

**PREDICTING MALFUNCTION IN QUASI STEADY
STATE ROTATING MACHINES**

Thesis submitted for the degree of
Doctor of Philosophy
at the University of Leicester

by

Christos Kitsos
Department of Engineering
University of Leicester

2007

UMI Number: U234251

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U234251

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

PREDICTING MALFUNCTION IN QUASI STEADY STATE ROTATING MACHINES

C. KITSOS

ABSTRACT

Quasi steady state rotating machines are machines usually designed to run at fixed load and speed. A good example of such a system is a dry vacuum pump. Dry vacuum pumps appeared in the mid-1980s in order to address problems caused by conventional fluid-sealed pumps and revolutionised many industries such as the semiconductor manufacturing industry. They dominate the market today providing enhancements to reliability, cleanliness and running costs. However, their working environment is often harsh, sometimes resulting in catastrophic faults. Continuously monitoring the state of the system and scheduling maintenance as appropriate is thus desirable.

Sliding mode techniques have been widely used in condition monitoring and fault detection schemes in recent years. Their main advantage is a fundamental robustness against certain kinds of parameter variations. They also enable faults and/or values of unmeasurable system parameters to be reconstructed.

The principal aim of this thesis is to apply sliding mode techniques in order to reduce the occurrence of unplanned pump stoppages, by monitoring appropriate subsystems and parameters, for the onset of cooling system failure, bearing failure and exhaust blockage. This is achieved using the concept of the equivalent injection signal that is necessary to maintain a sliding motion. Experimental test results acquired from the dry vacuum pump test-bed illustrate the usefulness of the approach for condition monitoring. Further, the method is cost effective since it requires only low cost temperature transducers and an exhaust pressure sensor that is already part of the typical sensor package for some pump processes.

The thesis concludes with ideas and recommendations regarding possible future work, including the application of fault classification techniques and the development of processes for generating an efficient and implementable code, suitable for the vacuum pumps' embedded control systems.

ACKNOWLEDGEMENTS

Numerous people helped me with this work, in a variety of direct and indirect ways. I shall try to be as accurate as possible, for any omissions or inaccuracies I apologise in advance.

I would like to start by thanking Prof. Barrie Jones who made this possible in the first place by considering me for this PhD program. I would also like to thank my supervisor Prof. Sarah Spurgeon for her guidance, support and encouragement throughout my time at Leicester.

My sincere thanks go to Dr. John Twiddle for helping to achieve this. I do not recall him ever denying to offer help and to answer any of my questions. John you have been a constant source of inspiration to me and it has been a great pleasure working with you.

I would like to acknowledge the kind support of BOC Edwards who provided the experimental test-bed and instrumentation. In particular, I wish to thank Jeremy Watson and Larry Marini for their advice and support.

ACKNOWLEDGEMENTS

I am extremely grateful to Dr. Fernando Schlindwein for the discussions we had about the project and for his friendship. It would have been impossible to keep my sanity and the opportunity he offered me to play football is truly appreciated.

The department of Engineering have provided an excellent environment for my research and many people from the staff contributed with their assistance. Thanks to my fellow students who also helped me with their friendship and lively discussions.

Thanks to my family who have been extremely understanding from the very beginning and supportive in many ways. Last, and definitely most, I want to thank Julie. Words cannot express my thanks for all your love and encouragement.

LIST OF PUBLICATIONS

Kitsos, C.; Spurgeon, S.K.; Twiddle, J.A.; Jones, N.B., *On sliding mode observer techniques for condition monitoring of rotary machines*, The 19th International Congress and Exhibition on Condition Monitoring and Diagnostic Engineering Management, Lulea, Sweden, 12-15 June, 2006.

Kitsos, C.; Spurgeon, S.K.; Twiddle, J.A.; Jones, N.B., *Fault diagnosis for rolling element bearings using sliding mode techniques*, The 6th UKACC Control Conference, Glasgow, Scotland, United Kingdom, 30 August-1 September, 2006.

Twiddle, J.A.; Kitsos, C.; Spurgeon, S.K.; Jones, N.B., *Detection of exhaust blockages in a dry vacuum pump using spectral estimation and sliding mode observer techniques*, The 19th International Congress and Exhibition on Condition Monitoring and Diagnostic Engineering Management, Lulea, Sweden, 12-15 June, 2006.

Twiddle, J.A.; Spurgeon, S.K.; Kitsos, C.; Jones, N.B., *A discrete-time sliding mode observer for estimation of auto-regressive model coefficients with an application in condition monitoring*, 9th International Workshop on Variable Structure Systems, Alghero, Sardinia, Italy, 5-7 June, 2006.

LIST OF PUBLICATIONS

Twiddle, J.A.; Spurgeon, S.K.; Kitsos, C., *An observer-based fault detection procedure incorporating an algebraic framework for identification and sliding modes*, Proceedings of the WCEAM CM 2007, 2nd World Congress on Engineering Asset Management (EAM) and the 4th International Conference on Condition Monitoring, Harrogate, UK, 11-14 June, 2007.

CONTENTS

Abstract	i
Acknowledgements	ii
List of publications	iv
Contents	vi
List of figures	xiii
CHAPTER 1 INTRODUCTION	1
1.1 Research motivation	1
1.2 Related work and research approach	2
1.3 Research objectives	5
1.4 Thesis organisation	6
CHAPTER 2 DRY VACUUM PUMPS AND RELATED FAULT DETECTION	
TECHNIQUES: A REVIEW	8
2.1 History of the vacuum pump	9
2.2 The role of vacuum pumps in industry	10
2.3 Dry vacuum pumps and the need for condition monitoring	10
2.3.1 Dry vacuum pumps	10
2.3.2 Condition monitoring of dry vacuum pumps	12

2.4	An overview of condition monitoring and fault diagnosis techniques	13
2.4.1	Fault detection and diagnosis terminology	14
2.4.2	Model-free techniques	15
2.4.3	Model-based techniques	16
2.4.4	Sliding mode techniques in fault diagnosis	20
2.5	Dry vacuum pump test-bed	21
2.5.1	Test-bed description	21
2.5.2	The roots/claw dry vacuum pump mechanism	22
2.6	Dry vacuum pump subsystems	27
2.6.1	The cooling system	27
2.6.2	The rolling element ball bearings	29
2.6.3	The exhaust system	29
2.7	The dry vacuum pump condition monitoring and fault detection system	30
2.7.1	Selection of suitable monitoring signals and sensors	30
2.7.2	Calibration of equipment	32
2.7.3	Data acquisition system	32
2.8	Conclusions	33
CHAPTER 3 SLIDING MODE TECHNIQUES		34
3.1	History of sliding mode techniques	36
3.2	The basic notion of VSS	38
3.2.1	A variable structure system example	39
3.3	Sliding mode theory	41
3.3.1	Switching surface design	41
3.3.2	Control law design	44

3.3.2.1	The single input case	45
3.3.2.2	The multivariable case	47
3.3.3	An illustrative example	50
3.4	The equivalent control method	52
3.5	Properties of the sliding motion	55
3.5.1	Reduction of order	55
3.5.2	Robustness property	55
3.6	The chattering problem	58
3.7	Sliding mode observers	62
3.7.1	Luenberger observer	62
3.7.2	Observer design via sliding mode techniques	64
3.8	Higher order sliding modes	66
3.9	Discrete time sliding mode	69
3.10	Conclusions	71
 CHAPTER 4 CASE STUDY 1: THE COOLING SYSTEM		73
4.1	Problem formulation	75
4.2	Heat-transfer models for the dry vacuum pump cooling system	76
4.3	Fault scenarios for a defective cooling system	81
4.4	Sliding mode observer development	82
4.5	Experimental set-up	84
4.6	Results	85
4.7	Conclusions	91

CHAPTER 5 AN ALGEBRAIC FRAMEWORK FOR IDENTIFICATION AND		
SLIDING MODES		93
5.1	An algebraic framework for identification	95
5.2	Mathematical development of the parameter estimation algorithm	98
5.3	Performance of the proposed estimation framework	103
5.3.1	Simulation of the model with time-invariant parameters	103
5.3.2	Simulation of the model with time-varying parameters	104
5.3.3	Estimation of the pump model parameters from captured pump data	105
5.3.4	Overview of parameter estimation results	106
5.4	The performance of the sliding mode observer scheme	107
5.5	Conclusions	109
CHAPTER 6 CASE STUDY 2: THE ROLLING ELEMENT		
BALL BEARINGS		110
6.1	Problem formulation	114
6.2	Heat-transfer models for the rolling element ball bearings	115
6.2.1	Low vacuum bearing temperature model	117
6.2.2	High vacuum bearing temperature model	119
6.3	Fault scenario for defective bearings	120
6.4	Sliding mode observer for parameter estimation and fault prediction	121
6.4.1	A first order sliding mode algorithm	122
6.4.2	A second order sliding mode algorithm	123
6.5	Experimental set-up	125
6.6	Results	126

