-1} \\ \vdots \\ z_{mk} \\ \vdots \\ z_{mk+LM_m-1} \end{bmatrix} - \begin{bmatrix} v_{1k} \\ \vdots \\ v_{1k+LM_1-1} \\ \vdots \\ v_{mk} \\ \vdots \\ v_{mk+LM_m-1} \end{bmatrix} - \mathbf{T}_{og} \begin{bmatrix} u_k \\ \vdots \\ u_{k+m_a-1} \end{bmatrix} - \mathbf{T}_{wg} \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+m_a-1} \end{bmatrix} \right) \quad 5.149$$

The estimate vector, $\hat{\mathbf{x}}_{GTO_k}$, is obtained as:

$$\hat{\mathbf{x}}_{GTO_k} = \widehat{\mathbf{O}}_g^{-1} \begin{bmatrix} z_{1k} \\ \vdots \\ z_{1k+LM_1-1} \\ \vdots \\ z_{mk} \\ \vdots \\ z_{mk+LM_m-1} \end{bmatrix} - \widehat{\mathbf{O}}_g^{-1} \widehat{\mathbf{T}}_{og} \begin{bmatrix} u_k \\ \vdots \\ u_{k+m_a-1} \end{bmatrix} \quad 5.150$$

Where \mathbf{O}_g^{-1} and $\widehat{\mathbf{O}}_g^{-1}$ are the inverse of the General Observability matrix and its estimated inverse. If the system has measurement, input and system matrices of equations (5.53), (5.61) and (5.62), and satisfies equations (5.115) and (5.116), then these matrices (\mathbf{O}_g^{-1} and $\widehat{\mathbf{O}}_g^{-1}$) are in block-diagonal form and the inverse is obtained by taking the inverse of each block's sub-Observability matrix individually as follows:

$$\mathbf{O}_g^{-1} = \begin{bmatrix} \mathbf{O}_{sub_1}^{-1} & \mathbf{0}_{LM_1 \times (n-LM_1)} & \mathbf{0}_{LM_2 \times (n-LMS_2-LM_2+1)} \\ \mathbf{0}_{LM_2 \times (LMS_2-1)} & \mathbf{O}_{sub_2}^{-1} & \mathbf{0}_{LM_2 \times (n-LMS_2-LM_2+1)} \\ & \vdots & \\ \mathbf{0}_{LM_m \times (LMS_m-1)} & \mathbf{O}_{sub_m}^{-1} & \end{bmatrix} \quad 5.151$$

$$\widehat{\mathbf{O}}_g^{-1} = \begin{bmatrix} \widehat{\mathbf{O}}_{sub_1}^{-1} & \mathbf{0}_{LM_1 \times (n-LM_1)} & \mathbf{0}_{LM_2 \times (n-LMS_2-LM_2+1)} \\ \mathbf{0}_{LM_2 \times (LMS_2-1)} & \widehat{\mathbf{O}}_{sub_2}^{-1} & \mathbf{0}_{LM_2 \times (n-LMS_2-LM_2+1)} \\ & \vdots & \\ \mathbf{0}_{LM_m \times (LMS_m-1)} & \widehat{\mathbf{O}}_{sub_m}^{-1} & \end{bmatrix}$$

Each sub-Observability matrix has the form of equation (5.44). Therefore, the estimated General Observability matrix is independent of the parameters. On the other hand, the

General System Toeplitz matrix for these systems is a zero matrix. This means that the input signal is not needed to obtain the alternative states and measurements as discussed in section (5.2). The difference between the states and the alternative states, $\hat{\mathbf{x}}_{TO_k}$ is a function of the noise as follows:

$$\mathbf{x}_k - \hat{\mathbf{x}}_{GTO_k} = -\widehat{\mathbf{O}}_g^{-1} \left(\begin{bmatrix} v_{1k} & \dots & v_{1k+LM_1-1} & \dots & v_{m_k} & \dots & v_{m_k+LM_m-1} \end{bmatrix}^T + \mathbf{T}_w \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+m_a-1} \end{bmatrix} \right) \quad 5.152$$

The alternative measurement, $\hat{\mathbf{z}}_{GTO_k}$, is then defined as follows:

$$\hat{\mathbf{z}}_{GTO_k} = \bar{\mathbf{H}} \hat{\mathbf{x}}_{GTO_k} = \bar{\mathbf{H}} \left(\mathbf{x}_k + \widehat{\mathbf{O}}_g^{-1} \left(\begin{bmatrix} v_{1k} \\ \vdots \\ v_{1k+LM_1-1} \\ \vdots \\ v_{m_k} \\ \vdots \\ v_{m_k+LM_m-1} \end{bmatrix} + \mathbf{T}_w \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+m_a-1} \end{bmatrix} \right) \right) \quad 5.153$$

Equation (5.153) can be rewritten as follows:

$$\hat{\mathbf{z}}_{GTO_k} = \bar{\mathbf{H}} \hat{\mathbf{x}}_{GTO_k} = \bar{\mathbf{H}} \mathbf{x}_k + \boldsymbol{\vartheta}_k \quad 5.154$$

Where $\boldsymbol{\vartheta}_k$ is the new measurement noise vector that has size of $\boldsymbol{\vartheta}_k \in \mathbb{R}^{n \times 1}$ and is defined as follows:

$$\boldsymbol{\vartheta}_k = \bar{\mathbf{H}} \left(\widehat{\mathbf{O}}_g^{-1} \left(\begin{bmatrix} v_{1k} \\ \vdots \\ v_{1k+LM_1-1} \\ \vdots \\ v_{m_k} \\ \vdots \\ v_{m_k+LM_m-1} \end{bmatrix} + \mathbf{T}_w \begin{bmatrix} \mathbf{w}_k \\ \vdots \\ \mathbf{w}_{k+m_a-1} \end{bmatrix} \right) \right) \quad 5.155$$

Note that if the measurement matrix has full rank, then $\boldsymbol{\vartheta}_k$ becomes equal to the measurement noise vector, \mathbf{v}_k , whereas $LM_1 = LM_2 = \dots = LM_m = 1$, $\bar{\mathbf{H}} = \mathbf{H}$, $\widehat{\mathbf{O}}_g = \mathbf{H}$ and $\mathbf{T}_w = \mathbf{0}_{n \times n}$. If the system has a single degree of freedom structure with only one measurement, then equations (5.147), (5.148), (5.149), (5.150) and (5.152) are reduced to equations (5.24), (5.26), (5.25), (5.32) and (5.34), respectively.

Once the alternative measurements are obtained, the SVSF can be used to estimate the states and the parameters as discussed in section (5.2.2). The combined algorithm between the General Observability and the General System Toeplitz matrix is referred to as the General Toeplitz/Observability SVSF and it has the algorithm described by equations (5.38), (5.39) and (5.40). The main difference between this method and the method described in section (5.2.2) is the procedure used to obtain the alternative states and measurements.

5.4 Measurement matrix treatment

In order to make the selection processes described in sections (5.3.3.2.3) and (5.3.3.2.2) applicable, the measurement matrix must have the structure of equation (5.53) (each measurement is only related to one state). Otherwise, the General Observability is obtained by trial and error (using the iterative method described in section (5.3.3.1)). The resultant General Observability matrix has several forms, and these forms have a high probability of being dependent on system parameters. In order to solve this problem, the following procedure (described by Fig 5.13) is used.

- The measurement matrix is divided into n column vectors as follows:

$$\mathbf{H} = \begin{bmatrix} \begin{bmatrix} h_{1,1} \\ \vdots \\ h_{m,1} \end{bmatrix} & \begin{bmatrix} h_{1,2} \\ \vdots \\ h_{m,2} \end{bmatrix} & \begin{bmatrix} h_{1,3} \\ \vdots \\ h_{m,3} \end{bmatrix} & \dots & \begin{bmatrix} h_{1,n} \\ \vdots \\ h_{m,n} \end{bmatrix} \\ \uparrow & \uparrow & \uparrow & & \uparrow \\ \text{vector 1} & \text{vector 2} & \text{vector 3} & & \text{vector } n \end{bmatrix} \quad 5.156$$

- A new vector is created that contains the locations of the zero columns in the measurement matrix. The first entry of the vector represents the location of the first zero column, the second entry of the vector represents the location of the second zero column, and so on. The resultant vector is referred to as the zeros location list, **ZeLoLi**, and it has sizes of $\mathbf{ZeLoLi} \in \mathbb{R}^{z_l \times 1}$. Where z_l is the number of zero columns in the measurement matrix.
- If $n - z_l \neq m$, then this procedure fails and the General Observability matrix can only be obtained by the iterative method of section (5.3.3.1).
- If $n - z_l = m$, then a new square matrix can be created from the measurement matrix by taking its non-zero columns. The new matrix is referred to as the reduced measurement matrix, \mathbf{H}_R and it has size of $\mathbf{H}_R \in \mathbb{R}^{m \times m}$.

- The state vector can be reduced by eliminating the states that do not have a measurement associated with them. The reduced vector is referred to as the reduced state vector, \mathbf{x}_{R_k} and it has size of $\mathbf{x}_{R_k} \in \mathbb{R}^{m \times 1}$. The measurement can be expressed using \mathbf{H}_R and \mathbf{x}_{R_k} as follows:

$$\mathbf{z}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k = \mathbf{H}_R\mathbf{x}_{R_k} + \mathbf{v}_k \quad 5.157$$

- The measurement vector can be rearranged to obtain a new measurement vector, \mathbf{z}_{new_k} , that has size of $\mathbf{z}_{new_k} \in \mathbb{R}^{m \times 1}$ and is defined as follows:

$$\mathbf{z}_{new_k} = \mathbf{H}_R^{-1}\mathbf{z}_k = \mathbf{I}_{m \times m}\mathbf{x}_{R_k} + \mathbf{H}_R^{-1}\mathbf{v}_k \quad 5.158$$

Each member of \mathbf{z}_{new_k} is related to one state of \mathbf{x}_{R_k} . Note that \mathbf{H}_R must be invertible, otherwise, equation (5.158) is invalid and this procedure fails.

- The identity matrix is then modified and expanded by adding the zero columns with locations obtained from the **ZeLoLi** (recovering the missing columns from the measurement matrix). The modified matrix is referred to as the new measurement matrix, \mathbf{H}_{new} , and it has size of $\mathbf{H}_{new} \in \mathbb{R}^{m \times n}$. Using the new measurement matrix, the new measurement vector is obtained as follows:

$$\mathbf{z}_{new_k} = \mathbf{H}_{new}\mathbf{x}_k + \mathbf{v}_{new_k} \quad 5.159$$

Where \mathbf{v}_{new_k} is the new measurement noise vector and it is defined as follows:

$$\mathbf{v}_{new_k} = \mathbf{H}_R^{-1}\mathbf{v}_k \quad 5.160$$

Each measurement in the resultant measurement vector is only related to one state. Therefore, procedures in sections (5.3.3.2.2) and (5.3.3.2.3) can be used as illustrated in the following example.

Example 5.14:

Further to Example 5.2, if the measurement matrix is defined as follows:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 4 & 0 & 5 & 0 & 6 & 0 \\ 7 & 0 & 8 & 0 & 10 & 0 \end{bmatrix} \quad 5.161$$

Then **ZeLoLi** = [2 4 6]^T and *zl* = 3. The reduced measurement matrix, \mathbf{H}_R , is obtained as follows:

$$\mathbf{H}_R = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 5 & 6 \\ 7 & 8 & 10 \end{bmatrix} \quad 5.162$$

and its inverse matrix is obtained as follows:

$$\mathbf{H}_R^{-1} = \begin{bmatrix} 2 & -2 & 1 \\ 2 & 3 & -2 \\ -3 & -1 & 1 \end{bmatrix} \quad 5.163$$

The new measurement matrix is then defined as follows:

$$\mathbf{H}_{new} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad 5.164$$

The new measurement vector is defined as follows:

$$\mathbf{z}_{new_k} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}_k + \mathbf{v}_{new_k} \quad 5.165$$

Note that the new measurement noise vector has different properties than the measurement noise vector.

The selection procedures of sections (5.3.3.2.2) and (5.3.3.2.3) are applied to systems that have a measurement matrix with the structure of equation (5.53). The new measurement vector, \mathbf{z}_{new_k} , provides a measurement matrix that qualifies this equation and relates each measurement to only one state. Therefore, it is used to obtain the General Observability matrix instead of the old measurement vector, \mathbf{z}_k .

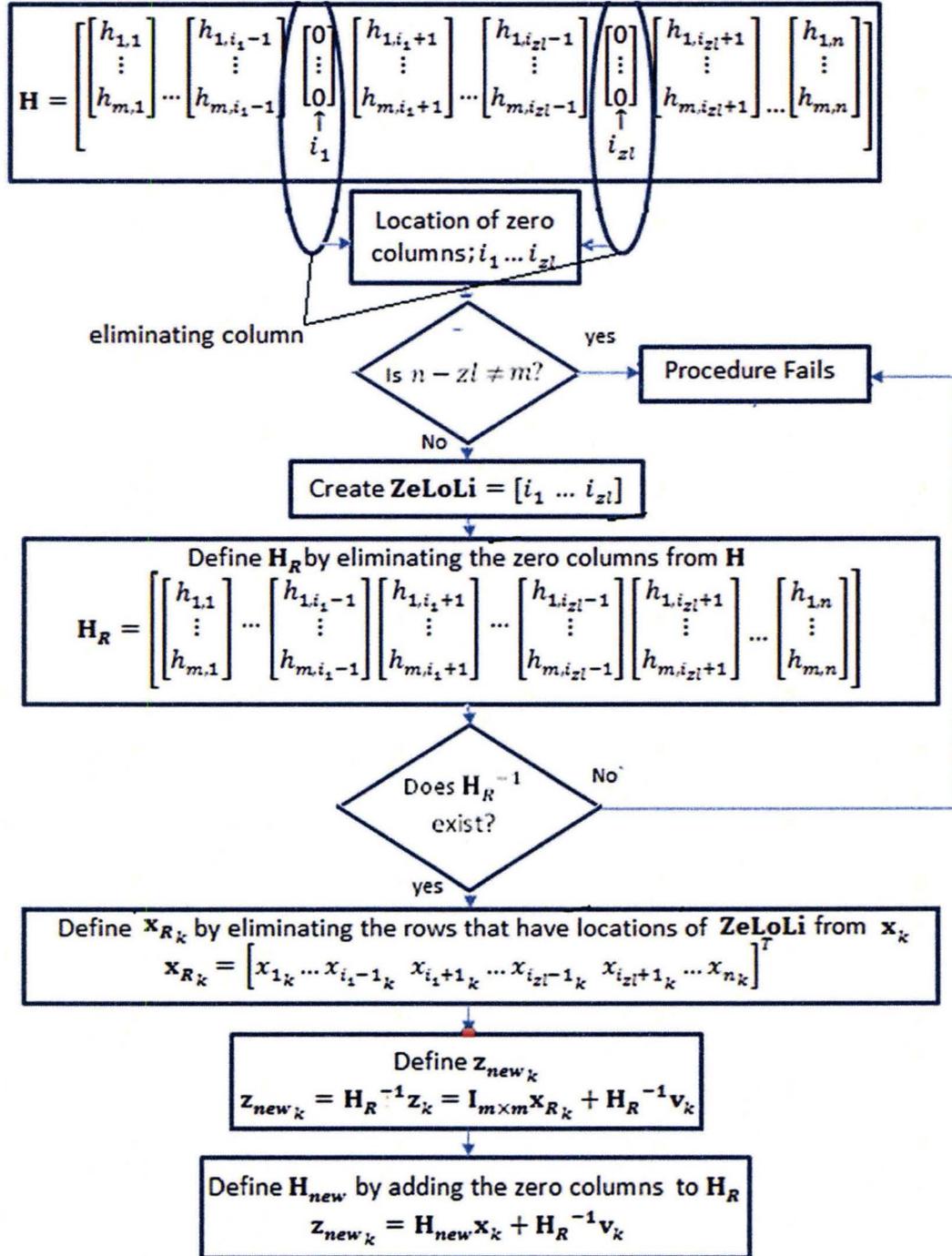


Fig 5.13: Transforming the measurement matrix to equation (5.53)

5.5 The General Toeplitz/Observability SVSF concepts for systems with measurement matrices that have partial rank

The SVSF is combined with the general system Toeplitz and Observability matrices to estimate the states and parameters. In this section, the combined algorithm is discussed in further detail. This section is an extended discussion of sections (4.2) and (5.2) to accommodate the SVSF concept for linear systems that have multiple degrees of freedom and/or multiple measurements, and have a partially ranked measurement matrix. The contributions and novelties of this research in this section are as follows:

- Combining the SVSF with the General System Toeplitz and Observability matrices to estimate the state vector. These matrices provide the SVSF with an alternative measurement vector that result in a measurement matrix with full rank. To avoid the effects of modeling errors on the alternative measurements, the system must be presented in its GOCF and satisfy the conditions of equations (5.115) and (5.116). Using the alternative measurement, the General Toeplitz /Observability SVSF algorithm become similar to the algorithm in section (4.1.3).
- Defining the resultant existence subspaces (the a priori and the a posteriori existence subspaces) created by the General Toeplitz/Observability SVSF, and developing mathematical formulas that describe them. This work is listed later in subsection (5.5.2).
- Developing chattering in the a priori and the a posteriori estimates mathematically and using their contents to refine the filter's model. These steps are discussed later in subsection (5.5.4).

5.5.1. The General Toeplitz/Observability SVSF

The SVSF needs the measurement matrix to have full rank. In section (5.1), the SVSF is otherwise combined with a method similar to the Luenberger algorithm in order to solve this limitation. However, this method has some limitations due to the Luenberger algorithm. In section (5.2), the Toeplitz/Observability SVSF is developed to overcome these limitations when the system is described in its observability canonical form. However, systems with multiple degrees of freedom and/or multiple measurements cannot be expressed in a standard Observability canonical form. Therefore, a new form referred to as the General Observability Canonical form (GOCF) is proposed in section (5.3). The Toeplitz/Observability SVSF (developed in section (5.2.2)) is modified by combining SVSF with the General System Toeplitz and Observability matrices. In this

section, the combined algorithm is referred to as the General Toeplitz/Observability SVSF and is used to estimate the states and parameters.

If the General System Toeplitz and Observability matrices are exactly known, then the vector $\hat{\mathbf{x}}_{GTO_k}$ can be extracted from the measurement and the input signals by using equation (5.150). $\hat{\mathbf{x}}_{GTO_k}$ would contain the actual state vector information blurred with the noise vectors as shown in equation (5.152). Using the resultant alternative measurement matrix, $\bar{\mathbf{H}}$, the vector $\hat{\mathbf{z}}_{GTO_k}$ can be used to compensate the missing $(n - m)$ measurements. The alternative measurements are then used with the SVSF as follows.

1 - Prediction Stage:

The a priori state estimate is obtained by equation (5.38) (described in section (5.2)) as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k|k-1} &= \hat{\mathbf{A}}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \hat{\mathbf{B}}_{k-1}u_{k-1} \\ \hat{\mathbf{z}}_{k|k-1} &= \bar{\mathbf{H}}\hat{\mathbf{x}}_{k|k-1}\end{aligned}\quad 5.38$$

2 - Update Stage:

A corrective gain is calculated and used for refining the a priori estimate into its a posteriori form as follows:

$$\begin{aligned}\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{e_{GTO_k}} \\ \hat{\mathbf{z}}_{k|k} &= \bar{\mathbf{H}}\hat{\mathbf{x}}_{k|k}\end{aligned}\quad 5.166$$

where $\mathbf{K}_{e_{GTO_k}}$ is the SVSF's gain for the general Toeplitz/Observability SVSF and is defined as follows:

$$\mathbf{K}_{e_{GTO_k}} = \bar{\mathbf{H}}^{-1} \left(\left| \mathbf{e}_{z_{GTO_k|k-1}} \right| + \gamma \left| \mathbf{e}_{z_{GTO_{k-1}|k-1}} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right) \quad 5.167$$

Where $\mathbf{e}_{z_{GTO_k|k-1}} = \hat{\mathbf{z}}_{GTO_k} - \hat{\mathbf{z}}_{k|k-1}$ and $\mathbf{e}_{z_{GTO_{k-1}|k-1}} = \hat{\mathbf{z}}_{GTO_{k-1}} - \hat{\mathbf{z}}_{k-1|k-1}$.

Note that if the system has a single degree of freedom with one measurement, then the general Toeplitz/Observability SVSF is reduced to the Toeplitz/Observability SVSF form, where $\mathbf{K}_{e_{GTO_k}}$, $\mathbf{e}_{z_{GTO_k|k-1}}$ and $\mathbf{e}_{z_{GTO_{k-1}|k-1}}$ become $\mathbf{K}_{e_{TO_k}}$, $\mathbf{e}_{z_{TO_k|k-1}}$ and $\mathbf{e}_{z_{TO_{k-1}|k-1}}$, respectively.

5.5.2. The General Toeplitz/Observability SVSF's existence subspaces

As discussed in section (4.3.1), the SVSF has two existence subspaces; the a priori and the a posteriori existence subspaces. These subspaces enclose the state trajectory and prevent the estimate in its a priori and a posteriori forms from leaving the state's neighbourhood. The widths of the existence subspaces are unknown and time varying. In chapter four, these widths were explored for systems with a measurement matrix that has full rank. In the following subsections, these two subspaces are considered in detail for systems with a measurement matrix that has partial rank. The systems are represented in their General Observability Canonical Form.

5.5.2.1. The a priori existence subspace for the General Toeplitz/Observability SVSF

In order to obtain the a priori existence subspace for the General Toeplitz/Observability SVSF, the following procedure is used.

The general a priori estimation error is defined by using equation (5.154) as follows:

$$\mathbf{e}_{z_{GTO_{k|k-1}}} = \hat{\mathbf{z}}_{GTO_k} - \hat{\mathbf{z}}_{k|k-1} = \bar{\mathbf{H}}\mathbf{x}_k + \boldsymbol{\vartheta}_k - \bar{\mathbf{H}}\hat{\mathbf{x}}_{k|k-1} \quad 5.168$$

Substituting equations (4.2) and (5.38) into equation (5.168) results in the following:

$$\mathbf{e}_{z_{GTO_{k|k-1}}} = \bar{\mathbf{H}}(\mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}u_{k-1} + \mathbf{w}_{k-1}) - \bar{\mathbf{H}}(\hat{\mathbf{A}}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} + \hat{\mathbf{B}}_{k-1}\mathbf{u}_{k-1}) + \boldsymbol{\vartheta}_k \quad 5.169$$

Rearranging equation (5.169), and using equation (5.38) the following is obtained:

$$\mathbf{e}_{z_{GTO_{k|k-1}}} = \bar{\mathbf{H}}(\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\hat{\mathbf{z}}_{GTO_k} + \mathbf{B}_{k-1}u_{k-1} + \mathbf{w}_{k-1} - \mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\vartheta}_{k-1}) - \bar{\mathbf{H}}(\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\hat{\mathbf{z}}_{k-1|k-1} + \hat{\mathbf{B}}_{k-1}\mathbf{u}_{k-1}) + \boldsymbol{\vartheta}_k \quad 5.170$$

Further to equation (5.167) and using equation (5.168), the estimated measurement is related to the alternative measurement as follows:

$$\hat{\mathbf{z}}_{k|k} = \hat{\mathbf{z}}_{k|k-1} + \left\{ \bar{\mathbf{H}}\mathbf{K}_{e_{GTO_k}} = \mathbf{e}_{z_{GTO_{k|k-1}}} + \boldsymbol{\gamma} \left| \mathbf{e}_{z_{GTO_{k-1|k-1}}} \right|^{\circ} \text{sgn} \left(\mathbf{e}_{z_{GTO_{k|k-1}}} \right) \right\} \quad 5.171$$

$$\hat{\mathbf{z}}_{k|k} = \hat{\mathbf{z}}_{GTO_{k-1}} + \boldsymbol{\gamma} \left| \mathbf{e}_{z_{GTO_{k-1|k-1}}} \right|^{\circ} \text{sgn} \left(\mathbf{e}_{z_{GTO_{k|k-1}}} \right)$$

Substituting equation (5.171) into equation (5.170) yields the following:

$$\mathbf{e}_{z_{GTO_{k|k-1}}} = \bar{\mathbf{H}}(\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\hat{\mathbf{z}}_{GTO_{k-1}} + \mathbf{B}_{k-1}u_{k-1} + \mathbf{w}_{k-1} - \mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\vartheta}_{k-1}) - \bar{\mathbf{H}}(\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\hat{\mathbf{z}}_{GTO_{k-1}} + \hat{\mathbf{B}}_{k-1}\mathbf{u}_{k-1}) + \boldsymbol{\vartheta}_k - \bar{\mathbf{H}}\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\gamma} \left| \mathbf{e}_{z_{GTO_{k-2|k-2}}} \right|^{\circ} \text{sgn} \left(\mathbf{e}_{z_{GTO_{k-1|k-2}}} \right) \quad 5.172$$

Rearranging equation (5.172) gives the following:

$$\begin{aligned} \mathbf{e}_{z_{GTO}_{k|k-1}} &= (\bar{\mathbf{H}}\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1} - \bar{\mathbf{H}}\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1})\hat{\mathbf{z}}_{GTO_{k-1}} + (\bar{\mathbf{H}}\mathbf{B}_{k-1} - \bar{\mathbf{H}}\hat{\mathbf{B}}_{k-1})\mathbf{u}_{k-1} \\ &\quad + \bar{\mathbf{H}}\mathbf{w}_{k-1} + \boldsymbol{\vartheta}_k - \bar{\mathbf{H}}\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\vartheta}_{k-1} \\ &\quad - \bar{\mathbf{H}}\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\gamma} \left| \mathbf{e}_{z_{GTO}_{k-2|k-2}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO}_{k-1|k-2}} \right) \end{aligned} \quad 5.173$$

Further to equation (4.40), and using equation (5.161), the general a posteriori estimation error is related to the initial condition as follows:

$$\mathbf{e}_{z_{GTO}_{k-1|k-1}} = \hat{\mathbf{z}}_{GTO_{k-1}} - \hat{\mathbf{z}}_{k-1|k-1} = -\boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{z_{GTO}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO}_{k-1|k-2}} \right) \quad 5.174$$

Substituting equation (5.174) into equation (5.173) yields the following:

$$\begin{aligned} \mathbf{e}_{z_{GTO}_{k|k-1}} &= (\bar{\mathbf{H}}\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1} - \bar{\mathbf{H}}\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1})\hat{\mathbf{z}}_{GTO_{k-1}} + (\bar{\mathbf{H}}\mathbf{B}_{k-1} - \bar{\mathbf{H}}\hat{\mathbf{B}}_{k-1})\mathbf{u}_{k-1} \\ &\quad + \bar{\mathbf{H}}\mathbf{w}_{k-1} + \boldsymbol{\vartheta}_k - \bar{\mathbf{H}}\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\vartheta}_{k-1} \\ &\quad - \bar{\mathbf{H}}\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{z_{GTO}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO}_{k-1|k-2}} \right) \end{aligned} \quad 5.175$$

The a priori existence subspace is then defined as follows:

$$\begin{aligned} \mathbf{e}_{x_{GTO}_{k|k-1}} &= \Delta\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\hat{\mathbf{z}}_{GTO_{k-1}} + \Delta\mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1} - \mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\vartheta}_{k-1} \\ &\quad - \hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{z_{GTO}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO}_{k-1|k-2}} \right) \end{aligned} \quad 5.176$$

From equation (5.176), the terms $(\Delta\mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\hat{\mathbf{z}}_{GTO_{k-1}})$ and $(\Delta\mathbf{B}_{k-1}\mathbf{u}_{k-1})$ capture the influence of the modeling errors in the system and input matrices. The term $(\mathbf{w}_{k-1} - \mathbf{A}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\vartheta}_{k-1})$ quantifies the impact of the system and measurement noise. The last term in equation (5.176), $(\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\gamma}^{k-1} \left| \mathbf{e}_{z_{GTO}_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO}_{k-1|k-2}} \right))$, describes the effects of the uncertainty in initial conditions and its impact on the a priori existence subspace. According to the latter term, the effect of the error in initial conditions decays in time at a rate of $\hat{\mathbf{A}}_{k-1}\bar{\mathbf{H}}^{-1}\boldsymbol{\gamma}^{k-1}$, and becomes negligible as $k \rightarrow \infty$. Then, the width of the a priori existence subspace becomes a function of the uncertainties, noise, and modeling errors.

5.5.2.2. The a posteriori existence subspace for the General Toeplitz /Observability SVSF

As discussed earlier in section (4.3.1.2), the difference between the width of the a priori existence subspace and the SVSF's gain is equal to the width of the a posteriori

existence subspace, and has a different sign from the a priori existence subspace. The a posteriori existence subspace is derived in this section.

The a posteriori estimation error at time k is defined as follows:

$$\begin{aligned}
 \mathbf{e}_{z_{GTO_k|k}} &= \hat{\mathbf{z}}_{GTO_k} - \hat{\mathbf{z}}_{k|k} = \hat{\mathbf{z}}_{GTO_k} - \hat{\mathbf{z}}_{k|k-1} - \bar{\mathbf{H}}\mathbf{K}\mathbf{e}_{GTO_k} \\
 \mathbf{e}_{z_{GTO_k|k}} &= \mathbf{e}_{z_{GTO_k|k-1}} - \bar{\mathbf{H}}\mathbf{K}\mathbf{e}_{GTO_k} \\
 \mathbf{e}_{z_{GTO_k|k}} &= \mathbf{e}_{z_{GTO_k|k-1}} - \left(\left| \mathbf{e}_{z_{GTO_k|k-1}} \right| + \gamma \left| \mathbf{e}_{z_{GTO_{k-1}|k-1}} \right| \right) \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right) \\
 &\rightarrow \mathbf{e}_{z_{GTO_k|k}} = -\gamma \left| \mathbf{e}_{z_{GTO_{k-1}|k-1}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right)
 \end{aligned} \tag{5.177}$$

Equation (5.177) can be written in a recursive way as follows:

$$\begin{aligned}
 \mathbf{e}_{z_{GTO_k|k}} &= -\gamma \left| \mathbf{e}_{z_{GTO_{k-1}|k-1}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right) \\
 &= -\gamma \left| -\gamma \left| \mathbf{e}_{z_{GTO_{k-2}|k-2}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_{k-1}|k-2}} \right) \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right) \\
 &= -\gamma^2 \left| \mathbf{e}_{z_{GTO_{k-2}|k-2}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right) \\
 &\quad \vdots \\
 &= -\gamma^k \left| \mathbf{e}_{z_{GTO_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right)
 \end{aligned} \tag{5.178}$$

The a posteriori existence subspace is then defined as follows:

$$\mathbf{e}_{x_{GTO_k|k}} = -\gamma^k \left| \mathbf{e}_{z_{GTO_{0|0}} \right| \circ \text{sgn} \left(\mathbf{e}_{z_{GTO_k|k-1}} \right) - \boldsymbol{\vartheta}_k \tag{5.179}$$

Equation (5.179) shows that the a posteriori existence subspace decays with time until the estimated output converges to the alternative measurement which contains the true trajectory blurred with alternative measurement noise. Therefore, the estimation is sensitive to the alternative measurement noise.

5.5.3. The smoothing boundary layer

As discussed in section (4.3.1), a smoothing function with a known smoothing boundary layer is used in the corrective action in order to eliminate the chattering signals, and to reduce the sensitivity to noise. The corrective action is interpolated based on the ratio between the amplitude of the general a priori measurement estimation error and the smoothing boundary layer's width. If the ratio is larger than one, then the corrective action is applied with its full amplitude. The SVSF's gain is then defined as follows:

$$\mathbf{K}_{\mathbf{e}_{GTO_k}} = \bar{\mathbf{H}}^{-1} \left(\left| \mathbf{e}_{z_{GTO_{k|k-1}}} \right| + \gamma \left| \mathbf{e}_{z_{GTO_{k-1|k-1}}} \right| \right) \circ \mathbf{sat} \left(\mathbf{e}_{z_{GTO_{k|k-1}}}, \boldsymbol{\Psi}_k \right) \quad 5.180$$

Where $\boldsymbol{\Psi}_k$ is a vector consisting of n –smoothing boundary layer widths at time k (one for each state), and \mathbf{sat} is a vector of the saturation functions defined by equations (4.47) and (4.48). The gain can be expressed by using the positive diagonal matrix \mathbf{S}_{at_k} (described by equation (4.50)) as follows:

$$\mathbf{K}_{\mathbf{e}_{GTO_k}} = \bar{\mathbf{H}}^{-1} \mathbf{S}_{at} \left(\mathbf{e}_{z_{GTO_{k|k-1}}}, \boldsymbol{\Psi}_k \right) \left(\left(\left| \mathbf{e}_{z_{GTO_{k|k-1}}} \right| + \gamma \left| \mathbf{e}_{z_{GTO_{k-1|k-1}}} \right| \right) \circ \mathbf{sgn} \left(\mathbf{e}_{z_{GTO_{k|k-1}}} \right) \right) \quad 5.181$$

The width of the smoothing boundary layer must be chosen carefully as it must be larger than the uncertain dynamics to smooth the a posteriori estimate. However, it cannot be very large as the general estimation error increases when the width increases. As mentioned before, this research takes an interest in systems described in their GOCF and have a measurement matrix that satisfies equations (5.115) and (5.116).

5.5.4. The chattering amplitudes and its information content

If the system is described in its GOCF, the measurement matrix satisfies equations (5.115) and (5.116), and the smoothing boundary layer has a zero-width, then the a priori chattering vector has zero rows at the locations described by the following equation:

$$e_{z_{GTO_{i_k|k-1}}} = 0 \text{ for} \quad 5.182$$

$$i = 1, \dots, (LMS_2 - 2), \dots, LMS_d, \dots, (LMS_{d+1} - 2), \dots, LMS_m, \dots, (n - 1)$$

The rest of the rows contain the system modeling errors blurred by the alternative measurement and system noise. As discussed in chapter four, the a priori Chattering signal can be used to:

- Point out modeling uncertainties once they occur.
- Provide the information that is needed to tune the model.

In this section, a method that is based on using the segment's mean of multiplication (cross correlation) between the a priori chattering signal and the alternative measurement and between the a priori chattering and the input is used to reconstruct the system modeling errors as follows.

The a priori chattering signal of the General Toeplitz/Observability SVSF is defined as follows:

$$\mathbf{Ch}_{k|k-1} = \begin{bmatrix} Ch_{1|k-1} \\ \vdots \\ Ch_{n|k-1} \end{bmatrix} \text{ where } Ch_{j|k-1} = \begin{cases} 0 & \left| e_{z_{GTO}i_{k|k-1}} \right| < \Psi_{j_k} \\ e_{z_{GTO}i_{k|k-1}} - \Psi_{j_k} & \left| e_{z_{GTO}i_{k|k-1}} \right| \geq \Psi_{j_k} \end{cases} \quad 5.183$$

For $j = 1 \dots n$

If the width of the smoothing boundary layer is set to zero, then equation (5.183) is reduced to the following:

$$\mathbf{Ch}_{k|k-1} = \mathbf{e}_{z_{GTO}i_{k|k-1}} \quad 5.184$$

Substituting equation (5.184) into equation (5.175) and assuming $\mathbf{Y}^{k-1} \rightarrow \mathbf{0}_{n \times n}$ lead to the following:

$$\begin{aligned} \mathbf{Ch}_{k|k-1} = & (\bar{\bar{\mathbf{H}}}\mathbf{A}_{k-1}\bar{\bar{\mathbf{H}}}^{-1} - \bar{\bar{\mathbf{H}}}\hat{\mathbf{A}}_{k-1}\bar{\bar{\mathbf{H}}}^{-1})\hat{\mathbf{z}}_{GTO_{k-1}} + (\bar{\bar{\mathbf{H}}}\mathbf{B}_{k-1} - \bar{\bar{\mathbf{H}}}\hat{\mathbf{B}}_{k-1})\mathbf{u}_{k-1} \\ & + \bar{\bar{\mathbf{H}}}\mathbf{w}_{k-1} + \boldsymbol{\vartheta}_k - \bar{\bar{\mathbf{H}}}\mathbf{A}_{k-1}\bar{\bar{\mathbf{H}}}^{-1}\boldsymbol{\vartheta}_{k-1} \end{aligned} \quad 5.185$$

If $\bar{\bar{\mathbf{H}}}$ is the identity matrix, then the a priori chattering equation becomes as follows:

$$\mathbf{Ch}_{k|k-1} = \Delta\mathbf{A}_{k-1}\hat{\mathbf{z}}_{GTO_{k-1}} + \Delta\mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1} + \boldsymbol{\vartheta}_k - \mathbf{A}_{k-1}\boldsymbol{\vartheta}_{k-1} \quad 5.186$$

Multiplying both sides by \mathbf{u}_{k-1}^T gives the following:

$$\begin{aligned} \mathbf{Ch}_{k|k-1}\mathbf{u}_{k-1}^T = & \Delta\mathbf{A}_{k-1}\hat{\mathbf{z}}_{GTO_{k-1}}\mathbf{u}_{k-1}^T + \Delta\mathbf{B}_{k-1}\mathbf{u}_{k-1}\mathbf{u}_{k-1}^T + \mathbf{w}_{k-1}\mathbf{u}_{k-1}^T \\ & + \boldsymbol{\vartheta}_k\mathbf{u}_{k-1}^T - \mathbf{A}_{k-1}\boldsymbol{\vartheta}_{k-1}\mathbf{u}_{k-1}^T \end{aligned} \quad 5.187$$

Equation (5.187) describes the term $\mathbf{Ch}_{k|k-1}\mathbf{u}_{k-1}^T$ at each time step. Taking a segment of this signal with length of d (starts with the time step l and ends with the time step $l + d - 1$) and calculating the mean of $\mathbf{Ch}_{k|k-1}\mathbf{u}_{k-1}^T$ of that segment give the following:

$$\begin{aligned} \frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{Ch}_{i|i-1}\mathbf{u}_{i-1}^T & = \frac{1}{d} \sum_{i=l}^{l+d-1} \Delta\mathbf{A}_{i-1}\hat{\mathbf{z}}_{GTO_{i-1}}\mathbf{u}_{i-1}^T + \frac{1}{d} \sum_{i=l}^{l+d-1} \Delta\mathbf{B}_{i-1}\mathbf{u}_{i-1}\mathbf{u}_{i-1}^T \\ & + \frac{1}{d} \sum_{i=l}^{l+d-1} (\mathbf{w}_{i-1}\mathbf{u}_{i-1}^T + \boldsymbol{\vartheta}_i\mathbf{u}_{i-1}^T - \mathbf{A}_{i-1}\boldsymbol{\vartheta}_{i-1}\mathbf{u}_{i-1}^T) \end{aligned} \quad 5.188$$

This segment will be referred to as the *information extraction segment* (IES), to separate it from the segments used to obtain the General Observability matrix. The system and the input matrices are assumed to be constant inside the IES; hence the modeling errors are constant and defined as follows:

$$\Delta\mathbf{A}_{i-1}|_{i=l\dots l+d-1} = \Delta\mathbf{A}_l \text{ and } \Delta\mathbf{B}_{i-1}|_{i=l\dots l+d-1} = \Delta\mathbf{B}_l \quad 5.189$$

Substituting equation (5.189) into equation (5.188) gives the following:

$$\begin{aligned} \frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{C} \mathbf{h}_{i|i-1} u_{i-1}^T &= \frac{1}{d} \Delta \mathbf{A}_l \sum_{i=l}^{l+d-1} \hat{\mathbf{z}}_{GTO_{i-1}} u_{i-1}^T + \frac{1}{d} \Delta \mathbf{B}_l \sum_{i=l}^{l+d-1} u_{i-1} u_{i-1}^T \\ &+ \frac{1}{d} \sum_{i=l}^{l+d-1} (\mathbf{w}_{i-1} u_{i-1}^T + \boldsymbol{\vartheta}_i u_{i-1}^T - \mathbf{A}_{i-1} \boldsymbol{\vartheta}_{i-1} u_{i-1}^T) \end{aligned} \quad 5.190$$

Note that as $d \rightarrow \infty$, the vectors $\frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{w}_{i-1} u_{i-1}^T$, $\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_i u_{i-1}^T$ and $\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_{i-1} u_{i-1}^T$ approaches their expectations. If the noise vectors are white, then:

$$\begin{aligned} \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{w}_{i-1} u_{i-1}^T \right) &= \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_i u_{i-1}^T \right) \\ &= \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_{i-1} u_{i-1}^T \right) = \mathbf{0}_{n \times 1} \end{aligned} \quad 5.191$$

Substituting equation (5.191) into equation (5.190) gives the following:

$$\begin{aligned} \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{C} \mathbf{h}_{i|i-1} u_{i-1}^T \right) &= \frac{1}{d} \lim_{d \rightarrow \infty} \left(\Delta \mathbf{A}_l \sum_{i=l}^{l+d-1} \hat{\mathbf{z}}_{GTO_{i-1}} u_{i-1}^T + \Delta \mathbf{B}_l \sum_{i=l}^{l+d-1} u_{i-1} u_{i-1}^T \right) \end{aligned} \quad 5.192$$

Now, the a priori chattering is multiplied by the alternative measurement $\hat{\mathbf{z}}_{GTO_{k_1-1}}^T$ as follows:

$$\begin{aligned} \mathbf{C} \mathbf{h}_{k|k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T &= \Delta \mathbf{A}_{k-1} \hat{\mathbf{z}}_{GTO_{k-1}} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T + \Delta \mathbf{B}_{k-1} u_{k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \\ &+ \mathbf{w}_{k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T + \boldsymbol{\vartheta}_k \hat{\mathbf{z}}_{GTO_{k_1-1}}^T - \mathbf{A}_{k-1} \boldsymbol{\vartheta}_{k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \end{aligned} \quad 5.193$$

Where k_1 is a time step that can be chosen as follows:

$$k > k_1 + m_a \text{ OR } \{\Delta k = k - k_1\} > m_a \quad 5.194$$

Where m_a is the maximum LM_i (segment length in the General Observability matrix). This assumption causes the following:

$$E \left[\mathbf{w}_{k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \right] = E \left[\boldsymbol{\vartheta}_k \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \right] = E \left[\boldsymbol{\vartheta}_{k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \right] = \mathbf{0}_{n \times n} \quad 5.195$$

Proof:

Using equations (5.154), the alternative measurement at time $k_1 - 1$ is defined as follows:

$$\hat{\mathbf{z}}_{GTO_{k_1-1}} = \bar{\mathbf{H}}\mathbf{x}_{k_1-1} + \boldsymbol{\vartheta}_{k_1-1} \quad 5.196$$

where $\boldsymbol{\vartheta}_{k_1-1}$ is the alternative measurement noise at time $k_1 - 1$ and is defined by using equation (5.155) as follows:

$$\boldsymbol{\vartheta}_{k_1-1} = \bar{\mathbf{H}} \left(\bar{\mathbf{O}}_g^{-1} \begin{pmatrix} v_{1_{k_1-1}} \\ \vdots \\ v_{1_{k_1+LM_1-2}} \\ \vdots \\ v_{m_{k_1-1}} \\ \vdots \\ v_{m_{k_1+LM_m-2}} \end{pmatrix} + \mathbf{T}_w \begin{bmatrix} \mathbf{w}_{k_1-1} \\ \vdots \\ \mathbf{w}_{k_1+m_a-2} \end{bmatrix} \right) \quad 5.197$$

These equations show that $\hat{\mathbf{z}}_{GTO_{k_1-1}}$ and $\boldsymbol{\vartheta}_{k_1-1}$ are functions of the measurement and the system noise vectors from time step $(k_1 - 1)$ to time step $(k_1 + m_a - 2)$ as shown in Fig 5.14. Note that the cross correlation between $\hat{\mathbf{z}}_{GTO_{k_1-1}}$ and \mathbf{w}_{k-1} vanishes if they are white signals, and their occurrence time steps have a difference that is larger than the maximum segment length in the General Observability matrix (equation (5.194)). This means that there is no overlap between these vectors as shown in Fig 5.14, and the expectation of equation (5.195) becomes as follows:

$$E \left[\mathbf{w}_{k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \right] = E \left[\mathbf{w}_{k-1} \boldsymbol{\vartheta}_{k_1-1}^T \right] = E \left[\mathbf{w}_{k-1} \begin{bmatrix} \mathbf{w}_{k_1-1} \\ \vdots \\ \mathbf{w}_{k_1+m_a-2} \end{bmatrix}^T \left(\bar{\mathbf{H}} \bar{\mathbf{O}}_g^{-1} \mathbf{T}_w \right)^T \right] = \mathbf{0}_{n \times n} \quad 5.198$$

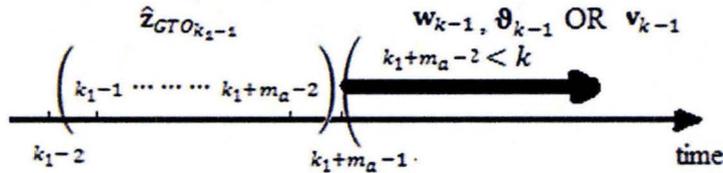


Fig 5.14: The time steps of the measurement and system noise vectors that are used to obtain the alternative measurement vector at time $k_1 - 1$ compared to the measurement and system noise vectors at time $k - 1$.