6.6.1	Fault detection	126
6.6.2	A comparison of first and second order sliding mode observer performance	129
6.7	Conclusions	131
CHAPTER 7 CASE STUDY 3: THE EXHAUST SYSTEM		133
7.1	Problem formulation	136
7.2	Sliding mode observer design	138
7.3	Experimental set-up	142
7.3.1	Hardware set-up	142
7.3.2	Signal pre-processing	143
7.3.3	Implementation of the observer	145
7.4	Results	146
7.5	Conclusions	152
CHAPTER 8 CONCLUDING REMARKS AND FUTURE WORK		154
8.1	Concluding remarks	154
8.2	Recommendations for future work	157
References		160
Appendices		178
Appendix A Details of dry vacuum pumps		179
A.1	Introduction	179
A.2	Basic vacuum theory	179

A.3	A review of dry vacuum pump designs	181
A.3.1	Piston design	181
A.3.2	Screw compressor design	182
A.3.3	Scroll design	182
A.3.4	Roots lobe design	183
A.4	Conclusion	184
Appendix B Instrumentation and data acquisition		185
B.1	Introduction	185
B.2	iGX dry vacuum pump instrumentation list	185
B.3	Data acquisition and hardware calibration tests	188
B.3.1	Inverter current signal output calibration	188
B.3.1.1	Objective	188
B.3.1.2	Method	188
B.3.1.3	Results	188
B.3.1.4	Conclusion	188
B.3.2	Simultaneous recording of temperature and inverter signals from the iGX dry vacuum pump	189
B.3.2.1	Objective	189
B.3.2.2	Problem formulation	189
B.3.2.3	Recommendations	192
B.3.2.4	Implementation and testing	193
B.3.2.5	Conclusion	195
B.3.3	Calibration of Lucas 4AM air flow meter	195
B.3.3.1	Objective	195

B.3.3.2	Problem formulation	196
B.3.3.3	Type 55D41/42 calibration equipment information	197
B.3.3.4	Test procedure	198
B.3.3.5	Signals	198
B.3.3.6	Results	198
B.3.3.7	Discussion	200
B.3.3.8	Conclusion	201
B.3.4	Cooling water flow meter	201
B.3.4.1	Objective	201
B.3.4.2	Problem formulation	202
B.3.4.3	Results	202
Appendix C Nominal model validation results		204
C.1	Introduction	204
C.2	Cooling system heat transfer model parameter estimation	204
C.3	Bearings heat transfer model parameter estimation	209
Appendix D Sliding mode observer stability		215
D.1	Introduction	215
D.2	System stability	215
Appendix E Heat transfer model developments		218
E.1	Introduction	218
E.2	The iGX dry vacuum pump cooling circuit design	218
E.3	Preliminary results	220

LIST OF FIGURES

CHAPTER 2

Figure 2-1: General scheme of model-based fault detection	17
Figure 2-2: The iGX dry vacuum pump test-bed	22
Figure 2-3: The roots mechanism	23
Figure 2-4: The roots mechanism during its operation	24
Figure 2-5: The claw mechanism	25
Figure 2-6: A single compression cycle for the claw type concept. (a) to (f) show the intake sequence and (g) to (l) the discharge sequence	26
Figure 2-7: Graph of the attainable pressure ratio as a function of the outlet pressure	27

CHAPTER 3

Figure 3-1: Asymptotically stable VSS consisting of two unstable structures	40
Figure 3-2: Phase portrait of the second order system	52
Figure 3-3: Discontinuous and equivalent control	54
Figure 3-4: Comparison of the equivalent control with the average control	54
Figure 3-5: Phase portrait with disturbance case	57
Figure 3-6: Comparison of the discontinuous and smooth control	60
Figure 3-7: Signum, saturation and signum-like functions	61

Figure 3-8: Smooth control signal using the signum-like function 62

CHAPTER 4

Figure 4-1: Block diagram of vacuum pump cooling system heat balance 77

Figure 4-2: Measured and estimated body temperature T_{B1} , model error
and the corresponding equivalent injection signal 86

Figure 4-3: Measured and estimated data for T_{B1} , T_{B2} and T_o under
a fault condition (\dot{m}_c) 87

Figure 4-4: Error in temperature between the measured and estimated
value for T_{B1} , T_{B2} and T_o observers under a fault condition (\dot{m}_c) 88

Figure 4-5: Injection signals ν_{B1} , ν_{B2} and ν_o for coolant flow failure 89

Figure 4-6: Component parameter estimate $\Delta \dot{m}_c$ for observers T_{B1} and T_o 90

Figure 4-7: Component parameter estimate Δh_c for observers T_{B2} and T_o 90

Figure 4-8: Component parameter estimate Δk for observers T_{B1} and T_{B2} 91

CHAPTER 5

Figure 5-1: Estimates for time-invariant parameters a and b 103

Figure 5-2: Parameter estimates in the event of a simulated 10%
deviation in coolant flow rate 104

Figure 5-3: Parameter estimated in the event of a simulated temperature
sensitive parameter b . The parameter undergoes an exponential
rise of 15% 105

Figure 5-4: Parameter estimates from captured data 106

Figure 5-5: Recorder coolant flow rate, \dot{m}_c and observer fault signal, $\Delta \dot{m}_c$ 108

LIST OF FIGURES

Figure 5-6: Recorded coolant flow rate, \dot{m}_c and observer fault signal, $\Delta \dot{m}_c$ in the event of a 40% reduction in coolant flow rate	109
CHAPTER 6	
Figure 6-1: Dry vacuum pump cooling system diagram	116
Figure 6-2: Error in temperature between LV and HV end models and corresponding measured experimental data under fault-free conditions	127
Figure 6-3: Measured and estimated data for the T_{LV} and T_{HV} under fault conditions	128
Figure 6-4: Component parameter estimate Δk_i under fault conditions	129
Figure 6-5: Observer injection signal for the classical (CSM) and higher order sliding mode (HOSM) schemes	130
Figure 6-6: Observer error for CSM and HOSM at $T_s = 0.1s$	131
CHAPTER 7	
Figure 7-1: Block diagram of a PLL	144
Figure 7-2: RMS state error as a function of sampling frequency	146
Figure 7-3: Yule-Walker AR model power spectrum	147
Figure 7-4: Ensemble-averaged exhaust pressure amplitude vs. rotor angle, showing the variation in amplitudes with increasing volume of fill	148
Figure 7-5: AR estimates of PR vs. pump revolutions for a series of silencer blockages. Note that spectra are calculated every 20 revolutions	148
Figure 7-6: The PR increases with the volume of silencer obstruction. These average results are the means of the four data series	

LIST OF FIGURES

in Figure 7-5	149
Figure 7-7: RMS of the observer injection signal increases with the volume of silencer obstruction	150
Figure 7-8: Comparison of PR values calculated using FFT, AR and sliding mode observer methods	151
Figure 7-9: Spectra calculated from sliding mode observer generated AR model coefficients for clear silencer	151
APPENDICES	
Figure A-1: Schematic arrangement of a four-stage piston pump	181
Figure A-2: The screw type mechanism	182
Figure A-3: Scroll type design	183
Figure A-4: Schematic of a three stage roots dry pump	183
Figure B-1: Dry vacuum pump and Instrumentation. See Table B-1 for key to instruments. Note that the exhaust pressure sensor and inlet mass flow meter were not fitted at the time of the photograph	187
Figure B-2: Inverter signal output calibration data	189
Figure B-3: Analogue output connections	190
Figure B-4: Effect of inverter output connections on thermocouples readings	191
Figure B-5: LM35 with R-C damper	193
Figure B-6: Effect of inverter output connection on IC readings	194
Figure B-7: Comparison on IC and thermocouple pump body sensor readings	195
Figure B-8: Schematic showing type 55D41/42 calibration equipment	197
Figure B-9: Plot of air flow meter output voltage vs. air flow rate	200
Figure B-10: Plot of the inlet pressure vs. air flow rate	201

LIST OF FIGURES

Figure B-11: Block diagram of cooling water pulse converter	202
Figure B-12: Graph of the cooling water pulse	203
Figure B-13: Graph of the cooling water flow rate	203
Figure C-1: Graph of regression data and best-fit line for parameter c	206
Figure C-2: Plot of measured and estimated T_B open loop data with coolant water on	207
Figure C-3: Plot of measured and estimated T_B	208
Figure C-4: Plot of the estimated parameter a_1 from the sliding mode observer	210
Figure C-5: Relationship between a_1 and ΔT for open/close valve	211
Figure C-6: Comparison between observer and function data for open valve	212
Figure C-7: Comparison between observer and function data for close valve	213
Figure C-8: Measured and estimated T_{LV} for low vacuum bearing temperature model	213
Figure C-9: Measured and estimated T_{HV} for high vacuum bearing temperature model	214
Figure E-1: Dry vacuum pump cooling circuit diagram, design No. 3	219
Figure E-2: Vacuum pump and coolant temperature at stage 1	220
Figure E-3: Vacuum pump and coolant temperature at stage 2	221

CHAPTER 1

INTRODUCTION

1.1 RESEARCH MOTIVATION

The rapid progress of technology has a profound effect on the design of machines, which have become increasingly complex. Further, consistent demands from customers for high performance, safety, robustness and reliability have all led in an intensely competitive marketplace. This is relevant in the vacuum pump industry, for example, where these demands have promoted research and development in vacuum pump technology.

Vacuum pumps have relied on the existence of some liquid (water, oil) to provide a seal between the high and low pressure regions in the pump mechanism for more than 300 years [Hablanian 1988]. This use of liquid, however, often involves significant maintenance work, increased running costs and causes air or water pollution [Zakrzewski 1988]. Dry vacuum pumps emerged in the mid-1980s in order to address problems caused by the conventional liquid-sealed pumps [Troup 1991]. These systems contain no sealing or lubricating fluids between the rotors and stator, thus eliminating the risk of contamination of the process chamber by fluid back-streaming from the

pump. Other advantages include increased safety, reliability and reduced cost of operation [Duval 1989]. Nevertheless, the harsh nature of many processes where dry vacuum pumps are used creates a need for continuous condition monitoring [Troup 1989], [Lessard 2000]. For instance, in chemical vapour deposition semiconductor processes, the high value of the raw materials means that judgment of whether a replacement pump is necessary before processing a new batch is of critical importance because untimely pump failure will result in total loss of the batch.

Obviously, condition monitoring and fault diagnosis has enormous benefits for any vacuum system in terms not only of safety, but also with the overall reliability and availability. As indicated in [Duval 1987], vacuum pump monitoring equipment must deliver early warnings in order to avoid failures and unscheduled down time of the system. It is also reported that monitoring of vacuum pumps on a regular basis is probably the best solution to increase the up time of the system. This move towards condition monitoring of vacuum pumps in order to enable preventive maintenance has also been identified by Troup and Dennis [Troup 1991].

1.2 RELATED WORK AND RESEARCH APPROACH

Classical condition monitoring and fault diagnosis techniques are largely based on simple limit checks or on simple signal-based methods such as spectrum analysis. The great advantage of these methods is their simplicity, but the current rising requirements of modern on-board condition monitoring systems limits their application. A potential way to satisfy these requirements is to employ model-based fault detection techniques. These techniques are based on mathematical process models or signal models, where many characteristic process quantities, parameters and variables are taken into account

in order to construct the model. Such models represent the dynamic behaviour of a system and theoretically any detected changes from normal features (nominal values) can infer a fault. Thus, model-based techniques offer the advantage of an in-depth fault diagnosis, since the source of faults can be better detected.

Several survey papers have been written on condition monitoring and fault diagnosis. Frank [Frank 1990] reviewed the most common techniques of model-based residual generation using parameter identification and state estimation methods. It is concluded that the quality of the available mathematical model plays an important role in model-based methods and the support of a model-free diagnostic method may often be unavoidable. A comprehensive survey of supervision, fault detection and fault diagnosis is presented by Isermann [Isermann 1997]. The survey begins by introducing an overall scheme of a knowledge-based method and then concentrates on model-based methods that are based on parameter estimation, parity equations or state observers. It also considers the task of determining fault details such as size, location and time of detection. If a-priori knowledge between faults and symptoms is known then it is reported that inference methods like fault trees can be used. In contrast, without a-priori knowledge of faults and symptoms, classification methods like fuzzy clusters are utilised. Isermann [Isermann 2005] briefly introduced model-based methods and considered their application for an actuator, a passenger car and a combustion engine. It is reported that in all cases the model-based methods perform well, but some effort is needed to obtain mostly non-linear dynamic models.

Although model-based fault detection methods have numerous benefits, the need for a mathematical model is a clear constraint. Often effects such as disturbances,

nonlinearities and modelling errors obscure the effect of faults and they result in false or missed alarms. Therefore, to ensure that no false alarms occur, the designed condition monitoring and fault detection scheme must be robust. The sliding mode technique [Edwards 1998] is an inherently robust methodology that has been recognised as a candidate methodology for condition monitoring and fault diagnosis. Apart from its robustness property, the technique also offers order reduction, disturbance rejection and fairly simple implementation. The fault detection method employs observers for state and parameter estimation. Specifically, faults can be detected by analysis of the so-called equivalent injection signal and appropriate manipulation of this signal can effectively either reconstruct the fault signal or monitor parameters whose variation may be used to infer the health of a system. Publications in the literature on sliding mode observer-based fault detection schemes can be found in [Bhatti 1999], [Edwards 2000] and [Goh 2002].

This thesis focuses on sliding mode techniques and addresses the development of a condition monitoring and fault detection scheme for a quasi steady-state rotating machine. A quasi steady-state machine refers to machines designed to mostly run with fixed load and rotating speed. Such a machine is a dry vacuum pump provided by BOC-Edwards. Clearly, a vacuum pump with a reliable and cost-effective diagnostic capability will be very desirable for potential customers and thus of commercial advantage to BOC-Edwards.

1.3 RESEARCH OBJECTIVES

The main objective of this thesis is to develop a practical on-board condition monitoring and fault detection system for a dry vacuum pump. The successful operation of this pump relies on the performance of various components, e.g., cooling system, bearings, exhaust system, and thus a detailed understanding of the behaviour of these elements is required. Further, the harsh nature of many processes where dry vacuum pumps are used means that although key factors that affect pump reliability are internal, instrumentation located within the flow paths of the pump and connecting pipe-work has a short life expectancy. Thus, the developed condition monitoring scheme must be capable of accurately and reliably inferring faults from a set of instruments that can be located externally.

Another objective of this thesis is to study sliding mode techniques and to assess their application in real-time condition monitoring and fault detection. Within this context, also to explore the recently developed higher order sliding mode concept and to compare the performance of a second order algorithm with a first order one. Finally, the design of the proposed condition monitoring scheme requires a nominal mathematical model. If this fault detection scheme is to be routinely implemented in industry, automated methods are needed to establish nominal parameter values. In this thesis, an algorithm for estimation of linear time-invariant model parameters is developed, implemented and tested.

1.4 THESIS ORGANISATION

This chapter has introduced the problem of interest and the motivation behind conducting this research. It also presents the approach of the thesis and has established the broad objectives. The rest of the thesis is organised as follows.

Chapter 2 provides a brief history of vacuum pumps and discusses their role in the industry. Various methods of condition monitoring and fault detection are reviewed and their strengths and limitations are considered. It describes in detail the dry vacuum pump test bed and identifies possible monitoring signals and instrumentation for the diagnostic scheme.

Chapter 3 introduces sliding mode theory. It begins with a brief history of the sliding mode methodology and provides the main design steps of a sliding mode algorithm. The properties of the technique are presented and the problem of removing undesirable discontinuous signals is discussed. The design of a sliding mode observer is described and the possibility of using such observers for fault detection is considered. It also considers an alternative way of removing undesirable discontinuous signals with the use of higher order sliding modes. This new class of sliding modes not only preserves the main features of classical sliding modes, but provides an improved tracking (sliding) accuracy under sliding motion when compared to classical sliding modes.

Chapter 4 presents the results from the first case study of this thesis. This includes the development of heat transfer models for the water cooling system and the design of a sliding mode observer for parameter estimation and fault detection. The experimental set-up is fully described and a discussion of the experimental results is given.

Chapter 5 investigates the possibility of using an algebraic framework for parameter identification. In particular, it evaluates an estimation algorithm as a means of rapidly estimating parameters of first order dynamic models used in the sliding mode observer based condition monitoring scheme that is considered in this thesis.

Chapters 6 and 7 contain the results obtained from the remaining case studies. The first case study involves the use of a sliding mode observer for the detection of overheated bearings. It also examines the use of a second order sliding mode algorithm and demonstrates that the algorithm is capable of smoothing the discontinuous nature of signals. The second case study details the development of a condition monitoring and fault detection scheme for the exhaust system. This involves the use of spectral estimation and sliding mode observer techniques.

Finally, chapter 8 presents the conclusions of this thesis and provides recommendations regarding potential future work.