The same idea holds for the alternative measurement noise vector at time $k - 1$ and k . Thereby, $E \left[\boldsymbol{\vartheta}_k \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \right] = E \left[\boldsymbol{\vartheta}_{k-1} \hat{\mathbf{z}}_{GTO_{k_1-1}}^T \right] = \mathbf{0}_{n \times n}$.

Taking an IES of the signal $\mathbf{Ch}_{k|k-1}\hat{\mathbf{z}}_{GTO_{k-1}}^T$ with length d (starts with the time step l and ends with the time step $l + d - 1$) and calculating its mean give the following:

$$\begin{aligned} \frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{Ch}_{i|i-1}\hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T &= \frac{1}{d} \Delta \mathbf{A}_l \sum_{i=l}^{l+d-1} \hat{\mathbf{z}}_{GTO_{i-1}} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T + \frac{1}{d} \Delta \mathbf{B}_l \sum_{i=l}^{l+d-1} u_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \\ &+ \frac{1}{d} \sum_{i=l}^{l+d-1} \left(\begin{array}{c} \mathbf{w}_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \\ + \boldsymbol{\vartheta}_i \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T - \mathbf{A}_{i-1} \boldsymbol{\vartheta}_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \end{array} \right) \end{aligned} \quad 5.199$$

Note that as $d \rightarrow \infty$, the matrices $\frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{w}_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T$, $\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_i \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T$ and $\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T$ approach their respective expectations. If the noise vectors are white, then the following is obtained by using equation (5.195).

$$\begin{aligned} \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{w}_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \right) &= \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_i \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \right) \\ &= \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \boldsymbol{\vartheta}_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \right) = \mathbf{0}_{n \times n} \end{aligned} \quad 5.200$$

Substituting equation (5.200) into equation (5.199) gives the following:

$$\begin{aligned} \lim_{d \rightarrow \infty} \left(\frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{Ch}_{i|i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \right) &= \frac{1}{d} \Delta \mathbf{A}_l \lim_{d \rightarrow \infty} \left(\sum_{i=l}^{l+d-1} \hat{\mathbf{z}}_{GTO_{i-1}} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \right) \\ &+ \frac{1}{d} \Delta \mathbf{B}_l \lim_{d \rightarrow \infty} \left(\sum_{i=l}^{l+d-1} u_{i-1} \hat{\mathbf{z}}_{GTO_{i-\Delta k-1}}^T \right) \end{aligned} \quad 5.201$$

Equations (5.192) and (5.201) can be rewritten as follows:

$$\begin{aligned} \lim_{d \rightarrow \infty} \left(\mathbf{C}_{\mathbf{Ch}_{d|d-1}, u_{d-1}} \right) &= \Delta \mathbf{A}_l \lim_{d \rightarrow \infty} \left(\mathbf{C}_{\hat{\mathbf{z}}_{GTO_{d-1}}, u_{d-1}} \right) + \Delta \mathbf{B}_l \lim_{d \rightarrow \infty} \left(\mathbf{C}_{u_{d-1}, u_{d-1}} \right) \\ \lim_{d \rightarrow \infty} \left(\mathbf{C}_{\mathbf{Ch}_{d|d-1}, \hat{\mathbf{z}}_{GTO_{d-\Delta k-1}}} \right) &= \Delta \mathbf{A}_l \lim_{d \rightarrow \infty} \left(\mathbf{C}_{\hat{\mathbf{z}}_{GTO_{d-1}}, \hat{\mathbf{z}}_{GTO_{d-\Delta k-1}}} \right) + \Delta \mathbf{B}_l \lim_{d \rightarrow \infty} \left(\mathbf{C}_{u_{d-1}, \hat{\mathbf{z}}_{GTO_{d-\Delta k-1}}} \right) \end{aligned} \quad 5.202$$

Where $\mathbf{C}_{\mathbf{a}_b, \mathbf{c}_d}$ is the segment's mean of the multiplication (cross-correlation) between \mathbf{a}_c and \mathbf{b}_e^T and it is defined as follows:

$$\mathbf{C}_{\mathbf{a}_c, \mathbf{b}_e} = \frac{1}{d} \sum_{i=l}^{l+d-1} \mathbf{a}_c \mathbf{b}_e^T \quad 5.203$$

The modeling errors ($\Delta \mathbf{A}_l$ and $\Delta \mathbf{B}_l$) can be obtained by solving the system of equations shown in equation (5.203) and have the following form:

$$\Delta \mathbf{A}_l = \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, u_{d-1}}}) \mathbf{M} - \lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, \hat{z}_{GTO_{d-\Delta k-1}}}}) \lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, \hat{z}_{GTO_{d-\Delta k-1}}}})^{-1} \right) \times \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, u_{d-1}}}) \mathbf{M} - \mathbf{I}_{n \times n} \right)^{-1}$$

$$\Delta \mathbf{B}_l = \left[- \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, u_{d-1}}}) \mathbf{M} - \left[\lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, \hat{z}_{GTO_{d-\Delta k-1}}}) \times \lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, \hat{z}_{GTO_{d-\Delta k-1}}}})^{-1} \right] \right) \lim_{d \rightarrow \infty} (\mathbf{C}_{u_{d-1}, u_{d-1}})^{-1} \right. \\ \left. \times \left(\lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, u_{d-1}}}) \mathbf{M} - \mathbf{I}_{n \times n} \right)^{-1} \lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, u_{d-1}}}) \right]$$
5.204

$$\text{Where } \mathbf{M} = \lim_{d \rightarrow \infty} (\mathbf{C}_{u_{d-1}, u_{d-1}})^{-1} \lim_{d \rightarrow \infty} (\mathbf{C}_{u_{d-1}, \hat{z}_{GTO_{d-\Delta k-1}}}) \lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, \hat{z}_{GTO_{d-\Delta k-1}}}})^{-1}.$$

If the input is a white signal (i.e. $E(u_k) = 0$ and $E(u_k u_{k-1}^T) = 0$), then the term $\lim_{d \rightarrow \infty} (\mathbf{C}_{u_{d-1}, \hat{z}_{GTO_{d-\Delta k-1}}})$ vanishes (this is valid as long as $\Delta k > m_a$). In such a case, the modeling errors are obtained as follows:

$$\Delta \mathbf{A}_l = \lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, \hat{z}_{GTO_{d-\Delta k-1}}}}) \lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, \hat{z}_{GTO_{d-\Delta k-1}}}})^{-1}$$

$$\Delta \mathbf{B}_l = \left(- \left[\lim_{d \rightarrow \infty} (\mathbf{C}_{\text{Ch}_{d|d-1, \hat{z}_{GTO_{d-\Delta k-1}}}) \times \lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, \hat{z}_{GTO_{d-\Delta k-1}}}})^{-1} \right] \lim_{d \rightarrow \infty} (\mathbf{C}_{\hat{z}_{GTO_{d-1}, u_{d-1}}}) \right) \lim_{d \rightarrow \infty} (\mathbf{C}_{u_{d-1}, u_{d-1}})^{-1}$$
5.205

5.5.5. Application of the General Toeplitz/Observability SVSF to a three degrees of freedom system

The General Toeplitz/Observability SVSF has been tested on the three degrees of freedom system described in Example 5.2.

5.5.5.1. Simulation setup

If the system described in Example 5.2 has the following parameters:

$$k_1 = 20 \frac{kN}{m}, k_2 = 1 \frac{kN}{m}, k_3 = 1 \frac{kN}{m}, k_4 = 400 \frac{kN}{m}, c_1 = 2 \frac{kN \cdot sec}{m}, c_2 = 2 \frac{kN \cdot sec}{m}, c_3 = 2 \frac{kN \cdot sec}{m}, c_4 = 2 \frac{kN \cdot sec}{m}, b_1 = 2, b_2 = 1, b_3 = 1.5, M_1 = 10kg, M_2 = 5kg \text{ and } M_3 = 500 kg$$
5.206

and assuming the sampling time is equal to $T_s = 0.001 \text{ sec}$, then the system has the following model:

$$\begin{bmatrix} x_{1_{1k+1}} \\ x_{1_{2k+1}} \\ x_{2_{1k+1}} \\ x_{2_{2k+1}} \\ x_{3_{1k+1}} \\ x_{3_{2k+1}} \end{bmatrix} = \begin{bmatrix} 1 & 0.001 & 0 & 0 & 0 & 0 \\ -42.1 & 0.4 & 1 & 0.2 & 40 & 0.2 \\ 0 & 0 & 1 & 0.001 & 0 & 0 \\ 0.2 & 0.4 & -0.4 & 0.2 & 0.2 & 0.4 \\ 0 & 0 & 0 & 0 & 1 & 0.001 \\ 0.8 & 0.004 & 0.002 & 0.004 & -0.802 & 0.992 \end{bmatrix} \begin{bmatrix} x_{1_{1k}} \\ x_{1_{2k}} \\ x_{2_{1k}} \\ x_{2_{2k}} \\ x_{3_{1k}} \\ x_{3_{2k}} \end{bmatrix} + \begin{bmatrix} 0 \\ 0.002 \\ 0 \\ 0.001 \\ 0 \\ -0.0015 \end{bmatrix} u_k + \begin{bmatrix} w_{1_{1k}} \\ w_{1_{2k}} \\ w_{2_{1k}} \\ w_{2_{2k}} \\ w_{3_{1k}} \\ w_{3_{2k}} \end{bmatrix} \quad 5.207$$

The measurement vector is defined as follows:

$$\begin{bmatrix} z_{1k} \\ z_{2k} \\ z_{3k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_{1_{1k}} & x_{1_{2k}} & x_{2_{1k}} & x_{2_{2k}} & x_{3_{1k}} & x_{3_{2k}} \end{bmatrix}^T + \begin{bmatrix} v_{1k} \\ v_{2k} \\ v_{3k} \end{bmatrix} \quad 5.208$$

The measurement and system noise vectors are white with a noise-to-signal ratio of 1% of their corresponding states, and with variances of $\sigma_v^2 = 10^{-20} \times \begin{bmatrix} 1.25 & 0 & 0 \\ 0 & 4.635 & 0 \\ 0 & 0 & 0.32 \end{bmatrix}$

and $\sigma_w^2 = 10^{-20} \times \begin{bmatrix} 1.4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.3 \times 10^5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4.8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.2 \times 10^4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.35 & 0 \\ 0 & 0 & 0 & 0 & 0 & 512 \end{bmatrix}$, respectively. The first derivative of

v_{1k} , v_{2k} and v_{3k} have a noise-to-signal ratio of 5% of the states $x_{1_{2k}}$, $x_{2_{2k}}$ and $x_{3_{2k}}$, respectively. The first derivative of $w_{1_{1k}}$, $w_{2_{1k}}$ and $w_{3_{1k}}$ have a noise-to-signal ratio of 5% of the states $x_{1_{2k}}$, $x_{2_{2k}}$ and $x_{3_{2k}}$, respectively. The SVSF's coefficient matrix γ has a value of $\gamma = 0.2 \times \mathbf{I}_{6 \times 6}$ and the SVSF's smooth boundary layer is designed as $\Psi = 2 \times [10^{-5} \ 10^{-3} \ 10^{-5} \ 10^{-2} \ 10^{-5} \ 10^{-3}]$. The input consists of a random signal superimposed on step changes. The simulation time consists of ten regions. The system has some parametric changes (faults) in some of these regions. The faults include changes in at least one of the following:

- Spring stiffness.
- Damping coefficients.
- Masses.
- Input gains.

These changes and their start and end time steps are listed in tables (5.10) and (5.11). Note that the modeling errors vary between 80 – 320% of their actual values as shown in table 5.11.

Region #	Beginning time step	Ending time step	Region #	Beginning time step	Ending time step
1	1	11664	6	70525	81858
2	11665	25510	7	81859	98361
3	25511	40209	8	98362	113810
4	40210	60080	9	113811	126038
5	60081	70524	10	126039	142413

Table 5.10: The beginning and the ending of the changes.

region	1	2	3	4	5	6	7	8	9	10
	error %									
a_{21}	80	80	80	80	80	80	80	38	80	80
a_{22}	80	80	80	80	80	80	80	87	90	87
a_{23}	80	80	80	80	80	100	80	80	80	80
a_{24}	80	80	80	80	80	80	80	80	100	100
a_{25}	80	80	80	80	80	80	80	100	80	80
a_{26}	320	320	320	320	320	320	320	320	320	320
a_{41}	80	80	80	80	80	100	80	80	80	80
a_{42}	80	80	80	80	80	80	80	80	100	100
a_{43}	80	80	80	80	60	100	80	80	80	80
a_{44}	80	80	80	80	93	80	80	80	96	96
a_{45}	80	80	80	80	100	100	80	80	80	80
a_{46}	120	120	120	120	120	120	120	120	120	120
a_{61}	80	80	96	80	80	80	80	100	80	80
a_{62}	80	80	96	80	80	80	80	80	80	80
a_{63}	80	80	96	80	100	100	80	80	80	80
a_{64}	80	80	96	80	100	80	80	80	100	100
a_{65}	80	80	96	80	80	80	80	95	80	80
a_{66}	0.645	0.645	4	0.645	0.241	0.645	0.645	0.645	0.241	0.241
b_2	100	100	100	100	100	100	100	100	100	100
b_4	100	100	100	100	100	100	100	100	100	100
b_6	100	100	100	100	100	100	100	100	100	100

Table 5.11: The % of the error between the system and the filter parameters for the simulation regions.

The system in equation (5.207) is represented by its GOCF, and the measurement matrix is described by equation (5.208) which satisfies the conditions of equations (5.115) and (5.116). Therefore, the General Observability/Toeplitz SVSF can be used to estimate

the states and its chattering signal can be used to point out and extract modeling errors. The general Observability matrix is obtained using the algorithm in section (5.3.3.2.2). From Example 5.8, the locations of the measured states, LMS_i , are set to the values of 1, 3 and 5. The segments' lengths are defined in equation (5.121) as follows:

$$\begin{aligned} LM_1 &= LMS_2 - LMS_1 = 3 - 1 = 2 \\ LM_2 &= LMS_3 - LMS_2 = 5 - 3 = 2 \\ LM_3 &= n + 1 - LMS_3 = 6 + 1 - 5 = 2 \end{aligned} \quad 5.121$$

The General Observability matrix is defined by equation (5.122) as follows:

$$\mathbf{O}_g = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_1\mathbf{A} \\ \mathbf{H}_2 \\ \mathbf{H}_2\mathbf{A} \\ \mathbf{H}_3 \\ \mathbf{H}_3\mathbf{A} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} & 0 & 0 & 0 & 0 \\ \begin{bmatrix} 1 & T_s \end{bmatrix} & 0 & 0 & 0 & 0 \\ 0 & 0 & \begin{bmatrix} 1 & 0 \end{bmatrix} & 0 & 0 \\ 0 & 0 & \begin{bmatrix} 1 & T_s \end{bmatrix} & 0 & 0 \\ 0 & 0 & 0 & 0 & \begin{bmatrix} 1 & 0 \end{bmatrix} \\ 0 & 0 & 0 & 0 & \begin{bmatrix} 1 & T_s \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \mathbf{O}_{sub_1} & \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \mathbf{O}_{sub_2} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 2} & \mathbf{0}_{2 \times 2} & \mathbf{O}_{sub_3} \end{bmatrix} \quad 5.122$$

The General System Toeplitz matrix is a zero matrix.

5.5.5.2. Simulation results

The results of the General Toeplitz/Observability SVSF are shown in Figs (5.15) to (5.41). These results show that the method gives good (the average absolute percentage errors of the states and parameters are less than 4%) robust performance. Although the second sub-system in the ninth region is unstable, the filter tracks the estimates with acceptable level of accuracy (the maximum absolute percentage errors of the states and parameters - except those of region 4 - are less than 2%), and the parameters are reconstructed by using the chattering signal. The filter detects parametric changes once they occur, and then uses a data segment consisting of 500 sampling points from the chattering to reconstruct the modeling errors as discussed in section (5.5.4). The sample's size qualifies the law of large numbers as discussed in section (4.3.4.2.2). Once the filter's model is tuned, the filter continues estimating the states. During the model tuning and reconstruction process, the results show that the stability is not affected by modeling errors.

As discussed in section (4.3.4.2.2), the detection process is sensitive to the width of the smoothing boundary layer. The results of this simulation show that there is no occurrence of false alarms. However, delays in detecting modeling errors have been noticed at the beginning of regions 3, 4 and 5. Table 5.12 shows that the maximum delay period is around 20 time steps, and it occurs at the fourth time interval that starts at time

step 40,210. These delays mean that the width of the smoothing boundary layer is quite large and needs to be reduced. For regions 1, 2, 6, 7, 8, 9 and 10, the results show that the algorithm detects parametric changes once they occur.

Region #	Time step of fault being detected	Region #	Time step of fault being detected
1	1	6	70525
2	11665	7	81859
3	25521	8	98362
4	40230	9	113811
5	60085	10	126039

Table 5.12: The estimated beginning of each region for the tuning process simulation.

The estimated states are sensitive to the noise vectors and their derivatives. In order to reconstruct the modeling errors with acceptable levels of accuracy, these vectors should have noise-to-signal ratios that are less than 10%. If the ratios increase, the estimated error increases and the parameters need more iterations to converge to their actual values. For example, the noise vectors' ratios exceed 10% in the fourth region, and it has been noticed that the maximum absolute error increased to 75% for a_{41} and a_{45} . Although the errors increase in these two parameters, the filter remains stable.

The root mean square error of the estimated states and parameters (described by equation (5.209)) are summarized in table 5.13. Note that no other published methods (up to this thesis) can replicate these results.

$$RMSE_j = \frac{1}{length(x)} \sqrt{\sum_{i=1}^{length(x)} (y_{ji} - \hat{y}_{ji})^2} \text{ for } y = x_{1_1}, x_{1_2}, x_{2_1}, x_{2_2}, x_{3_1}, x_{3_2}, a_{21}, a_{22}, \tag{5.209}$$

$a_{23}, a_{24}, a_{25}, a_{26}, a_{41}, a_{42}, a_{43}, a_{44}, a_{45}, a_{46}, a_{61}, a_{62}, a_{63}, a_{64}, a_{65}, a_{66}, b_2, b_4$ and b_6

	RMSE		RMSE		RMSE		RMSE
x_{1_1}	8.7×10^{-15}	a_{22}	1.7×10^{-06}	a_{43}	2.8×10^{-06}	a_{64}	9.1×10^{-09}
x_{1_2}	2.64×10^{-11}	a_{23}	2.26×10^{-07}	a_{44}	1.3×10^{-05}	a_{65}	2.7×10^{-04}
x_{2_1}	1.7×10^{-14}	a_{24}	5.4×10^{-07}	a_{45}	5.8×10^{-05}	a_{66}	2.6×10^{-08}
x_{2_2}	4.6×10^{-10}	a_{25}	4.47×10^{-02}	a_{46}	6.6×10^{-06}	b_2	1.74×10^{-10}
x_{3_1}	1×10^{-14}	a_{26}	1.1×10^{-07}	a_{61}	2.7×10^{-04}	b_4	2.13×10^{-11}
x_{3_2}	4.7×10^{-11}	a_{41}	6.35×10^{-05}	a_{62}	7.5×10^{-09}	b_6	9.1×10^{-11}
a_{21}	4.33×10^{-02}	a_{42}	2.1×10^{-06}	a_{63}	6.7×10^{-09}		

Table 5.13: The root mean square error of the estimated states and parameters obtained from the General Toeplitz/Observability SVSF's tuning process.

The proposed algorithm has some limitations which can be summarized as follows:

- The measurement matrix must specify certain conditions.
- The system must be described in its GOCF.
- The system and measurement noise vectors must have small noise-to-signal ratios, i.e. their derivatives have noise-to-signal ratios that are less than 10%. Higher ratios increase the errors in the constructed parameters.
- The estimates are sensitive to the noise amplitude and its derivatives.

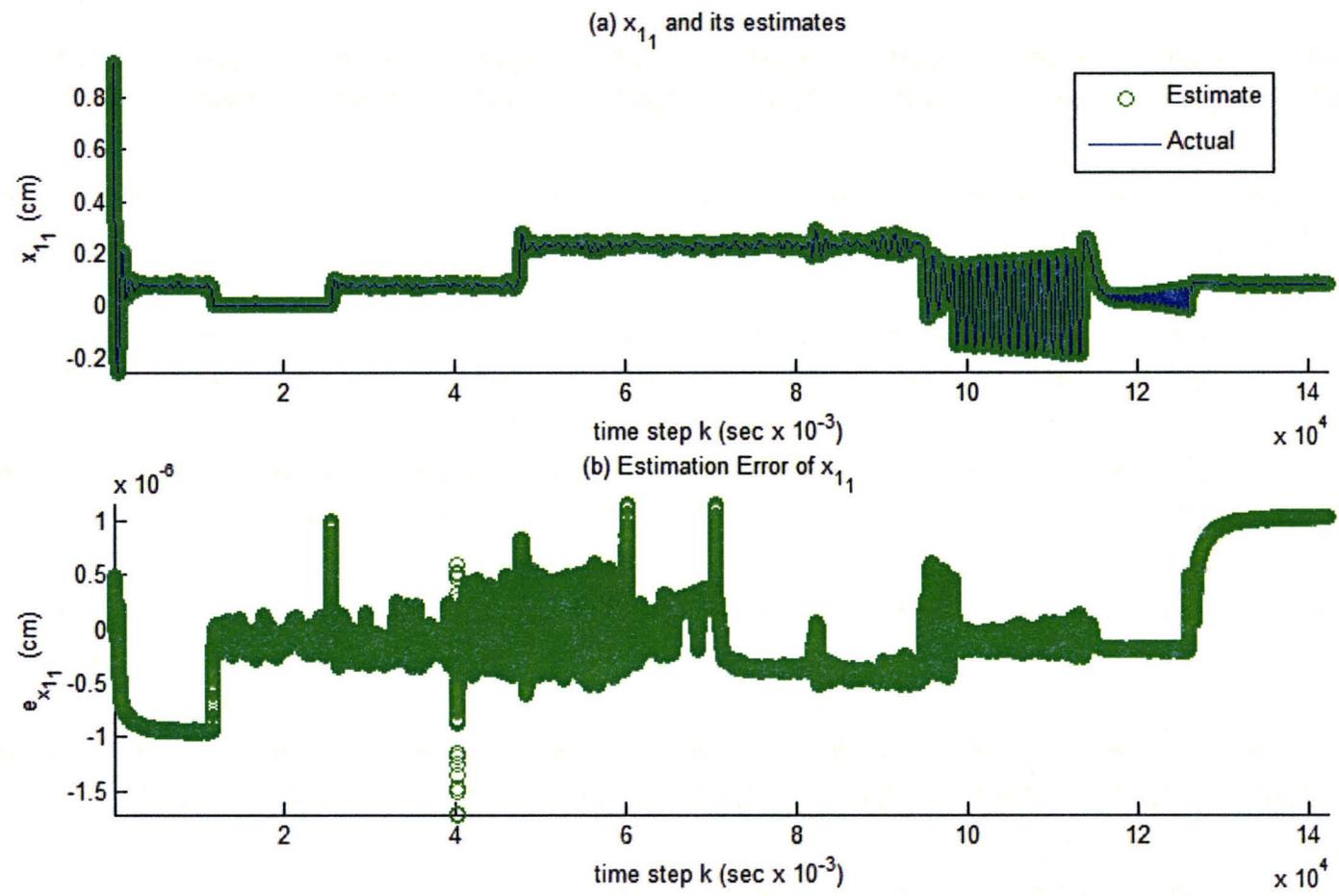


Fig 5.15: (a) x_{1_1} 's actual and estimated values, (b) estimation error of x_{1_1} for the General Toeplitz/Observability SVSF's tuning process.

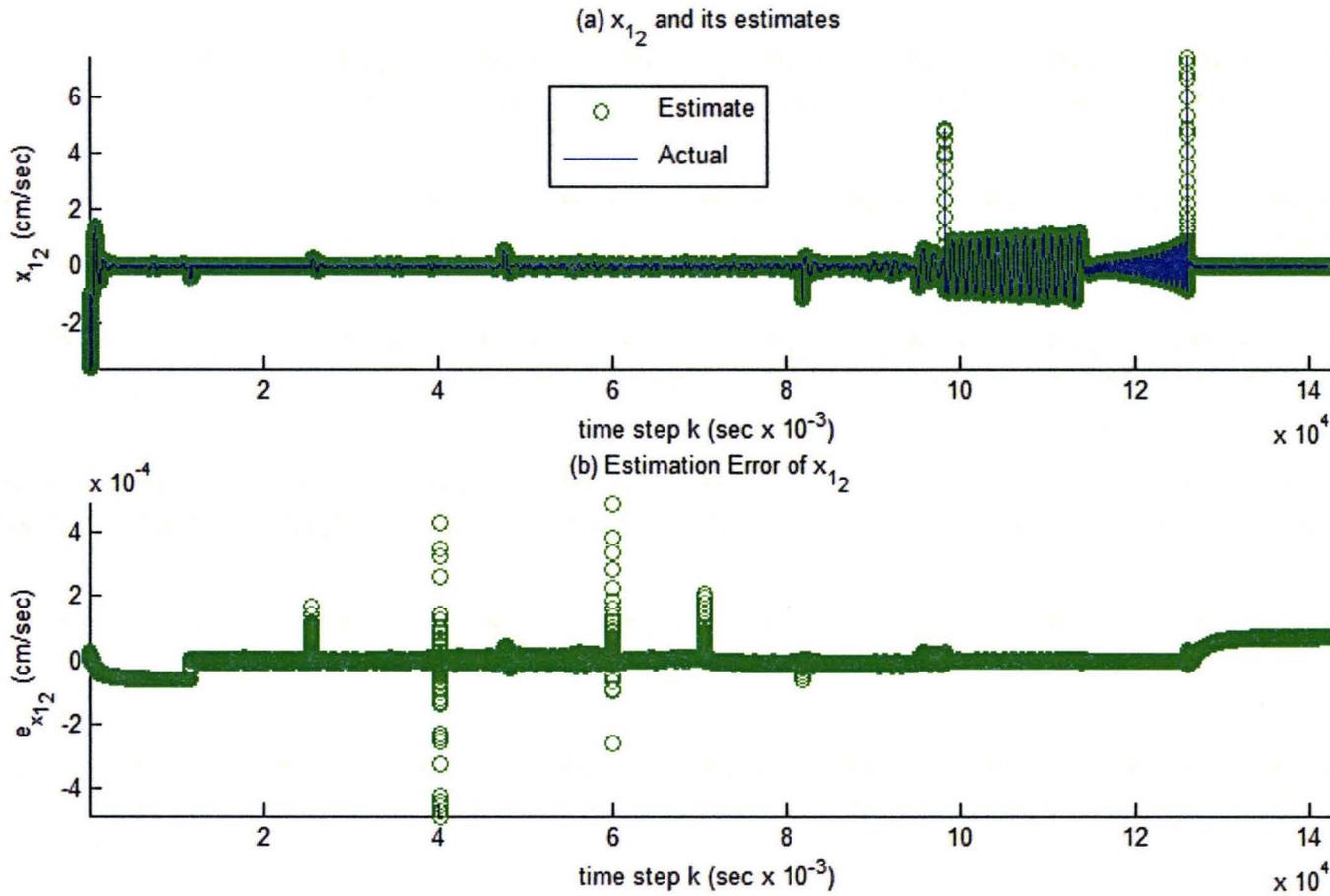


Fig 5.16: (a) $x_{1,2}$'s actual and estimated values, (b) estimation error of $x_{1,2}$ for the General Toeplitz/Observability SVSF's tuning process.

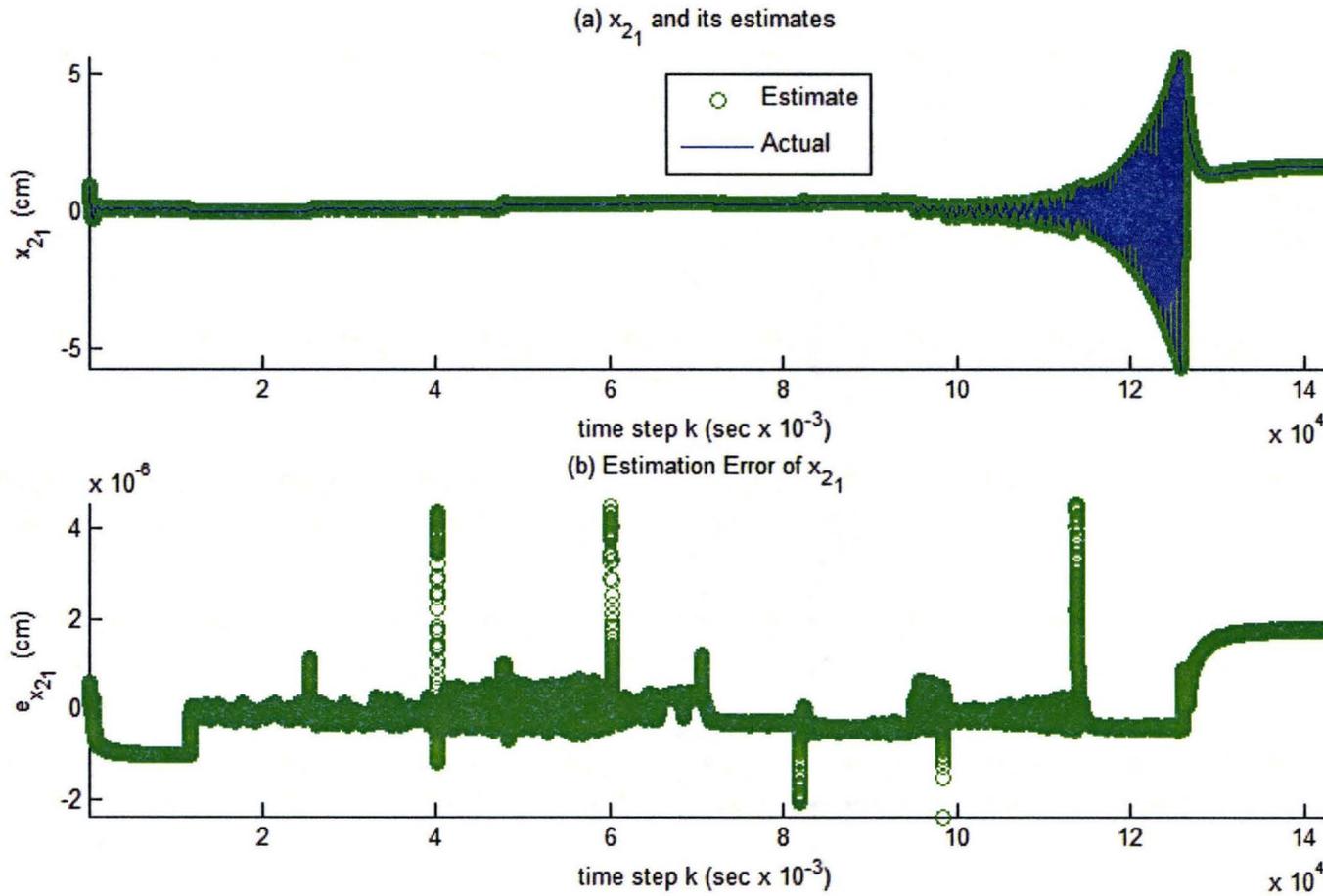


Fig 5.17: (a) x_{21} 's actual and estimated values, (b) estimation error of x_{21} for the General Toeplitz/Observability SVSF's tuning process.

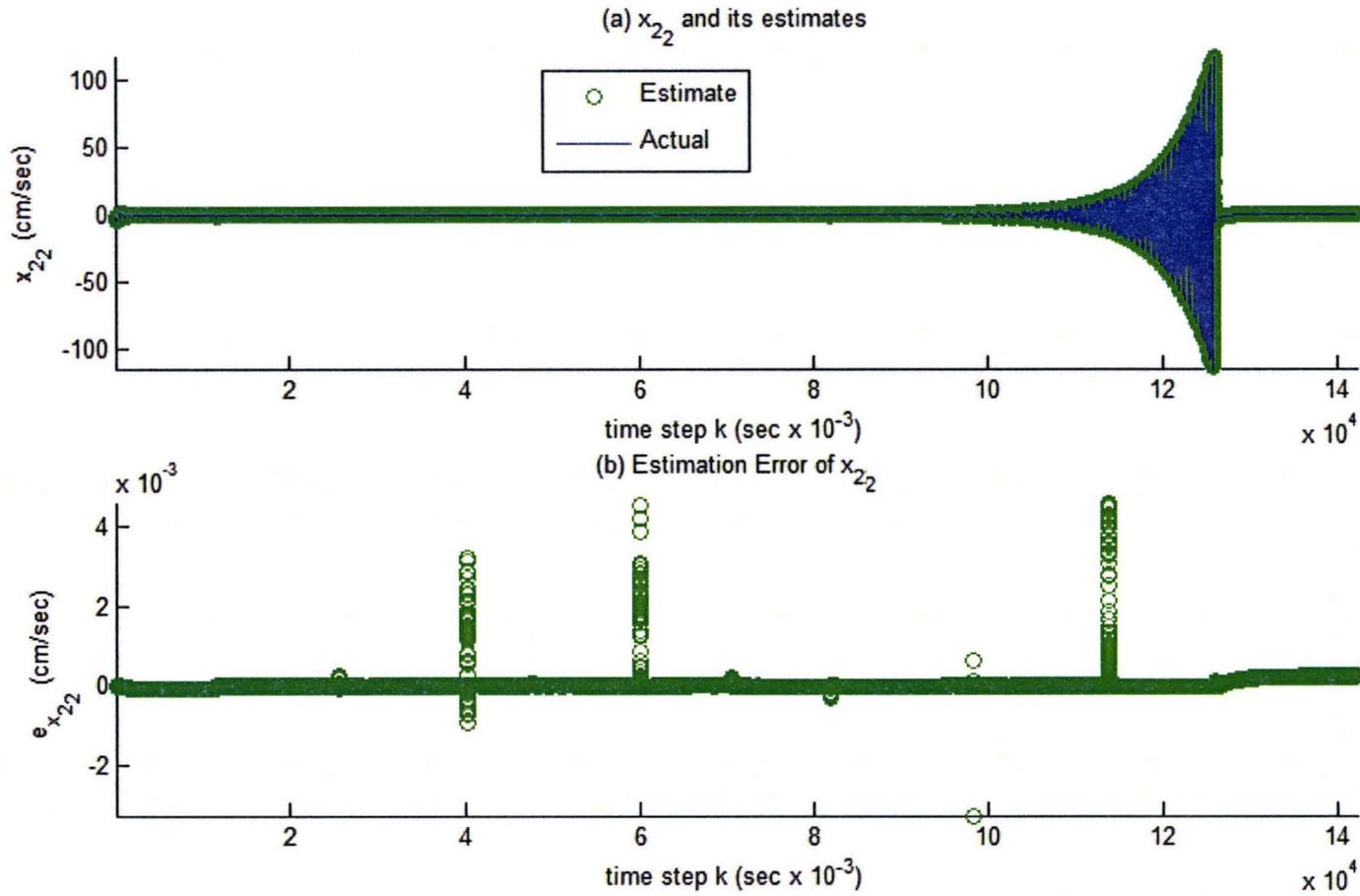


Fig 5.18: (a) x_{2_2} 's actual and estimated values, (b) estimation error of x_{2_2} for the General Toeplitz/Observability SVSF's tuning process.

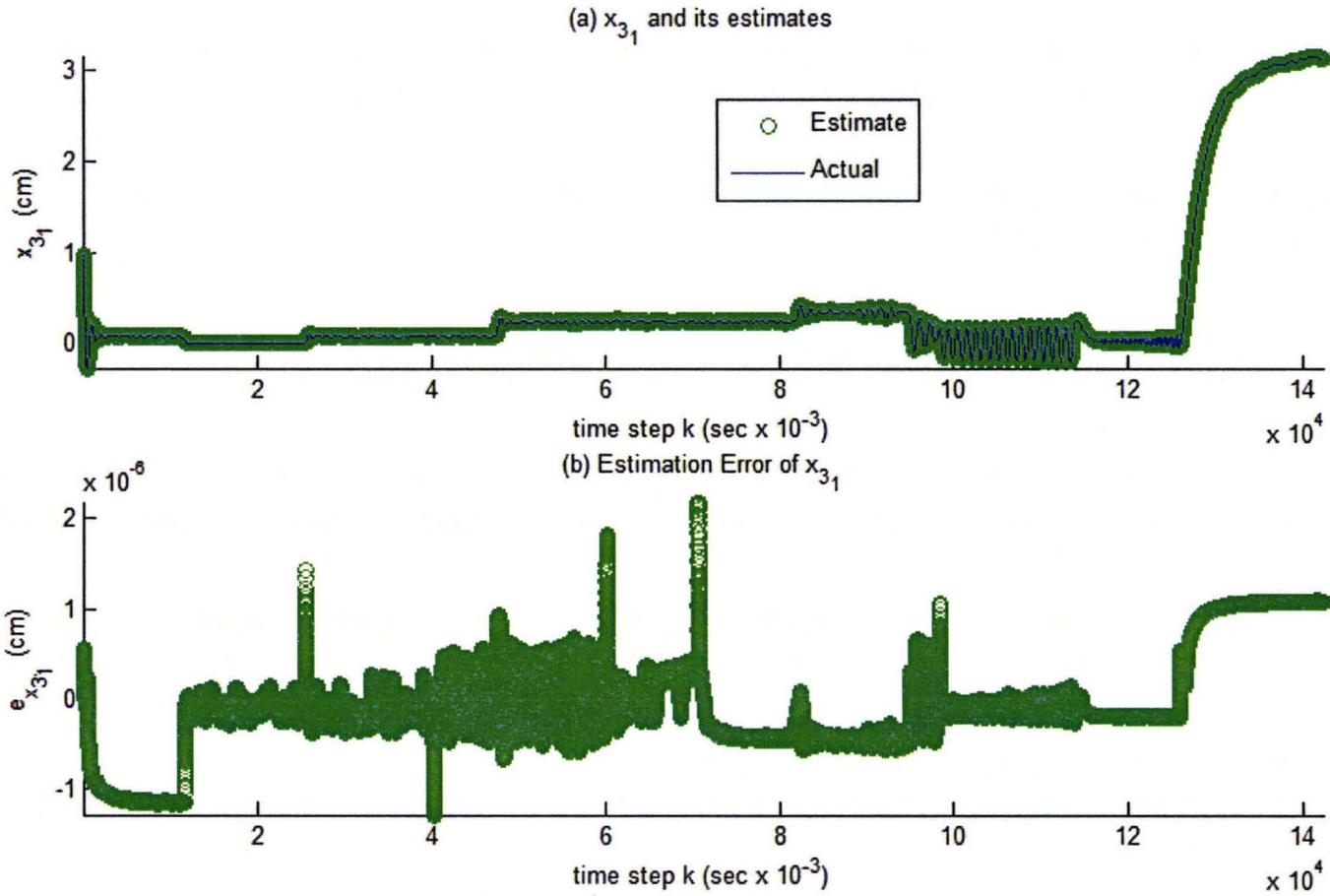


Fig 5.19: (a) x_{3_1} 's actual and estimated values, (b) estimation error of x_{3_1} for the General Toeplitz/Observability SVSF's tuning process.

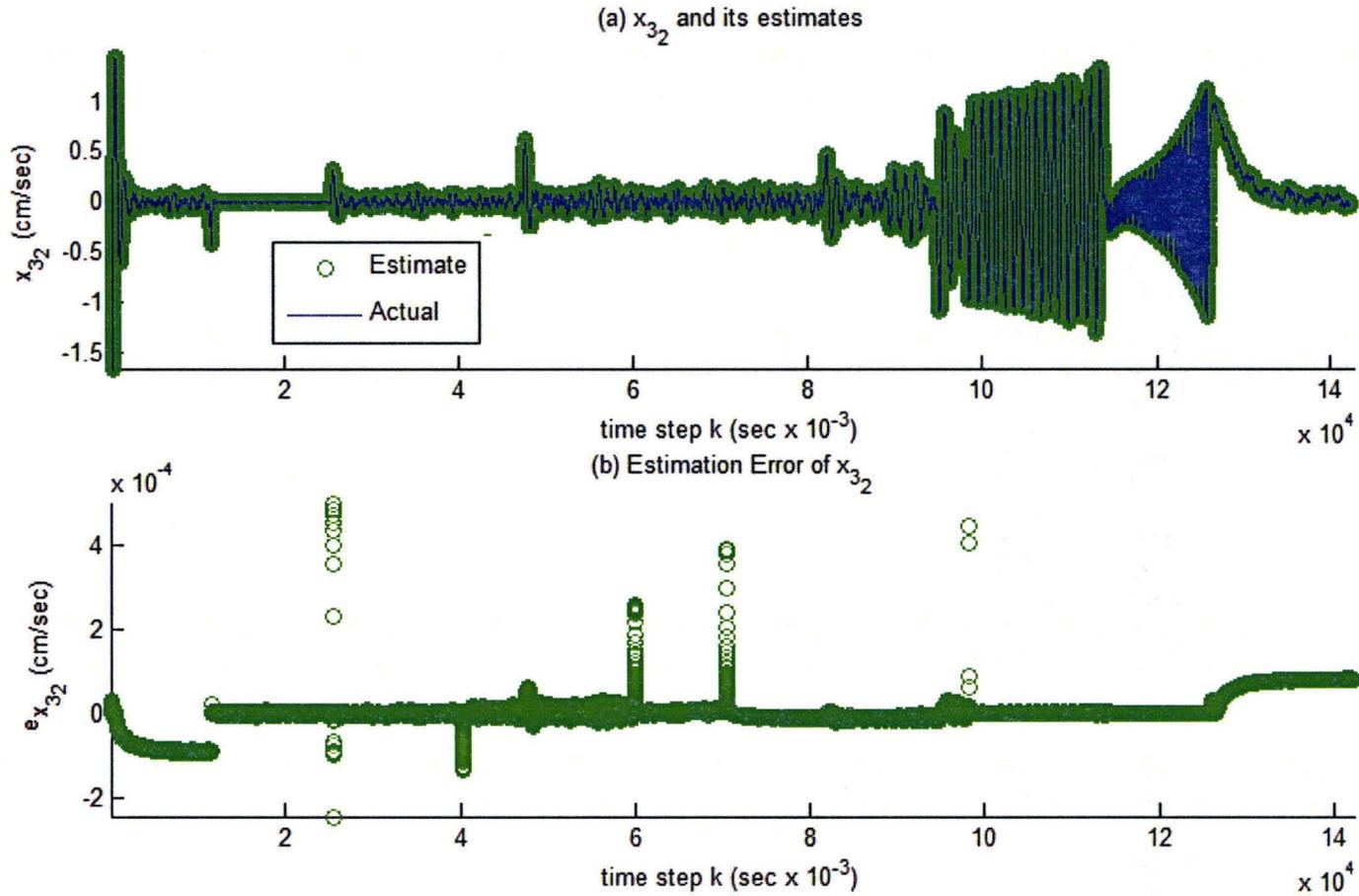


Fig 5.20: (a) x_{32} 's actual and estimated values, (b) estimation error of x_{32} for the General Toeplitz/Observability SVSF's tuning process.

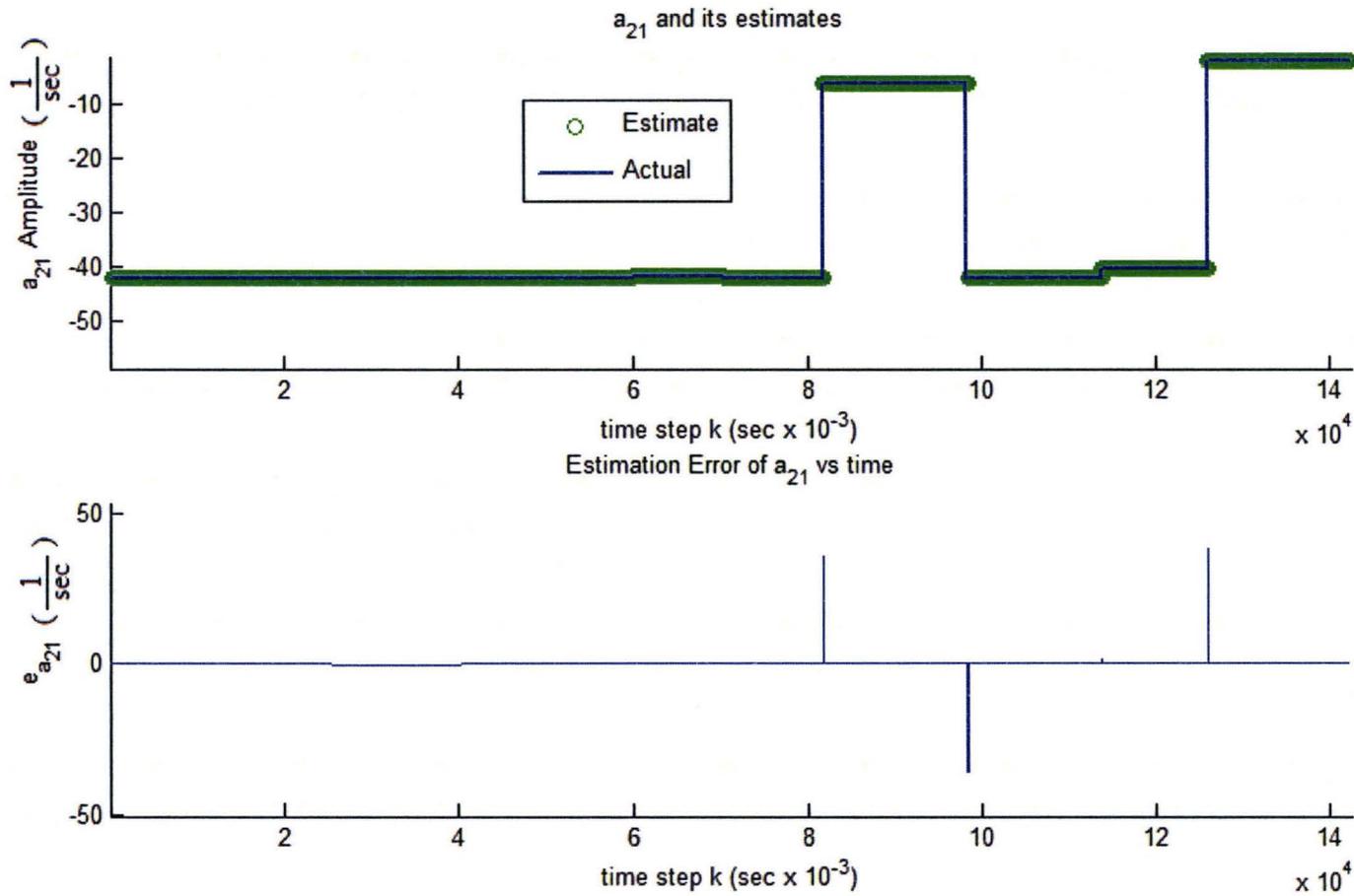


Fig 5.21: (a) a_{21} 's actual and estimated values, (b) estimation error of a_{21} for the General Toeplitz/Observability SVSF's tuning process.

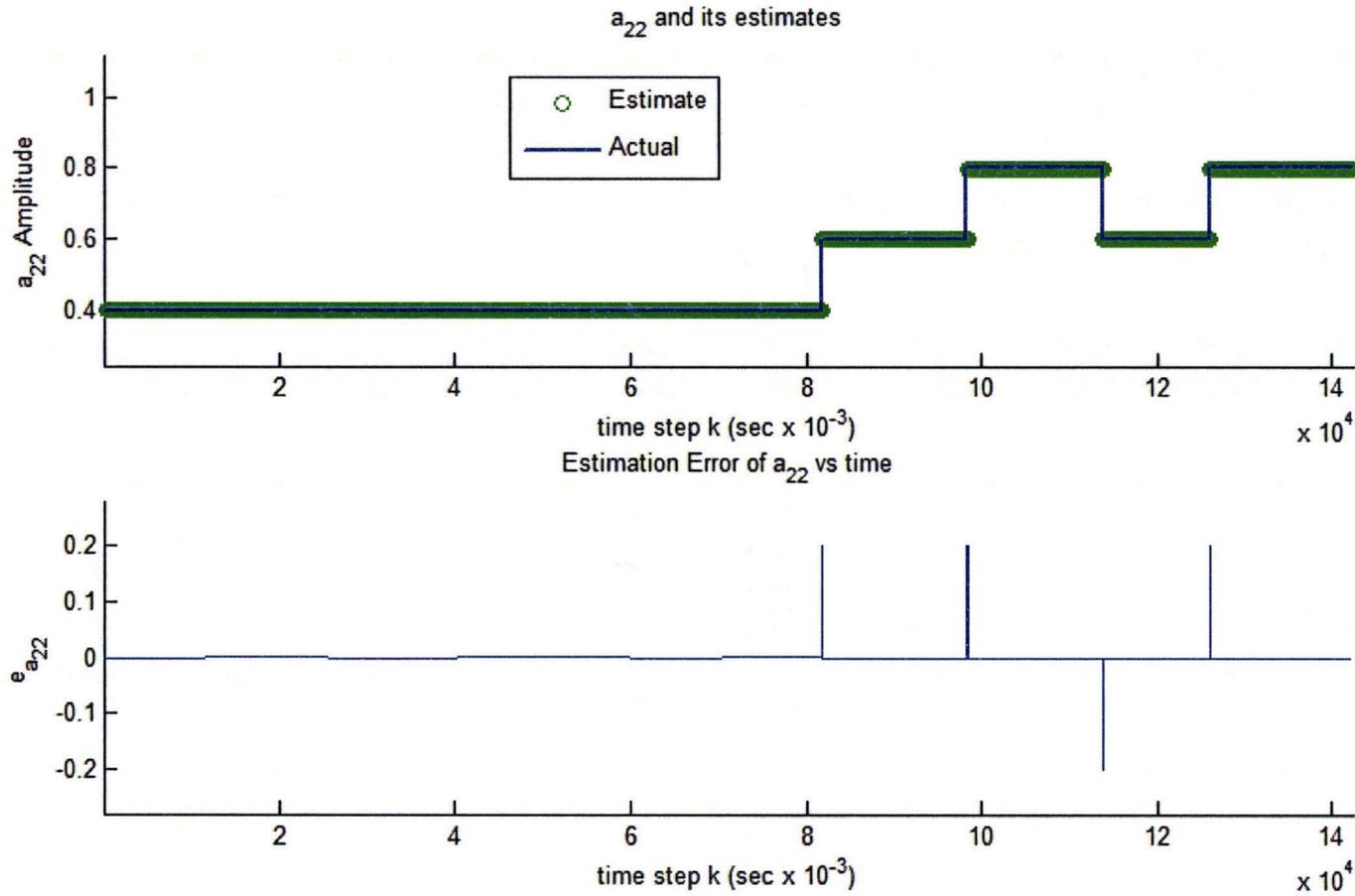


Fig 5.22: (a) a_{22} 's actual and estimated values, (b) estimation error of a_{22} for the General Toeplitz/Observability SVSF's tuning process.

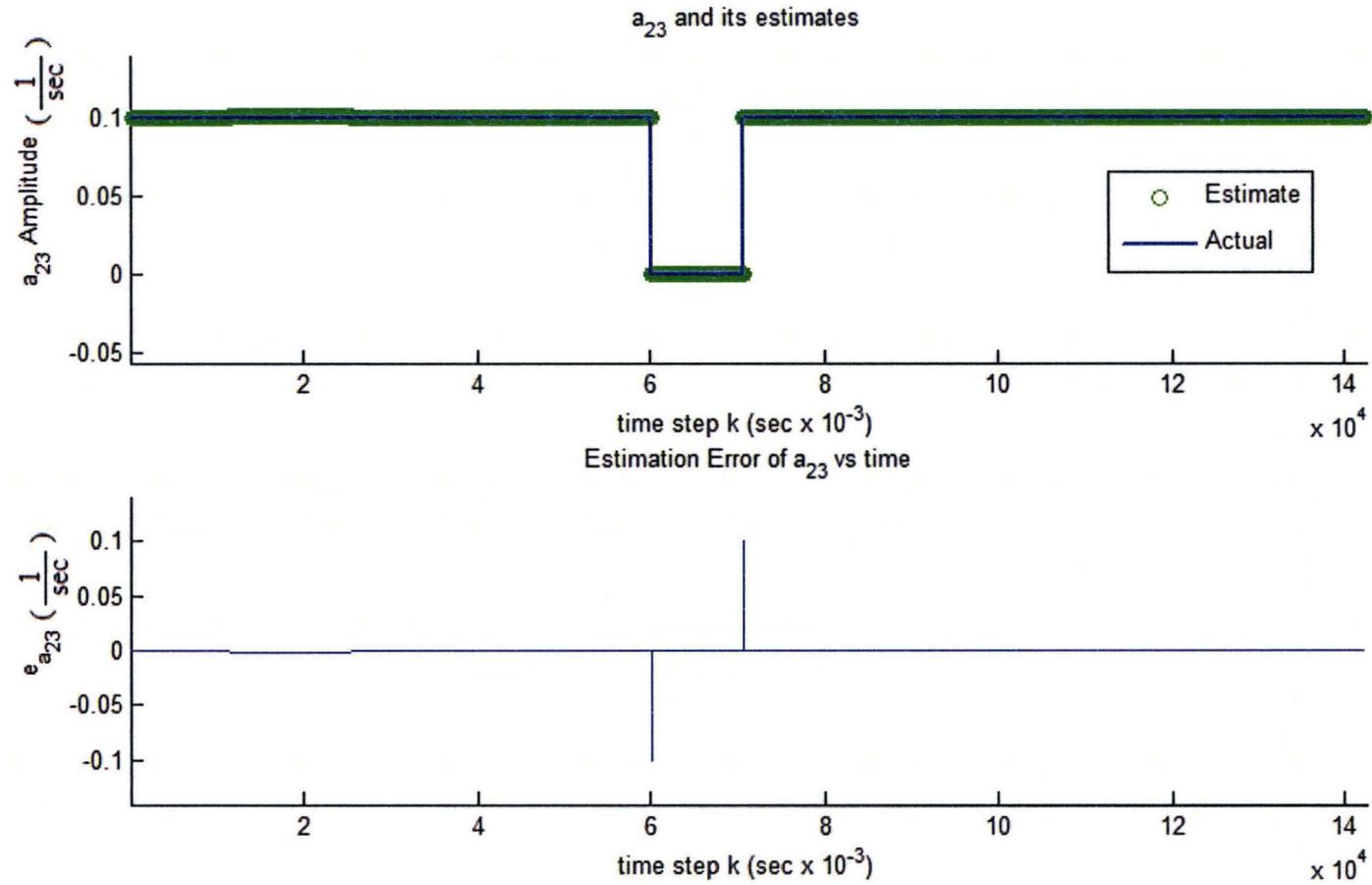


Fig 5.23: (a) a_{23} 's actual and estimated values, (b) estimation error of a_{23} for the General Toeplitz/Observability SVSF's tuning process.

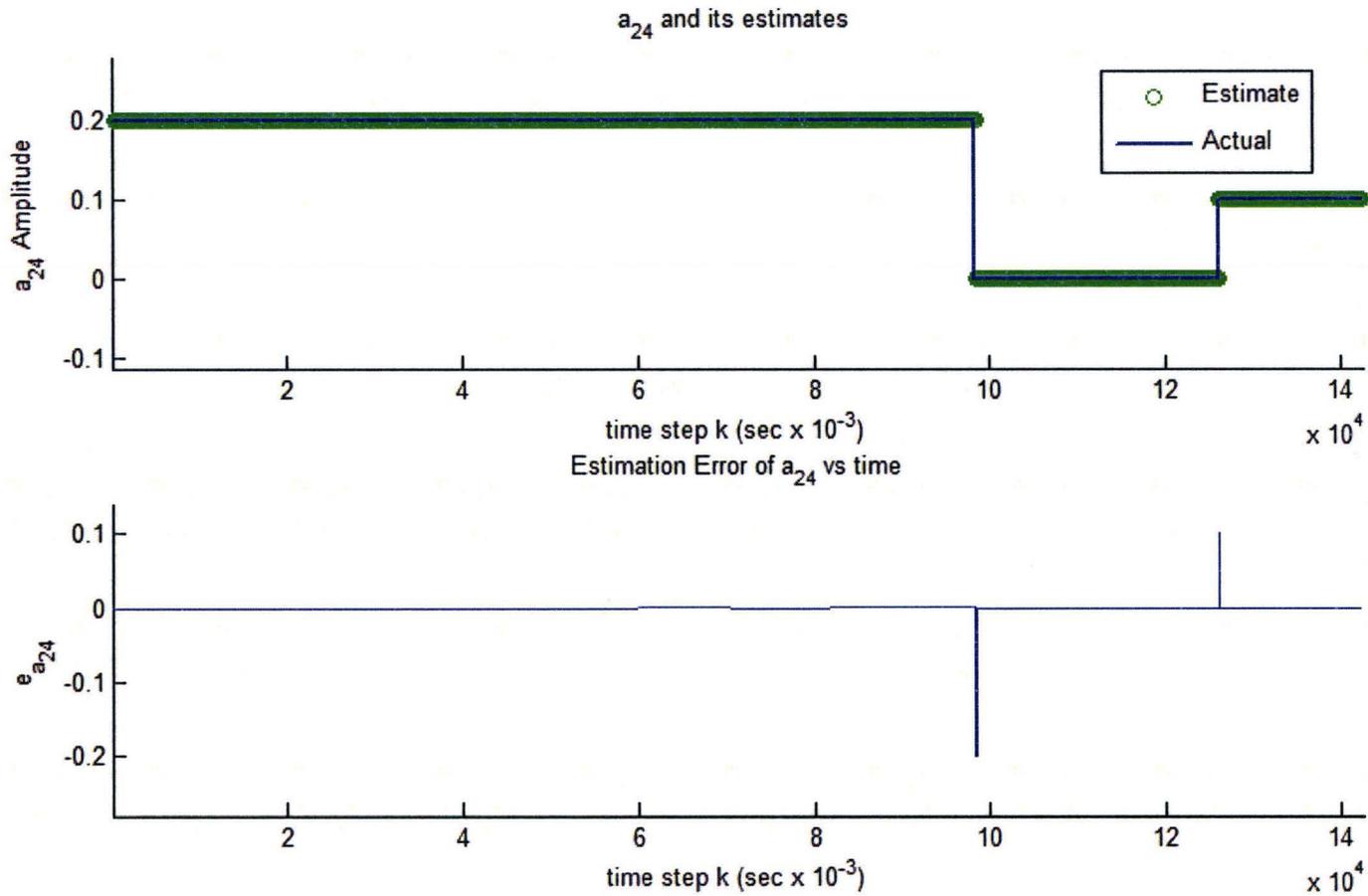


Fig 5.24: (a) a_{24} 's actual and estimated values, (b) estimation error of a_{24} for the General Toeplitz/Observability SVSF's tuning process.

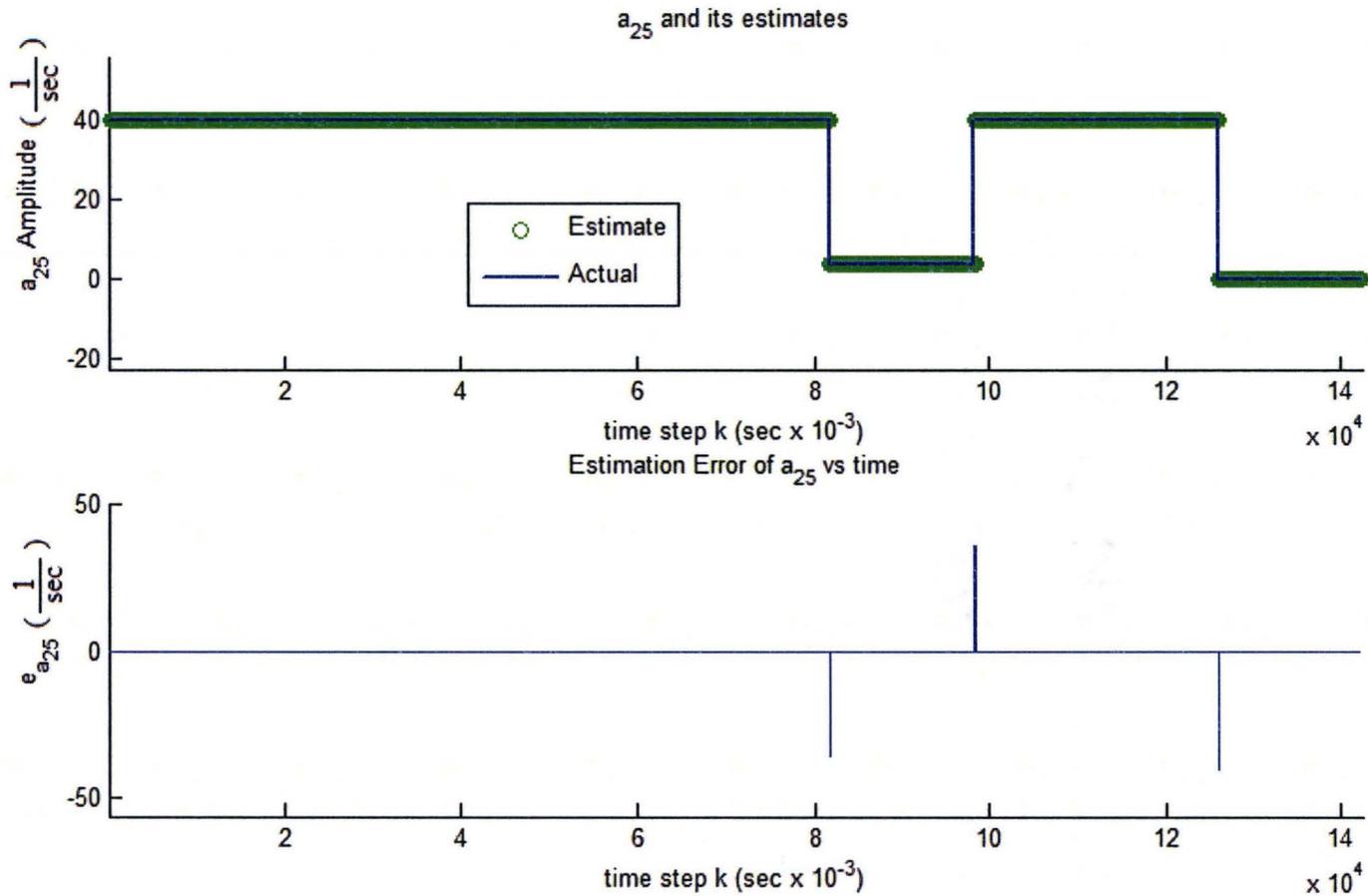


Fig 5.25: (a) a_{25} 's actual and estimated values, (b) estimation error of a_{25} for the General Toeplitz/Observability SVSF's tuning process.

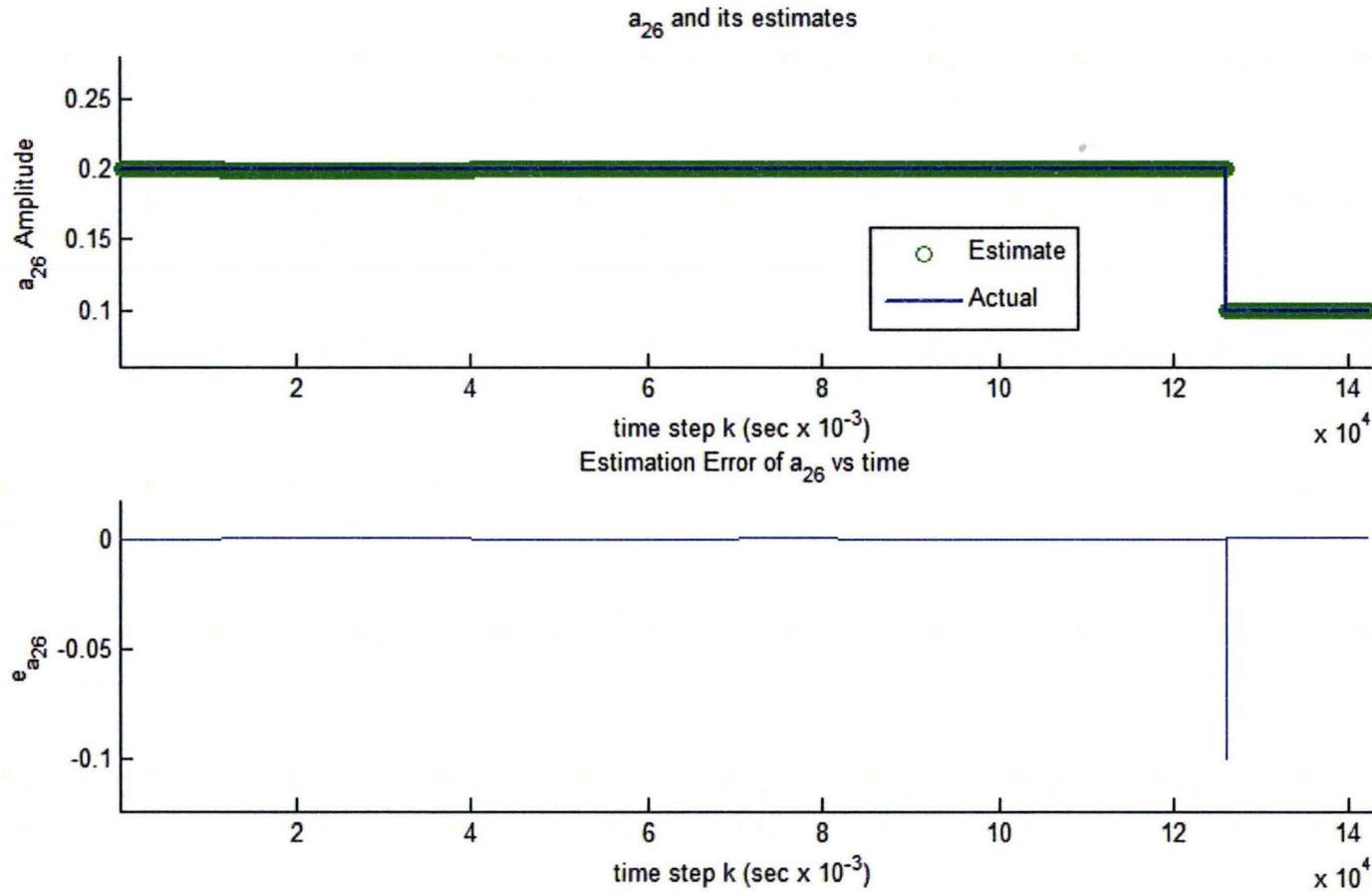


Fig 5.26: (a) a_{26} 's actual and estimated values, (b) estimation error of a_{26} for the General Toeplitz/Observability SVSF's tuning process.

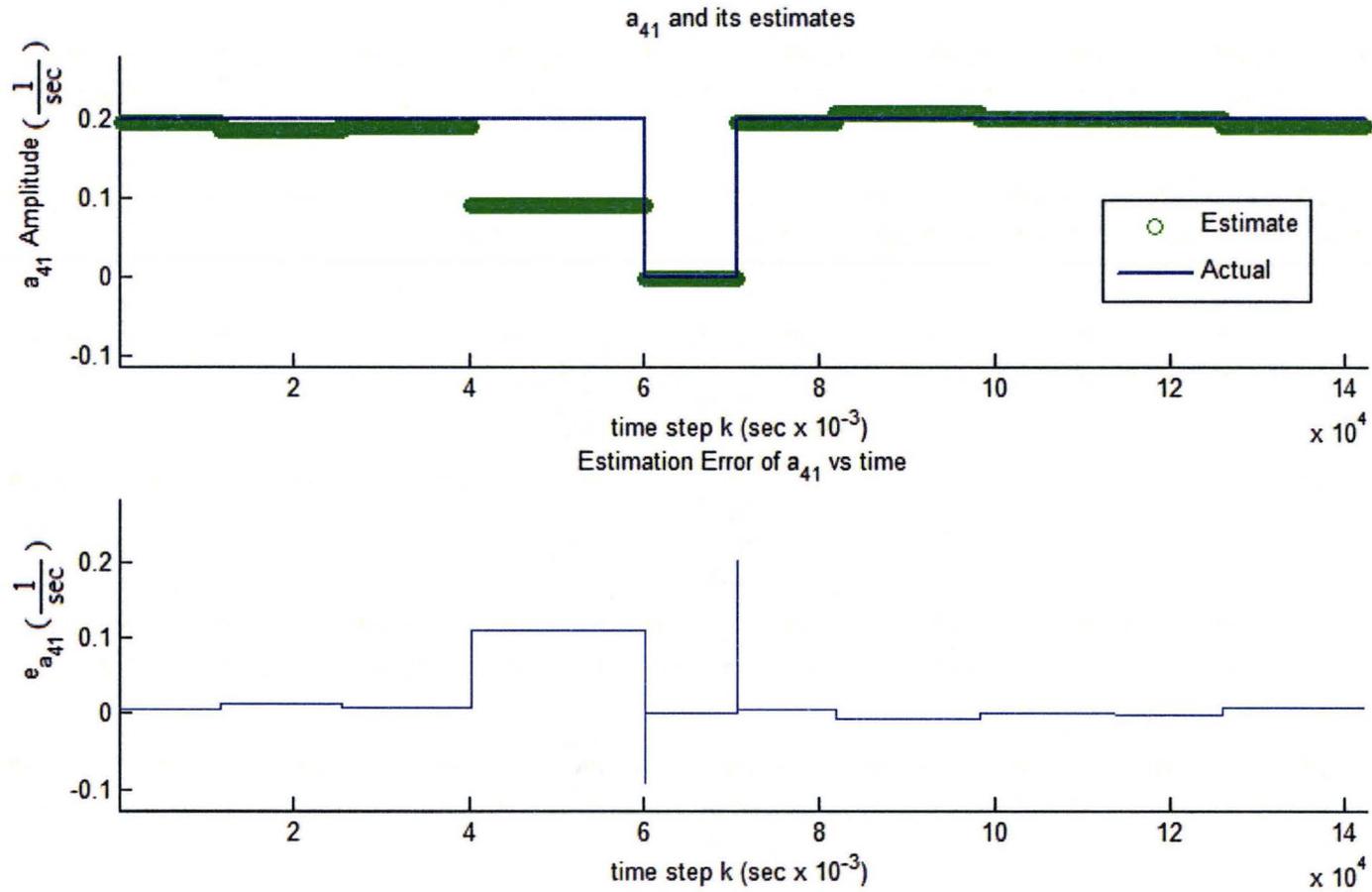


Fig 5.27: (a) a_{41} 's actual and estimated values, (b) estimation error of a_{41} for the General Toeplitz/Observability SVSF's tuning process.

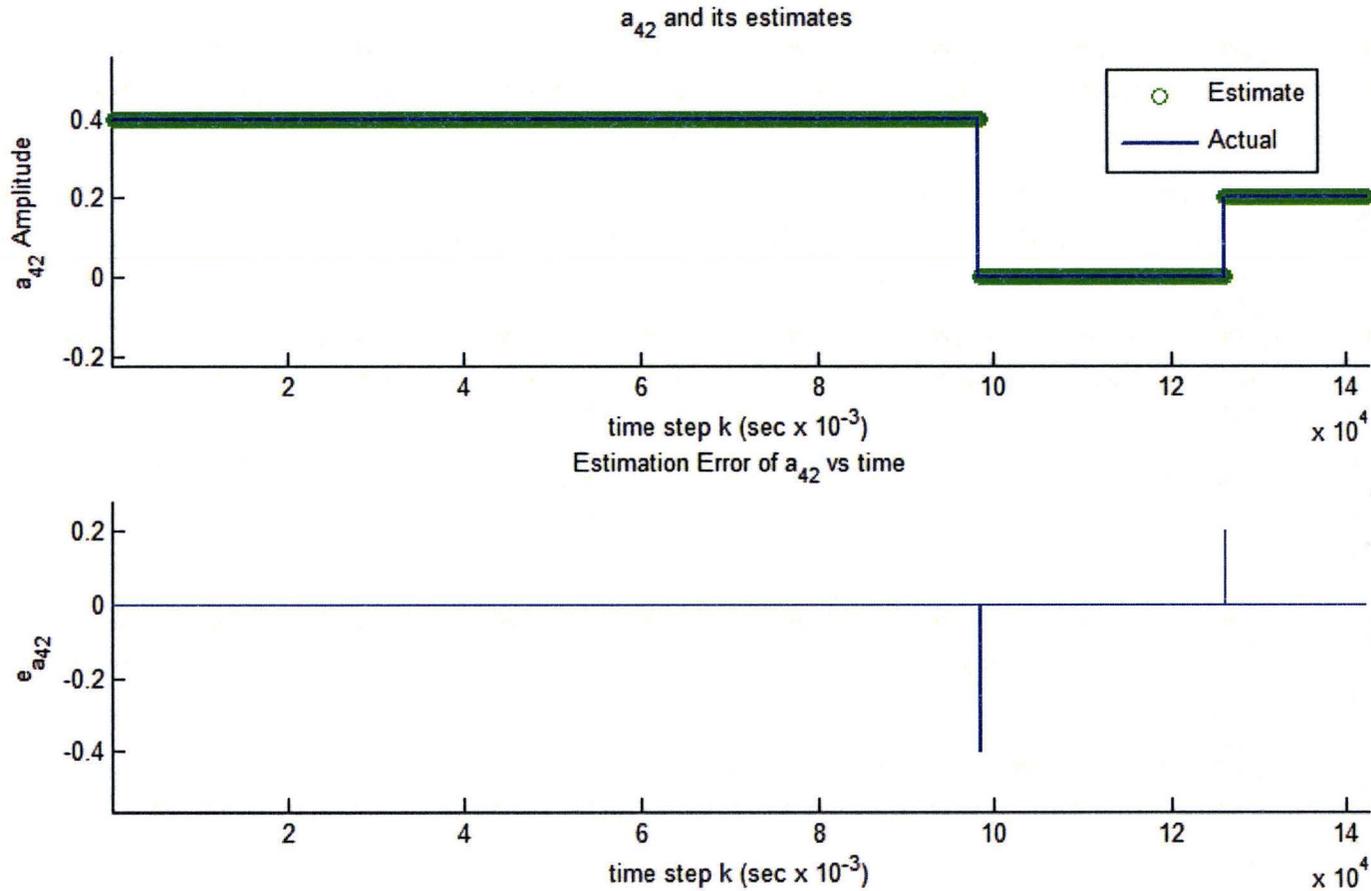


Fig 5.28: (a) a_{42} 's actual and estimated values, (b) estimation error of a_{42} for the General Toeplitz/Observability SVSF's tuning process.

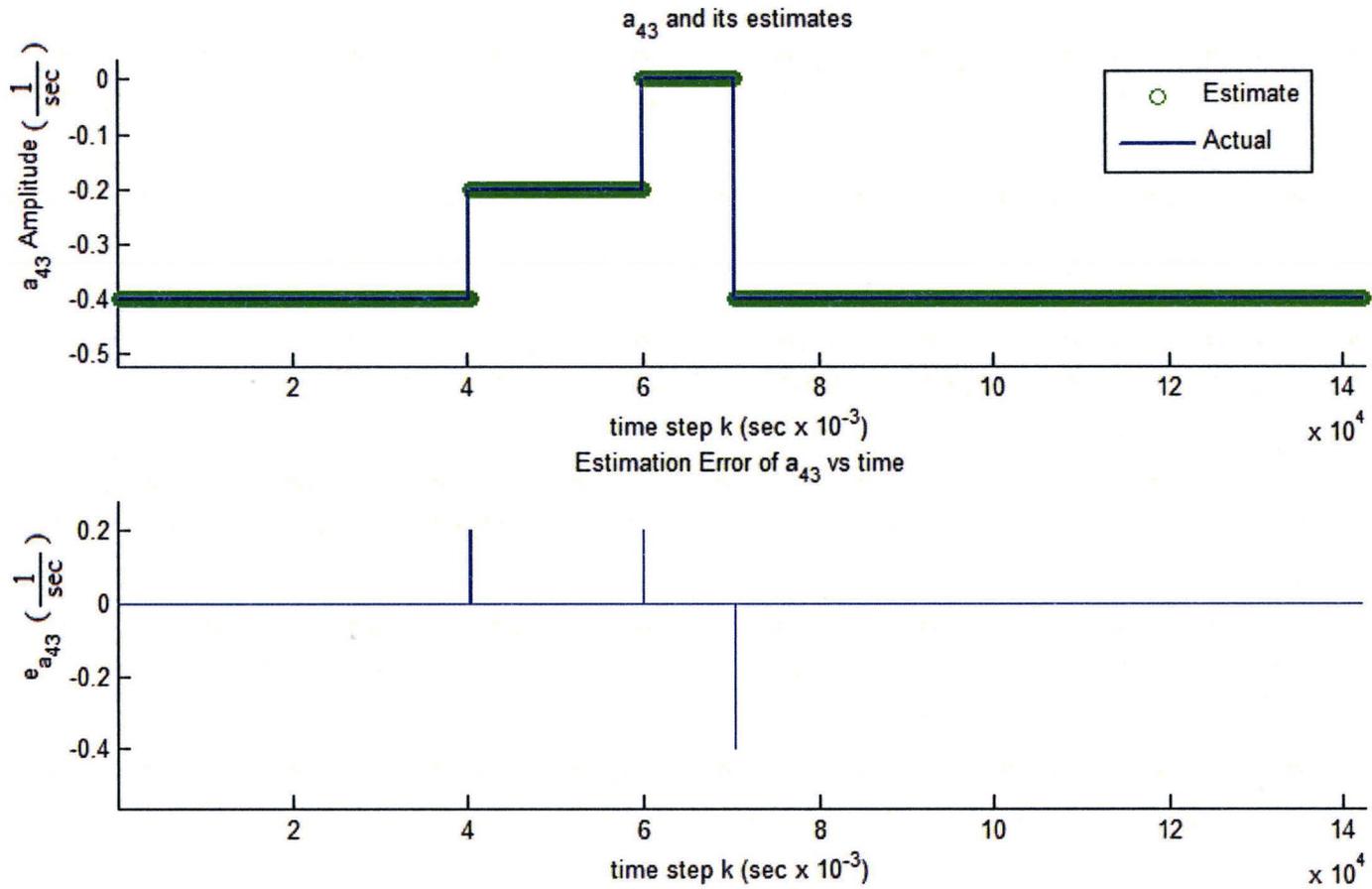


Fig 5.29: (a) a_{43} 's actual and estimated values, (b) estimation error of a_{43} for the General Toeplitz/Observability SVSF's tuning process.