CHAPTER 2

DRY VACUUM PUMPS AND RELATED FAULT DETECTION TECHNIQUES: A REVIEW

The main objectives of this chapter are to explain the working principles of vacuum pump technology and to review commonly used fault detection techniques.

The chapter is organised as follows: A brief history of vacuum pumps and their role in the industry nowadays are outlined in sections 2.1 and 2.2 respectively. Section 2.3 explains why it is likely in the future that most of the vacuum pumps used in association with high vacuum technology will be dry. It also clearly states the requirement for condition monitoring and fault diagnosis. A review of relevant techniques for condition monitoring and fault detection is presented in section 2.4. Section 2.5 provides a detailed description of the dry vacuum pump test-bed used in this thesis. The following section describes the subsystems of the dry vacuum pump for which monitoring schemes will be developed to establish the onset of faults. Section 2.7 covers the instrumentation requirements and describes the data acquisition system.

2.1 HISTORY OF THE VACUUM PUMP

The history of vacuum pumps begins in the 17th century when Otto von Guericke, a mayor of a town called Magdeburg in Germany, invented the first vacuum pump and used it to investigate the properties of vacuum. Guericke's pump comprised a piston and a cylinder and was designed to evacuate air out of vessels. In one experiment, Guericke demonstrated the force of atmospheric pressure by joining two large hollow metal hemispheres and pumping the air out of the enclosure. After the air was removed from the sphere two teams of eight horses attached to each hemisphere were unable to pull them apart, demonstrating how good was Guericke's pump [Harris 2001].

In the next two centuries improvements and modifications were made to Guericke's piston pump; however the basic design remained the same. During this period the attainable pressure by the various pump designs was decreased only slightly and vacuum pumps were mainly used for laboratory work or as a source of entertainment, since vacuum was a novel subject.

Vacuum technology made rapid advances when the incandescent lamp industry emerged at the turn of the 20th century. New pumping methods and designs such as rotary mechanical pumps, diffusion pumps and molecular pumps were invented capable of achieving higher vacuum and pumping speeds [Hablanian 1984].

In the last 50 years the use of vacuum pumps has increased and they are now commonplace across a wide range of industries. Hablanian [Hablanian 2003] reviewed the progress made over this time for different designs of vacuum pumps and reported

the need for a cleaner vacuum environment which has led to the development of liquid-free vacuum systems.

2.2 THE ROLE OF VACUUM PUMPS IN INDUSTRY

The need and benefits of performing various processes under a vacuum has led to the firm establishment of vacuum pump systems in industry. For example, as part of the circuit-building process in the semiconductor industry, chemical vapor deposition (CVD) is used to deposit dielectric layers of SiO_2 (i.e. silicon dioxide) onto the wafers. During the process, a controlled vacuum environment is needed to prevent the formation of particles.

Reasons for using a vacuum in a process or a physical measurement are given by Harris [Harris 2001]. Among others, it is reported that vacuum is utilised in order to achieve a pressure difference, to decrease energy transfer and to produce clean surfaces.

Furthermore, various industrial applications that require vacuum processing are presented. This indicates the dependence and importance that vacuum technology has on modern industry.

2.3 DRY VACUUM PUMPS AND THE NEED FOR CONDITION MONITORING

2.3.1 DRY VACUUM PUMPS

Conventional vacuum pump designs utilize some form of liquid (e.g. oil, mercury) to seal the tight passages between the inlet and discharge regions. The presence of liquid, apart from the sealing function, also provides lubrication and cooling for the system. Nonetheless, clear drawbacks arise from the usage of such vacuum pumps in many

manufacturing applications. For instance, the cost of operation due to the expensive fluids and filters can be very high. Further, the contamination of the pumping chamber by back-streaming of the fluid and the disposal of the dangerous contaminated working fluid are some other disadvantages, as mentioned by Duval [Duval 1989].

Dry vacuum pumps appeared in the mid-1980s in order to address problems caused by fluid-sealed vacuum pump technologies. The term 'dry pump' is used to describe a positive displacement vacuum pump (i.e. a machine where fluid is drawn into a finite space bounded by mechanical parts and is then sealed in it until eventually it is allowed to flow out) which is able to discharge the pumped gases to the atmosphere and is free of lubricants or sealing fluids in the pumping module.

It is evident that through the absence of fluid, dry pumping is presenting significant advantages and a number of publications are available on the subject. Bachmann and Kuhn [Bachmann 1990] compared a dry pump with an oil-sealed rotary vane pump for a specific application. It is concluded that the dry pump requires considerably less careful handling and longer periods of maintenance-free operation are achieved.

Zakrzewski et al [Zakrzewski 1988] reviewed developments in the areas of reliability and cleanliness for oil-free mechanical pumps. Vacuum cleanliness is evaluated by using a mass spectrometer and it is observed that dry pumps are at least 1000 times cleaner than oil-sealed pumps.

Today, a variety of different dry vacuum pump mechanisms like screw compressors, orbital scroll compressors, piston pumps, roots and claw types are available (see appendix A for details). Each one of the above configurations has different features and

advantages. Thus, the choice of a particular mechanism should be made according to the requirements of a certain application [Hablanian 1990].

2.3.2 CONDITION MONITORING OF DRY VACUUM PUMPS

Vacuum pumps have a significant role in industry (see section 2.2) and are regarded as critical components in many applications. Lessard [Lessard 2000] introduced some of the most demanding processes in the semiconductor industry like etching and ion implantation. It is concluded that the combination of reactive chemistry and solid deposition in the pump demands for extraordinary measures to be taken to prevent failures of the system. Such failures will cause the loss of a valuable batch of wafers as well as the loss of a considerable amount of time while the machine is repaired.

BOC-Edwards studied the failures of various components such as seals, shafts and gears. Thus, a large archive of useful information has been produced, which helped the company to improve the reliability of their pump models. Nevertheless, customers demand higher levels of reliability, safety, cleanliness and serviceability. These demands have created a need for the development of a condition monitoring and fault diagnosis system.

Any diagnostic scheme designed to provide early detection of faults under all conditions has clear advantages. Enhancement of safety is obtained since a catastrophic failure may cause risk for human life. Major breakdowns are avoided, the overall reliability is improved and the probability that the system is operational when required (i.e. the availability) is increased.

The trend towards monitoring of key parameters to enable preventive maintenance is identified by Troup and Dennis [Troup 1991]. Duval [Duval 1987] reported the use of intelligent pumping systems, where sensors are controlled by microprocessors or microcomputers in order to deliver status information. Another area of interest is networked condition monitoring of vacuum pumps. Connecting a group of pumps together through an online network, effects real-time monitoring.

2.4 AN OVERVIEW OF CONDITION MONITORING AND FAULT DIAGNOSIS TECHNIQUES

From the very beginning of industry, a natural concern about the condition of machines existed. Originally the detection of malfunctions was carried out by a human operator and was based on biological senses like listening for unusual sounds, looking for changes in shape and touching to feel heat or vibration. Later, with the introduction of measuring devices, common parameters (e.g. temperature) were monitored. A serious problem with this method was that these devices were also proved prone to malfunction, which resulted in false alarms. Today, the use of computers permits the development of more sophisticated and powerful methods that can detect faults in the process earlier, and can locate them better, even in complex systems [Gertler 1998].

Detailed overviews of fault detection and diagnosis methods are available in the following books [Basseville 1985], [Viswanadham 1987], [Patton 1989], [Gertler 1998], [Chen 1999] [Isermann 2005]. Moreover, a number of surveys have been published on the subject [Gertler 1988], [Frank 1990], [Isermann 1984 1997 2005]. Gertler [Gertler 1988] pointed out that the approaches to the problem of fault detection and diagnosis can be classified into two strategies: model-based and model-free methods. It is also

reported that the isolability, sensitivity and robustness are major quality issues in the selection of an algorithm. Frank [Frank 1990] proposed the combination of both strategies, since it allows the evaluation of all available information and knowledge of a system. An attempt to review some fault detection methods is made by Isermann [Isermann 1984]. In this article emphasis is given on methods for monitoring unmeasurable quantities such as process parameters and process state variables. Isermann and Balle [Isermann 1997] evaluated a number of publications in the area of model-based fault detection and diagnosis. Based on their evaluation, parameter estimation and observer-based techniques are the most frequently applied methods for fault detection. Further, the growing trend towards fuzzy logic-based [Miguel 1996], [Twiddle 2001] and neural network-based [Jones 2000], [Parikh 2001] methods is also stated. Finally, an introduction to model-based fault detection followed by three applications for an actuator, a passenger car and a combustion engine is given by Isermann [Isermann 2005].