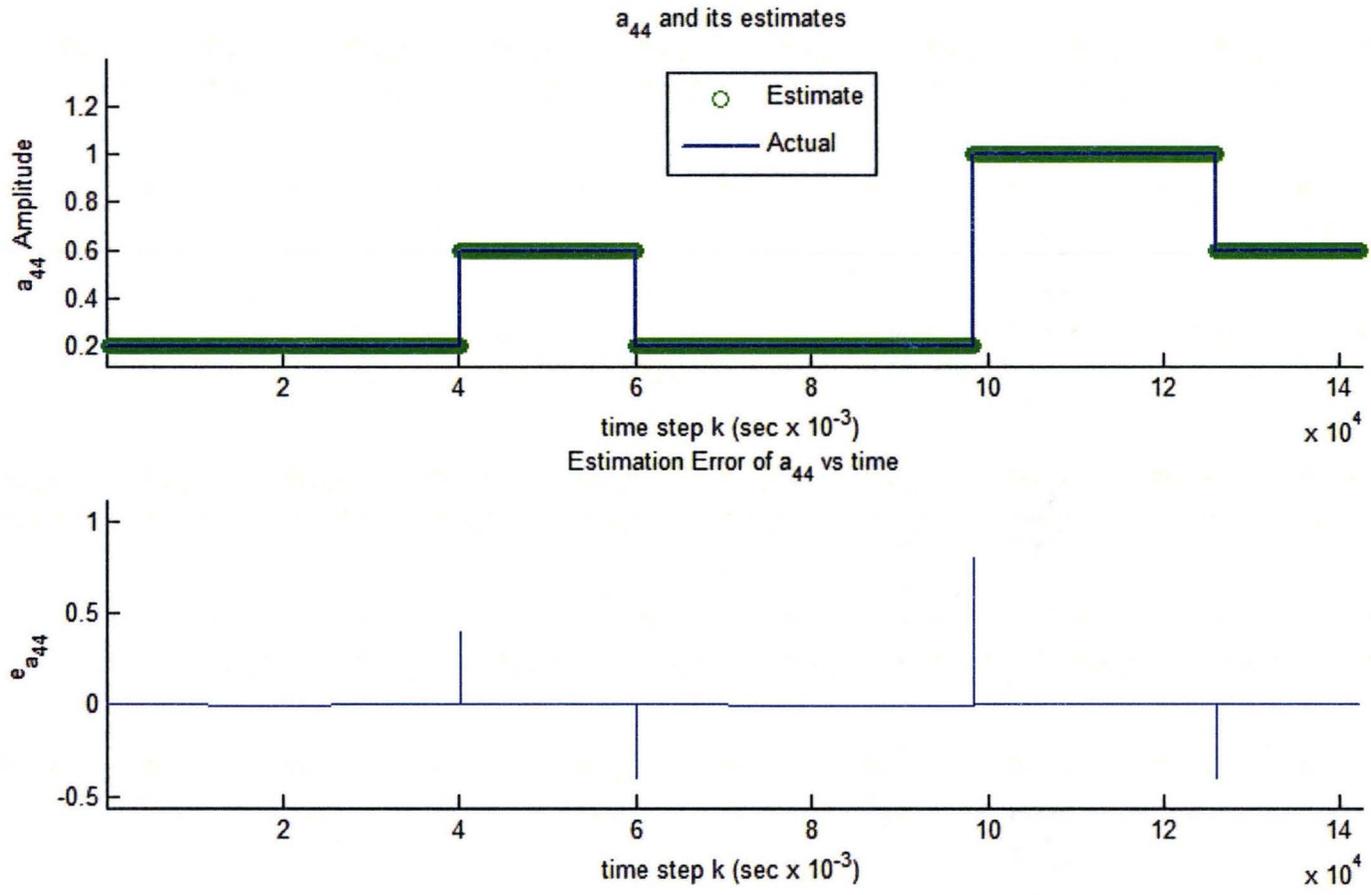


Fig 5.30: (a) a_{44} 's actual and estimated values, (b) estimation error of a_{44} for the General Toeplitz/Observability SVSF's tuning process.

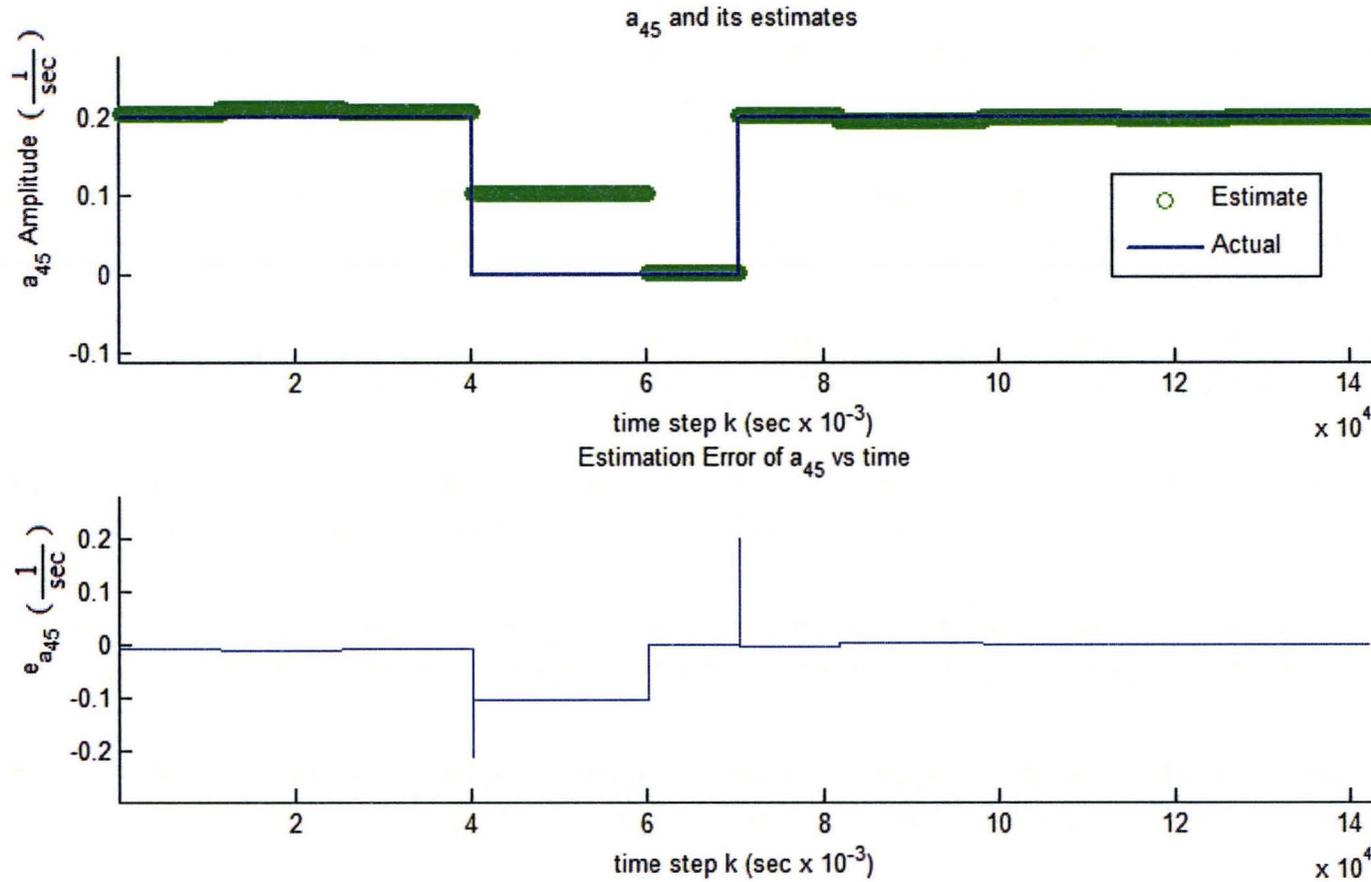


Fig 5.31: (a) a_{45} 's actual and estimated values, (b) estimation error of a_{45} for the General Toeplitz/Observability SVSF's tuning process.

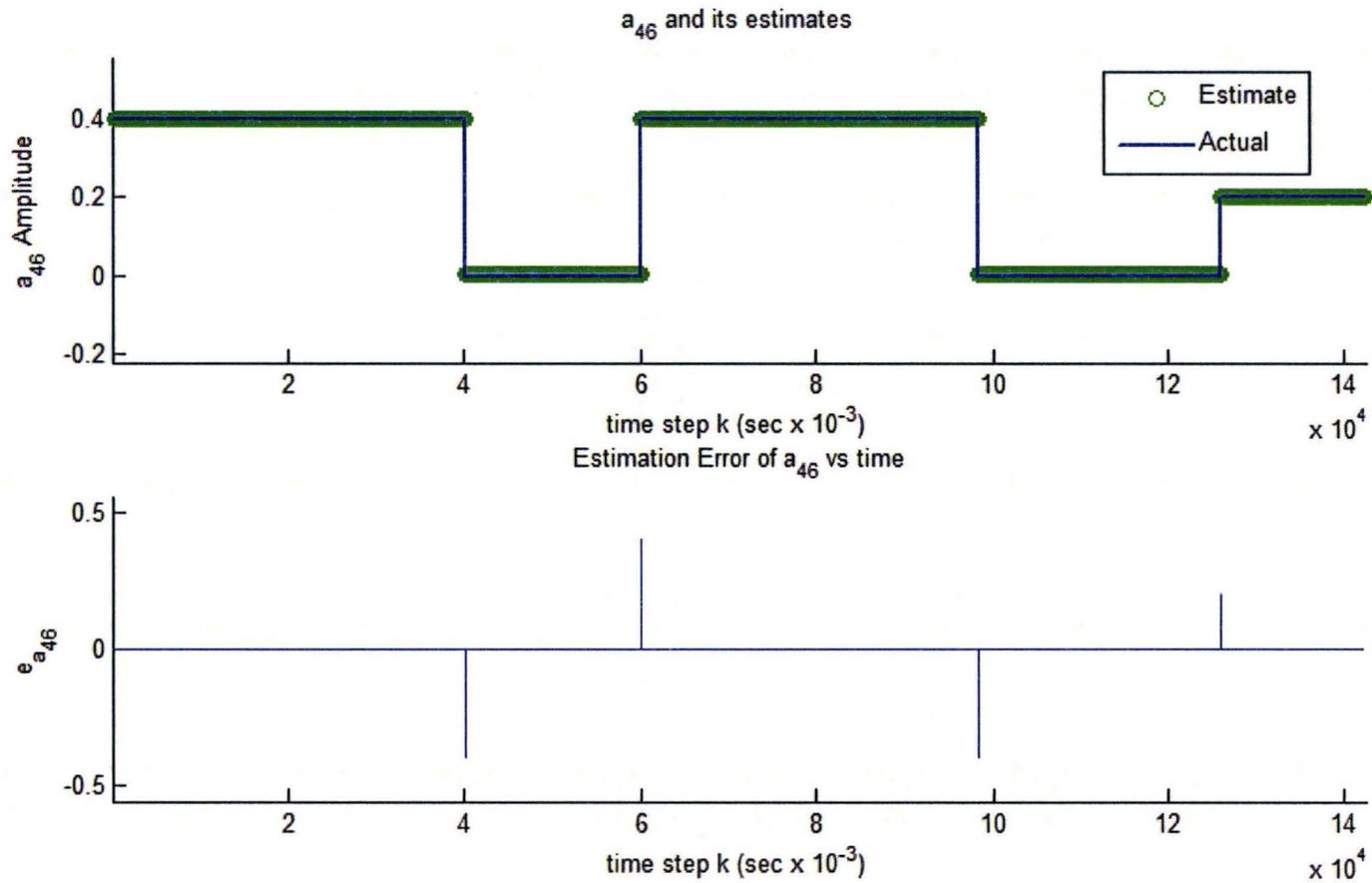


Fig 5.32: (a) a_{46} 's actual and estimated values, (b) estimation error of a_{46} for the General Toeplitz/Observability SVSF's tuning process.

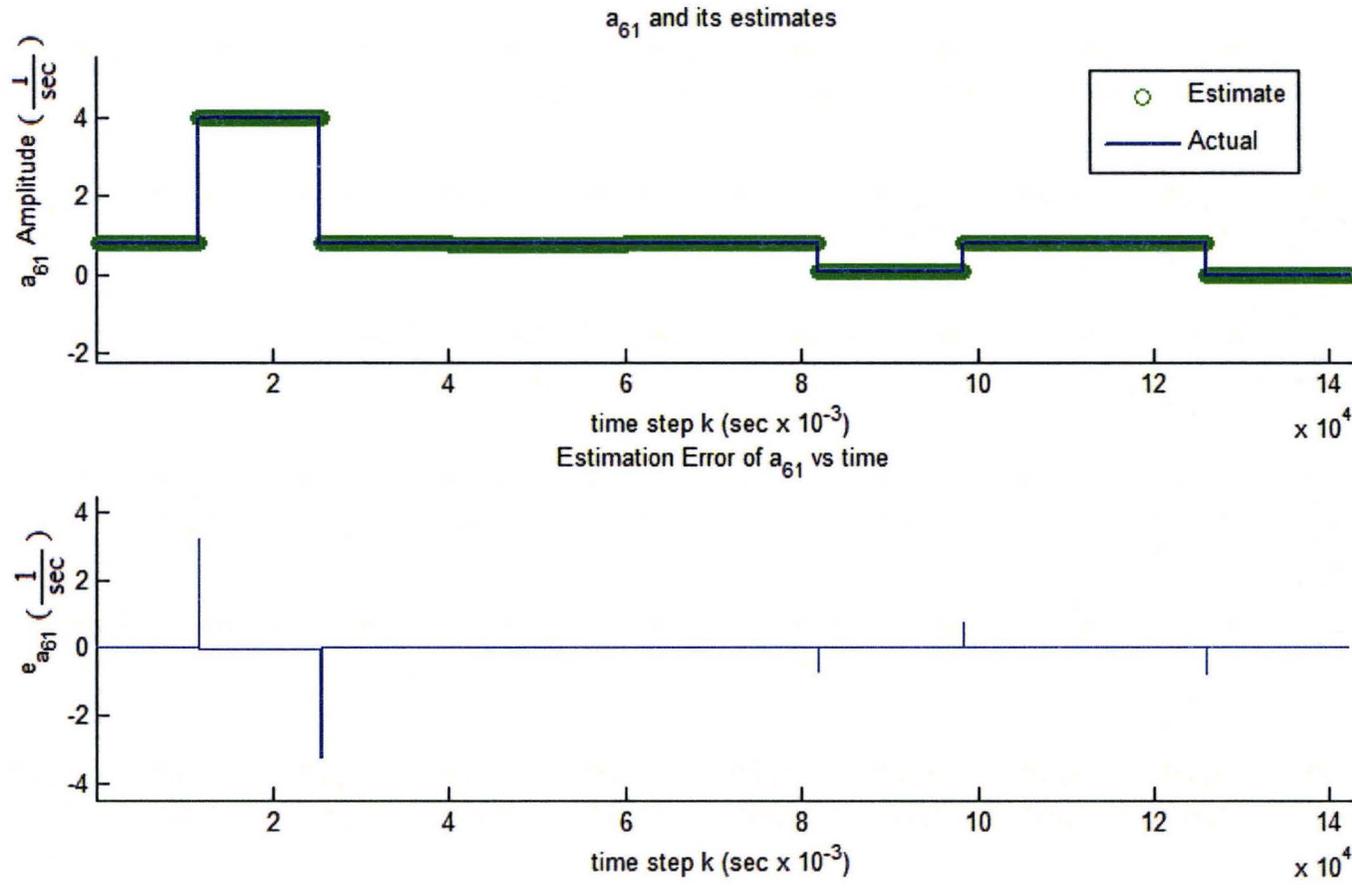


Fig 5.33: (a) a_{61} 's actual and estimated values, (b) estimation error of a_{61} for the General Toeplitz/Observability SVSF's tuning process.

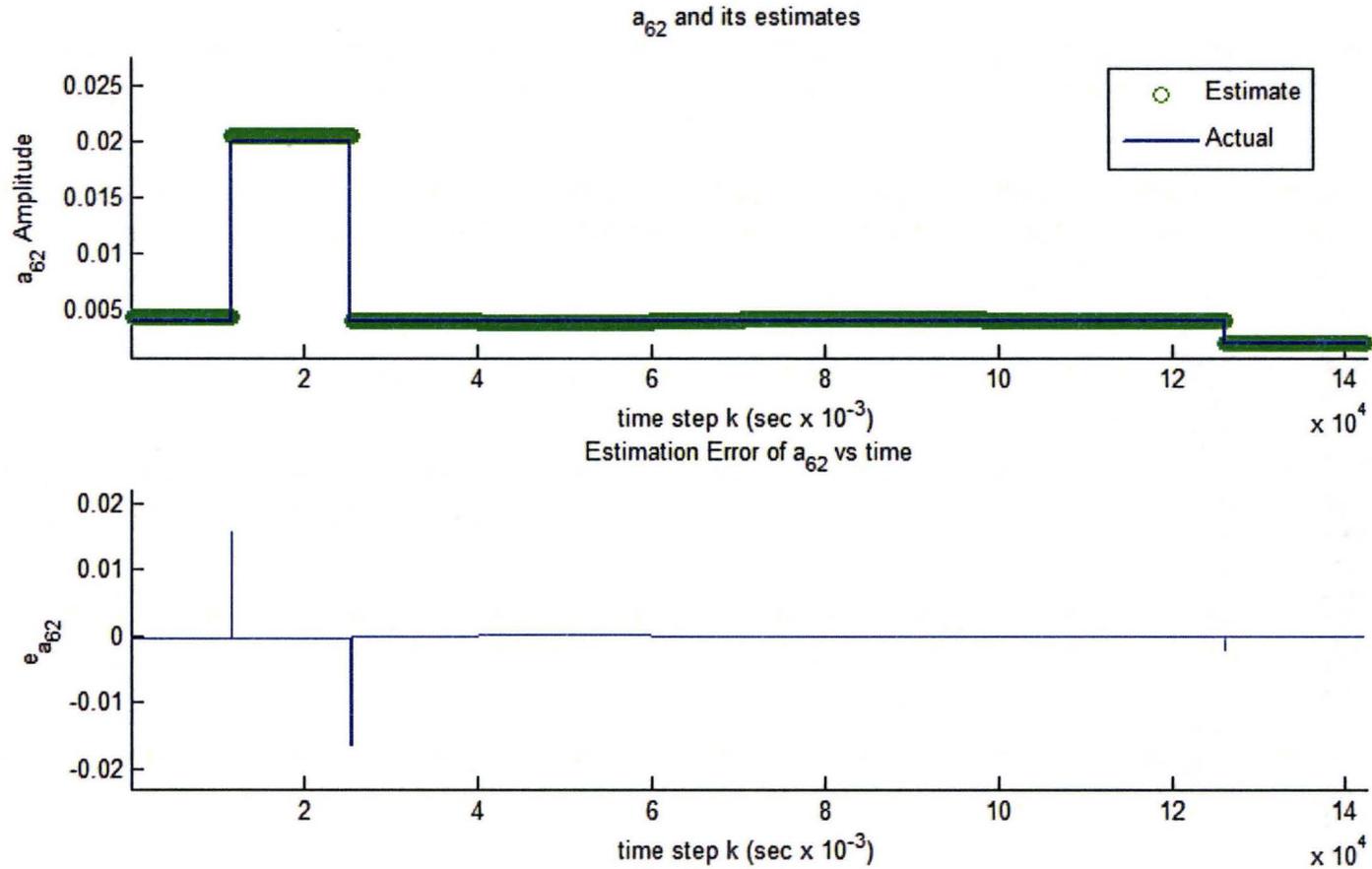


Fig 5.34: (a) a_{62} 's actual and estimated values, (b) estimation error of a_{62} for the General Toeplitz/Observability SVSF's tuning process.

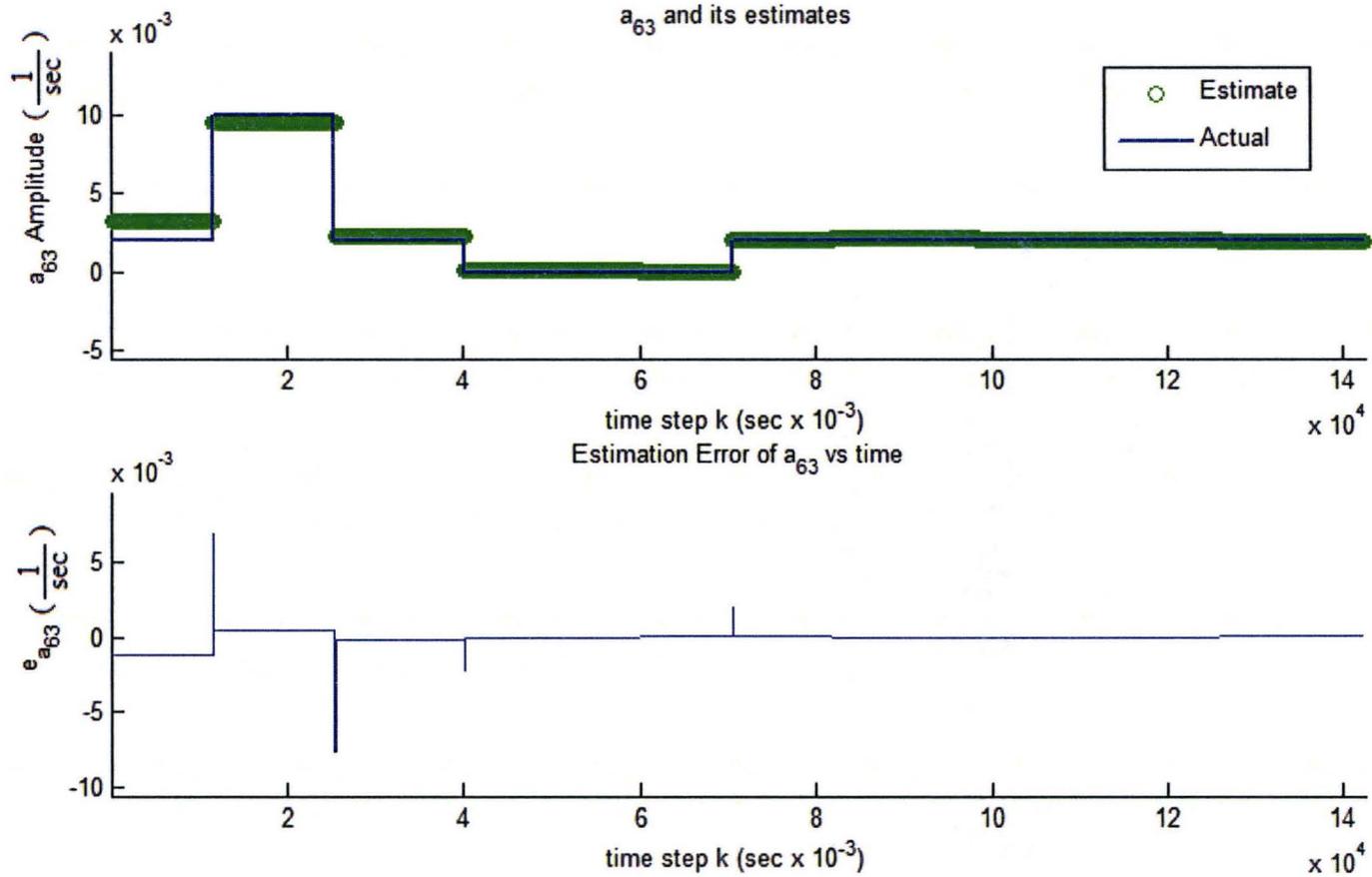


Fig 5.35: (a) a_{63} 's actual and estimated values, (b) estimation error of a_{63} for the General Toeplitz/Observability SVSF's tuning process.

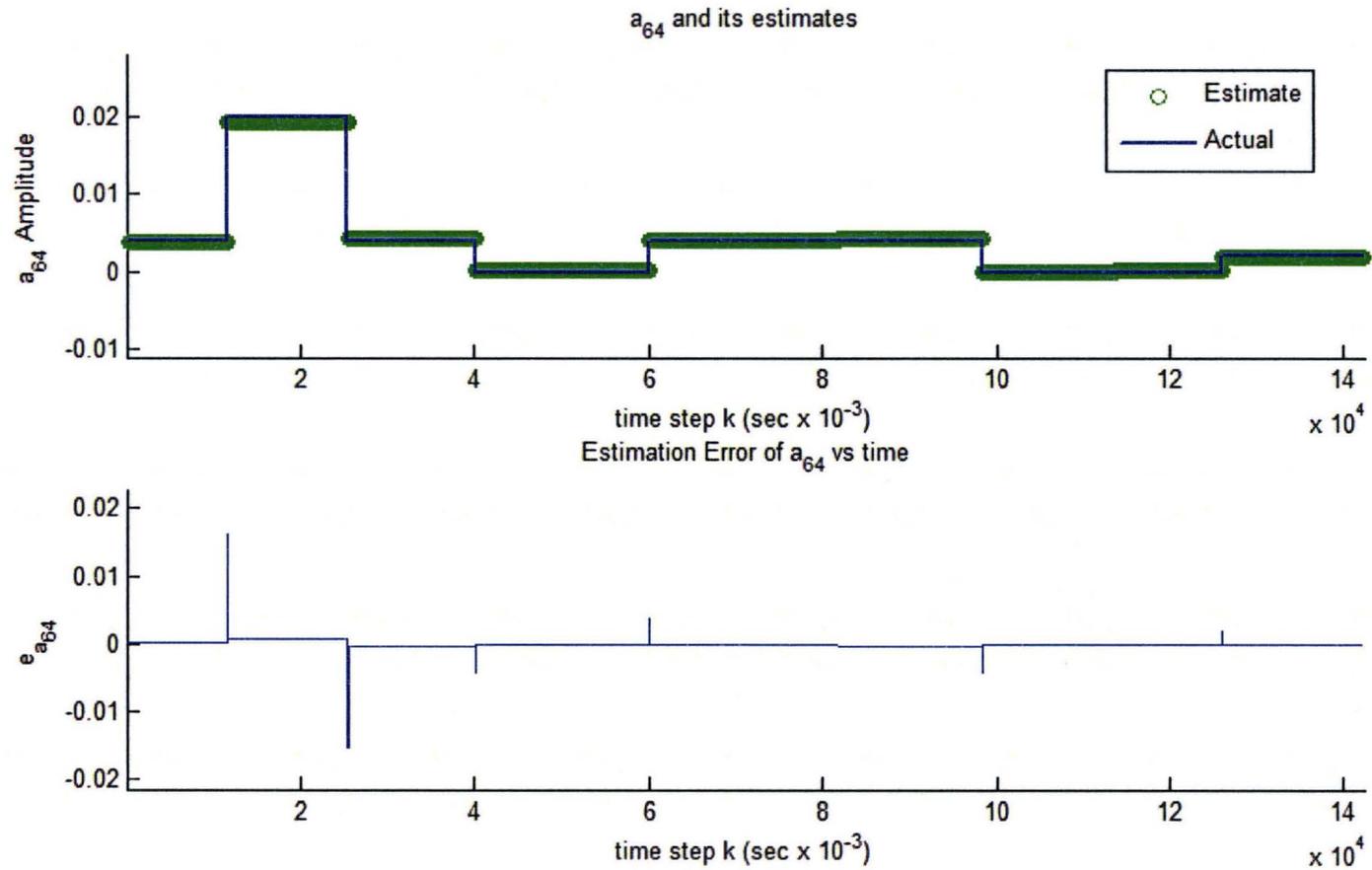


Fig 5.36: (a) a_{64} 's actual and estimated values, (b) estimation error of a_{64} for the General Toeplitz/Observability SVSF's tuning process.

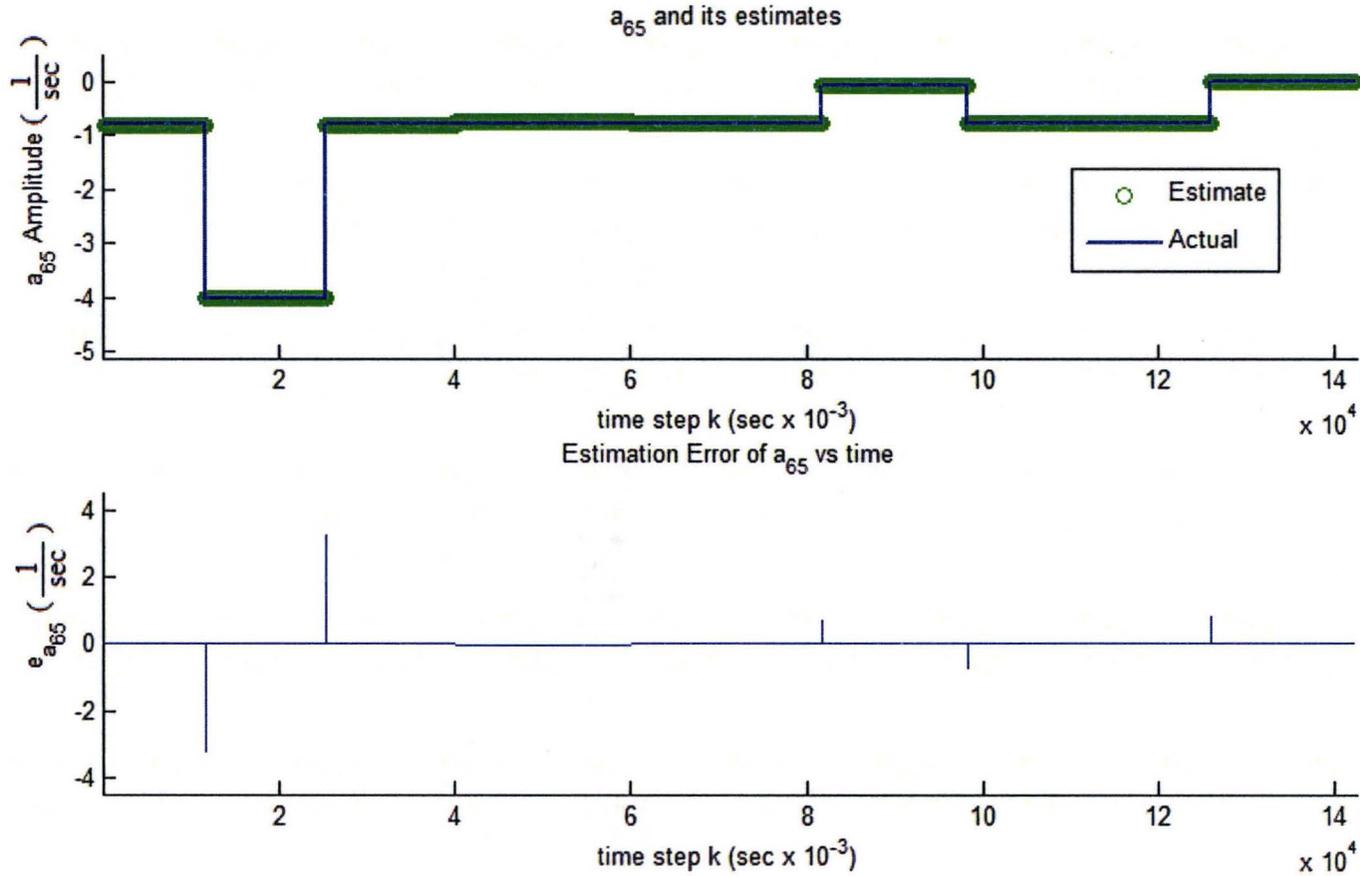


Fig 5.37: (a) a_{65} 's actual and estimated values, (b) estimation error of a_{65} for the General Toeplitz/Observability SVSF's tuning process.

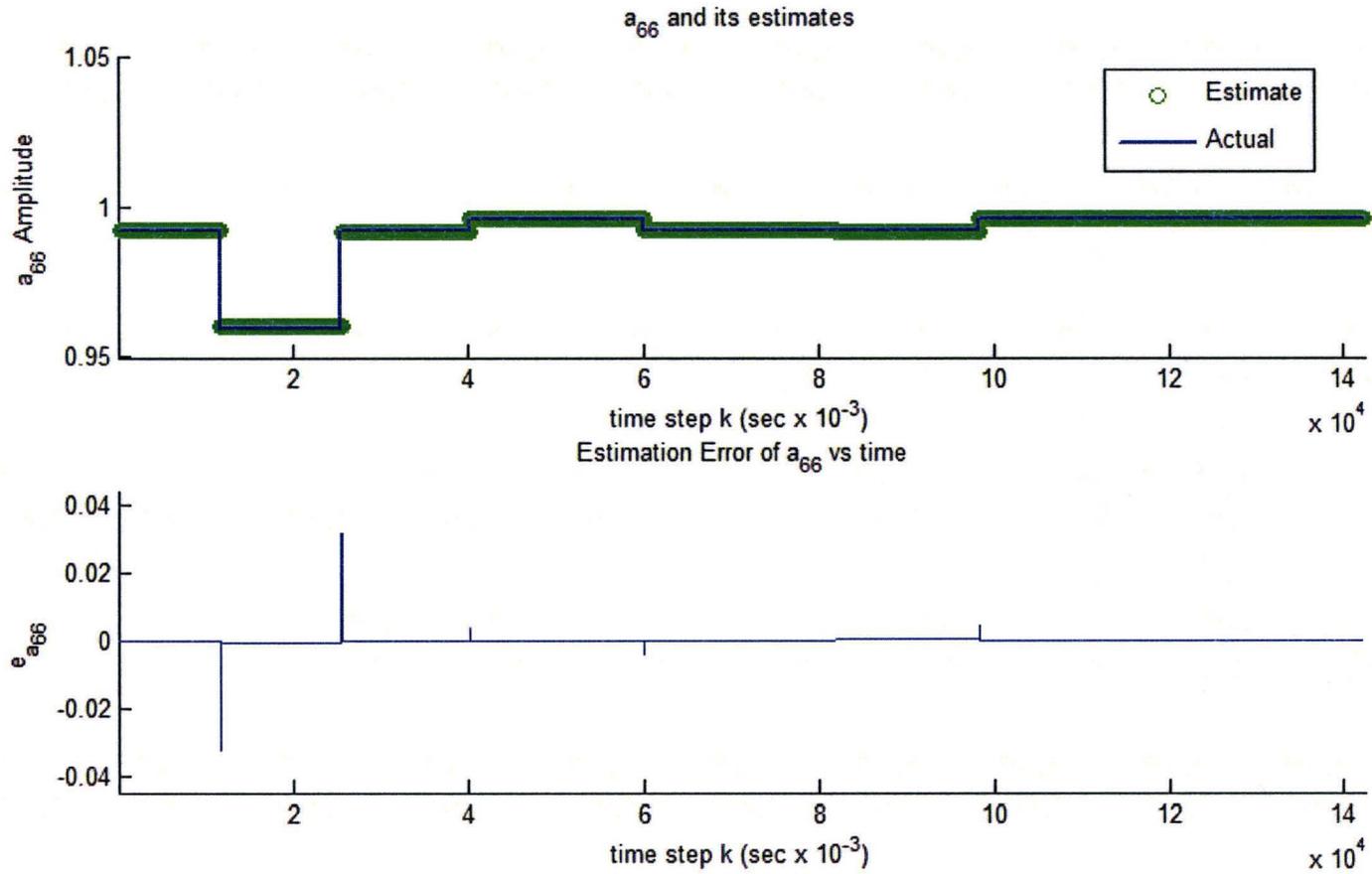


Fig 5.38: (a) a_{66} 's actual and estimated values, (b) estimation error of a_{66} for the General Toeplitz/Observability SVSF's tuning process.

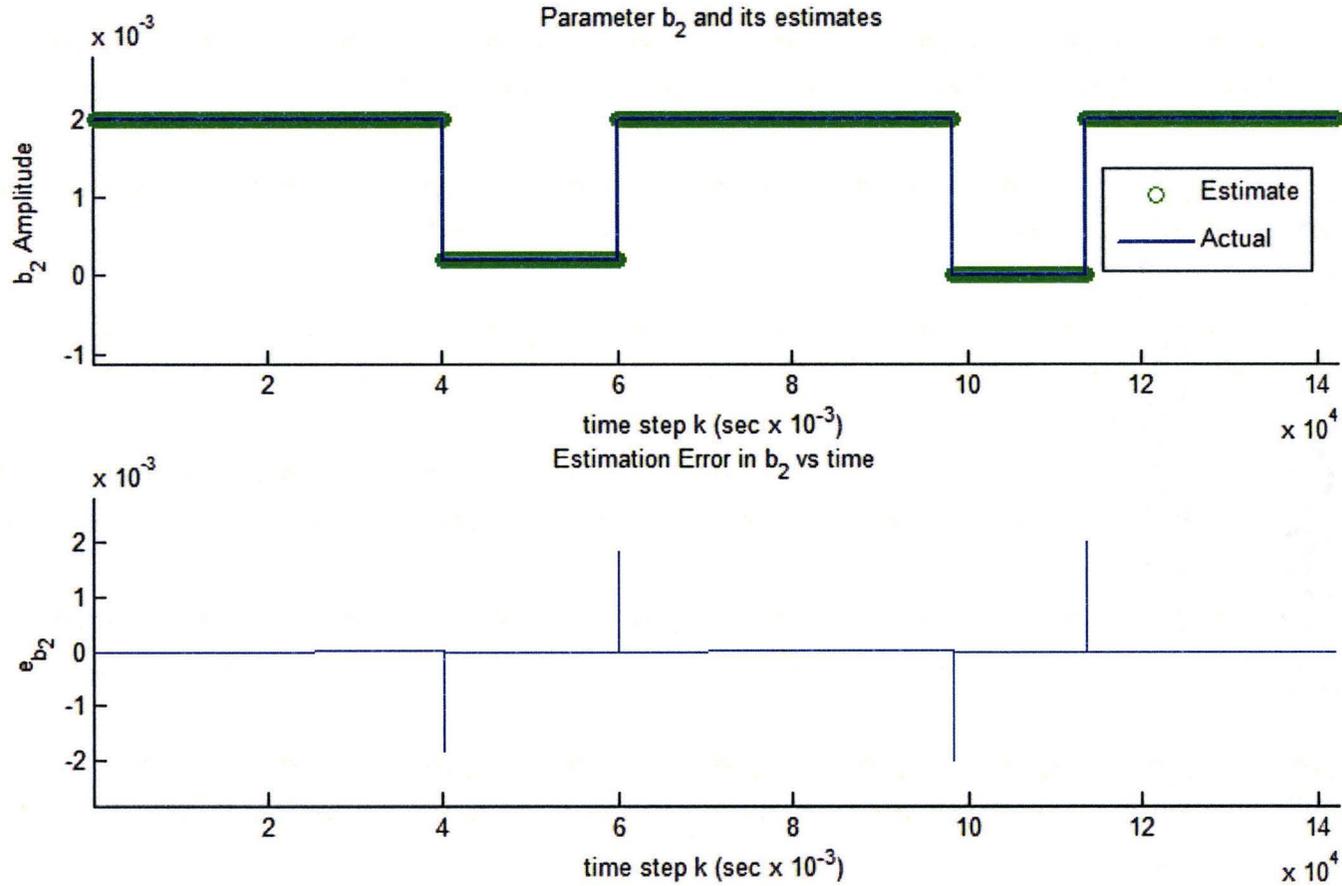


Fig 5.39: (a) b_2 's actual and estimated values, (b) estimation error of b_2 for the General Toeplitz/Observability SVSF's tuning process.

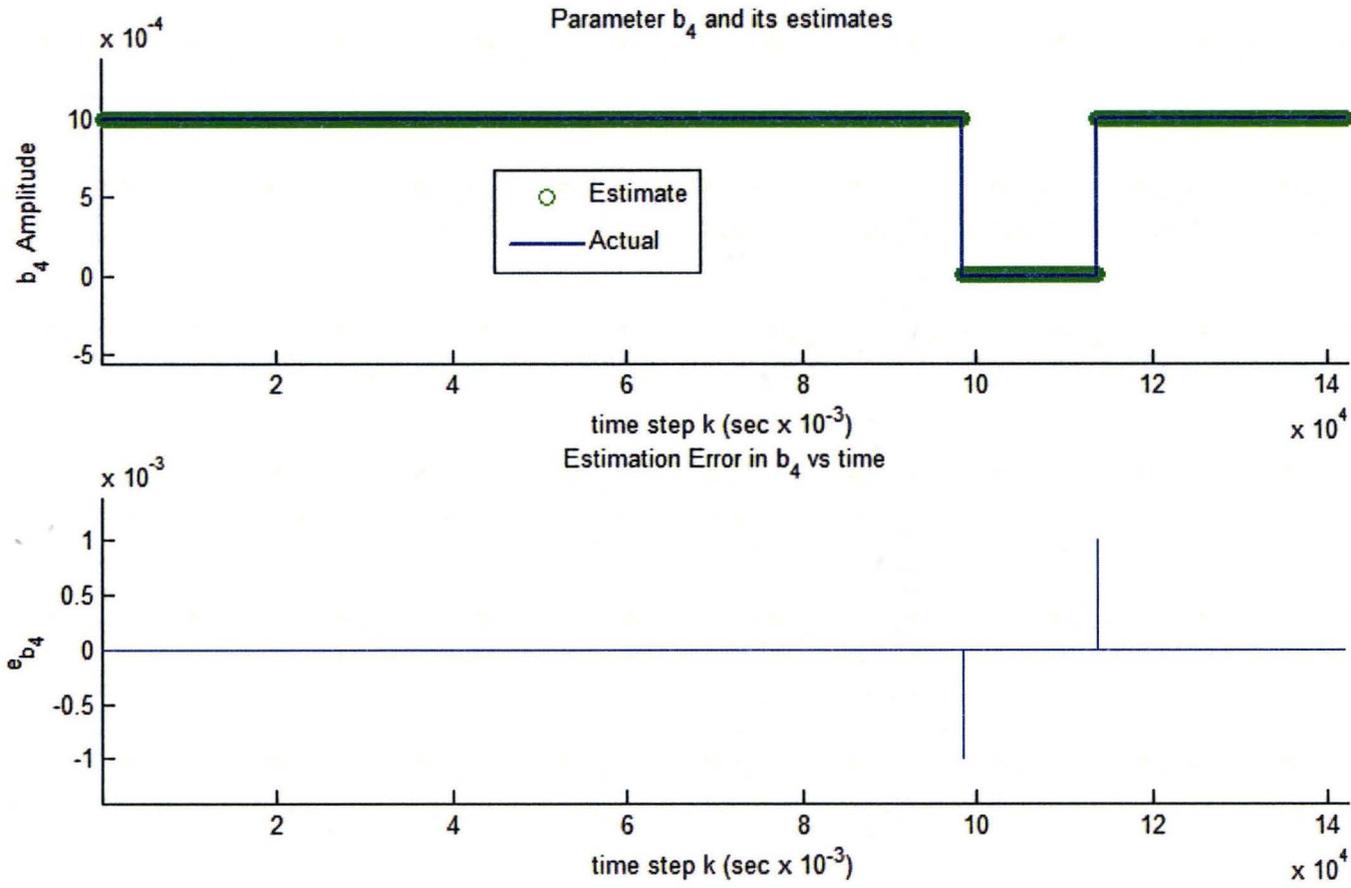


Fig 5.40: (a) b_4 's actual and estimated values, (b) estimation error of b_4 for the General Toeplitz/Observability SVSF's tuning process.

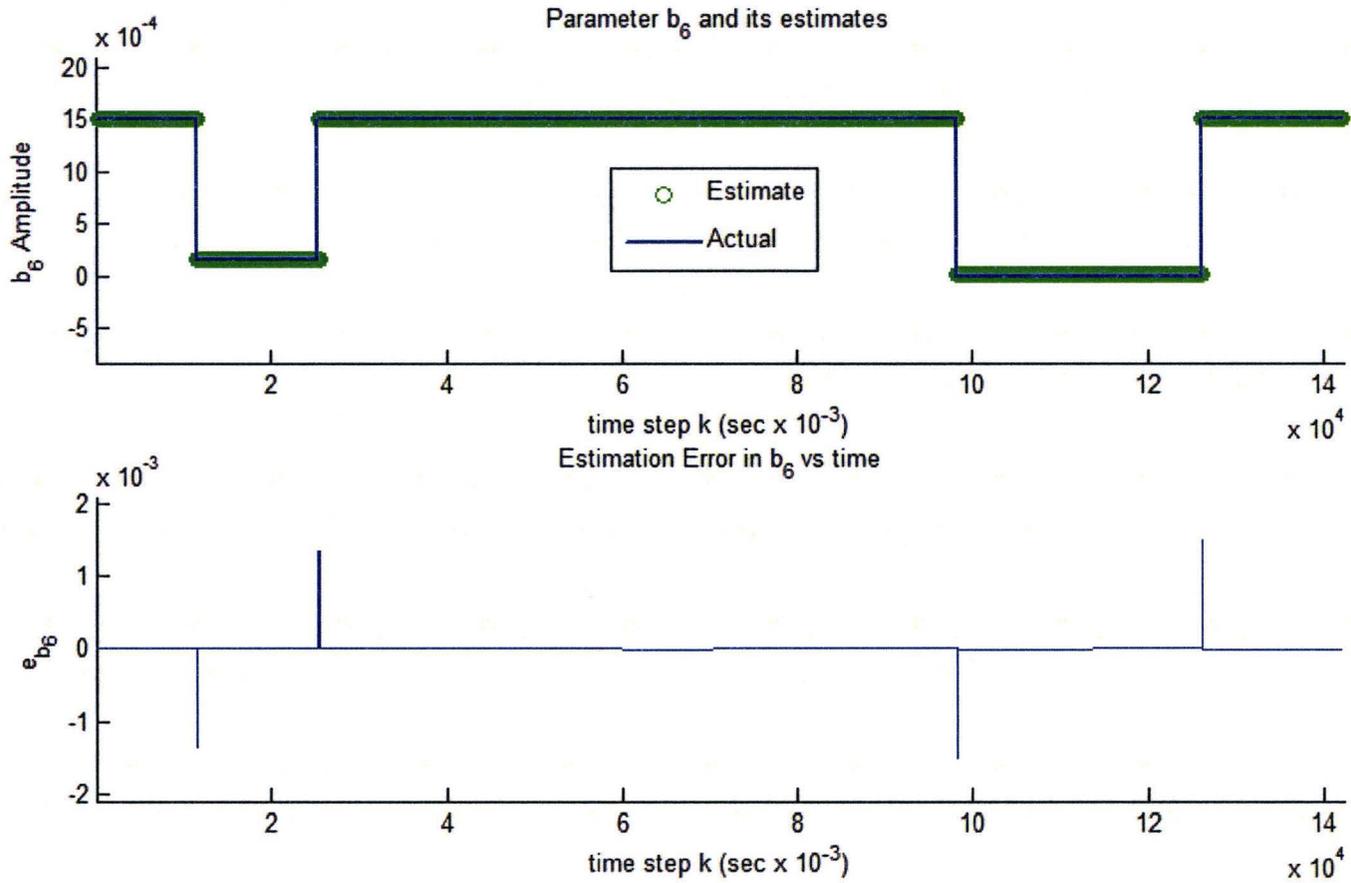


Fig 5.41: (a) b_6 's actual and estimated values, (b) estimation error of b_6 for the General Toeplitz/Observability SVSF's tuning process.

Chapter Six:

Iterative Bi-Section/Shooting Method for Parameter Estimation

6.1. Introduction

In this chapter, a novel parameter estimation technique, referred to as the Iterative Bi-Section/Shooting method (IBSS), is proposed as a searching technique to obtain the best parameters that fit given a known system structure. The chapter presents three methods as follows:

- The iterative Bi-Section/Shooting method – An exhaustive searching mechanism to estimate the system's parameters.
- The iterative Bi-Section/Shooting method with the KF – To estimate the system's states and parameters.
- The iterative Bi-Section/Shooting method with the SVSF – To estimate the system's states and parameters presenting an improved performance given the other two methods.

Here, the IBSS method is combined with the SVSF to adapt and refine the filter's model and hence reduce the modeling errors. This combined method is referred to as the IBSS/SVSF strategy. For comparison purposes, the IBSS is combined with the KF and the resultant algorithm is compared with the IBSS/SVSF. Before presenting these new methodologies, the basic concepts are explained.

6.2. The Bi-Section Method

The Bi-Section method, also known as the binary-search method, is a well-known **numerical root-finder** for the equation $f(x) = 0$ and it is based on the following theorem:

Theorem 6.1 zeros of continuous functions:

If the function $f(x)$ is continuous over the interval (a_i, a_e) , and it satisfies $f(a_i)f(a_e) < 0$ then there is **at least** one point a in the interval with the property of $f(a) = 0$ as shown in

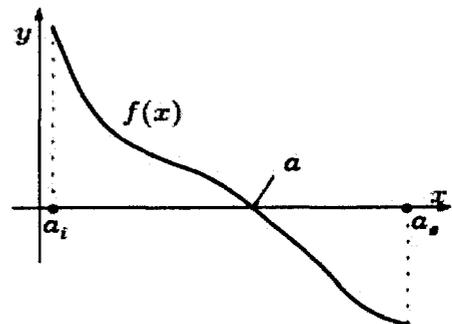


Fig 6.1: The Bi-Section's Principle, (Kaw & Kalu, 2010).

Fig 6.1, [(Kaw & Kalu, 2010) and (Quarteroni, Sacco, & Saleri, 2000)]. The Bi-Section method finds the point, a , by using the following algorithm, (Kaw & Kalu, 2010):

The algorithm starts by defining an interval that contains the root (a) of the equation $f(x) = 0$, as follows:

$$a \in (a_i, a_e) \quad 6.1$$

Where a_i and a_e are the interval boundaries and they are chosen to satisfy the following condition:

$$f(a_i)f(a_e) < 0 \quad 6.2$$

Then the middle of the interval, a_m , is taken and the corresponding function's sign is examined. The point a_m divides the interval into two subintervals, and based on its function sign, one of these sub-intervals is chosen to be a new interval for the next iteration as follows:

$$\text{If } f(a_i)f(a_m) < 0 \text{ then the new interval is defined as } (a_i, a_m) \quad 6.3$$

Else the new interval is defined as (a_m, a_e) .

The interval is then divided in half iteratively until the width of the interval becomes smaller than a threshold and the root is considered to be the half of the final interval. This algorithm is summarized in Figs (6.2) and (6.3).

If multiple zeros exist inside the interval, then only one of the zeros will be obtained depending on the interval size and its boundary values as shown in Fig 6.4. Therefore, this method is stable as it always converges

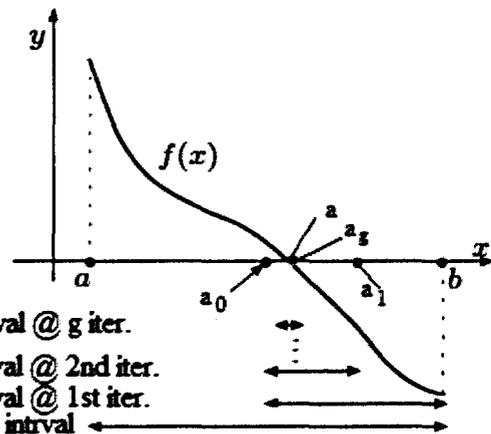


Fig 6.2: The Bi-Section Steps, (Kaw & Kalu, 2010).

to one of the zeros. Moreover, the level of accuracy is controllable which has a maximum absolute value equal to half of the last interval (threshold). However, its disadvantages are its slow rate of convergence and its sensitivity to the size of the interval, (Kaw & Kalu, 2010). Due to its stability and simplicity, this method has been used in many applications; i.e. computing the H_∞ Norm of transfer functions in (Boyd, Balakrishnan, & Kabamba, 1989), and in system identification as in (Moore, 1984).

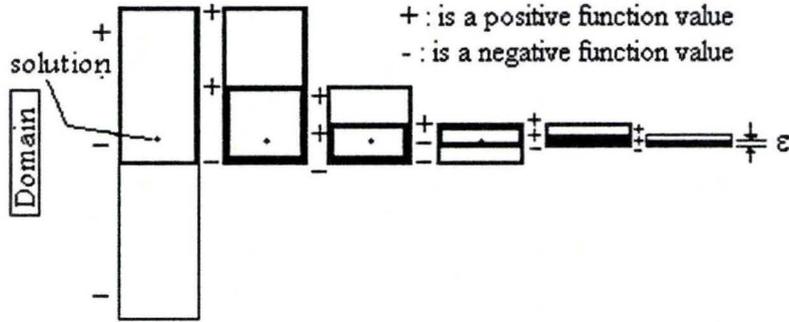


Fig 6.3: Reducing the solution interval using Bi-Section method.

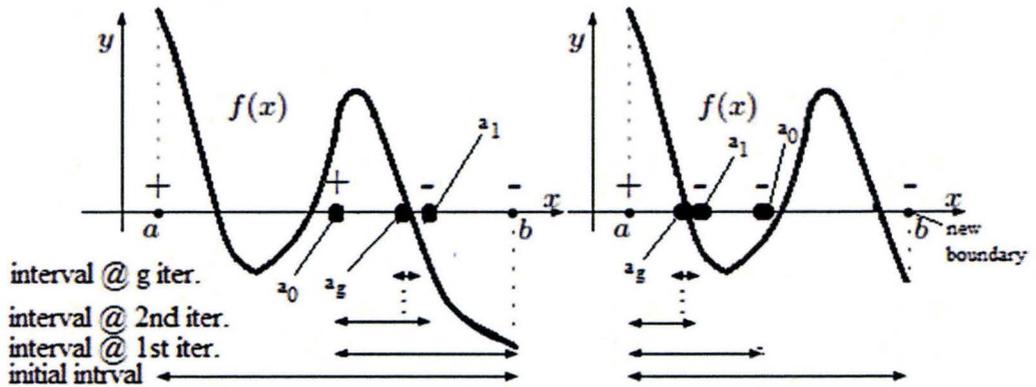


Fig 6.4: Multiple zeros and interval effects.

6.3. The Shooting Method

According to (Kaw & Kalu, 2010), the Shooting method is a numerical technique used to solve a differential equation with boundary conditions (at time t_f) defined as follows:

$$\sum_{i=0}^n x^{(i)}(t) = x^{(n)}(t) + x^{(n-1)}(t) + \dots + x(t) = x_p(t) \tag{6.4}$$

For the boundaries $[x(t_f) \dots x^{(n-1)}(t_f)]^T = \mathbf{x}_f$

and convert it to an initial condition problem (at time t_0) as follows:

$$\sum_{i=0}^n x^{(i)}(t) = x^{(n)}(t) + x^{(n-1)}(t) + \dots + x(t) = x_p(t) \tag{6.5}$$

For initial condition $[x(t_0) \dots x^{(n-1)}(t_0)]^T = \mathbf{x}_0$

Where x_p is the input.

The Shooting method has the same idea of hitting a target by a canon projectile. If a canon is used to hit a target, the muzzle angel must be adjusted properly; otherwise the

projectile does not hit its target. If the adjustment process is done manually, then several trials will be done to achieve the proper angle as shown in Fig 6.5. The first trial is done by adjusting the muzzle by an initial angel, and then shooting the projectile. According to the projectile’s final destination and its difference from the target location, the muzzle angel is adjusted, and another trial is done. The angle is adjusted several times (the trial is repeated) until the projectile hits its target. Similarly, the Shooting method starts by guessing initial conditions, \hat{x}_0 , for the system (e.g. the muzzle initial angle), then finding the solutions of the system’s equation for the entire domain up to the final values, \hat{x}_f as shown in Fig 6.6. By comparing the resultant final values with their actual values, x_f , the initial guess is then refined and the process is repeated iteratively to minimize the error in the final (boundary) values. Once the error becomes smaller than a threshold value, iteration stops and the solution is adapted.

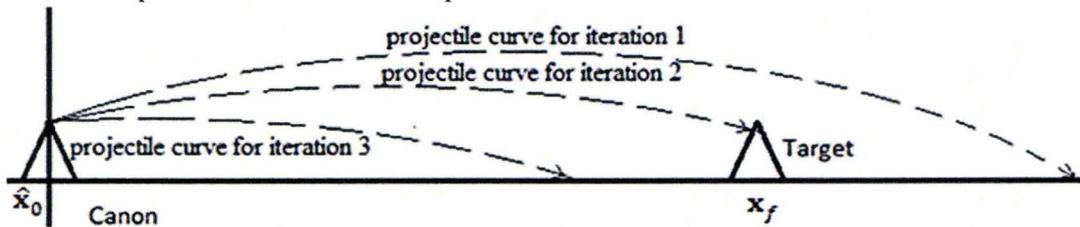


Fig 6.5: Adjusting the angle of the Canon Muzzle manually to hit a target.

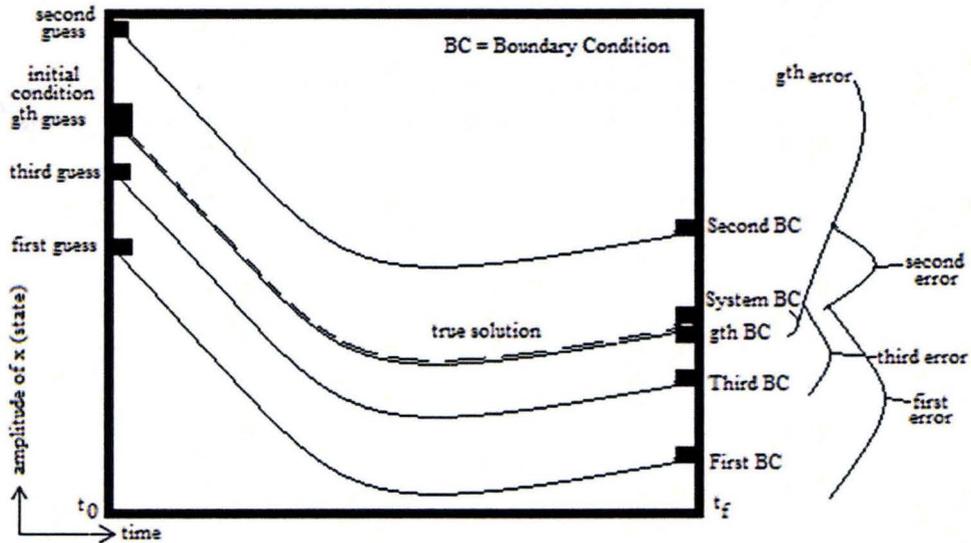


Fig 6.6: The Shooting method for one dimensional ordinary differential equation problem.

6.4. The Iterative Bi-Section/Shooting Method

The Bi-Section/Shooting method is a combination used to iteratively extract model parameters from the measurements for systems in which only model structure is known. The maximum number of parameters that could be estimated for an n^{th} order system using this method is $n + 1$. The system's parameters are assumed to be constant during the operation and they are divided into two groups; the first group is of size n and is obtained by the Shooting and Bisection methods, and the second group is extracted based on measurements and the parameters from the first group. It has a size equal to the number of observable parameters. The second group is extracted from noisy measurements. Therefore, they are stochastic variables with variances that are functions of noise and modeling errors pertaining to the parameters in the first group. For example, in second order systems, the gain and the natural frequency are the first group parameters, and the damping ratio is the second group parameter. The damping ratio may be extracted from measurements if the gain and the natural frequency are known. The estimate of the damping ratio becomes a stochastic variable with variance that is a function of the noise and the variance in the latter two parameters as shown in Fig 6.7. The figure shows that if the modeling errors are reduced, then the estimated parameter's variance is reduced, and its minimum value is obtained once the modeling errors become zero. Mathematically, the minimum variance of the curve in Fig 6.7 has a derivative with a zero value and the sign of the derivative changes across this point. By studying the sign of the derivative, the minimum variance could be obtained using the Bi-Section method.

Parameter estimation is performed by defining a search interval (for each range of group one parameters), estimating group two parameters and obtaining their variances. Based on the variance of group two parameters, the intervals of group one parameters are reduced by using the Bi-Section method until a threshold is reached. For example, to implement this method for a first order system with two model parameters; i.e. X and Y , the following process is used.

One of the two parameters is chosen first; i.e. X . An interval is specified for this parameter and five different values are arbitrarily chosen; i.e. X_1 to X_5 . These values are uniformly distributed along the interval and they are assumed to be the Shooting method's initial guesses.

For each of these values, and using data segments of measurements and input, the second parameter, Y_i , may be extracted. Note that the extracted parameter Y_i is not constant and is a temporal function. Thus, five variance values of the second parameter Y_i are obtained

one for each X_1 to X_5 from the Shooting method. As discussed earlier, each variance is a function of the noise as well as error in the first parameter estimate X_i .

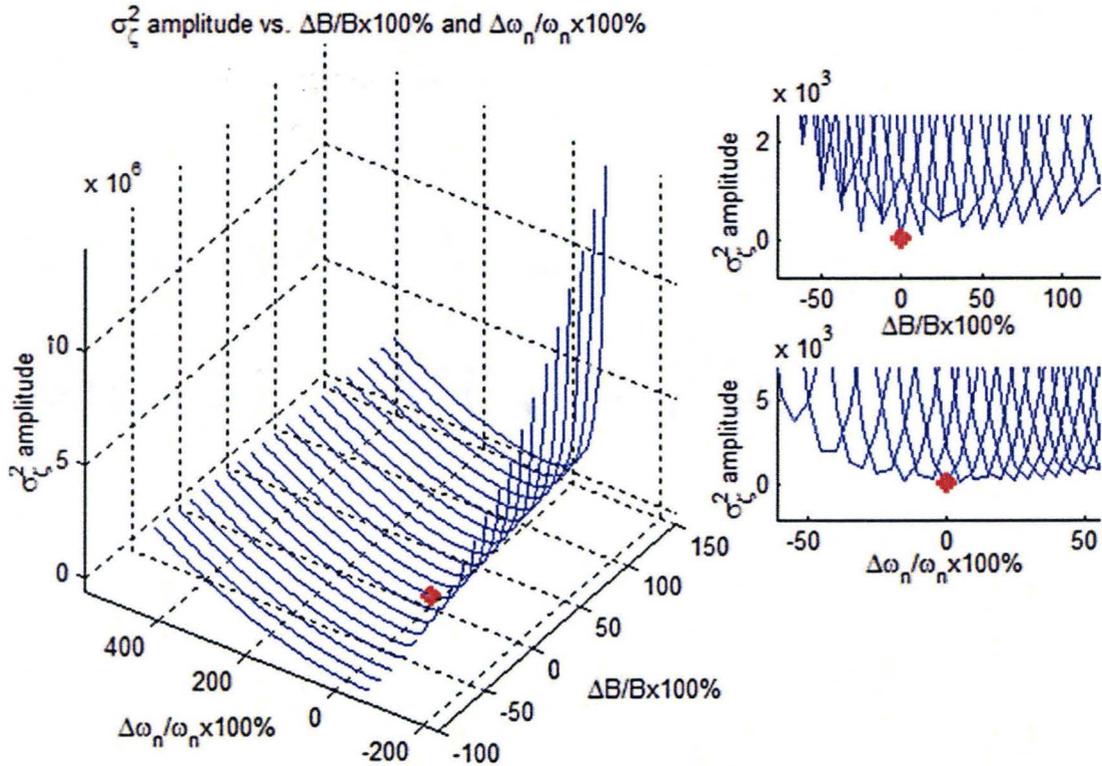


Fig 6.7: The affection of Modeling errors on the extracted parameter, ξ , for a second order system.

The variance values are distributed as shown in Fig 6.8. Note that these points have a parabolic shape. Taking the derivative of the shape function (to obtain the root of the derivative that represent the minimum variance) and using the Bi-Section method, a new subinterval is obtained.

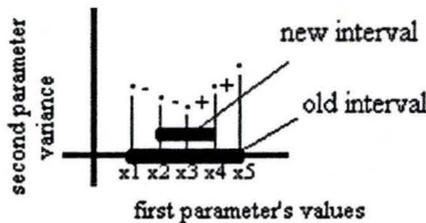


Fig 6.8: Refining the interval using the Bisection Method.

The algorithm is iteratively repeated until the width of the subinterval becomes smaller than a threshold. At each iteration, the first parameter is assumed to be half of the resultant interval, and the second parameter is chosen to be the mean of the corresponding extracted vector.

For higher order system, the algorithm becomes nested and more complicated, and the number of nested loops depends on the size of the first group and is defined as follows (assuming that each sub-loop consists of five loops):

$$\text{number of loops} = 5^{(n-1)} + 1 \quad 6.6$$

These loops are used to search for the best combination of the first group parameters based on the second group profile (variance). For example, Second order systems have three parameters; ω_n , ξ and B , and the IBSS requires six loops; 5-inner and 1-outer loops. Third order systems have four parameters and the IBSS requires twenty six loops; 25-inner and 1-outer.

The algorithm of IBSS is demonstrated by the following example.

Example 6.1

This example demonstrates the application of the IBSS algorithm to a second order system. The parameters are divided into two groups as previously discussed. The first group consists of the gain and the natural frequency, and their estimation is performed by an outer loop and five inner loops. The second group consists of the damping ratio. Assuming the outer loop is related to the gain B and the inner loops are related to the natural frequency ω_n . The computation loops are illustrated in Fig 6.9 and as follows.

OUTER LOOP:

- 1- The algorithm selects one of the parameters from the first group, e.g. B , for the outer loop. It relies on the availability of upper and lower bounds for it. If the parameter is the gain B , then the lower and upper bounds are defined as: B_1 and B_5 , respectively.
- 2- The algorithm defines five intermediate values within the above range, i.e. B_1 to B_5 , where B_1 is the lowest value, B_5 is the highest value and the values of B_2 to B_4 are evenly distributed between B_1 and B_5 .

- 3- For each value of B_i , an inner loop is created from the other parameter in group one, e.g. ω_n . Therefore, five inner loops are created. Each one of the inner loops has the following process:

INNER LOOP $i, i = 1 \dots 5$:

- The algorithm selects the other parameter from the first group (which is in this example ω_n) to construct the inner loop. Upper and lower bounds for it are denoted as: ω_{i_5} and ω_{i_1} , respectively.
- The algorithm defines five intermediate values within the above range, i.e. ω_{i_1} to ω_{i_5} , where ω_{i_1} is the lowest value, ω_{i_5} is the highest value, and the values of (ω_{i_2} to ω_{i_4}) are evenly distributed between ω_{i_1} and ω_{i_5} .
- Each value of ω_{i_j} is assumed to be an initial guess for the shooting method, where i and j denote the outer and inner loops, respectively. The pair (B_i, ω_{i_j}) are assumed to be the values of the unknown parameters B and ω_n . Note that if B and ω_n are known, the system satisfies the Observability condition for the estimation of the remaining parameters (ξ) using the measurement. Therefore, the third parameter, $\hat{\xi}_{i_j}$, can be extracted by using the measurement (z), the input (u), the sampling time T_s and the assigned pair of (B_i, ω_{i_j}) through filtering or by using the inverse model as follows:

$$\hat{\xi}_{i_j k-1} = \frac{B_i u_{k-1} - \omega_{i_j}^2 T_s z_{1k-1} - (z_{2k} - z_{2k-1})}{2\omega_{i_j} T_s z_{2k-1}} \quad 6.7$$

Where $\hat{\xi}_{i_j k-1}$ is the extracted damping ratio at time $k - 1$ using the pair (B_i, ω_{i_j}) .

- $\hat{\xi}_{i_j} = [\hat{\xi}_{i_j k-d+1} \dots \hat{\xi}_{i_j k}]$ is a stochastic variable segment that is a function of the system and measurement noise as well as modeling errors. The variance of each $\hat{\xi}_{i_j}$, $\sigma_{\hat{\xi}_{i_j}}$, is calculated for each corresponding ω_{i_j} which results in five values distributed as shown in Fig 6.8 (note the parabolic shape). The derivative of these values (differences between two successive variance points) is obtained and its sign is examined. Using the Bi-Section

method, a new subinterval is created from the old interval by reassigning the interval boundaries; ω_{i_1} or/and ω_{i_5} as shown in table 6.1. Note that only one minimum value of the variance exists in the interval. The extreme cases (case 1 and case 5 in table 6.1) treat the location of the minimum variance value to be close to the interval boundaries.

- e. After defining the new interval, steps (b) to (d) are repeated iteratively until the width of the resultant interval of natural frequency is smaller than a threshold, ε .
- f. Once the loop stops, the natural frequency, ω_{n_i} , of that loop is assumed to be the half of the final resultant interval.

END OF THE INNER LOOP i

- 4- The algorithm uses the measurements, the input, the sampling time and the pairs (B_i, ω_{n_i}) to obtain the damping ratio $\hat{\xi}_i$ for each inner loop as follows:

$$\hat{\xi}_i = \frac{B_i u_{k-1} - \omega_{n_i}^2 T_s z_{1k-1} - (z_{2k} - z_{2k-1})}{2\omega_{n_i} T_s z_{2k-1}} \quad 6.8$$

- 5- Equation (6.8) is used to estimate the damping ratio for each time step such that $\hat{\xi}_i = [\hat{\xi}_{i_{k-d+1}} \dots \hat{\xi}_{i_k}]$. The variance of $\hat{\xi}_i$, $\sigma_{\hat{\xi}_i}$, is calculated for each corresponding pair (B_i, ω_{n_i}) and results in five values distributed as shown in Fig 6.8. Similar to step (d), the derivative (difference) of these values is obtained and its sign is examined using the Bi-Section method. A new subinterval is then created for the gain B by reassigning the interval boundaries; B_1 or/and B_5 as discussed in step (d) and as shown in table 6.1 (by replacing ω_{i_j} with B_i).

- 6- After defining the new interval, steps (2) to (5) are repeated iteratively until the width of the resultant interval of the gain B is smaller than a threshold, ρ .

END OF THE OUTER LOOP

Case #	Sign $\sigma_{\xi_{i_5}} - \sigma_{\xi_{i_4}}$	Sign $\sigma_{\xi_{i_4}} - \sigma_{\xi_{i_3}}$	Sign $\sigma_{\xi_{i_3}} - \sigma_{\xi_{i_2}}$	Sign $\sigma_{\xi_{i_2}} - \sigma_{\xi_{i_1}}$	New ω_{i_1}	New ω_{i_5}	
1	+	+	+	+	Old ω_{i_1}	Old ω_{i_2}	
2	+	+	+	-	Old ω_{i_1}	Old ω_{i_3}	
3	+	+	-	-	Old ω_{i_2}	Old ω_{i_4}	
4	+	-	-	-	Old ω_{i_3}	Old ω_{i_5}	
5	-	-	-	-	Old ω_{i_4}	Old ω_{i_5}	

Table 6.1: The new interval boundaries using the Bi-Section method.

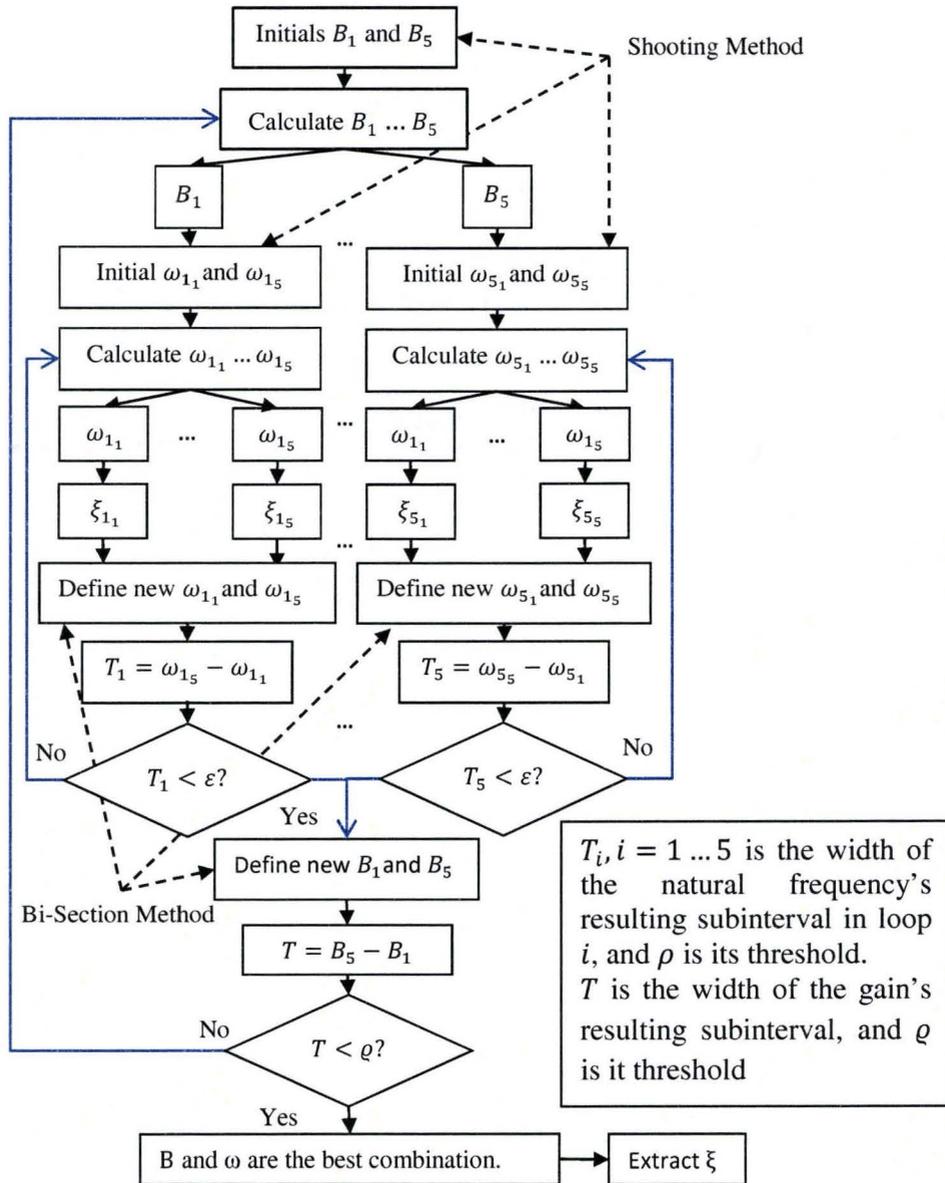


Fig 6.9: The IBSS algorithm for second order system.

Once the outer loop stops, the gain, \hat{B} , is assumed to be half of its final interval. One more inner loop is done using the gain \hat{B} and the steps (a) to (f) to obtain $\hat{\omega}_n$. The damping ratio, $\hat{\xi}$, is then obtained by using the measurement, the input, the pair $(\hat{B}, \hat{\omega}_n)$ and the inverse model as follows:

$$\hat{\xi} = \frac{1}{d} \sum_{j=k-d+1}^k \frac{\hat{B}u_{j-1} - \hat{\omega}_n^2 T_s z_{1j-1} - (z_{2j} - z_{2j-1})}{2\hat{\omega}_n T_s z_{2j-1}} \quad 6.9$$

Increasing the number of parameters results in a more nested algorithm, where for each inner loop there will be five sub-inner loops (four parameters), and for each sub-inner loop there will be a further five sub-sub-inner loops (five parameters), and so on. The algorithm's computational time grows exponentially when the number of parameters increases. This algorithm is only suitable for systems with low levels of complexity (first and second order systems).

6.5. The Iterative Bi-Section/Shooting Method combined with the SVSF and the KF

The IBSS is used to refine the estimated model. However, it does not estimate the states. Therefore, the IBSS needs to be combined with a filter such as the KF or the SVSF in order to estimate the states. This combination can also be used to estimate observable parameters. In this study, the combination of the SVSF with the IBSS is presented and is compared to the KF with the IBSS.

6.5.1. The Iterative Bi-Section/Shooting KF

The Iterative Bi-Section/Shooting method is combined with the Kalman Filter (IBSS/KF) to estimate the states and the parameters using data segments of the measurement signal. The segments are needed to excite the IBSS. As mentioned earlier, the parameters in the segments are assumed to be constant, otherwise the IBSS method will be misled. The IBSS element will be used to refine the filter's model to reduce the errors and then the KF is used to obtain the states. The main short coming of the IBSS/KF is that if the system's parameters change, the IBSS/KF will not know when this change has occurred. Therefore, a segment of the measurements is taken at a time, in which the parameters are assumed to be constant. The segment is processed using the IBSS to obtain the parameters. In the next step, the KF estimates the states as shown in Fig 6.10 and as follows:

- 1- The measurement signal is divided into segments. The parameters of the system are assumed to be constant within the segment. Each segment will be processed individually by the IBSS and the KF. The last time step value in a segment is considered as an initial condition to the next segment (for the KF).
- 2- The IBSS is applied to the segment to achieve the best value of the filter's model parameters. The cost function (or criterion of the goodness) is the

lowest variance in the estimation of the damping ratio, ξ , (in case of second order system) which can be extracted using equation (6.9).

- 3- The KF estimates the states of the system in the segment based on the new model's parameters from the IBSS.
- 4- Steps 1 to 3 are repeated for all data segments.

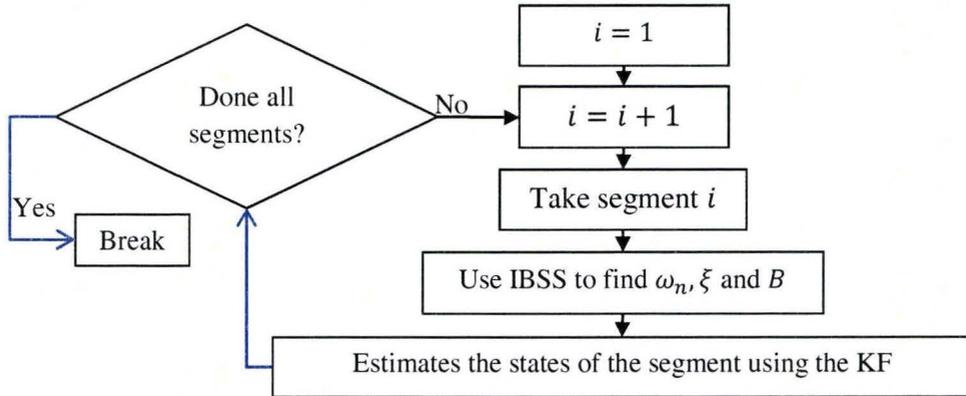


Fig 6.10: The IBSS/EKF algorithm for second order systems.

6.5.2. The Iterative Bi-Section/Shooting SVSF

The advantage of the combination of the IBSS with the SVSF is that the secondary indicators of performance of the SVSF can very accurately determine when a physical parameter has changes. This is very advantageous since the IBSS requires that parameters are constant during the interval that they are estimated. This provides a dynamic segmentation ability to the IBSS/SVSF that is not possible with the IBSS/KF. The combined IBSS/SVSF provides a remarkable algorithm that enables the estimation of both the states and the model parameters for low order systems. The combined robust stability of the SVSF and the interval definition can lead to a stable overall process.

The Iterative Bi-Section/Shooting method is combined with the SVSF (IBSS/SVSF) to estimate the states and the parameters. Moreover, the SVSF's secondary indicators of performance are used to detect parametric changes in the system once they occur and pass that information to the IBSS for interval selection. As discussed in chapter four, if chattering occurs when the boundary layer is set to have a width that is a function of the upper bounds of uncertainties, then this means that the upper bound has been breached and at least one of the parameters changed. Hence, chattering can provide a good indicator of the system changes. Once chattering occurs, the IBSS refines the estimated model. The SVSF then uses the refined model to continue estimating the states

until chattering condition re-occurs. The combined algorithm is summarized in Fig 6.11 and as follows.

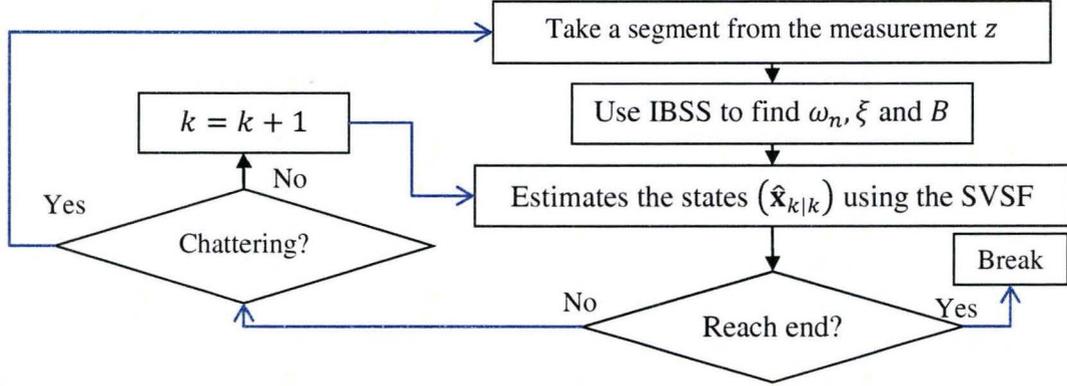


Fig 6.11: The IBSS/SVSF algorithm.