2.4.1 FAULT DETECTION AND DIAGNOSIS TERMINOLOGY

Before briefly describing the different approaches to fault detection and diagnosis it is useful to introduce some fundamental concepts. This is important because in this field the terminology is not consistent. Isermann and Balle [Isermann 1997] suggested several definitions based on the discussions of a technical committee. Some of the related terminology to this research is provided below.

- *Fault*: An unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable/usual/standard condition.

- *Failure*: A permanent interruption of a system's ability to perform a required function under specified operating conditions.
- *Fault detection*: Determination of the faults present in a system and the time of detection.
- *Monitoring*: A continuous real time task of determining the conditions of a physical system, by recording information, recognising and indicating anomalies in the behaviour.
- *Residual*: A fault indicator, based on a deviation between measurements and model equation based computations.

2.4.2 MODEL-FREE TECHNIQUES

It is usual when only poor or imprecise mathematical models are available to apply model-free methods. As the name suggests, these techniques do not rely on a mathematical model of the plant. However, the availability of a large amount of historical data is rather crucial. These methods are also known as black box methods and they range from limit checking and special sensors to physical redundancy and spectrum analysis.

- (1) *Limit checking*. A very simple and straightforward technique, limit checking [Pouliezos 1994] compares plant measurements with the use of a computer to preset limits. By exceeding a threshold value, a fault is indicated. Usually, two levels of limits exist: the first level gives only a warning whereas the second triggers an emergency action. A drawback of the limit checking method is that the threshold value should be set quite conservatively in order to take into account the normal input variations. Further, isolating a single component fault

may be difficult, since a fault may propagate to many plant variables, thus setting off a confusing series of alarms.

- (2) *Special sensors*. Special sensors are already part of the typical sensor package for many machines. Basically, they perform limit checking in hardware and may measure special variables (e.g. pressure, vibration, sound).
- (3) *Physical redundancy*. This technique utilises multiple sensors to measure the same physical quantity. The measurements are then compared and any significant discrepancy between them indicates a fault. Clearly, physical redundancy involves extra hardware cost and extra weight which is a major concern in many applications.
- (4) *Spectrum analysis*. Spectrum analysis for planned and predicted maintenance is an established technique [Thanagasundram 2007]. In this approach, frequency domain or spectrum analysis is performed on monitored variables such as plant vibration and pressure [Mathew 1984], [Tandon 1999]. Under normal conditions these variables will generate a unique signature. However, any deviation from this signature is a powerful sign of an abnormal condition.

2.4.3 MODEL-BASED TECHNIQUES

A variety of different model-based techniques applicable to the problem of fault detection and diagnosis exist. Most of these techniques are based on the use of analytical rather than physical redundancy. The essence of analytical redundancy in fault detection is to compare the actual behaviour of the monitored plant against the behaviour predicted from a mathematical model. The resulting differences among them are called the residuals. Nominally, the generated residuals are zero; in the presence of faults and noise, they deviate from zero. Therefore, the residuals need to be analysed to

reach a diagnostic decision. The architecture of model-based fault detection techniques is depicted in Figure 2-1.

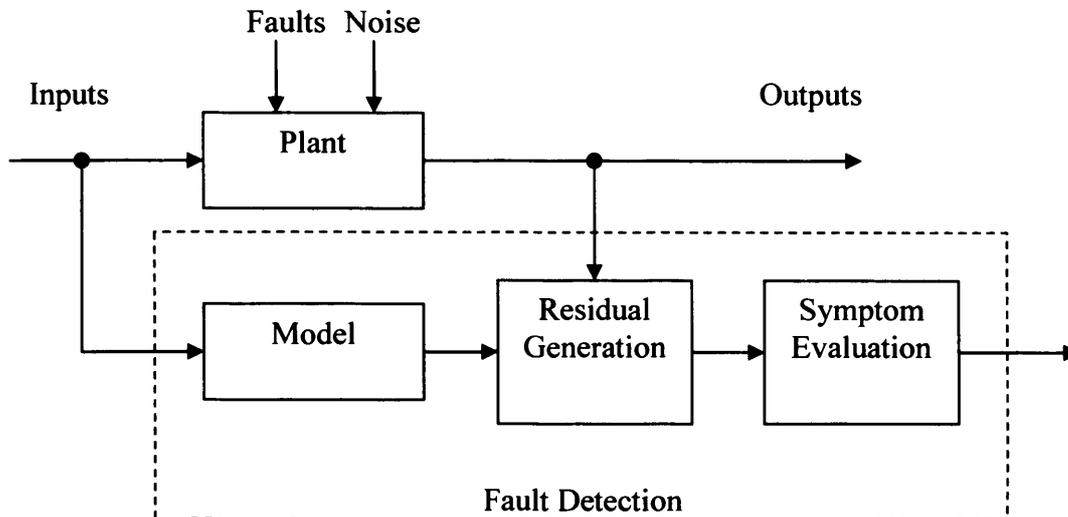


Figure 2-1: General scheme of model-based fault detection

When selecting a model-based fault detection system, several aspects should be considered [Isermann 1984]. The relationship between the performance of a method and the appearance of faults in the system of interest is pertinent. For example, there are methods designed to detect abrupt faults which are not suitable for the detection of incipient faults. Moreover, the ability to distinguish between different and multiple faults is an important but a difficult requirement. Lastly, practical aspects such as the relative complexity and the computational effort of an algorithm should be considered. In general, the selection of a technique depends on the plant and the experience of the designer.

The model-based schemes are basically signal processing techniques using state estimation, parity relations, parameter estimation and so on. Below, the main concepts are briefly outlined.

- (1) *Kalman filter*. In 1960 R.E. Kalman presented a novel approach [Kalman 1960] for a recursive solution of the discrete-data linear filtering problem. Essentially, the Kalman filter is a set of mathematical equations that provides a recursive means to estimate the state of a process, in a way that minimises the mean of the squared error. Mehra and Peshon [Mehra 1971] and Willsky [Willsky 1976] first looked into the possibility of applying the Kalman filter to the fault detection problem. The residual in Kalman filters is called the innovation and is defined as the difference between the measured output and the estimate acquired by the filter. When the model is accurate and the noise is white with zero mean, then in fault free systems, the innovation is also white with zero mean. A nonzero innovation therefore will denote the presence of faults.
- (2) *Diagnostic observers*. In certain systems the states are not always available or sometimes are difficult to measure. A common way of overcoming this difficulty is to use an observer [Luenberger 1971]. According to Gertler [Gertler 1998] the original idea of utilising observers for the fault detection problem came from [Beard 1971] and [Jones 1973]. In this approach, either state observers or output observers can be applied. Classical state observers are utilised if the faults can be modelled as state variable changes, whereas output observers are employed whenever the reconstruction of state variables is not required or it is prohibitively expensive to measure them and only output information is available [Edwards 1998]. In both cases the generated residuals can be examined for the likelihood of faults. Other contributors in the area of fault detection by diagnostic observers include [Ding 1994], [Frank 1994], [Yang 1995] and [Chen 1996].

- (3) *Parity relations.* The early contributions to the parity relations technique have been made by Chow and Willsky [Chow 1984]. The basic idea of this methodology is to take a fixed model and to run it in parallel to the process, thereby forming an output error. The residuals, also called parity equations, are suitable for the detection of faults. Patton and Chen [Patton 1994] provided a review of parity space approaches to fault diagnosis for aerospace systems. A nonlinear parity equation residual generation scheme for the problem of diagnosing sensor and actuator faults in an internal combustion engine is presented by Krishnaswami et al [Krishnaswami 1994]. Finally, Muenchhof [Muenchhof 2006] demonstrated the use of computationally inexpensive parity equations for the detection of faults in a hydraulic linear servo axis.
- (4) *Parameter estimation.* Parameter estimation methods were originally developed for system identification. However, since on dynamic systems many faults are reflected in the model parameters, the technique can also be applied to the problem of fault detection [Isermann 1984], [Patton 1989], [Zhang 1994] and [Muenchhof 2004]. Primarily, a reference mathematical model has to be developed in a fault free situation. For linear processes this reference model can be differential equations where the parameters have some physical meaning. In many practical applications though the model parameters are partly known or not known at all. In such cases the parameters can be computed with parameter estimation methods (e.g. least squares estimates) by measuring the input and the output signals. Fault detection is accomplished by means of observing any deviations from the reference model.

2.4.4 SLIDING MODE TECHNIQUES IN FAULT DIAGNOSIS

The analytical methods described up to this point refer to linear models, whilst most practical systems are nonlinear in nature. For this reason, linearization of the model dynamics is often carried out. Nonetheless, linear approximations may prove to be poor and hence the effectiveness of the fault diagnosis system is greatly reduced.

Another strategy is to extend the residual generation techniques to nonlinear systems. Alcorta-Garcia and Frank [Alcorta-Garcia 1997] concluded that a complete solution to the fault detection problem for nonlinear models is still open. This is because of the difficulties encountered when estimating the state or the measurement vector of the nonlinear system, even if the nonlinearities are known.

The sliding mode technique is an inherently robust nonlinear methodology which has been used for condition monitoring and fault diagnosis in recent years. The principles of the sliding mode theory are introduced in detail in Chapter 3 of this thesis.

Jones et al [Jones 2000] presented a number of engineering and biomedical applications that have used sliding mode observers in order to recreate fault signals. A non-linear sliding mode observer is used for the detection of possible faults in a diesel engine coolant system by Goh et al [Goh 2002]. Furthermore, a particular sliding mode observer for fault detection and isolation is produced by Edwards et al [Edwards 2000]; the novelty of the approach is that the observer attempts to reconstruct the fault signals rather than detect the presence of a fault through a residual signal. The proposed observer is designed to maintain the sliding motion, even in the presence of faults, which are detected by analysing the equivalent output injection signal.

An alternative sliding mode observer scheme is adopted by Hermans and Zarrop [Hermans 1996], where in the event of a fault, the sliding motion breaks. The idea is demonstrated on a non-linear thermodynamic plant under feedback control and mismatches between the real plant and a simplified linear model are monitored.

2.5 DRY VACUUM PUMP TEST-BED

2.5.1 TEST-BED DESCRIPTION

A dry vacuum pump provided by BOC-Edwards is used as the experimental platform for this research. The iGX (the BOC-Edwards designated name for their manufactured pump) dry vacuum pump is driven by a three-phase induction motor and has a rotational speed of approximately 6300 rpm. It can achieve a peak pumping speed of $100 \text{ m}^3 \times \text{h}^{-1}$ and is capable of producing an ultimate pressure of 10^{-2} mbar. The system is enclosed in a soundproof enclosure, since low noise levels are one of the most important environmental requirements this day.

Additionally, the platform is supplied with a general purpose inverter in order to control or vary several parameters such as torque, rotor speed and power. For example, the pumping load near atmospheric pressure is high and the inverter enables the pump to start up under controlled conditions without surpassing safe limits (rotor speed) that can damage the system. Besides this control action, the inverter allows a number of operating parameters to be monitored or recorded such as the current. Finally, a 20 litres stainless steel vessel is fitted on the top of the inlet of the vacuum pump. A control valve (diaphragm type) is attached to the vessel that permits the user to control the vacuum by varying the pump inlet pressure. The iGX dry vacuum pump is illustrated in Figure 2-2.



Figure 2-2: The iGX dry vacuum pump test-bed

2.5.2 THE ROOTS/CLAW DRY VACUUM PUMP MECHANISM

Modern fluid-sealed mechanical pumps can produce high compression ratios of the order 10^5 just with a single stage. In contrast, a dry design in order to match this performance needs to have several sets or stages of rotors in series [Hablanian 1988].

The iGX dry vacuum pump consists of a cast iron stator that forms the outer housing of the pumping module. At its heart, the iGX employs the following design concept; one roots type stage and four stages of claw type impellers. The five rotor pairs are mounted on common shafts and are held in phase without any contact by a pair of timing gears. Moreover, at either end of the machine a set of ceramic bearings is fitted. Both the timing gears and bearings are isolated from the rotors and require lubrication. However, these lubricated parts of the pump mechanism can affect the system since they represent a potential source of contamination for the dry areas. Another problem is lubricant degradation caused by aggressive gases back-streaming through the pump. A

solution to these problems is given by installing high performance dynamic seals on the rotating shafts [Davis 2000].

A diagram of the roots principle is shown in Figure 2-3. The ‘figure-of-eight’ rotors are synchronised and rotate in opposite directions inside the stator. Due to the shape of the rotors and their synchronisation, they remain tangential to each other and tangential to the stator. The clearances between the rotors and the casing are very small (approx. 0.1 to 0.5 mm) in order to ensure minimal back-leakage of the pumped gas.

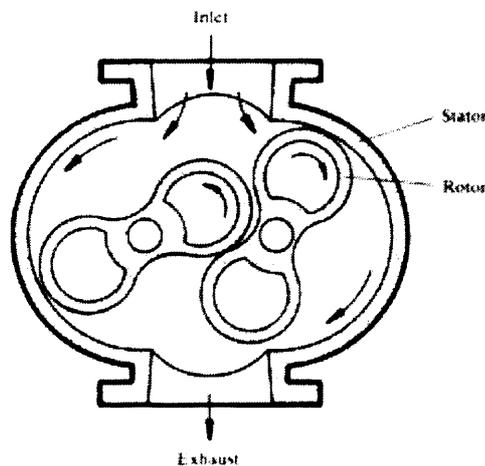


Figure 2-3: The roots mechanism (adapted from Harris [Harris 2001])

During the operation of the vacuum pump, a volume of gas is trapped between each rotor and the wall of the stator. This trapped gas is then transferred without internal compression to the second stage of the pump through a channel (see Figure 2-4).

Pumps based on the roots principle can obtain very large displacement speeds and are proved to be very effective when delivering against low-pressure differentials [Wycliffe 1987]. A drawback of the design is that it is not suited to deliver atmospheric pressure and complex designs such as gas recirculation need to be added in order to do so.

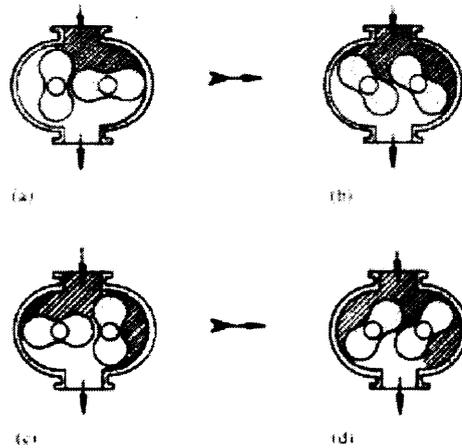


Figure 2-4: The roots mechanism during its operation (adapted from Harris [Harris 2001])

The claw mechanism is comprised of cylindrical rotors that have a big depression followed by a protruding claw (see Figure 2-5). The claw type rotors are mounted in reverse orientation so that during operation each claw enters the depression in the mating rotor. Also, the inlet and exhaust ports of the claw mechanism are arranged horizontally rather than vertically as in the roots design. This allows direct communication between the claw stages and is minimising the gas path through the pump. BOC-Edwards improved the performance of the iGX vacuum pump with their patented feature of inverted claws on alternate stages. Reversing, for example, the fourth claw stage of the pump, then the outlet of the third claw stage is directly in line with the inlet of the fourth stage. This feature has many advantages such as good vacuum performance and provides an easy passage for the particles through the pump.

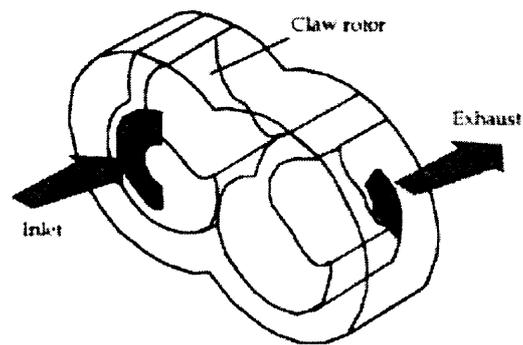


Figure 2-5: The claw mechanism (adapted from Harris [Harris 2001])

Unlike the roots mechanism, in the claw type stages compression of the pumped gas is taking place. A single cycle of the claw mechanism can be described as follows (see Figure 2-6): The inlet and exhaust ports are closed at the beginning and the gas is expanded and compressed in the swept volume. The expanded gas is separated from the compressed gas by the rotors 'sealing' with the stator. While the expansion and contraction of the gas continues the inlet port is first opened, followed by the exhaust port. During that period gas enters the inlet port and leaves the exhaust port. After both ports are closed, expansion and contraction of the gas continues until the mixing of the gas in the 'carry over volume' takes place. Finally, after the mixing of the gas, the cycle repeats itself.