- 1- A SVSF with an appropriate boundary layer is used to estimate the states as discussed in chapter four, and the a priori chattering is monitored.
- 2- Once chattering occurs, the IBSS takes a segment of the measurement and processes it to obtain the parameters and to refine the filter's model.
- 3- The SVSF then continues to obtain the estimates until another chattering condition occurs.
- 4- Steps 1 to 3 are repeated.

6.6. Application of the IBSS/SVSF and the IBSS/KF to an electro-hydrostatic actuator

In this study, the proposed algorithm has been tested on an electro-hydrostatic actuator (EHA) described in chapter four. The EHA can be described by a third order model defined in its discretized state space form as follows:

$$\begin{bmatrix} x_{1_{k+1}} \\ x_{2_{k+1}} \\ x_{3_{k+1}} \end{bmatrix} = \begin{bmatrix} 1 & T_s & 0 \\ 0 & 1 & T_s \\ 0 & -\omega_n^2 T_s & 1 - 2\xi\omega_n T_s \end{bmatrix} \begin{bmatrix} x_{1_k} \\ x_{2_k} \\ x_{3_k} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ BT_s \end{bmatrix} u_k + \begin{bmatrix} w_{1_k} \\ w_{2_k} \\ w_{3_k} \end{bmatrix} \quad (6.10)$$

$$\mathbf{z}_{k+1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{k+1} + \mathbf{v}_{k+1}$$

Where $\omega_n = \sqrt{2 \frac{\beta A_E^2}{MV_0}}$, $B = \frac{2D\beta A_E}{MV_0}$ and $\xi = \frac{1}{2\sqrt{2}} \left(\frac{B_E V_0 + LM\beta}{\sqrt{MV_0} \sqrt{\beta A_E^2}} \right)$. The parameter β is the

effective bulk modulus and its value is defined in the range $(1 \times 10^8 - 3 \times 10^8) Pa$. The effective bulk modulus is made to change randomly several times. As the effective bulk modulus changes, the parameters B , ω_n and ξ change. The number of parameters in the EHA system is similar to the number of parameters of the second order system. Therefore, the IBSS algorithm described in section (6.4) is used. The output signal has been divided into segments with a length equal to 200 time steps for the IBSS/KF. In the IBSS/SVSF, segmentation is not required. Here, the estimates remain within the smoothing boundary layer (which is large enough to encompass the existence subspace when there are no modeling errors). If a system parameter (model) changes, the filter estimates exit their smoothing boundary layer thus inducing chattering. This provides a very effective mechanism for detecting change in the system model at its inception and is utilized in the SVSF/IBSS formulation. As such, instead of taking segments continuously, a data segment of length 200 time steps is taken once chattering is detected. The changes in the parameters are randomly made and each change will last for more than 20000 time steps. Within the segment, the parameters are assumed to be constant. The IBSS will attempt to estimate the filter parameters ω_n , ξ and B , while the SVSF or the KF estimates the system states. The sampling time is 0.001 s. ω_n , B , and ξ randomly vary between 100 to 400 Hz, 1 to 100 $\frac{m}{rad \times sec}$ and 0 to 1, respectively (since the bulk modulus has changed).

The process and measurement noise vectors are assumed to be white with a noise-to-signal ratio of 10% of their corresponding states, and with variances of $\sigma_w^2 = \sigma_v^2 =$

$$\begin{bmatrix} 5 \times 10^{-14} & 0 & 0 \\ 0 & 1 \times 10^{-12} & 0 \\ 0 & 0 & 1 \times 10^{-8} \end{bmatrix}$$

For the KF, the system and measurement noise covariance matrices are defined as $\mathbf{Q} = \mathbf{R} = \sigma_w^2$, and the initial error covariance matrix has a value of $\mathbf{P}_0 = \mathbf{I}_{3 \times 3}$. For the SVSF, the coefficient matrix $\boldsymbol{\gamma}$ has a value of $\boldsymbol{\gamma} = 0.1 \times$

$\mathbf{I}_{3 \times 3}$. The smoothing boundary layer is designed to have width of $\boldsymbol{\Psi} = \begin{bmatrix} 3 \times 10^{-6} \\ 1.2 \times 10^{-5} \\ 6 \times 10^{-3} \end{bmatrix}$. The

input consists of a random signal superimposed on step changes as shown in Fig 6.12.

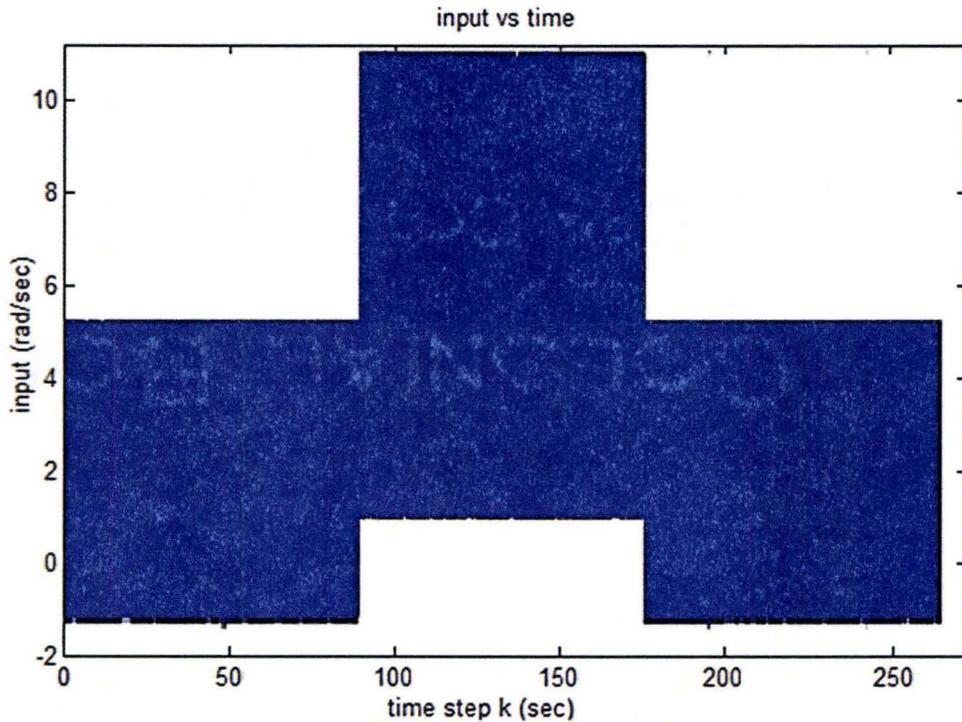


Fig 6.12: The input signal to the IBSS simulation.

6.6.1. Simulation results

6.6.1.1. The IBSS/KF

The results of the IBSS/KF are shown in Figs (6.13) to (6.18). The figures show that the IBSS/KF gives good stable and robust performance although modeling errors are present (the average absolute percentage errors of the states and parameters; i.e. $\frac{1}{\text{length}(y_i)} \sum_{i=1}^{\text{length}(y_i)} |(y_i - \hat{y}_i)/y_i| \times 100\%$ for y_i described by equation (6.11), are less than 4%). The IBSS provides the KF with a tuned model while the KF uses this model to estimate the states. The system and measurement noise affect the results as the estimation error increases when the noise amplitudes increase.

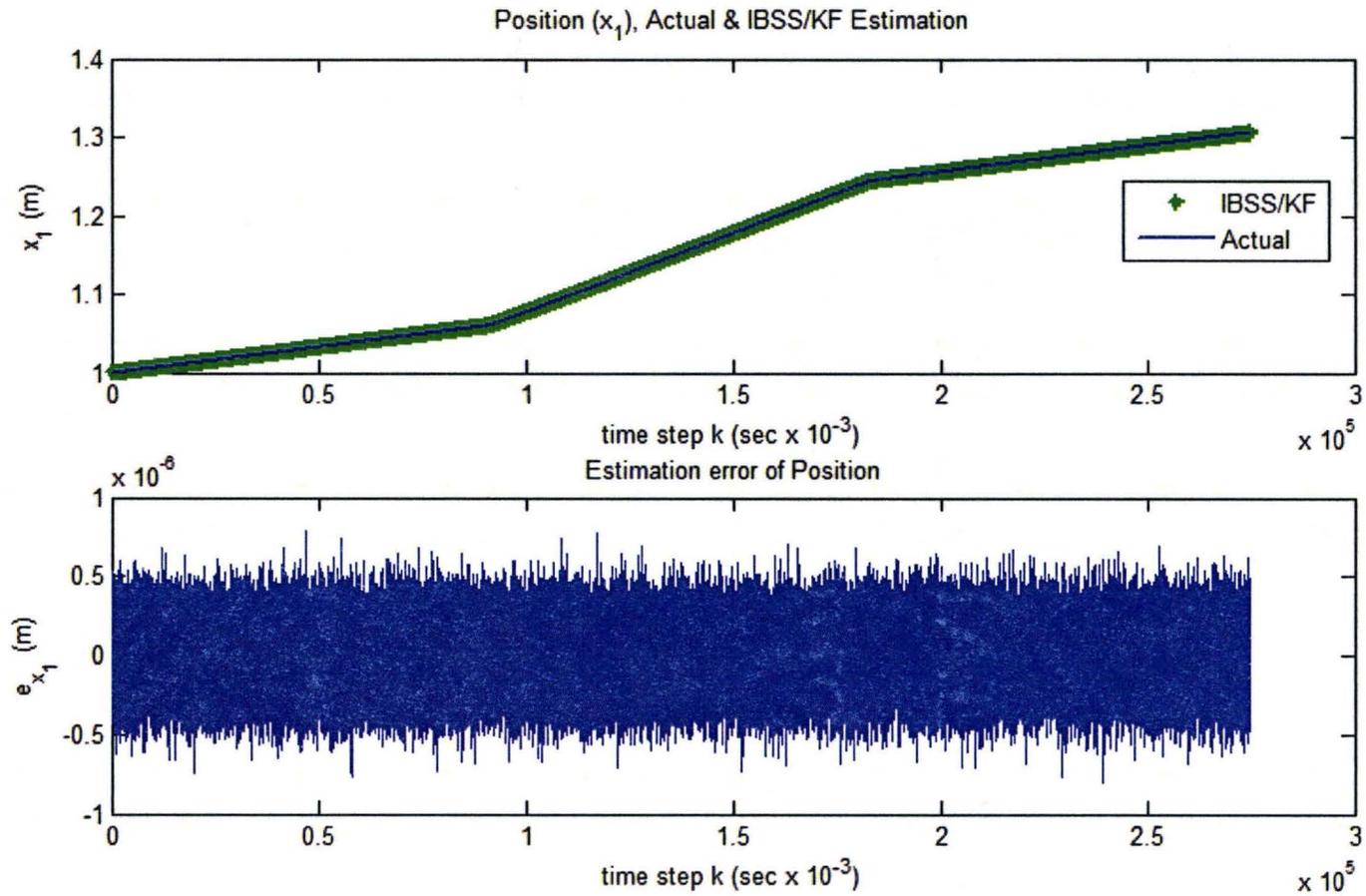


Fig 6.13: (a) x_1 's actual and estimated values, (b) estimation error of x_1 obtained by using the IBSS/KF.

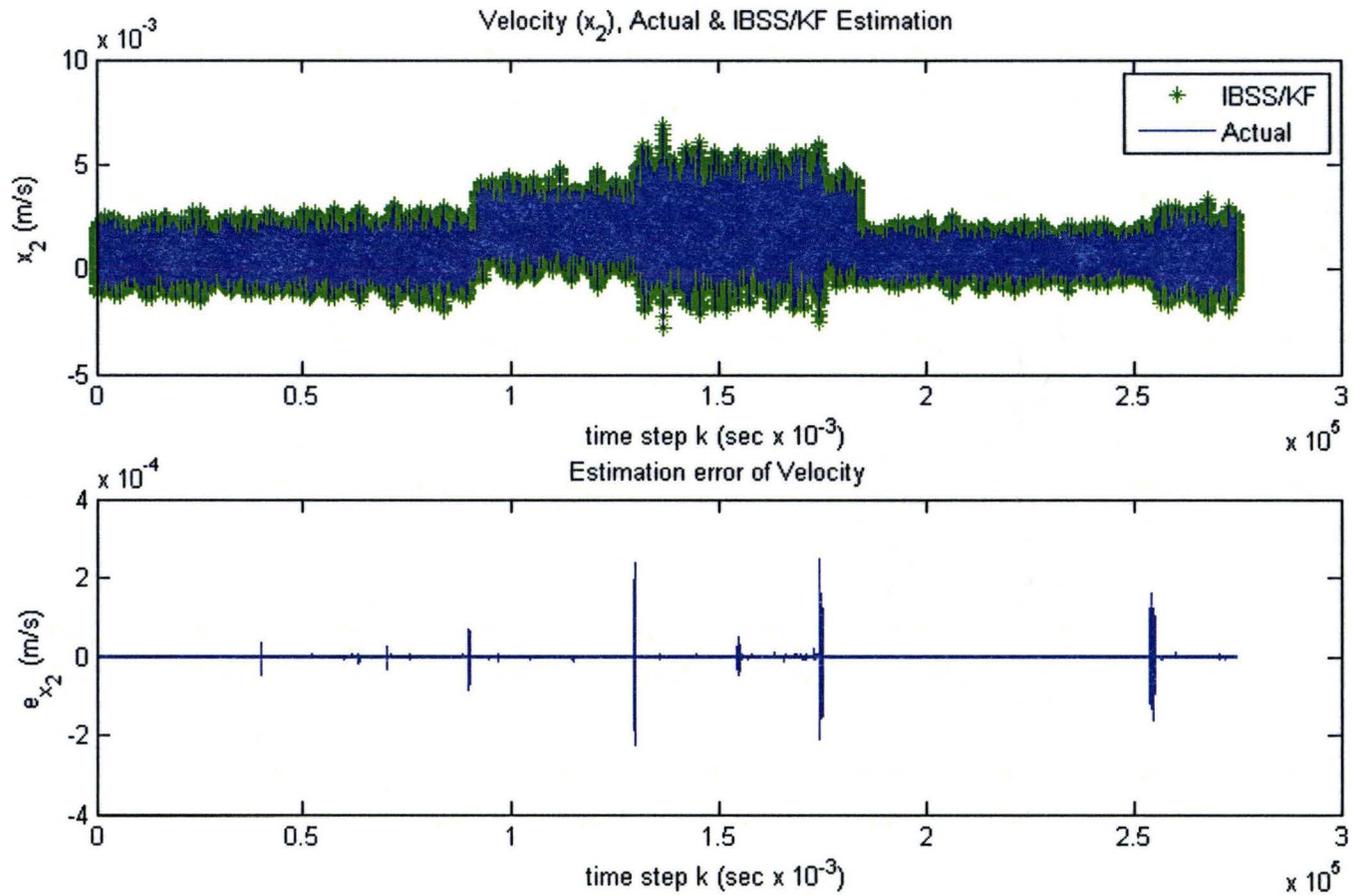


Fig 6.14: (a) x_2 's actual and estimated values, (b) estimation error of x_2 obtained by using the IBSS/KF.

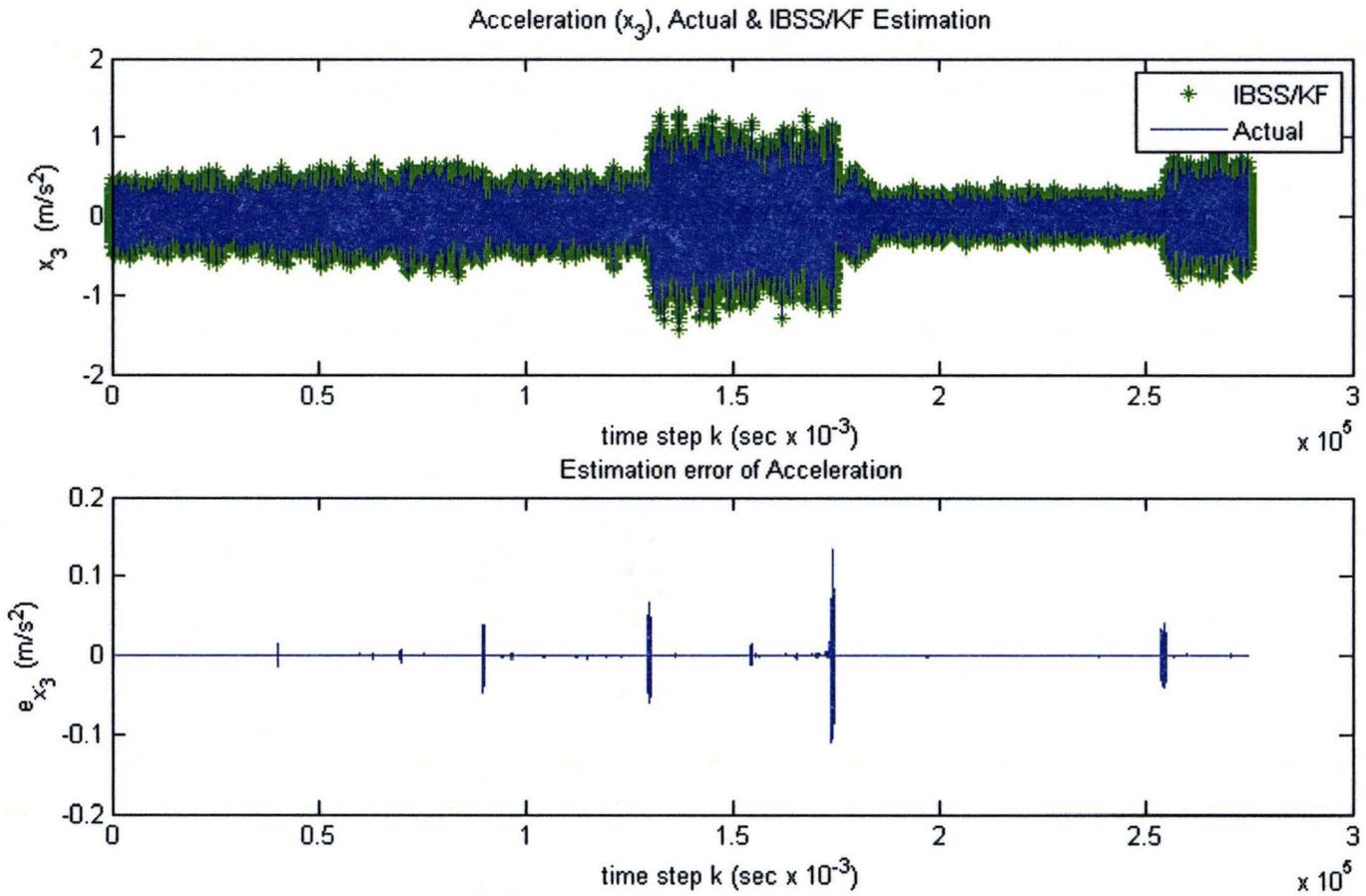


Fig 6.15: (a) x_3 's actual and estimated values, (b) estimation error of x_3 obtained by using the IBSS/KF.

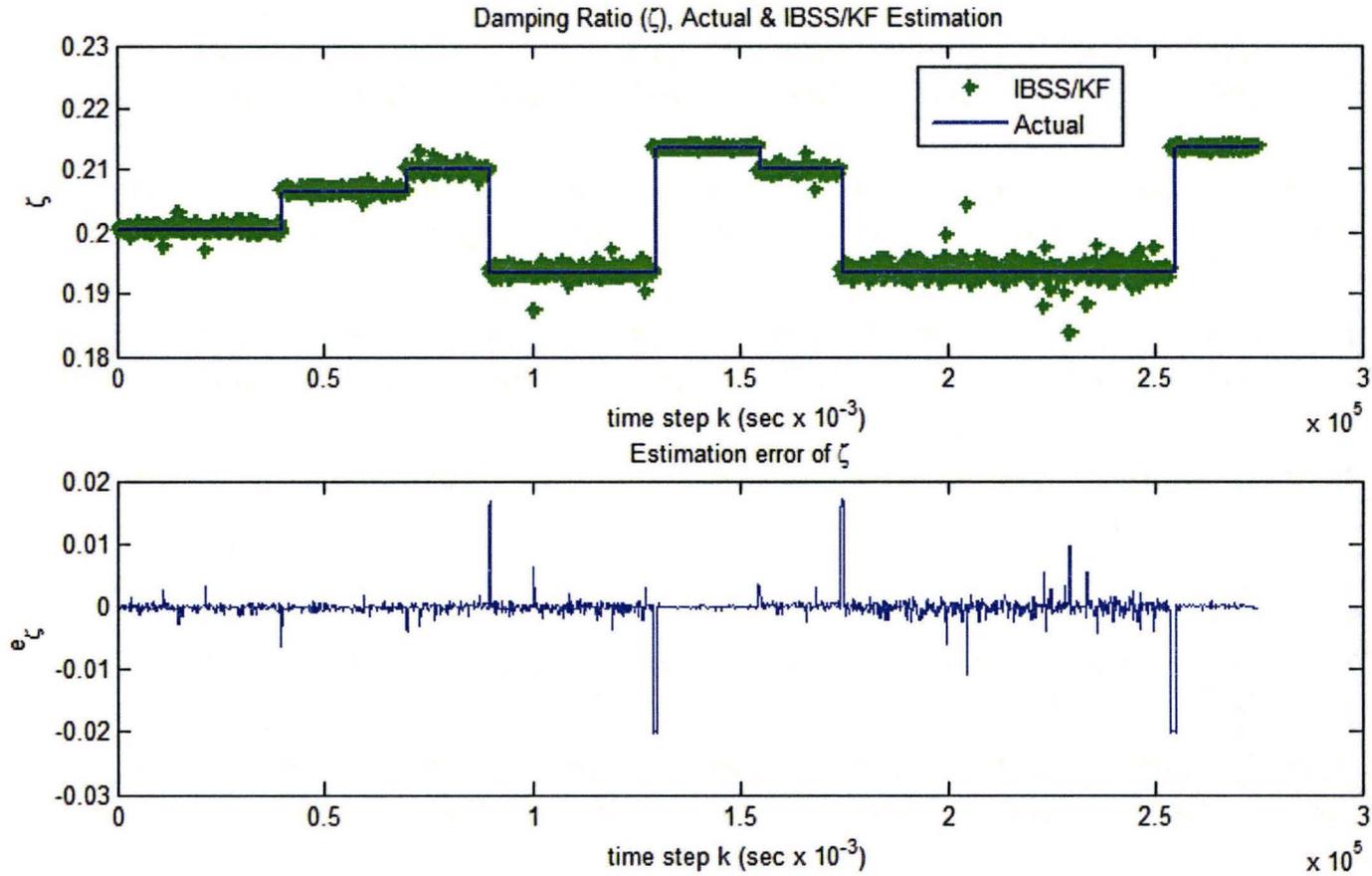


Fig 6.16: (a) ξ 's actual and estimated values, (b) estimation error of ξ obtained by using the IBSS/KF.

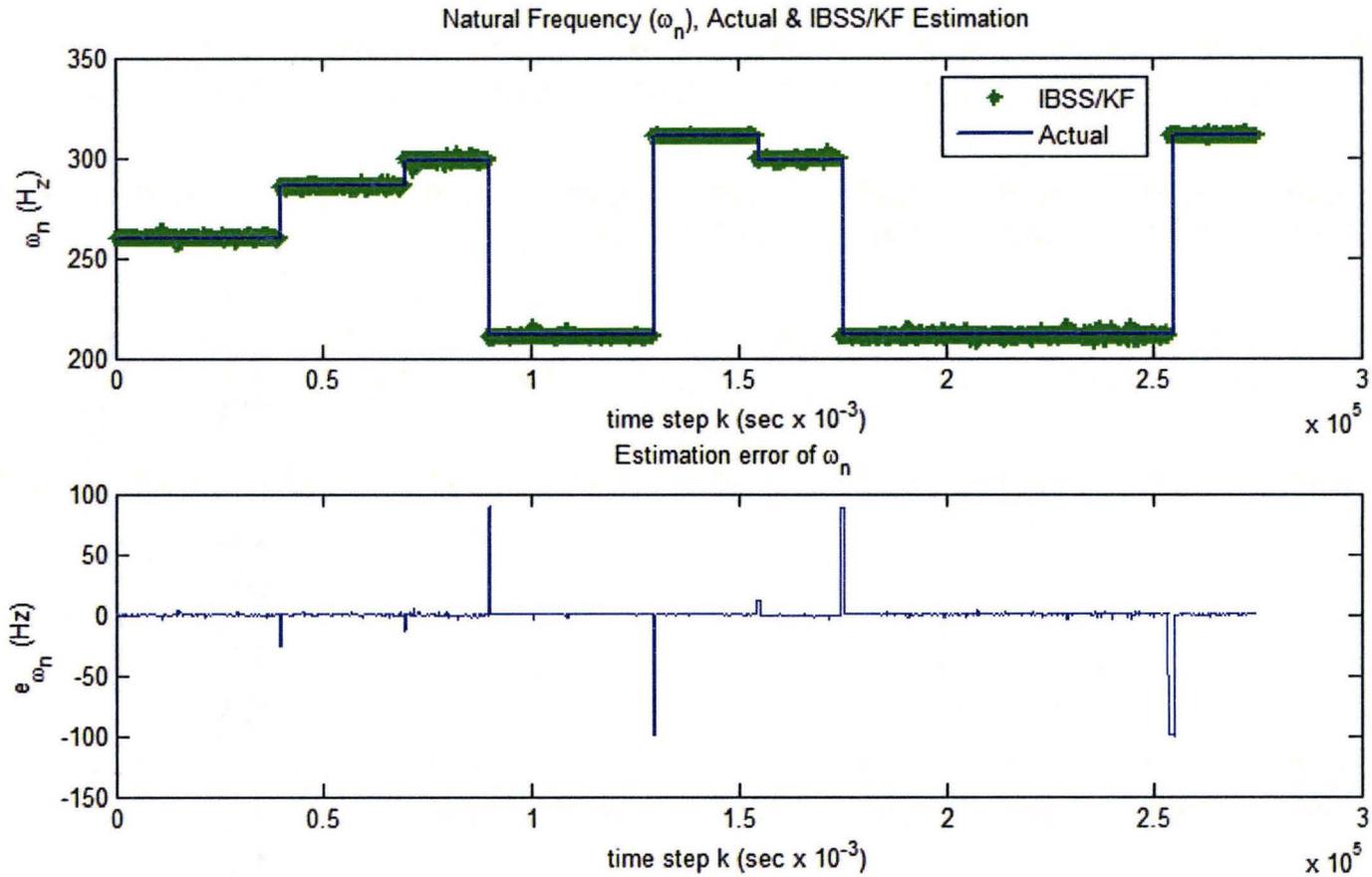


Fig 6.17: (a) ω_n 's actual and estimated values, (b) estimation error of ω_n obtained by using the IBSS/KF.

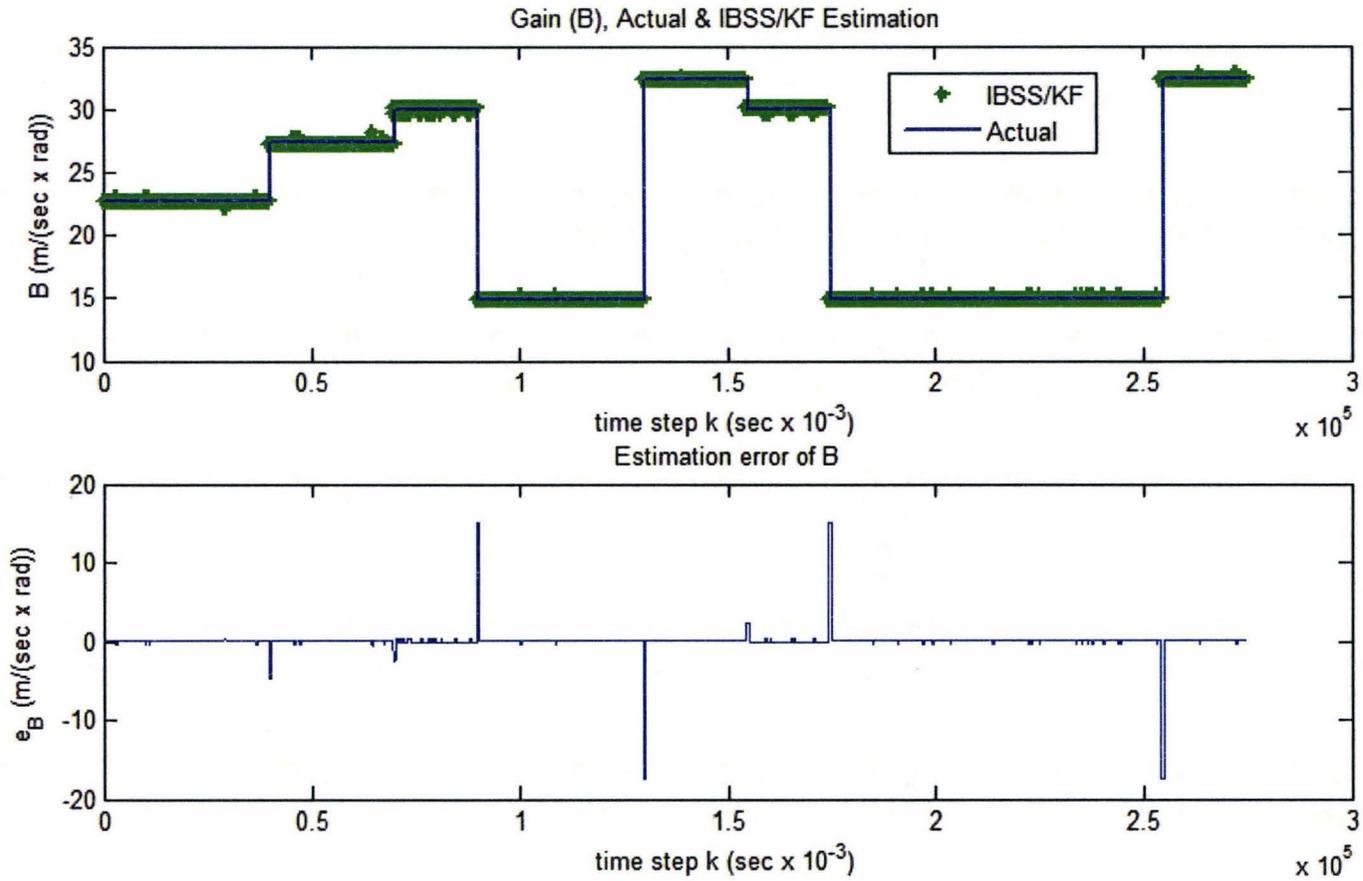


Fig 6.18: (a) gain's actual and estimated values, (b) estimation error of gain obtained by using the IBSS/KF.

6.6.1.2. The IBSS/SVSF

The results of the IBSS/SVSF are shown in Figs (6.19) to (6.24). The figures show that the IBSS/SVSF similarly to the IBSS/KF gives good stable and robust performance although modeling errors are present (the average absolute percentage errors of the states and parameters; i.e. $\frac{1}{length(y_i)} \sum_{i=1}^{length(y_i)} |(y_i - \hat{y}_i)/y_i| \times 100\%$ for y_i described by equation (6.11), are less than 4%).

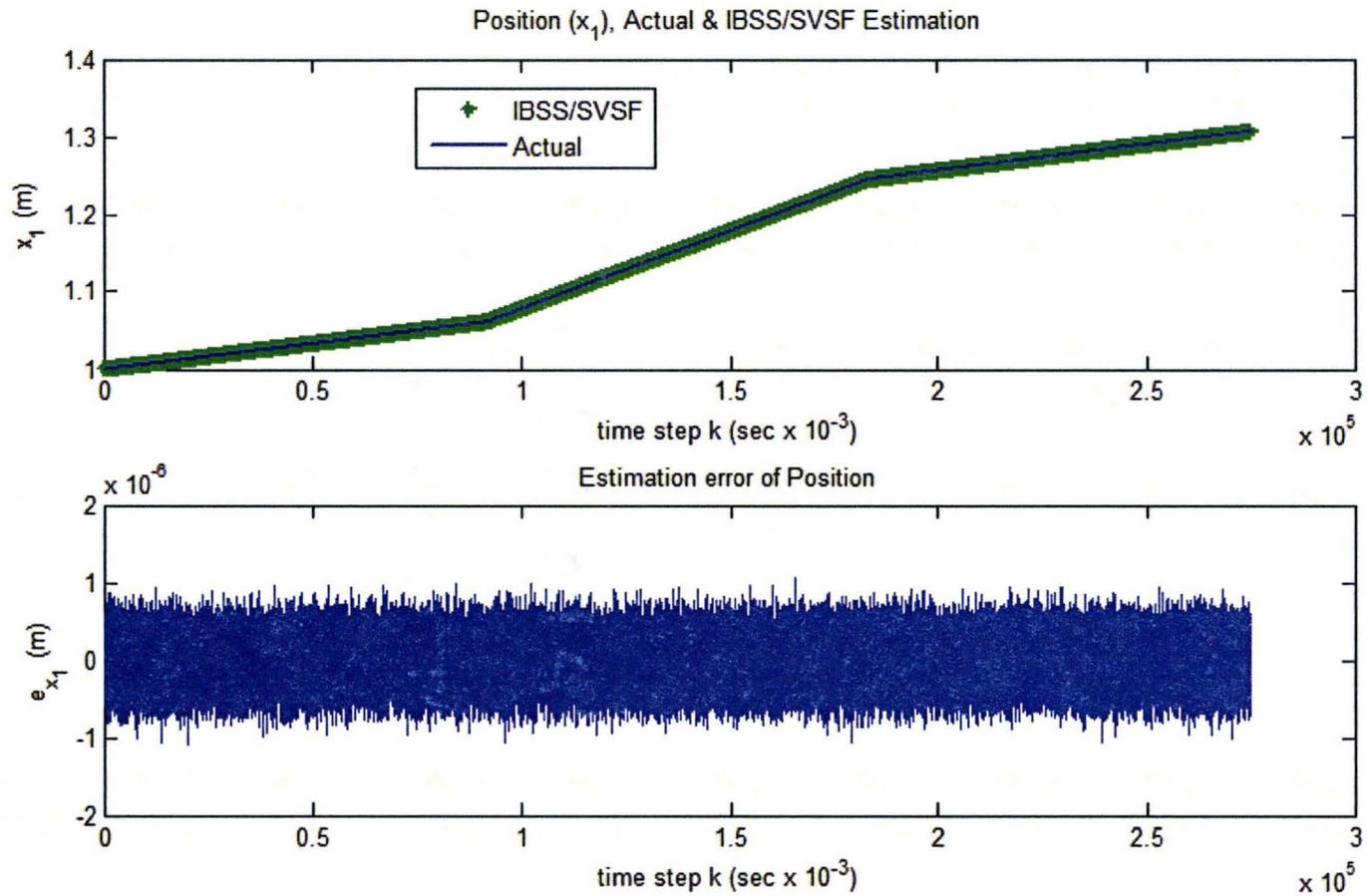


Fig 6.19: (a) x_1 's actual and estimated values, (b) estimation error of x_1 obtained by using the IBSS/SVSF.

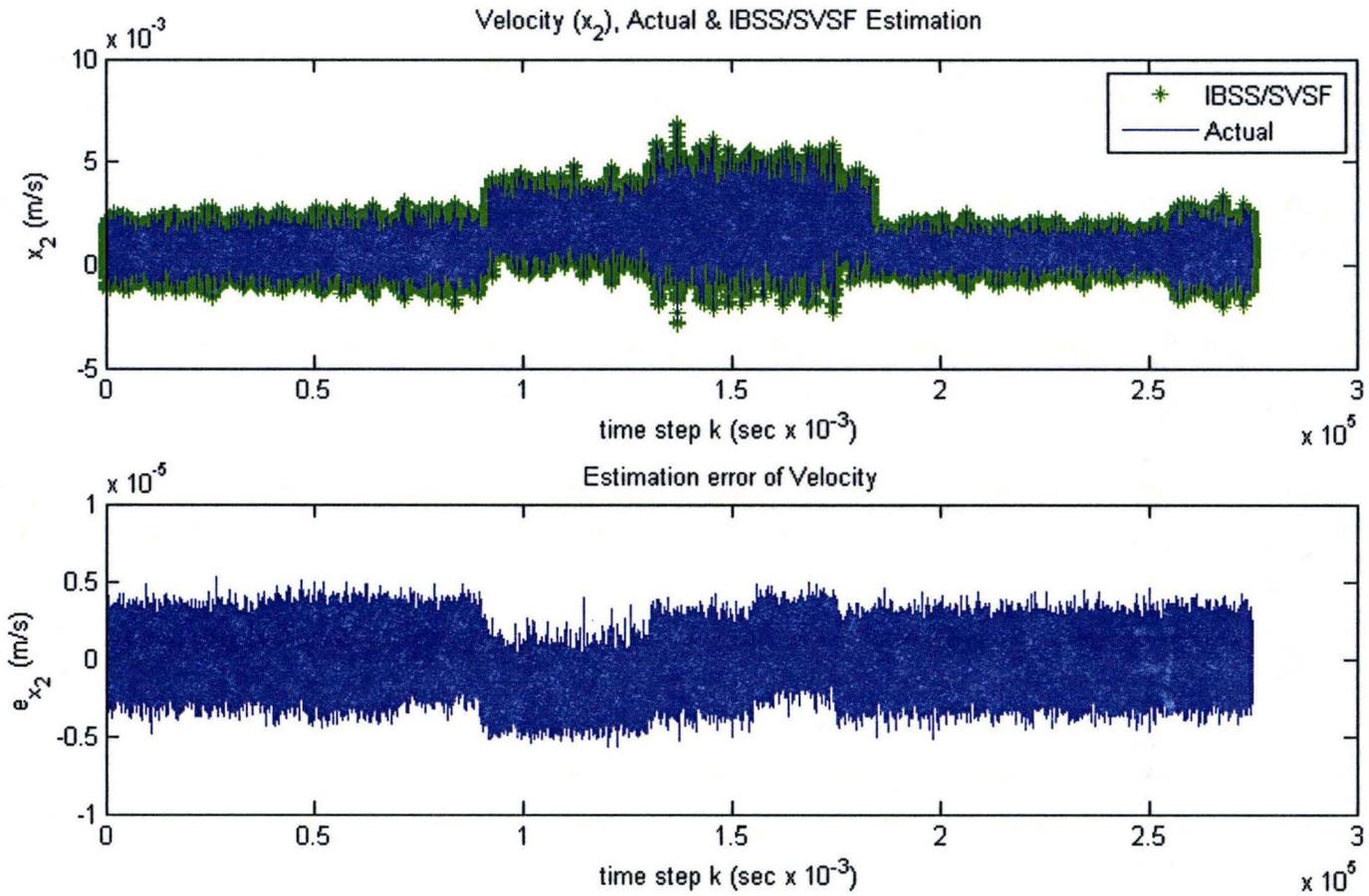


Fig 6.20: (a) x_2 's actual and estimated values, (b) estimation error of x_2 obtained by using the IBSS/SVSF.