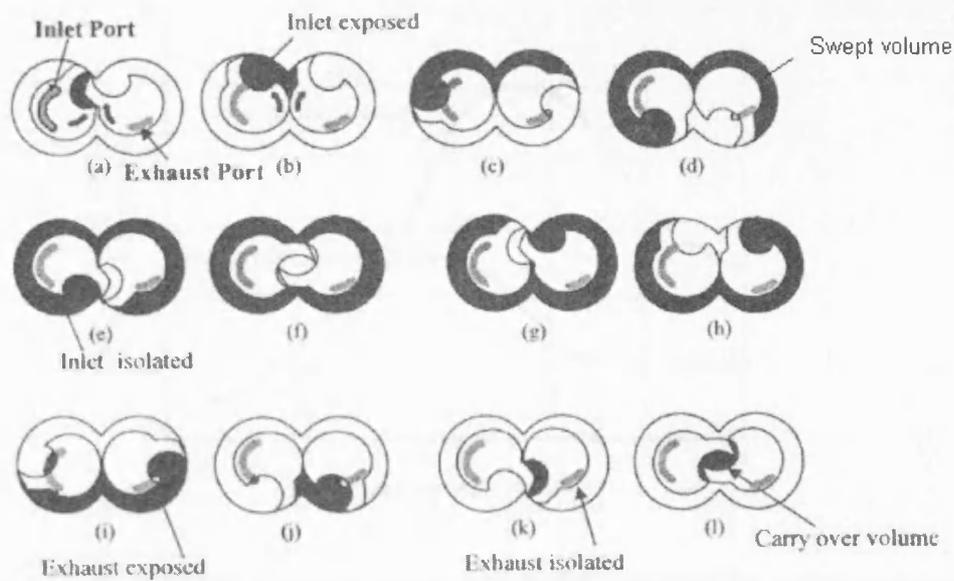


Figure 2-6: A single compression cycle for the claw type concept. (a) to (f) show the intake sequence and (g) to (l) the discharge sequence (adapted from Harris [Harris 2001])

The rationale for combining these two different rotor designs in the iGX dry vacuum pump is better understood by plotting the curve of the attainable pressure ratio as a function of the outlet pressure (see Figure 2-7). The graph shows that the claw type mechanism is more efficient at high pressures and can deliver to atmospheric pressure. Conversely, the roots principle is more efficient in the low pressure region and provides a better compression ratio. Therefore, by combining the two designs optimum performance is obtained.

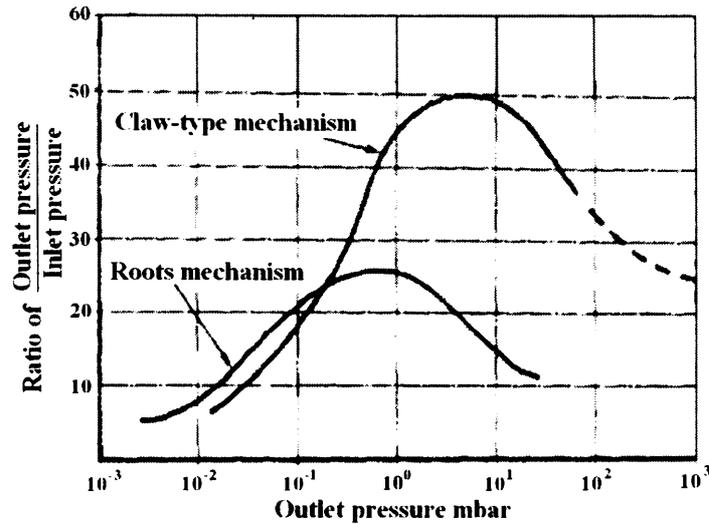


Figure 2-7: Graph of the attainable pressure ratio as a function of the outlet pressure
(adapted from Wycliffe [Wycliffe 1987])

2.6 DRY VACUUM PUMP SUBSYSTEMS

Having described the dry vacuum pump test-bed and its mechanism, this section describes the subsystems selected for condition monitoring and fault diagnosis. Specifically, their function within the vacuum pump and their characteristics are introduced.

2.6.1 THE COOLING SYSTEM

Dry vacuum pumps operate at high rotational speeds in order to obtain high pumping speeds and although the rotating parts do not touch either each other or the stator walls, heat is developing in the bearings and seals. Heat is also generated during the compression of the pumped gas and is transmitted to the rotors and the stator housing. While the stator can lose some of the generated heat through convection and radiation to the atmosphere, the rotors mostly lose heat through conduction from the shaft to the bearings (a small amount of heat is dissipated through convection and radiation to the

stator). Therefore, this necessitates the usage of some form of cooling to remove the heat from the pump and prevent excessive thermal growth.

The iGX dry vacuum pump dissipates the generated heat with the employment of a water cooling system that surrounds the working volume. This water 'jacket' is a stainless steel tube, approximately 7 meters in length, attached externally to the pump. The water pipe first wraps around the induction motor to dissipate the heat developed from it and then encloses the five rotor stages of the vacuum pump. Furthermore, the cooling system consists of a 'butterfly' valve and a flow meter to control and measure the flow rate respectively.

Another characteristic of the cooling system is the ability to regulate the temperature of the vacuum pump. Troup and Dennis [Troup 1991] reported that for every type of process the best operating temperature must be found. For condensable vapours, it is important that high temperatures throughout the pump are present to minimize deposition. However, with increasing temperatures, most common pumped solvents are volatile. Therefore, temperature adjustment is required to ensure that the gas temperature is above the condensation point and below the auto-ignition point. This is achieved by installing temperature controlled valves on the cooling circuit.

During the course of this research, the cooling circuit has evolved and three designs have been deployed with different routes for the cooling water. The initial design was the simplest one; the coolant was distributed around the vacuum pump without the use of any valves. On the second design a single thermostatic valve was installed, which maintained the pump stator temperature within some set limits. The final cooling

circuit, which was the most complicated, had two temperature controlled valves located at either end of the vacuum pump. Diagrams of the three cooling circuits are presented in later chapters.

2.6.2 THE ROLLING ELEMENT BALL BEARINGS

A very common and important component used in rotating machinery is the bearings. Ball bearings, roller bearings and magnetic bearings are only some of the available types of bearings used today.

The iGX dry vacuum pump has a set of rolling element ball bearings at each side to support its shafts. The bearings are comprised of an inner race, an outer race and have ceramic balls as their rolling element. Because of their design, ball bearings can run at high speeds and support both radial and axial loads. The use of ceramic balls in place of conventional steel balls has also improved their performance. Ceramic balls are lighter, run faster, exhibit lower vibration levels and have lower operating temperature.

2.6.3 THE EXHAUST SYSTEM

The function of the exhaust system of the dry vacuum pump is to efficiently direct the pumped gases away to the atmosphere. It contains flexible pipe-work, which is connected to a fume cupboard for safety reasons. Moreover, the exhaust system contains two absorptive silencers, a 'cigar-box' shaped silencer situated under the pump body and a cylindrical one that is beside the pump. The purpose of these silencers is to minimise the noise levels.

2.7 THE DRY VACUUM PUMP CONDITION MONITORING AND FAULT DETECTION SYSTEM

The preceding sections have described in detail the dry vacuum pump test-bed and identified the need for a fault detection and diagnosis system. In this section the requirements of this system are considered.

2.7.1 SELECTION OF SUITABLE MONITORING SIGNALS AND SENSORS

The success of any fault detection and diagnosis system depends strongly on the selection of appropriate monitoring signals and transducers. The first priority is to review the fault symptoms of the project related subsystems (see section 2.6 for details) in order to decide which signals it is most suitable to monitor. A thorough fault/symptom data sheet is supplied by BOC-Edwards. The data sheet includes useful information on various pump failure mechanisms and outlines their effect on the system. A summarised version of the data sheet for the project related subsystems is presented in Table 2-1.

From this table it can be observed that a typical fault symptom for both the cooling system and the ball bearings is vibration. Vibration analysis is an effective technique for condition monitoring of ball bearings and other types of machinery [Mathew 1984], [Tandon 1999]. Analysis of vibration measurements can be performed in the time or frequency domain and techniques such as trending, limit checking, or various forms of pattern or feature recognition are used to infer faults.

Table 2-1. Typical fault symptoms for related subsystems

Subsystems	Typical Fault Symptoms
Cooling system	<p>High temperatures: motor, stator and bearings.</p> <p>Low coolant flow.</p> <p>Bearing vibration.</p> <p>Oil level (coolant leaks into system).</p> <p>High cooling water temperature differential.</p>
Ball bearings	<p>Vibration at defect frequencies.</p> <p>Noise.</p> <p>High bearing temperature.</p> <p>High motor current.</p>
Exhaust system	<p>High exhaust pressure.</p> <p>High motor current.</p>

Although the effectiveness and value of vibration measurement methods is acknowledged, the complexity of developing implementable schemes can produce problems. In the analysis of vibration, high bandwidth signal processing means that on-line diagnostics may not be commercially viable. Further, frequently high order anti-aliasing filters need to be incorporated in the scheme to remove unwanted signal components.

Another typical fault symptom for the dry vacuum pump is high temperatures.

Temperature is a key system performance parameter that can give advanced warning of problems and is relatively straightforward to monitor. Besides, it is noticed that the nature of the signal dynamics is slow and low sample rate models are adequate.