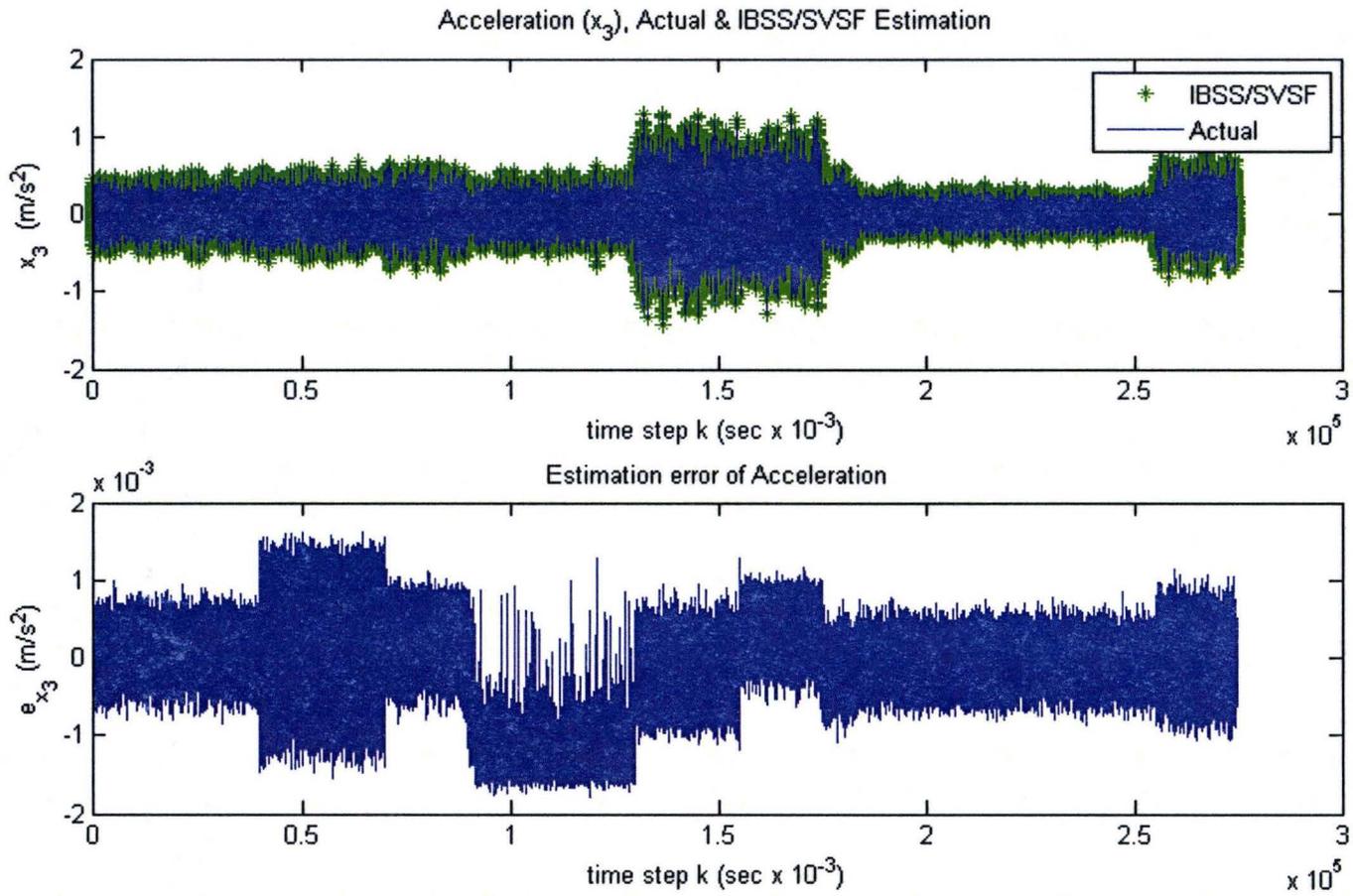


Fig 6.21: (a) x_3 's actual and estimated values, (b) estimation error of x_3 obtained by using the IBSS/SVSF.

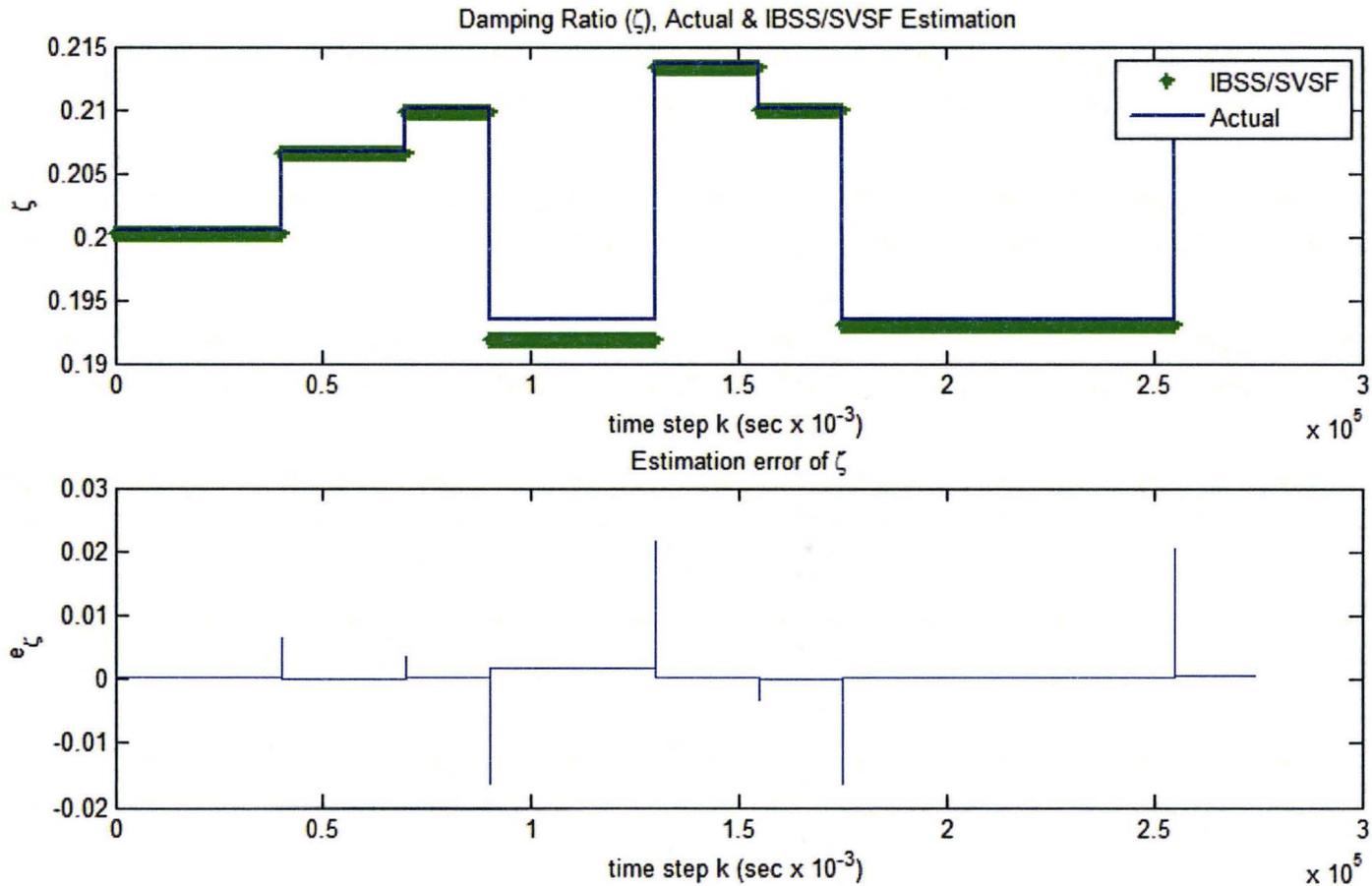


Fig 6.22: (a) ζ 's actual and estimated values, (b) estimation error of ζ obtained by using the IBSS/SVSF.

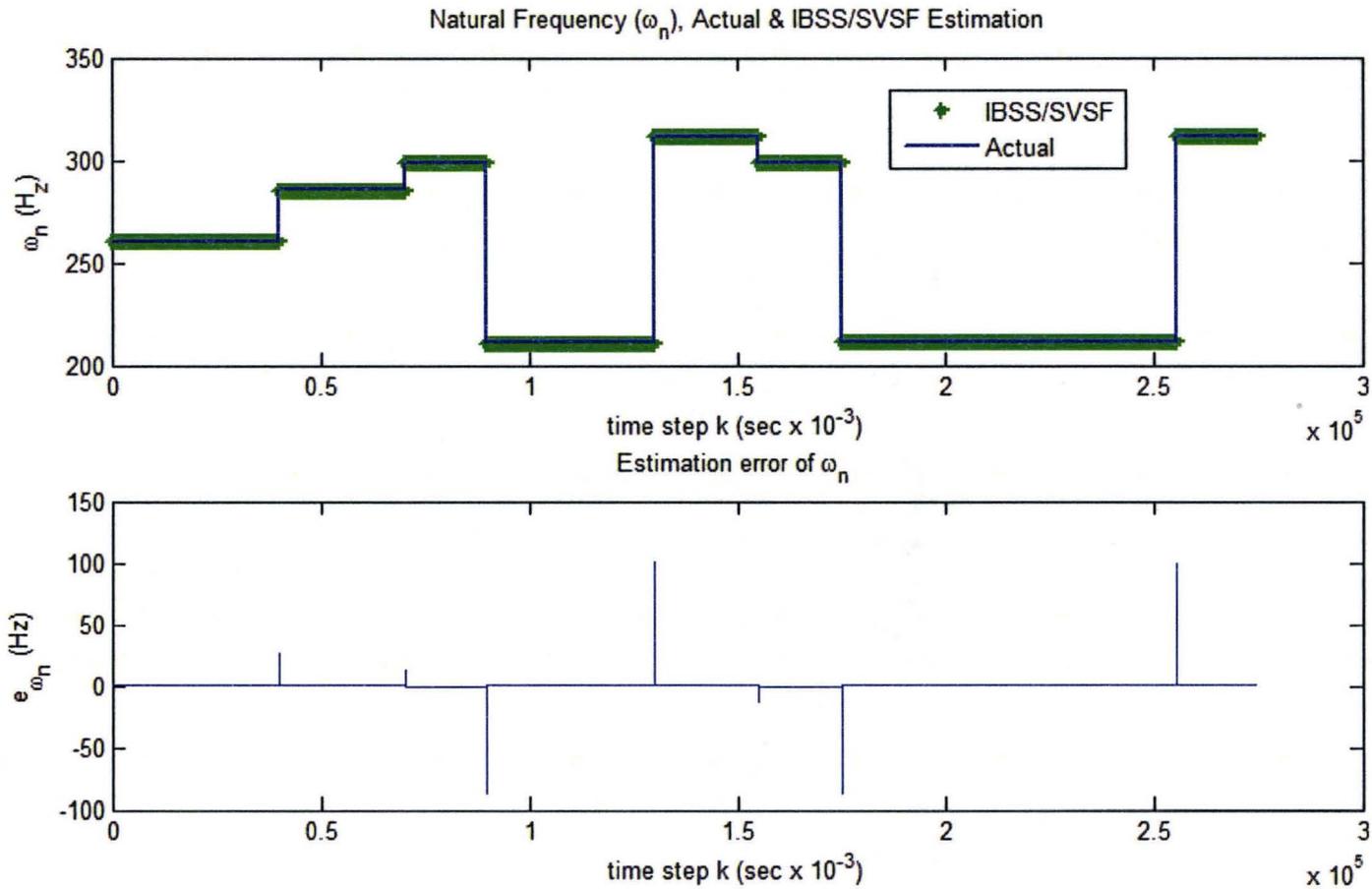


Fig 6.23: (a) ω_n 's actual and estimated values, (b) estimation error of ω_n obtained by using the IBSS/SVSF.

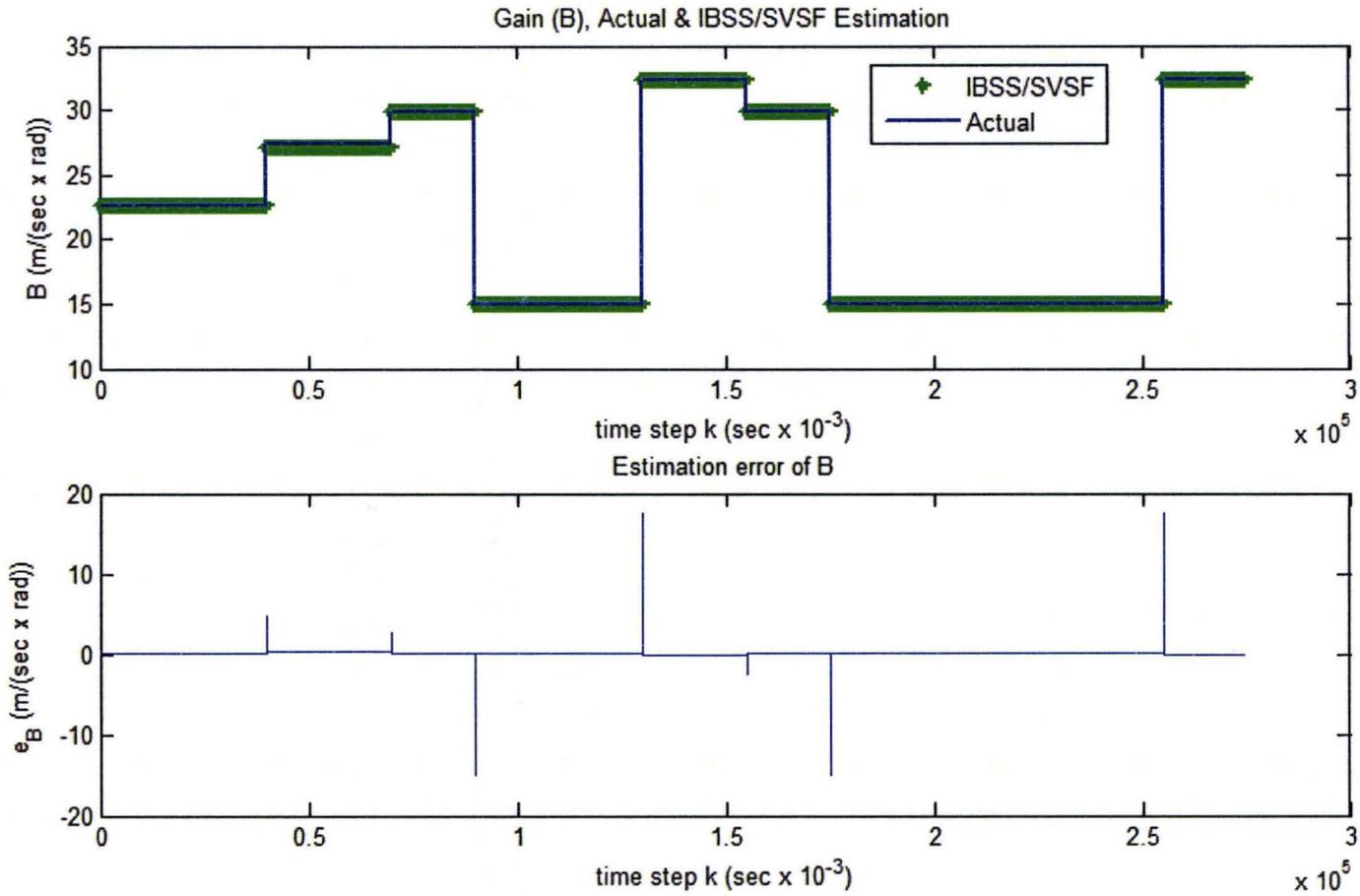


Fig 6.24: (a) gain's actual and estimated values, (b) estimation error of gain obtained by using the IBSS/SVSF.

6.6.1.3. Discussion

The two methods; IBSS/SVSF and IBSS/EKF, are compared in terms of:

- The root mean square error ($RMSE_j$), which is defined as follows:

$$RMSE_j = \frac{1}{length(x)} \sqrt{\sum_{i=1}^{length(x)} (y_{j_i} - \hat{y}_{j_i})^2} \text{ for } y = x_1, x_2, x_3, \xi, \omega_n \text{ and } B \quad 6.11$$

- The maximum absolute error ($MaxError_j$), which is equal to

$$MaxError_j = \max(|y_{j_i} - \hat{y}_{j_i}|) \text{ for } y = x_1, x_2, x_3, \xi, \omega_n \text{ and } B \quad 6.12$$

- The variance in the error ($VarError_j$) which is equal to

$$VarError_j = \frac{1}{length(x)-1} \sum_{i=1}^{length(x)} \left(y_{j_i} - \hat{y}_{j_i} - \frac{\sum_{i=1}^{length(x)} (y_{j_i} - \hat{y}_{j_i})}{length(x)} \right)^2 \text{ for } y = x_1, x_2, x_3, \xi, \omega_n \text{ and } B \quad 6.13$$

Table 6.2 summarizes the comparison.

	IBSS – KF	IBSS - SVSF
Computation Time	539.17 sec	42.8 sec
Position $RMSE$	1.16×10^{-5}	5.23×10^{-10}
Velocity $RMSE$	2.5×10^{-8}	2.8×10^{-9}
Acceleration $RMSE$	6.16×10^{-6}	8.92×10^{-7}
Damping Ratio $RMSE$	4.1×10^{-6}	1.5×10^{-6}
Natural Frequency $RMSE$	2×10^{-2}	1.2×10^{-3}
Gain $RMSE$	3.45×10^{-3}	1.86×10^{-4}
Position $MaxError$	1.3	1×10^{-6}
Velocity $MaxError$	3.8×10^{-3}	5.5×10^{-6}
Acceleration $MaxError$	5.8×10^{-1}	1.6×10^{-3}
Damping Ratio $MaxError$	2×10^{-2}	2×10^{-2}
Natural Frequency $MaxError$	100.26	99.97
Gain $MaxError$	17.42	17.45

Position <i>VarError</i>	3.7×10^{-5}	7.5×10^{-14}
Velocity <i>VarError</i>	1.72×10^{-10}	2.15×10^{-12}
Acceleration <i>VarError</i>	1.04×10^{-5}	2.16×10^{-7}
Damping Ratio	4.7×10^{-6}	5.9×10^{-7}
Natural Frequency <i>VarError</i>	109	4.5×10^{-1}
Gain <i>VarError</i>	3.27	9.4×10^{-3}
Identify the Changing in the Parameters	No	Yes
Computation Complexity	Complex	Simple
Memory needed	Large	Low
Segment Length	Effects	No-Effects

Table 6.2: Comparison between the IBSS/EKF and the IBSS/SVSF.

The IBSS/KF and the IBSS/SVSF yield good results on estimating the states and the parameters (the mean absolute percentage errors of the IBSS/KF and the IBSS/SVSF are less than 4%). However, there are some differences between the two algorithms. When the IBSS/KF is applied, the system cannot identify the positions where the parameters change. The algorithm divides the measurement signal into small segments and assumes that the segment is small enough, such that no changes happen within it. When a change happens within the segment, the error increases causing poor results for the parameters in that segment (the absolute percentage error may reach 120%) as shown in Fig 6.25. On the other hand, the IBSS/SVSF does not have this problem. Using the secondary indicators of performance allows the segment lengths to be adapted according to the time instance of change in the parameters.

Further to Fig 6.25, the IBSS/KF and the IBSS/SVSF are both able to estimate the states and the parameters. However, their results differ in terms of adaptation, variance in the error and the time needed to estimate the parameters and the states, as shown in table 6.2. The amplitude of noise affects the IBSS/KF more than the IBSS/SVSF, and the profiles of the estimated parameters are smoother in IBSS/SVSF than those of the IBSS/KF, as shown in Fig 6.26.

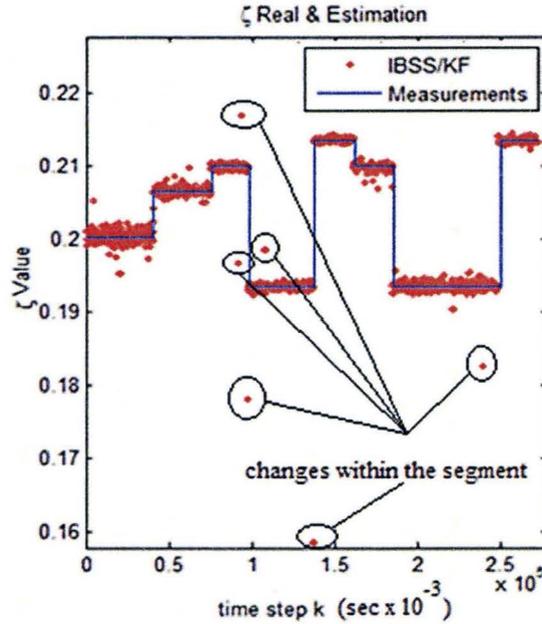


Fig 6.25: The error when changes happen within the segment.

The IBSS/SVSF requires less time to estimate the states and the parameters compared to the IBSS/KF. Taking a segment each time and analyzing it takes a longer time than taking one segment per interval and analyzing it. This causes the IBSS/KF to take more than twelve times what is need for the IBSS/SVSF, as shown in table 6.2. This restricts using the IBSS/KF in on-line applications. Also, this problem leads to the use of a large amount of memory due to calculation of segmentation and analysis of data. On the other hand, the IBSS/SVSF needs less memory and computing resources.

Changing the segment length does not affect the IBSS/SVSF, while it greatly impacts the IBSS/KF. For example, reducing the segment length to 100 time steps makes the IBSS/SVSF be 1.3 times faster than the reported value in table 6.2 without affecting its performance. However, this reduction can potentially reduce the overall *RMSE* of the IBSS/KF as the segments that have parametric changes within them become smaller. When the changes happen inside a segment, it means that the model is incorrectly estimated because it will be based on two partially different system models. If the segment is small, then the error effects of that segment becomes negligible in the overall *RMSE*. However, this is obtained at the expense of the computational time which is almost tripled.

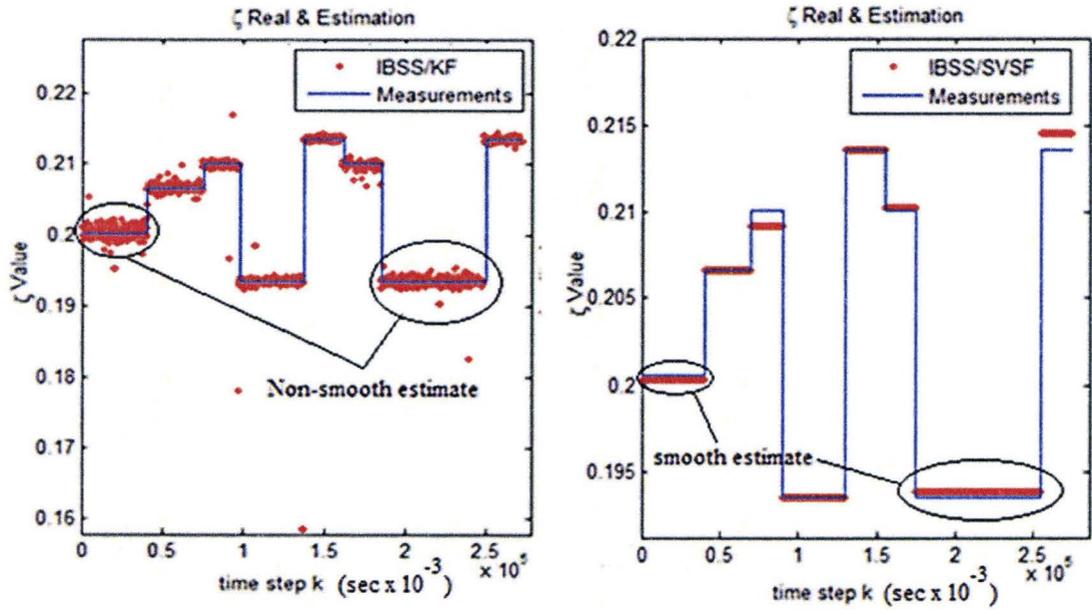


Fig 6.26: Comparison between the IBSS/KF and the IBSS/SVSF in term of the damping ratio profile.

Chapter Seven:

Summary and Conclusion Remarks

This thesis is concerned with the development of robust and model-based parameter and state estimation techniques for their application to fault detection. In model-based fault detection system parameters are monitored to give an indication of system state of health by comparing them against thresholds, (Tang & Wang, 2010). Faults usually result in un-modeled uncertainties in the filter model leading to numerical instability for this type of application. Therefore, filters must be robust to uncertainties and remain stable. One of the most robust types of filter that can be used for fault detection applications is the Sliding Mode Observer (SMO). Their equivalent output error injection component (filtering the output error signal by using low-pass filter) can be used to detect and extract modeling errors. However, this type of observer is severely impacted by noise and the usage of the digital filtering on the equivalent output error can result in excessive loss of information.

A recently developed version of the SMO formulated in a predictor corrector form, referred to as the Smooth Variable Structure Filter (SVSF), was proposed in (Habibi, 2007). The SVSF is based on the Sliding Mode Control concept. It defines a hyperplane and then applies a discontinuous corrective action that forces the estimates to switch back and forth across that plane within an existence subspace. The discontinuous action gives the filter its robustness. However, it results in chattering. To remove chattering, the discontinuous action is refined by using a saturation function with an associated fixed-width smoothing boundary layer. The SVSF has two sets of indicators of performance; the primary indicators of performance, which are the estimation errors, and the secondary indicators of performance, which are the a priori chattering signal that can occur if the smoothing boundary layer width is smaller the existence subspace width. The latter is the major SVSF characteristic that is used in this thesis. The SVSF and its forms are discussed in chapters four and five including its proof of stability and its gain derivation.

The application of the SVSF to a linear system with a full rank measurement matrix is explored in chapter four. It is noted that the SVSF has two existence subspaces. One as previously reported for the a priori estimate and one newly determined for the a posteriori estimate. These subspaces are mathematically formulated in sections (4.3.1.1) and (4.3.1.2). The a priori existence subspace's equation shows that its width is a function of all uncertainties including system and measurement noise, modeling errors and errors

in the initial conditions. Conversely, the mathematical formula of the a posteriori existence subspace shows that its width is a function of measurement noise and the errors in the initial conditions. The latter term becomes negligible in both subspaces after a few time steps (depending on the SVSF's coefficient matrix, γ). This means that modeling errors in initial condition do not affect the SVSF's a posteriori estimate.

The characteristics of the smoothing boundary layer are investigated. The proposed smoothing boundary layer in (Habibi, 2007) has a fixed-width. In this thesis, it is proven that if the smoothing boundary layer has a time-varying width linked to the a priori output estimation error, then the filter performance may be improved. In section (4.3.2.1), a novel strategy is presented to obtain a time varying smoothing boundary layer by using two SVSFs. The first filter is used to maintain the filter's stability and robustness. The a posteriori estimate of this filter is propagated to the second filter to obtain the time varying a posteriori estimate. The second filter is used to provide a time varying smoothing boundary layer to the first filter based on the diagonal elements of the error covariance matrix. This method has been tested on an electro-hydrostatic actuator in section (4.3.4.1). The results show that the proposed method is stable and has performance that exceeds that of the standard SVSF.

If the smoothing boundary layer is designed properly, then chattering is removed. However, if fault-driven modeling errors are added such that the amplitude of the output's a priori estimation error grows larger than the width of the smoothing boundary layer, then chattering will be observed in the a priori estimate, (Habibi, 2007). The a priori chattering can be used to determine the source and the amplitude of modeling uncertainties. In section (4.3.3), it is proven that the a priori chattering contains the modeling errors information. Three methods are proposed to extract these modeling errors from the a priori chattering signal (assuming that the system and the measurement noise vectors are stationary). These methods are based on the law of large numbers, which states that the segment's mean and variance of a stationary signal approach its expectation and covariance as long as the segment's length is large. The first method involves taking n segments of the a priori chattering, the input, and the measurement, and then calculating the segments' means. According to the law of large numbers, the white noise vectors vanish in these calculations. The modeling errors are then obtained by using a simple algebraic algorithm (section (4.3.3.1)). The second method involves taking a segment of the a priori chattering, the input, and the measurement, and then calculating the segment's covariance between the a priori chattering and the measurement, and then between the a priori chattering and the input. Using these two equations, the modeling errors are

extracted (section (4.3.3.2)). This method requires a better knowledge of the noise and an input signal that is white. The third method is similar to the second method except that it uses the cross-correlation instead of the covariance calculation to remove the above mentioned short comings (section (4.3.3.3)). The proposed methods only consider cases involving abrupt step changes. An opportunity for further research is to relax this limitation and consider parameter estimation involving gradual and continuous change. The monitoring and reconstruction method is tested on an electro-hydrostatic actuator that is made to have nine parametric changes with errors of up to 340%. The results show the following:

- The SVSF is a robust and stable filter.
- The detection process is sensitive to the width of the smoothing boundary layer. Compared to the noise uncertainties and the system matrix, if the width of the smoothing boundary layer is
 - o Small, then a false alarm may be detected.
 - o Large, then a delay in detecting modeling errors may occur.

In most real applications, the measurement matrix does not have a full rank. In such a case, the SVSF is combined with a reduced order technique that is similar to the Luenberger method. The reduced order technique extracts the hidden states from the measurements, and then uses them to compensate for the missing measurements. In order to apply the Luenberger method, the system must be observable and differentiable. Using the Luenberger method has limitations due to its compatibility with an observer rather than a filter and its sensitivity to noise.

To overcome the Luenberger method's limitations, a novel algorithm referred to as the Toeplitz/Observability SVSF is proposed in section (5.2.2). The system Toeplitz and the Observability matrices could be used to extract the hidden states by using previously stored values of the measurement and the input signals. The Toeplitz/Observability SVSF is applicable to systems with one measurement signal and requires that the system Toeplitz and the Observability matrices are exactly known. If modeling errors are present, then the system must be presented in its Observability canonical form. If the system qualifies this condition, then the resulting system Toeplitz and Observability matrices are not linked to system parameters; thereby the estimation error is a function of the system and measurement noise vectors and their derivatives. The Toeplitz/Observability SVSF is tested on an electro-hydrostatic actuator in section (5.2.3) and compared to the KF, demonstrating its superior performance in the presence of modeling uncertainties.

The Toeplitz/Observability SVSF is modified in chapter five to accommodate systems with multiple-degrees of freedom and/or multiple-measurements. A new canonical form using newly defined matrices referred to as the General System Toeplitz and the General Observability matrices is proposed referred to as the General Observability Canonical Form. Each measurement should represent one state only and at least the first state of each sub-system must be measured. If these conditions are satisfied, then the General System Toeplitz and/or the General Observability matrices are independent of the system parameters. In this thesis, two algorithms are developed to obtain the General Observability matrix.

The General Observability and the General System Toeplitz matrices are used for formulating a new method referred to as the General Toeplitz/Observability SVSF. The new method extends the extraction of modeling uncertainties from the chattering signals to observable systems with multiple measurements.

A novel iterative parameter estimation technique, referred to as the Iterative Bi-Section/Shooting method (IBSS), is proposed. The IBSS is a searching technique used to obtain model parameters for systems in which only the model structure is known. The IBSS is further combined with the SVSF and the KF. These methods are applied to an electro-hydrostatic actuator that is made to have changing parameters. The results show the superior performance of the IBSS/SVSF. Thus the SVSF/IBSS allow the extraction of all parameters and states in this order system using only the measurement vector.

The work presented in this thesis allows the extraction of modeling uncertainties using the SVSF concepts in state and parameter estimation. It is fundamentally important to white box diagnostics and health monitor.

Appendix A:**Nomenclature**

Italic-upper case letters are used to denote matrices and vectors, while their elements are denoted by italic lower case letters with subscripts i and/or j .

SYMBOL	COMMENTS	SIZE
$^{-1}, +$	Notation denoting an inverse and a pseudo inverse, respectively.	
$(\cdot)_{\max}$	The upper bound of the element inside the bracket.	
S'	The matrix/vector S after applying the transformation matrix T .	
$a^{(b)}$	The b^{th} derivative of a .	
$ _{\text{ABS}}$	Absolute value.	
\wedge	Estimation value.	
$A \circ B$	Schur product between A and B .	
T	Matrix Transpose.	
A, \hat{A}	The time-invariant system matrix and its estimate, respectively.	$n \times n$
A_k, \hat{A}_k	The time-variant system matrix at time k and its estimate, respectively.	$n \times n$
B, \hat{B}	The time-invariant input matrix and its estimate, respectively.	$n \times 1$
B_k, \hat{B}_k	The time-variant input matrix at time k and its estimate, respectively.	$n \times 1$
β	The input matrix of a continuous system.	$n \times 1$
$Ch_{k k-1}$	The a priori chattering vector at time k .	$n \times 1$

$\mathbf{Ch}_{k k}$	The a posteriori chattering vector at time k .	$n \times 1$
\mathbf{C}_{a_i, b_j}	Segment's cross correlation between \mathbf{a}_i and \mathbf{b}_j .	
\mathbf{d}_k	The uncertainties vector of the SVSF at time k .	$n \times 1$
$diag(\mathbf{a})$	Create a diagonal matrix with \mathbf{a} 's elements on its diagonal.	
Δa	Difference between a 's actual and estimated values.	
$ENDSUB_i$	Location of the last row in the i^{th} sub-system according to the system matrix.	
$\mathbf{e}_{z_{k k}}, \mathbf{e}_{z_{k k-1}}$	The a posteriori and a priori output's estimation error vectors at time k , respectively.	$m \times 1$
$\mathbf{e}_{x_{k k}}, \mathbf{e}_{x_{k k-1}}$	The a posteriori and a priori state's estimation error vectors at time k , respectively.	$n \times 1$
$\mathbf{e}_{z_{GTO_{k k}}}, \mathbf{e}_{z_{GTO_{k k-1}}}$	The a posteriori and a priori output's estimation error vectors of the General Toeplitz/Observability SVSF at time k , respectively.	$n \times 1$
$\mathbf{e}_{x_{GTO_{k k}}}, \mathbf{e}_{x_{GTO_{k k-1}}}$	The a posteriori and a priori state's estimation error vectors of the General Toeplitz/Observability SVSF at time k , respectively.	$n \times 1$
$\mathbf{e}_{z_{TO_{k k}}}, \mathbf{e}_{z_{TO_{k k-1}}}$	The a posteriori and a priori output's estimation error vectors of the Toeplitz/Observability SVSF at time k , respectively.	$n \times 1$
$\mathbf{e}_{x_{TO_{k k}}}, \mathbf{e}_{x_{TO_{k k-1}}}$	The a posteriori and a priori state's estimation error vectors of the Toeplitz/Observability SVSF at time k , respectively.	$n \times 1$
$E(a)$	The expectation operator of the element a .	
$\mathbf{f}, \hat{\mathbf{f}}$	The non-linear system matrix and its estimate, respectively.	$n \times 1$
\mathbf{f}''	The transformed non-linear system matrix.	$n \times 1$
Γ	The system matrix of a continuous system.	$n \times n$
\mathbf{Y}	The (VSF and EVSF)'s positive constant matrix.	$n \times m$
γ	The SVSF's positive constant matrix.	$n \times n$
$\mathbf{H}, \hat{\mathbf{H}}$	The time-invariant output matrix and its estimate, respectively.	$m \times n$
\mathbf{H}_{i_k}	The time-variant i^{th} portion of the output matrix at time k .	$1 \times n$

$\mathbf{H}_k, \hat{\mathbf{H}}_k$	The time-variant output matrix at time k and its estimate, respectively.	$m \times n$
$\bar{\mathbf{H}}_k$	The alternative measurement matrix at time k.	$n \times 1$
i, j	Subscripts used to identify elements of matrices and vectors.	1×1
$\mathbf{I}_{n \times n}$	The identity matrix with dimensions of $n \times n$.	$n \times n$
$\mathbf{J}_k, \mathbf{J}_{2k}$	Cost functions.	
k	Time step value.	1×1
\mathbf{K}_{VSF_k}	The correction gain of the Variable Structure Filter at time k.	$n \times 1$
\mathbf{K}_{EVSF_k}	The correction gain of the Extended Variable Structure Filter at time k.	$n \times 1$
\mathbf{K}_{SVSF_k}	The correction gain of the Smooth Variable Structure Filter at time k for systems that have measurement matrix with full rank.	$n \times 1$
$\mathbf{K}_{SVSF-II_k}$	The correction gain of the Smooth Variable Structure Filter at time k for systems that have measurement matrix with partial rank.	$n \times 1$
\mathbf{K}_{eTO_k}	The correction gain of the Toeplitz/Observability SVSF at time k.	$n \times 1$
\mathbf{K}_{eGTO_k}	The correction gain of the General Toeplitz/Observability SVSF at time k.	$n \times 1$
$\lim_{b \rightarrow c} a$	The value of a when b approaches c.	
LMS_i	Location of the measured state represented by the measurement i.	1×1
LM_i	The segment length of the measurement i needed for extraction purposes.	1×1
m	Number of measurements.	1×1
m_i	Number of the measurements related to the block i.	1×1
$MaxError$	The maximum absolute error.	1×1
m_a	The maximum measurement segment length (LMS).	1×1
\mathbf{M}_k	A Lyapunov function.	$n \times n$
n	System's number of states.	1×1

n_i	Number of states in the sub-system i in a multi-degree of freedom system.	1×1
$NSUB$	Number of sub-systems in a multi-degree of freedom system.	1×1
$\mathbf{O}, \widehat{\mathbf{O}}$	The Observability matrix and its estimate, respectively.	$n \times n$
$\mathbf{O}_{sub_i}, \widehat{\mathbf{O}}_{sub_i}$	The i^{th} sub-Observability matrix and its estimate, respectively.	
$\mathbf{O}_g, \widehat{\mathbf{O}}_g$	The General Observability matrix and its estimate, respectively.	$n \times n$
\mathbf{O}_{g_i}	The General Observability matrix at the i^{th} iteration in the modified selection procedure.	
$\omega_n, \widehat{\omega}_n$	The system's Natural Frequency and its estimate, respectively.	1×1
$\mathbf{P}_{zz k-1}, \mathbf{P}_{zz k}$	The a priori and a posteriori output's error covariance matrices at time k , respectively.	$m \times m$
$\mathbf{P}_{k k-1}, \mathbf{P}_{k k}$	The a priori and a posteriori error covariance matrices at time k , respectively.	$n \times n$
$\mathbf{P}_k, \mathbf{P}_0$	The error covariance matrix at time k and its initial condition, respectively.	$n \times n$
$\boldsymbol{\theta}_k$	The alternative measurement noise vector at time k .	$n \times 1$
$\boldsymbol{\Psi}_k$	The smoothing boundary layer vector at time k .	$n \times 1$
$\boldsymbol{\Psi}_{tv_k}$	The time-varying smoothing boundary layer vector at time k .	$n \times 1$
$\widehat{\boldsymbol{\Phi}}_k$	The Jacobian of the estimated system matrix at time k .	$n \times n$
\mathbf{Q}_k	The process noise covariance matrix at time k .	$n \times n$
\mathbf{R}_k	The measurements noise covariance matrix at time k .	$m \times m$
$\mathbb{R}^{a \times b}$	Space dimension of size $a \times b$, a is the number of rows and b is the number of columns.	1×1
$RMSE$	The root mean square error.	1×1
$\mathbf{sat}(\mathbf{a}, \mathbf{b})$	The saturated function of \mathbf{a} using the boundary layer \mathbf{b} .	
$sat(a, b)$	The saturated function of element a using the boundary layer b .	1×1
$\mathbf{S}_{at}(\mathbf{a}, \mathbf{b})$	Is the absolute diagonal matrix of the function $\mathbf{sat}(\mathbf{a}, \mathbf{b})$.	

$\sum_{i=b}^c \mathbf{a}_i$	The summation of vector \mathbf{a} from time b to time c .	
$\sigma_{\mathbf{a}_i, \mathbf{b}_j}$	The covariance of \mathbf{a}_i and \mathbf{b}_j .	
$\text{sgn}(\mathbf{a})$	The sign function of the vector \mathbf{a} .	
$\text{sgn}(a)$	The sign function of the element a .	1×1
$STSUB_i$	Location of the first row in the i^{th} sub-system according to the system matrix.	1×1
\mathbf{T}	Coordinate transformation matrix.	$n \times n$
$\mathbf{T}_o, \hat{\mathbf{T}}_o$	The system Toeplitz matrix and its estimate, respectively.	$n \times n$
$\mathbf{T}_{og}, \hat{\mathbf{T}}_{og}$	The General System Toeplitz matrix and its estimate, respectively.	$n \times n$
\mathbf{T}_w	The Toeplitz matrix of the system noise	$n \times (n.m_a)$
\mathbf{T}_{wg}	The General Toeplitz matrix of the system noise.	$n \times (n.m_a)$
T_s	Sampling time.	1×1
u_k	The input at time k .	1×1
$U(t)$	The continuous input.	1×1
$\mathbf{v}_k, \mathbf{V}_{max}$	The measurement noise at time k and its upper bound, respectively.	$m \times 1$
$\mathbf{w}_k, \mathbf{W}_{max}$	The system noise at time k and its upper bound, respectively.	$n \times 1$
$\mathbf{x}_k, \mathbf{x}_0, \mathbf{x}_f$	The state vector at time k , and its initial and boundary conditions, respectively.	$n \times 1$
$\hat{\mathbf{x}}_{k k}, \hat{\mathbf{x}}_{k k-1}$	The a posteriori and a priori estimates at time k , respectively.	$n \times 1$
$\hat{\mathbf{x}}_{TO_{k k}}, \hat{\mathbf{x}}_{TO_{k k-1}}$	The a posteriori and a priori estimates of the Toeplitz/Observability SVSF at time k , respectively.	$n \times 1$
$\hat{\mathbf{x}}_{GTO_{k k}}, \hat{\mathbf{x}}_{GTO_{k k-1}}$	The a posteriori and a priori estimates of the General Toeplitz/Observability SVSF at time k , respectively.	$n \times 1$
\mathbf{y}_k	The transformed system's states .	$n \times 1$
$\hat{\mathbf{y}}_{k k}, \hat{\mathbf{y}}_{k k-1}$	The transformed a posteriori and a priori estimates at time k , respectively.	$n \times 1$

$\hat{\mathbf{y}}'_{k k}, \hat{\mathbf{y}}'_{k k-1}$	The revised form of the a posteriori and a priori estimates using the measurement vector at time k, respectively.	$n \times 1$
$\mathbf{z}_k, \mathbf{z}_0$	The output vector at time k and its initial value, respectively.	$m \times 1$
\mathbf{z}_{subi_k}	The measurement portion related to the block i at time k.	
$\hat{\mathbf{z}}_{k k}, \hat{\mathbf{z}}_{k k-1}$	The a posteriori and a priori output's estimation vectors at time k, respectively.	$m \times 1$
$\hat{\mathbf{z}}_{TO_{k k}}, \hat{\mathbf{z}}_{TO_{k k-1}}$	The a posteriori and a priori output's estimation vectors of the Toeplitz/Observability SVSF at time k, respectively.	$n \times 1$
$\hat{\mathbf{z}}_{GTO_{k k}}, \hat{\mathbf{z}}_{GTO_{k k-1}}$	The a posteriori and a priori output's estimation vectors of the General Toeplitz/Observability SVSF at time k, respectively.	$n \times 1$
$\mathbf{\Xi}_k$	The alternative measurement vector of the original Smooth Variable Structure Filter.	$n \times 1$
$\mathbf{0}_{a \times b}$	A matrix with dimension $a \times b$ and zero elements.	$a \times b$
$\xi, \hat{\xi}$	The system's Damping Ratio and its estimate, respectively.	1×1

Table A.1 - Nomenclature

Bibliography

- Ainsleigh, P. (2007). Performance Of Multiple-Model Filters And Parameter-Sensitivity Analysis For Likelihood Evaluation With Shock Variance Models. *IEEE Transactions on Aerospace and Electronic Systems*, 43(1), 135-149.
- Ali, Z., Deriche, M., & Landolsi, M. (2009). Sigma Point Kalman Filters For Multipath Xchannel Estimation In CDMA Networks. *Proceedings of the 2009 6th International Symposium on Wireless Communication Systems*, (pp. 423-427).
- Alspach, D., & Sorenson, H. (1972). Nonlinear Bayesian Estimation Using Gaussian Sum Approximations. *IEEE Transactions On Automatic Control*, AC-17.
- Ambadan, J., & Tang, Y. (2009). Sigma-Point Kalman Filter Data Assimilation Methods For Strongly Non-Linear Systems. *Journal of the Atmospheric Sciences*, 66(2).
- Anderson, B., & Moore, J. (1979). *Optimal Filtering*. Prentice-Hall.
- Arasaratnam, I., & Haykin, S. (2009). Cubature Kalman Filters. *IEEE Transactions on Automatic Control*, 54(6), 1254-1269.
- Arasaratnam, I., Haykin, S., & Elliott, R. (2007). Discrete-Time Nonlinear Filtering Algorithms Using Gauss-Hermite Quadrature. *Proceedings of the IEEE*, 95(5), 953-977.
- Athans, M., Wishner, R., & Bertolini, A. (1968). Suboptimal State Estimation For Continuous-Time Nonlinear Systems From Discrete Noisy Measurements. *IEEE Transactions on Automatic Control*, AC-13(5), 504-514.
- Aurora, C., & Ferrara, A. (2007). A Sliding Mode Observer for Sensorless Induction Motor Sped Regulation. *International Journal of Systems Science*, 38(11), 913 - 929.
- Aurora, C., & Ferrara, A. (2007). A Sliding Mode Observer for Sensorless Induction Motor Speed Regulation. *International Journal of Systems Science*, 38(11), 913 - 929.
- Aurora, C., & Ferrara, A. (2007, July 11 - 13). Design and Experimental Test of a Speed/Flux Sliding Mode Observer for Sensorless Induction Motors. *Proceedings of the 2007 American Control Conference* (pp. 5881 - 5886). New York: IEEE.

- Aurora, C., Ferrara, A., & Levant, A. (2001). Speed Regulation of Induction Motors: A Sliding Mode Observer-Differentiator Based Control Scheme. *Proceedings of the 40th IEEE Conference on Decision and Control* (pp. 2651 - 2656). Florida, USA: IEEE.
- Bandyopadhyay, B., Gandhi, P., & Kurode, S. (2009, September). Sliding Mode Observer Based Sliding Mode Controller for Sloss-Free Motion Through PID Scheme. *IEEE Transactions on Industrial Electronics*, 56(9), 3432 - 3442.
- Barakat, M. (2005). *Signal Detection And Estimation*. Norwood :Artech House.
- Barbot, J., Boukhobza, T., & Djemai, M. (1996). Sliding Mode Observer for Triangular Input Form. *Proceedings of the 35th Conference on Decision and Control* (pp. 1489 - 1490). Kobe, Japan: IEEE.
- Barker, A., Brown, D., & Martin, W. (1995). Bayesian estimation and the Kalman Filter. *Department of Computer Science, Virginia University*, 30(10), 55-77.
- Bar-Shalom, T., Li, X., & Kirubarajan, T. (2001). *Estimation With Applications To Tracking And Navigation – Theory, Algorithm And Software*. John Wiley & Sons, Inc.
- Bar-Shalom, Y., & Li, X. (1993). *Estimation and Tracking: Principles, Techniques and Software*. Norwood, MA: 1993 ARTECH HOUSE, INC.
- Bartolini, G., Damiano, A., Gatto, G., Marongiu, I., Pisano, A., & Usai, E. (2003). Robust Speed and Torque Estimation in Electrical Drives by Second Order Sliding Modes. *IEEE Transactions on Control Systems Technology*, 11(1), 84 - 90.
- Bartolini, G., Levant, A., Usai, E., & Pisano, A. (1999). 2-Sliding Mode with Adaptation. *Proceedings of the 7th Mediterranean Conference on Control and Automation (MED99)*, (pp. 2421 - 2429). Haifa, Israel.
- Bayard, D., & Kang, B. (2003). A High-Order Kalman Filter For Focal Plane Calibration Of NASA's Space Infrared Telescope Facility (SIRTF). *AIAA Guidance, Navigation and Control Conference*,. USA.
- Bernardo, J., & Smith, A. (1994). *Bayesian Theory*. John Wileys& Sons.

- Berndt, B., Evans, R., & Williams, K. (1998). *Gauss And Jacobi Sums*. New York: Wiley.
- Bierman, J. (1975). Measurement Updating Using The U-D Factorization. *Proceedings of the IEEE Conference on Decision and Control*, 337-346.
- Boyd, S., Balakrishnan, V., & Kabamba, P. (1989). A Bisection Method for Computing the H-infinity Norm of a Transfer Matrix and Related Problems. *Math. Control Signals Systems*, 207 - 221.
- Briët, J., & Harremoës, P. (2009). Properties Of Classical And Quantum Jensen-Shannon Divergence. 79(5).
- Brown, A., & Bowles, W. (1983). Measurement Updating Using The U-D Factorization Method When The measurement noise is correlated. 1983, (pp. 344-348).
- Budde, L. (1963). The Reduction Of An Arbitrary Real Square Matrix to Tridiagonal Form Using Similarity Transformations. *Mathematic of Computation*, 17, 433-437.
- Burger, G., Van Leeuwen, P., & Evensen, G. (1998). Analysis Scheme In The Ensemble Kalman Filter. *Monthly Weather Review*, 126(6), 1719-1724.
- Carrasco, R., Cipriano, A., & Carelli, R. (2005). Nonlinear State Estimation in Mobile Robots using a Fuzzy Observer. *in Processing of the 16th IFAC World Congress*.
- Chaal, H., Jovanovic, M., & Busawon, K. (2009). Sliding Mode Observer Based Direct Torque Control of a Brushless Doubly-Fed Reluctance Machine. *2009 IEEE Symposium on Industrial Electronics and Applications, ISIEA 2009 - Proceedings*. 2, pp. 866 - 871. Kuala Lumpur, Malaysia: IEEE.
- Chambers, L. (1999). *Practical Handbook Of Genetic Algorithms - Version 1*. Boca Raton.
- Chan, Y., Hu, A., & Plant, J. (1979). A Kalman Filter Based Tracking Scheme With Input Estimation. *IEEE Transactions On Aerospace and Electronic Systems*, AES-15, 237-244.

- Chang, C., & Athans, M. (1978). State Estimation For Discrete Systems With Switching Parameters. *IEEE Transactions on Aerospace and Electronic Systems, AES-14*(3), 418-425.
- Chao, P., & Shen, C. (2009, March). Sensorless Tilt Compensation for a Three-Axis Optical Pickup Using a Sliding Mode Controller Equipped With a Sliding Mode Observer. *IEEE Transactions on Control Systems Technology, 17*(2), 267 - 282.
- Chen, S., & Moskwa, J. (1997). Application of Nonlinear Sliding-Mode Observers for Cylinder Pressure Reconstruction. *Control Eng. Practice, 5*(8), 1115 - 1121.
- Chen, Z. (2003). Bayesian Filtering: From Kalman Filters To Particles And Beyond.
- Chu, D., Lin, W., & Tan, R. (2006). A Numerical Method For A Generalized Algebraic Riccati Equation,. *SIAM Journal on Control and Optimization, 45*(4).
- Cipra, T., Romera, R., & Rubio, A. (1992). Square Root Kalman Filter With Contaminated Observations. *Division de Economia, Universidad Carlos III de Madrid*.
- Colbia-Vega, A., de Leon-Morales, J., Fridman, L., Salas_Pena, O., & Mata-Jimenez, M. (2008). Robust Excitation Control Design Using Sliding-Mode Technique for Multimachine Power Systems. *Electric Power Systems Research, 1627* - 1634.
- Coumarbatch, C., & Gajic, Z. (2000). Exact Decomposition Of The Algebraic Riccati Equation Of Deterministic Multimodeling Optimal Control Problems. *IEEE Transactions On Automatic Control, 45*(4).
- Daryabor, A., & Momeni, H. (2008). A Sliding Mode Observer Approach to Chaos Synchronization. *International Conference on Control, Automation and Systems 2008* (pp. 1626 - 1629). Korea: Inst. of Elec. and Elec. Eng. Computer Society.
- Drakunov, S., & Utkin, V. (1995, December 13 - 15). Sliding Mode Observers: Tutorial. *Proceeding of the 34th Conference on Decision & Control, 4*, pp. 3379 - 3378. New Orleans: IEEE.
- Drakunov, S., Utkin, V., Zarei, S., & Miller, J. (1996). Sliding Mode Observers for Automotive Applications. *Proceedings of the 1996 IEEE International Conference on Control Applications* (pp. 344 - 346). Dearborn: IEEE.

- Dungate, D., Theobald, R., & Nurse, F. (1999). Higher-Order Kalman Filter To Support Fast Target Tracking In A Multi-Function Radar System. *IEE Colloquium Target Tracking: Algorithms and Applications*, 14/1-14/3.
- Edwards, C. (2004). A Comparison of Sliding Mode and Unknown Input Observers for Fault Reconstruction. *43rd IEEE Conference on Decision and Control* (pp. 5279 - 5284). Atlantis, Paradise Island, Bahamas: IEEE.
- Edwards, C., & Spurgeon, S. (1994). On the Development of Discontinuous Observers. *International Journal of Control*, 59, 1211 - 1229.
- Edwards, C., Hebden, R., & Spurgeon, S. (2005, November). Sliding Mode Observer for Vehicle Mode Detection. *Vehicle System Dynamics*, 43(11), 823 - 843.
- Edwards, C., Spurgeon, S., & Patel, C. T. (2007). *Mathematical Methods for Robust and Nonlinear Control* (Vol. 367/2007). Berlin: Springer Berlin / Heidelberg.
- Edwards, C., Spurgeon, S., & Patton, R. (2000). Sliding Mode Observers for Fault Detections. *Automatica*, 36, 541 - 553.
- Eherndorfer, M. (2007). A Review Of Issues In Ensemble-Based Kalman Filtering. *MeteorologischeZeitschrift*, 16(6), 795-818.
- Emel'yanov, S., Korovin, S., & Levant, A. (1996). High Order Sliding Modes in Control Systems. *Computational Mathematics and Modeling*, 7(3), 294 - 318.
- Evensen, G. (2003). The Ensemble Kalman Filter – Theoretical Formulation And Practical Implementation. *Ocean Dynamics*, 53, 343-367.
- Fathpour, N., & Jonckheere, E. (2002). Hamiltonian Structure Of The Algebraic Riccati Equation And Its Infinitesimal V-Stability. *Proceedings Fifteenth International Symposium on Mathematical Theory of Networks and Systems*, 1779-1792.
- Fei, J., & Batur, C. (2008, December). Adaptive Sliding Mode Control with Sliding Mode Observer for a Microelectromechanical Vibratory Gyroscope. *Proceedings of the Institution of Mechanical Engineers, Part I (Journal of Systems and Control Engineering)*, 222(18), 839 - 849.
- Fitzgerald, R. (1971). Divergence Of the Kalman Filter. *IEEE Transactions on Automatic Control*.

- Floquet, T., Edwards, C., & Spurgeon, S. (2007). On Sliding Mode Observers for Systems with Unknown inputs. *International Journal of Adaptive Control and Signal Processing*, 638 - 656.
- Gadsden, S., & Habibi, S. (2011). Derivation of an Optimal Boundary Layer Width for the Smooth Variable Structure Filter. *American Control Conference*. Submitted September 27, 2010 .
- Givens, W. (1958). Computation Of Plane Unitary Rotations Transforming A General Matrix To Triangular Form. *Journal of the Society for Industrial and Applied Mathematics*, 6, 26-50.
- Golub, G., & Van Loan, C. (1996). *Matrix Computations*. London: The Johns Hopkins Press Ltd.
- Grewal, M., & Andrews, A. (2001). *Kalman Filtering: Theory and Practice Using Matlab* (Second Edition ed.). USA: John Wiley & Sons Inc.
- Habibi I, S. (2008). Parameter Estimation Using a Combined Variable Structure and Kalman Filtering Approach. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, 130(5), 0510041 - 05100414.
- Habibi, S. (2005). Performance Measures of the Variable Structure Filter. *Transactions of the Canadian Society for Mechanical Engineering*, 29(2), 267 - 295.
- Habibi, S. (2006). The Extended Variable Structure Filter. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, 128(2), 341 - 351.
- Habibi, S. (2007). The Smooth Variable Structure Filter. *Proceedings of the IEEE*, 95(5), 1026 - 1059.
- Habibi, S., & Burton, R. (2002). The Variable Structure Filter. *2002 ASME International Mechanical Engineering Congress and Exposition*. 9, pp. 157 - 165. New Orleans, LA, United states: American Society of Mechanical Engineers.
- Habibi, S., & Burton, R. (2003). The Variable structure Filter. *Journal of Dynamic Systems, Measurement, and Control*, 125, 287 - 293.
- Habibi, S., & Burton, R. (2004). Parameter identification for a high performance hydrostatic actuation system using the variable structure filter concept.

Proceedings of the ASME Fluid Power Systems and Technology Division. 11, pp. 93 - 101. Anaheim, CA, United states: American Society of Mechanical Engineers.

- Habibi, S., & Burton, R. (2007). Parameter identification for a high-performance hydrostatic actuation system using the variable structure filter concept. *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME, 129(2)*, 229 - 235.
- Habibi, S., & Goldenberg, A. (2000). Design of a New High-Performance ElectroHydraulic Actuator. *IEEE/ASME Transactions on Mechatronics, 5(2)*, 158 - 164.
- Habibi, S., Burton, R., & Chinniah, Y. (2002). Estimation Using a New Variable Structure Filter. *Proceedings of the American Control Conference, 4*, pp. 2937 - 2942. Anchorage, AK, United states.
- Habibi, S., Burton, R., & Sampson, E. (2006). High Precision Hydrostatic Actuation Systems for Micro and Nanomanipulation of Heavy Loads. *Transactions of the ASME, 128*, 778 - 787.
- Hakiki, K., Mazari, B., Liazid, A., & Djaber, S. (2006). Fault Reconstruction Using Sliding Mode Observers. *American Journal of Applied Sciences, 3(1)*, 1669 - 1674.
- Hammarling, S. (1977). A Survey Of Numerical Aspects Of Plane Rotations. *Middlesex Polytechnic Report Maths*.
- Han, L. (2004). A Fuzzy Kalman Filtering Strategy For State Estimation. *Thesis, Department of Mechanical Engineering, University of Saskatchewan*.
- Hasan, S., & Husain, I. (2009). A Luenberger-Sliding Mode Observer for Online Parameter Estimation and Adaptation in High-Performance Induction Motor Drives. *45(2)*, 772 - 781.
- Hashimoto, H., Utkin, V., Xu, J., Suzuki, H., & Harashima, F. (1990, November 27 - 30). VSS Observer for Linear Time Varying System. *IECON '90. 16th Annual Conference of IEEE Industrial Electronics Society. 1*, pp. 34 - 39. New York: IEEE.

- Haskara, I., Ozguner, U., & Utkin, V. (1996, December 5 - 6). On Variable Structure Observers. *Proceedings, 1996 IEEE International Workshop on Variable Structure Systems. - VSS'96* - (pp. 193-198). Tokyo: IEEE.
- Haykin, S. (2002). *Adaptive Filtering Theory*. Prentice Hall.
- Henrici, P. (1964). *Elements Of Numerical Analysis*. Wiley.
- Heo, S., & Xu, Y. (1999). Constructing Cubature Formulae For Spheres And Balls. *Journal of Computational and Applied Mathematics*, 112(1), 95-119.
- Hernandez, J., & Barbot, J. (1996, September). Sliding Observer-Based Feedback Control for Flexible Joints Manipulator. *Automatica*, 32(9), 1243-1254.
- Ho, T. (2008). Switched IMM-EV algorithms For State Estimation Of Some Jump Markov Systems. *Proceeding of the 17th World Congress, The International Federation of Automatic Control*, (pp. 1396-1401).
- Ho, Y., & Lee, R. (1964). A Bayesian Approach To Problems In Stochastic Estimation And Control. *IEEE Transactions on Automatic Control*, 9(4), 333-339.
- Hongcai, Z., Youmin, Z., Zhibin, H., & Jingchao, L. (1992). An Adaptive U-D Factorized Extended Kalman Filter And Its Application To Flight Path Reconstruction. *Proceedings of the IEEE International Symposium on Industrial Electronics*.
- Hsu, H. (1995). *Signals And Systems*. Schaum's Series, McGRAW-HILL.
- Hyland, J. (2002). An Iterated-Extended Kalman Filter Algorithm For Tracking Surface And Sub-Surface Targets. *Oceans 2002 Conference and Exhibition. Conference Proceedings*, (pp. 1283-1290).
- Jihua, Z., Nanning, Z., Zejian, Y., & Qiang, Z. (2009). A SLAM Algorithm Based On The Central Difference Kalman Filter. *2009 IEEE Intelligent Vehicles Symposium*, Pages: 123-128.
- Jilkovand, V., & Li, X. (2002). On The Generalized Input Estimation. *Information & Security*, 9, 90-96.
- Jing, H., Jing, Q., Changfan, Z., & Cheng, L. (2008, June 25 - 27). Design of Sliding Mode Observer for Uncertain Nonlinear Systems. *Proceedings of the 7th World*

Congress on Intelligent Control and Automation, WCICA'08 (pp. 1720 - 1723).
Chongqing - China: IEEE.

- Johnston, L., & Krishnamurthy, V. (2001). An Improvement To The Interacting Multiple Model (IMM) Algorithm. *IEEE Transactions On Signal Processing*, 49(12), 2909-2923.
- Julier, S. (2003). The Spherical Simplex Unscented Transformation. *Proceedings of the American Control Conference*, (pp. 2430-2434).
- Kailath, T. (1974). A View Of Three Decades Of Linear Filtering Theory. *IEEE Transactions on Information Theory*, IT-20(Issue 2), 146-181.
- Kailath, T. (1980). *Linear Systems*. Prentice-Hall, In., Englewood Cliffs.
- Kalman, R. (1960). A New Approach To Linear Filtering And Prediction Problems. *ASME Journal of Basic Engineering*, 82, 35-45.
- Kalsi, K., Lian, J., Hui, S., & Zak, S. (2009). Sliding-Mode Observers for Uncertain Systems. *2009 American Control Conference* (pp. 1189 - 1194). USA: IEEE.
- Kalsi, K., Lian, J., Hui, S., & Zak, S. (2010). Sliding Mode Observers for Systems with Unknown Inputs: A High-Gain Approach. *Automatica*, 347 - 353.
- Kaminski, P., Bryson, A., & Schmidt, S. (1971). Discrete Square Root Filtering: A Survey Of Current Techniques. *IEEE Transactions on Automatic Control*.
- Kaw, K., & Kalu, E. (2010). *Numerical Methods with Applications* (Second Edition ed.).
- Khan, M., Spurgeon, S., & Levant, A. (2003). Simple Output-Feedback 2-Sliding Controller for Systems of Relative Degree Two. *European Control Conference ECC 2003* (pp. 114 - 118). Cambridge, UK: IEEE.
- Kim, I., Kim, M., & Youn, M. (2006, August). New Maximum Power Point Tracker Using Sliding-Mode Observer for Estimation of Solar Array Current in the Grid-Connected Photovoltaic System. *IEEE Transaction s on Industrial Electronics*, 53(4), 1027 - 1035.
- Kim, J., & Shin, D. (2005). Joint Estimation Of Time Delay And Channel Amplitude By Simplex Unscented Filter Without Assisted Pilot In CDMA Systems. *The 7th*

International Conference on Advanced Communication Technology, (pp. 233-240).

- Kim, S., Han, W., & Kim, S. (2004, June). Design of a New Adaptive Sliding Mode Observer for Sensorless Induction Motor Drive. *Electric Power Systems Research*, 70(1), 16 - 22.
- Kleinert, H. (2004). *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*. World Scientific (Singapore).
- Kreyszig, E. (2000). *Advance Engineering Mathematics*. Wiley, ISBN.
- Kurode, S., Bandyopadhyay, B., & Gandhi , P. (2009). Sliding Mode Observer for Estimation of Slosh States in a Moving Container. *Proceedings of the IEEE International Conference on Industrial Technology*. Churchill, VIC, Australia: Institute of Electrical and Electronics Engineers Inc., 445 Hoes Lane / P.O. Box 1331, Piscataway, NJ 08855-1331, United States.
- Lary, D., & Mussa, H. (2004). Using an Extended Kalman Filter Learning Algorithm For Feed-Forward Neural Networks To Describe Tracer Correlations. *Atmospheric Chemistry And Physics Discussion*, 3653-3667.
- Leu, G., & Baratti, R. (2000). An Extended Kalman Filtering Approach With A Criterion To Set Its Tuning Parameters – Application To a Catalytic Reactor. *Computers & Chemical Engineering*, 23(11-12), 1839-1849.
- Levant, A. (1998). Robust Exact Differentiation via Sliding Mode Technique. *Automatica*, 34(3), 379 - 384.
- Levant, A. (2003). Higher-Order Sliding Modes, Differentiation and Output-Feedback Control. *International Journal of Control*, 76, 924 - 941.
- Li, D., & Li, S. (2007). Real Time Sequential Kalman Filter Sensor Registration Algorithm. *CIE International Conference of Radar Proceedings*.
- Li, R., & Zhao, G. (2009). Position Sensorless Control for PMSM Using Sliding Mode Observer and Phase-Locked Loop. *2009 IEEE 6th International Power Electronics and Motion Control Conference* (pp. 1867 - 1870). Wuhan, China: IEEE.

- Li, X. (2000). Multiple-Model Estimation With Variable Structure – Part II: Model-Set Adaption. *IEEE Transactions on Automatic Control*, 45(11), 2047-2060.
- Li, X., & Bar-Shalom, Y. (1994). A Recursive Multiple Model Approach To Noise Identification. *IEEE Transactions on Aerospace and Electronic Systems*, 30(3), 671-684.
- Li, X., & Bar-Shalom, Y. (1996). Multiple-Model Estimation With Variable Structure. *IEEE Transactions on Automatic Control*, 41(4), Pages: 478-493.
- Li, X., & Zwang, Y. (2000). Multiple-Model Estimation With Variable Structure – Part V: Likely-Model Set Algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 36(2), 448-466.
- Li, X., Jikov, V., & Ru, J. (2005). Multiple-Model Estimation With Variable Structure – Part VI: Expected-Model augmentation. *IEEE Transactions on Aerospace and Electronic Systems*, 41(3), 853-867.
- Li, X., Zwang, Y., & Zwi, X. (1999). Multiple-Model Estimation With Variable Structure – Part IV: Design And Evaluation Of Model-Group Switching Algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 35(1), 242-254.
- Li, X., Zwi, X., & Zwang, Y. (1999). Multiple-Model Estimation With Variable Structure – Part III: Model-Group Switching Algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 35(1), 225-241.
- Madani, T., & Benallegue, A. (2007). Sliding Mode Observer and Backstepping Control for a Quadrotor Unmanned Aerial Vehicle. *Proceedings of the 2007 American Control Conference* (pp. 5887 - 5892). New York City, USA: IEEE.
- Matia, F., Jimenez, A., Rodriguez-Losada, D., & Al-Hadithi, B. M. (2004). A Novel Fuzzy Kalman Filter for Mobile Robots Localization. *Information Processing and Management of Uncertainty in Knowledge-Based Systems*, (pp. 167-174). Perugia (Italia).
- Maybeck, P. (1979). *Stochastic Models, Estimation, And Control. I.*
- McCann, R., & Husain, I. (1997, October 5 - 9). Application of a Sliding Observer for Switched Reluctance Motor Drives. *Proceedings of the 1997 IEEE Industry*

Applications Conference 32nd IAS Annual Meeting. Part 3 (of 3). 1, pp. 525 - 532.
New Orleans: IEEE, Piscataway, NJ, United States.

McCann, R., Islam, M., & Husain, I. (2001). Application of a Sliding-Mode Observer for Position and Speed Estimation in Switched Reluctance Motor Drives. *IEEE Transactions on Industry Applications*, 37(1), 51 - 58.

Mihaylova, L., Lefebvre, T., Stafetti, E., Btuyninckx, H., & De Schutter, J. (2002). Contact Transitions Tracking During Force-Controlled Compliant Motion Using An Interacting Multiple Model Estimator. *Information & Security, An International Journal*, 9, 114-129.

Milosavljevic, C. (1985). General conditions for the existence of a quasisliding mode on the switching hyperplane in discrete variable structure systems. *Automation and Remote Control*, 46(3/1), 307 - 314.

Monsees, G. (2002). Discrete-Time Sliding Mode Control. *Ph.D. Thesis - de Technische Universiteit Delft*.

Moore, R. (1984). A survey of interval methods for differential equations. *Proceedings of the 23rd IEEE Conference on Decision and Control*. 3, pp. 1529 - 1535. USA: IEEE.

Negenborn, R. (2003). Robot Localization And Kalman Filters – On Finding Your Position In A Noisy World.

Nguyen, H., & Walker, E. (n.d.). *A First Course In Fuzzy Logic*. Boca Raton.

Nikoukhah, R., Willsky, A., & Levy, B. (1992). Kalman Filtering And Riccati Equations For Descriptor Systems. *IEEE Transactions on Automatic Control*, 37(9).

Nise, N. (2007). *Control Systems Engineering* (Fifth ed.). USA: John Wiley & Sons, Inc.

Nrgaard, M., Poulsen, N., & Ravn, O. (2000). Ravn, New Developments In State Estimation For Nonlinear Systems. *Automatica*, 36(11), 1627-1638.

Ogle, T., & Blair, W. (2002). Derivation Of A Fixed-Lag, Alpha-Beta Filter For Target Trajectory Smoothing. *Proceedings of the Thirty-Fourth Southeastern Symposium on System Theory*, 26-30.

- Ogle, T., & Blair, W. (2004). Fixed-Lag Alpha-Beta Filter for Target Trajectory Smoothing. *IEEE Transactions on Aerospace and Electronic Systems*, 40(4), 1417-1421.
- Ormsby, C. Generalized Residual Multiple Model Adaptive Estimation Of Parameters and States. *Air Force Institute of Technology, Wright Patterson air Force Base*.
- Ormsby, C., Raquet, J., & Maybeck, P. (2006). A Generalized Residual Multiple Model Adaptive Estimator Of Parameters and States. *Mathematical and Computer Modelling*, 43(9 - 10), 1092 - 1113.
- Painter, J., Kerstetter, D., & Jowers, S. (1990). Reconciling Steady-State Kalman and alpha-Beta Filter Design. *IEEE Transactions on Aerospace and Electronic Systems*, 26(6), 986-991.
- Perruquetti, W., & Barbot, J. (2002). *Sliding Mode Control In Engineering*. New York: Marcel Dekker, Inc.
- Petrova, G. (2004). Cubature Formulae For Spheres, Simplices And Balls. *Journal of Computational and Applied Mathematics*, 162(2), 2004.
- Press, W., Teukolsky, S., Vetterling, W., & Flannery, B. (1992). *Numerical Recipes in C – The Art Of Scientific Computing*. New York: Cambridge University Press.
- Prestin, J., & Rosca, D. (2006). On Some Cubature Formulas On The Sphere. *Journal of Approximation Theory*, 142(1), 1-19.
- Qaiser, S., Bhatti, A., Samar, R., Iqbal, M., & Qadir, J. (2008). Higher Order Sliding Mode Observer Based Estimation of Precursor Concentration in a Research Reactor. *Proceedings - 4th IEEE International Conference on Emerging Technologies 2008, ICET 2008* (pp. 338 - 343). Rawalpindi, Pakistan: Inst. of Elec. and Elec. Eng. Computer Society.
- Quarteroni, A., Sacco, R., & Saleri, F. (2000). *Numerical Mathematics*. New York: Springer-Verlag New York, Inc.
- Rao, S., Buss, M., & Utkin, V. (2008). An Adaptive Sliding Mode Observer for Induction Machines. *2008 American Control Conference (ACC '08)* (pp. 1947 - 1951). Seattle, WA, USA: IEEE.

- Rao, S., Buss, M., & Utkin, V. (2009). Simultaneous State and Parameter Estimation in Induction Motors Using First- and Second- Order Sliding Modes. *IEEE Transactions on Industrial Electronics*, 56(9), 3369 - 3376.
- Reif, K. (1999). Nonlinear State Observation Using H Filtering Riccati Design. *IEEE Transactions on Automatic Control*, 44(1), 203-208.
- Rolink, M., Boukhobza, T., & Sauter, D. (2006, December 19). High Order Sliding Mode Observer for Fault Actuator Estimation and Its Application to the Three Tanks Benchmark. *Scientific Commons - hal-00121029*, 1, 1 - 7.
- Sacher, W., & Bartello, P. (2008). Sampling Errors In Ensemble Kalman Filtering Part I – Theory. *Monthly Weather Review*, 136(8).
- Sadati, N., & Ghaffarkhah, A. (2007). Polyfilter; A New State Estimation Filter For Nonlinear Systems. *International Conference on Control, Automation and Systems*, (pp. 2643-2647).
- Sadhu, S., Mondal, S., Srinivasan, M., & Ghoshal, T. (2006). Sigma Point Kalman Filter For Bearing Only Tracking. *Sigma Point Kalman Filter For Bearing Only Tracking*, 3769-3777.
- Saif, M., Chen, W., & Wu, Q. (2008). High Order Sliding Mode Observers and Differentiators-Application to Fault Diagnosis Problem. *Lecture Notes in Control and Information Sciences*, 375, 321 - 344.
- Sakoy, P., & Oke, P. (2008). A Deterministic Formulation Of The Ensemble Kalman Filter – An Alternative To Ensemble Square Root Filter. *Tellus, Series A (Dynamic Meteorology and Oceanography)*, 60(2), 361-371.
- Sanchez, E., Shibata, T., & Zadeh, L. (1997). *Genetic Algorithms And Fuzzy Logic Systems :Soft Computing Perspectives, Singapore ; . World Scientific Pub.*
- Schenkendorf, R., Kremling, A., & Mangold, M. (2009). Optimal Experimental Design With The Sigma Point Method. *IET Systems Biology*, 3(1), 10-23.
- Shen, X. (1995). Discrete H Filter Design With Application to Speech Enhancement. *1995 International Conference on Acoustics, Speech, and Signal Processing. Conference Proceedings*, 2, pp. 1504-1507.

- Shen, X., & Deng, L. (1997). Game Theory Approach To Discrete H Filter Design. *IEEE Transactions on Signal Processing*, 45(4), 1092-1095.
- Shojaie, K., Ahmadi, K., & Shahri, A. (2007). Effects Of Iteration In Kalman Filter Family For Improvement Of Estimation Accuracy In Simultaneous Localization And Mapping. *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*.
- Simon, D. (2001). From Here To Infinity. *Embedded Systems Programming*, 14(11), 20-32.
- Simon, D. (2003). Kalman Filtering For Fuzzy Discrete Time Dynamic Systems. *Applied Soft Computing Journal*, 3(3), 191-207.
- Simon, D. (2006). *Optimal State Estimation: Kalman, H [infinity] And Nonlinear Approaches*. Wiley-Interscience.
- Simon, D., & El-Sherief, H. (1996). Hybrid Kalman/MiniMax filtering in phase-locked loops. *Control Engineering Practice*, 4(5), 615-623.
- Slotine, J., Hedrick, J., & Misawa, E. (1986). On Sliding Observers for Nonlinear Systems. *Proceedings of the 1986 American Control Conference*. (pp. 1794 - 1800). Seattle, WA, USA: IEEE.
- Slotine, J., Hedrick, J., & Misawa, E. (1987, Sep). On Sliding Observers for Nonlinear Systems. *Transactions of the ASME. Journal of Dynamic Systems, Measurement and Control*, 109(3), 245-252.
- Sorenson, H. (1970). Least-Squares Estimation: From Gauss To Kalman. *IEEE Spectrum*, 7(7), 63-68.
- Spurgeon, S. (2008). Sliding Mode Observers: a Survey. *International Journal of Systems Science*, 39(8), 751-764.
- Spurgeon, S., Edwards, C., & Foster, N. (1996). Robust Model Reference Control Using a Sliding Mode Controller/Observer Scheme with application to a Helicopter Problem. *Proceedings. 1996 IEEE International Workshop on Variable Structure Systems. - VSS'96* - (pp. 36 - 41). Tokyo, Japan: IEEE.

- Spurgeon, S., Edwards, C., & Foster, N. (1996). Robust Model Reference Control Using a Sliding Mode Controller/Observer Scheme With Application to a Helicopter Problem. *Proceedings of the 1996 IEEE International Workshop on Variable Structure Systems, VSS'96* (pp. 36 - 42). Tokyo, Jpn: IEEE.
- Sun, P., & Marko, K. (1998). The Square Root Kalman Filter Training Of Recurrent Neural Networks. *IEEE International Conference on Systems*.
- Tan, C., & Edwards, C. (2003, January). Sliding Mode Observers for Robust Detection and Reconstruction of Actuator and Sensor Faults. *International Journal of Robust and Nonlinear Control*, 13, 443 - 463.
- Tan, C., & Edwards, C. (2004, July 20 - 23). Multiplicative Fault Reconstruction Using Sliding Mode Observers. *2004 5th Asian Control Conference (IEEE Cat. No.04EX904)*. 2, pp. 957 - 962. Melbourne, Vic., Australia: IEEE.
- Tan, C., & Edwards, C. Sliding Mode Observers for Robust Detection and Reconstruction of Actuator and Sensor Faults.
- Tang, X., Zhao, X., & Zhang, X. (2008). The Square-Root Spherical Simplex Unscented Kalman Filter For State And Parameter Estimation. *International Conference on Signal Processing Proceedings*, (pp. 260-263).
- Tang, Z., & Wang, G. (2010). Vibration Fault Detection and Diagnosis in Aircraft Power Plant Using Model-Based Technique. *2010 Chinese Control and Decision Conference (CCDC)* (pp. 2499 - 2502). Xuzhou, China: IEEE.
- Tylavsky, D., & Sohie, G. (1986). Generalization Of The Matrix Inversion Lemma. *Proceedings of the IEEE*, 74(7), 1050-1052.
- Utkin, V. (1977). Variable Structure Systems with Sliding Mode. *IEEE Transactions on Automatic Control*, AC-22(2), 212 - 222.
- Utkin, V. (1983). Variable Structure Systems: Present and Future. *Automation and Remote Control*, 44, 1105 - 1120.
- Utkin, V., Guldner, J., & Shi, J. (1999). *Sliding Mode Control in Electromechanical Systems*. Philadelphia, PA: Taylor & Francis Inc, USA.

- Van Der Heidan, F., Duin, R., de Ridder, D., & Tax, C. (2004). *Classification, Parameter Estimation And State Estimation – An Engineering Approach Using MATLAB*. England: John Wiley & Sons, Ltd.
- Van Der Merwe, R. (2004). Sigma Point Kalman Filters For Probabilistic Inference In Dynamic State-Space Models, Ph.D thesis, OGI School of Science & Engineering. Oregon Health & Science University.
- Van Der Merwe, R., & Wan, E. (2004). Sigma-Point Kalman Filters For Integrated Navigation, . *Proceedings of the Annual Meeting - Institute of Navigation*, 641-654.
- Walcott, B., & Zak, S. (1987, Feb). State Observation of Nonlinear uncertain Dynamical Systems. *IEEE Transactions on Automatic Control*, AC-32(2), 166 - 170.
- Walcott, B., Corless, M., & Zak, S. (1987, June). Comparative Study of State Observation Techniques. *International Journal of Control*, 45(6), 2109-2132.
- Wang, H., & Gregory, R. (1964). On the Reduction Of an Arbitrary Real Square Matrix to Tridiagonal Form. *Mathematic of Computation*, 501-505.
- Wang, L., Wang, L., Liao, C., & Liu, J. (2009). Sigma-Point Kalman Filter Application On Estimating Battery SOC. *5th IEEE Vehicle Power and Propulsion Conference*, 1592-1595.
- Wang, S. (2007, February). Integrated Control and Estimation Based on Sliding Mode Control Applied to Electrohydraulic Actuator. *Thesis*. Saskatoon, Saskatchewan, Canada: University of Saskatchewan.
- Wang, S., Burton, R., & Habibi, S. (2009). Filtering Controller in Sliding Mode: From the Estimation to Control. *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering*, 833 - 846.
- Wang, S., Habibi, S., & Burton, R. (2006). A Smooth Variable Structure Filter for State Estimation. *25th IASTED International Conference on Modelling, Identification, and Control, MIC 2006* (pp. 446 - 453). Lanzarote, Canary Islands, Spain: Acta Press, Building B6, Suite 101, 2509 Dieppe Avenue S.W., Calgary, AB, T3E 7J9, Canada.

- Welch, G., & Bishop, G. (2006). An Introduction To The Kalman Filter. *Department of Computer Science, University of North Carolina.*
- Wenzel, T., Burnham, K., Williams, R., & Blundell, M. (2004). Hybrid Genetic Algorithms/ Extended Kalman Filter Approach For Vehicle State And Parameter Estimation. *Proceedings of the International Conference Control 2004, University of Bath*, (pp. ID-104).
- Yager, R., & Zadeh, L. (1992). *An Introduction to fuzzy logic applications in intelligent systems*. Kluwer Academic, .
- Yan, Z., & Utkin, V. (2002). Sliding Mode Observers for Electric Machines - An Overview. *IECON 02 [Industrial Electronics Society, IEEE 2002 28th Annual Conference of the]*. 3, pp. 1842 - 1847. IEEE.
- Yannone, R. (1988). UD Factorization Applied To Airborne Kalman-Filter-Based Fusion. *IEEE Proceedings of the National Aerospace and Electronics Conference*, 326-333.
- Yo, J., & Kim, Y. (2003). Alpha-Beta-Tracking Index (α-β-γ) Tracking Filter. *Signal Processing*, 83(1), 169-180.
- Zhang, U., Gao, F., & Tian, L. (2008). INS/GPS Integrated Navigation For Wheeled Agricultural Robot Based On Sigma-Point Kalman Filter, 2008 Asia Simulation Conference . *7th International Conference on System Simulation and Scientific Computing*, (pp. 1425-1431).
- Zhang, Y., Zhou, D., & Duan, G. (2006). An Adaptive Iterated Kalman Filter. *IMACS Multiconference on "Computational Engineering in Systems Applications"*, CESA, 1727-1730.
- Zhao, L., Liu, Z., & Chen, H. (2009). Sliding Mode Observer for Vehicle Velocity Estimation With Road Grade and Bank Angles Adaption. *2009 IEEE Intelligent Vehicles Symposium*, (pp. 701 - 706). Xi'an, China.
- Zheng, J., Feng, Y., & Yu, X. (2008). Hybrid Terminal Sliding Mode Observer Design Method for Permanent Magnet Synchronous Motor Control System. *2008 International Workshop on Variable Structure Systems (VSS 2008)*, 106 - 111.

Zhou, Y., McLaughlin, D., & Entekhabi, D. (2006). Assessing The Performance Of The Ensemble Kalman Filter For Land Surface Data Assimilation. *Monthly Weather Review*, 134(8), 2128-2142.

Zupanski, M. (2005). Maximum Likelihood Ensemble Filter – Theoretical Aspects. *Monthly Weather Review*, 133(6), 1710-1726.