Beyond the choice of appropriate monitoring signals, the selection of suitable transducers is also important. The role of transducers is to provide the diagnostic algorithms with sufficient information of appropriate quality. The principal factors for their selection are cost, functionality and robustness. For instance, piezoelectric accelerometers used in vibration analysis are well known for their accuracy, low level noise and robustness. The transducers are powered by expensive and bulky condition amplifiers which creates a problem when multiple readings are required. Alternatively, thermocouples are cheap and can be incorporated within the normal footprint of the pumping system. A list of the transducers fitted on the dry vacuum pump and used for monitoring is included in appendix B.

2.7.2 CALIBRATION OF EQUIPEMENT

Once the monitoring signals and the sensors are decided, the next step is to install, test and calibrate the equipment. This is a nontrivial procedure because not all variables are directly measurable. The inverter current signal is one such variable. Details of all calibration tests and signal conversions are presented in appendix B.

2.7.3 DATA ACQUISITION SYSTEM

The data acquisition system used in this project consists of a dSpace™ DS1103 PPC controller board and a personal computer. The board provides a great selection of interfaces, including sixteen 16-bit multiplexed ADC channels, eight 16-bit DAC channels and 32-bit I/O digital channels. The transducers are linked to the board via a connector panel for easy access. Moreover, part of the dSpace™ package is the graphical user interface software called ControlDesk. The software permits the making of virtual instrumentation, displaying graphically variables and provides the necessary

connection to MatlabTM/SimulinkTM. MatlabTM is the platform used for analysing data and performing numeric calculations whereas, SimulinkTM is the tool used for mathematical modelling and offline simulation.

2.8 CONCLUSIONS

The subject of this study, a dry vacuum pump, has been introduced in this chapter. The significance of vacuum systems in industry and the advantages of dry pumps have been clearly stated. The function of the cooling system, ball bearings and exhaust system within the vacuum pump has also been described. Furthermore, a number of fault detection techniques have been reviewed and their benefits and drawbacks have been considered.

The application of the sliding mode technique to the problem of fault detection and diagnosis has then been introduced. This technique is the basis of this thesis and its theory and properties will be presented in detail in the following chapter.

CHAPTER 3

SLIDING MODE TECHNIQUES

Chapter 2 of the thesis has described a number of techniques for the problem of condition monitoring and fault diagnosis. One possible approach to this problem and the basis of this research is the so-called sliding mode methodology. Hence this chapter is aimed at introducing the fundamental principles of sliding mode techniques prior to the development of the fault diagnostic scheme for the iGX dry vacuum pump.

The basic philosophy behind the sliding mode method is to force the system states and then constrain them on a predefined surface in the state space, by the use of a switched control law that changes value according to a prescribed switching rule. The surface is referred to as the sliding surface and the system on the surface is said to be in the sliding mode. Once the system is in the sliding mode two main advantages are obtained. Firstly, the closed loop system becomes totally insensitive to a particular class of uncertainty and secondly a reduction in dynamic order is introduced.

Apart from these properties, the technique is also entirely nonlinear in nature. As a result, the formation of the theory of sliding modes is not restricted to the domain of

systems which are expressed through linear models, but rather offers a generic design framework applicable to wide classes of nonlinear systems.

The dry vacuum pump under consideration is a complex system with inherent nonlinear dynamics, which according to the above renders sliding mode techniques an appropriate candidate for the development of the fault diagnostic scheme. Moreover, the vacuum pump operates in an uncertain industrial environment and thus robustness to modelling mismatches and uncertainty in parameters is a key requirement.

Another important advantage of sliding mode techniques to the particular problem of condition monitoring and fault detection is that it allows signals or parameters that may be expensive or sometimes difficult to measure to be estimated. This is achieved by examining the associated equivalent injection signal. The equivalent injection signal represents the average behaviour of the discontinuous signal that is required to maintain the sliding motion and can be readily obtained by appropriate filtering of the discontinuous signal. At this point, it is reasonable to question whether this equivalent injection property can predict the right system parameters under faulty conditions since sliding modes are mainly characterised by their robustness. However, as will be demonstrated in later chapters the equivalent injection method ensures that the effects of faults are not hidden by the robust scheme and parameter estimation and hence fault detection can be achieved.

The chapter is organised as follows: A brief history of sliding mode techniques is outlined in section 3.1. Section 3.2 introduces the concept of variable structure systems and provides an illustrative example. The principal design steps of a sliding mode

algorithm are described in section 3.3. Section 3.4 covers the equivalent control method and 3.5 highlights the properties of sliding motion. The following section identifies ‘chattering’ as the main problem of sliding modes and proposes various solutions for its alleviation. Section 3.7 provides a generic design framework for the synthesis of a sliding mode observer. This design framework sets the stage for the use of sliding mode observers in the field of condition monitoring and fault detection as described in later chapters. A promising way of suppressing the problem of chattering with the use of higher order sliding modes is introduced in section 3.8. Finally, section 3.9 discusses the critical issue of digital implementation of sliding modes.

3.1 HISTORY OF SLIDING MODE TECHNIQUES

Sliding mode techniques are a special class of systems known as variable structure systems (VSS). The notion of VSS first appeared in the Russian literature in the late 1950’s but only became familiar to the rest of the world in the mid-1970’s when a survey paper was issued in English [Utkin 1977]. Since that time, the principle has generated a considerable interest in the research community. The theory of sliding mode techniques is well covered in the books by Utkin [Utkin 1992] and Edwards and Spurgeon [Edwards 1998]. Moreover, thorough reviews on the topic are provided in a number of journal papers [DeCarlo 1988], [Hung 1993], [Young 1999].

Sliding mode techniques have subsequently been developed for robust control of uncertain nonlinear systems such as electric drives, automotive, chemical and biomechanical systems [Utkin 1993], [Bhatti 1999], [Herrmann 2003], [Lim 2003]. Many of these designs assume full state information, which in practice is frequently not available. The estimation of unmeasurable states is commonly called observation and

the dynamical system that estimates the states is called an observer [Luenberger 1971]. Utkin [Utkin 1992] first considered the possibility of designing an observer with discontinuous parameters. Other developments in the area of sliding mode observers are discussed in [Slotine 1987], [Edwards 1994]. Recently the method has been successfully applied to the problem of fault detection and diagnosis [Alessandri 1999], [Edwards 2000].

A survey of the relevant literature demonstrates that the concept of sliding mode is formulated in the continuous time domain. A key issue on the design of continuous time sliding modes is that infinitely fast sampling for total invariance to uncertainty is required. However, this total invariance property is weakened when direct digital implementation is applied. In the 1980's, the discrete time sliding mode (DSM) theory has been developed to address the problems associated with digital implementation. Unlike its continuous time counterpart it has not been studied to the same extent. Research work in this area is presented in [Furuta 1990], [Gao 1995], [Koshkouei 2000 2002].

Besides the difficulties related with digital implementation, a possible drawback of the classical, first order sliding mode approach is the 'chattering effect' - high frequency switching of the input (control) signals. Higher order sliding modes (HOSMs) are a generalisation of classical sliding modes (CSMs) and have been created to tackle the problem of chattering. They also provide improved tracking (sliding) accuracy under sliding motion when compared to CSMs. In HOSMs, the switching function, together with some derivatives of the switching function, are required to become zero. The injection signal is thus designed to act on a derivative of the switching function which is

greater than one. A number of such HOSM algorithms are described in the literature [Levant 1993 2001 2003], [Bartolini 1998 2000], [Fridman 2002].

3.2 THE BASIC NOTION OF VSS

Conceptually the simplest way to explain the variable structure method is to compare it with the linear state regulator design [Utkin 1977]. Suppose that a linear state space model is defined by

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3.1)$$

where $x(t)$ denotes the state space vector and $u(t)$ is the control signal. A and B are matrices of appropriate dimensions. In this case the control signal $u(t)$ has the following fixed structure

$$u(t) = -Kx(t) \quad (3.2)$$

where K is the state feedback gain matrix and its constant parameters are selected according to various design techniques like pole placement or quadratic minimization [Ogata 1997]. Conversely, VSS are built around a set of continuous subsystems that with an appropriate switching logic can change over in the duration of the control. As a result, the nature of the control action is discontinuous. This characteristic of the variable structure technique provides an effective and robust means of controlling nonlinear systems. To illustrate the usefulness of the approach an elementary example is described in the next section.

3.2.1 A VARIABLE STRUCTURE SYSTEM EXAMPLE

Consider the following second order system

$$\ddot{x}(t) = u(t) \quad (3.3)$$

and assume that the feedback control law is given by

$$u(t) = -\bar{E}x(t) \quad (3.4)$$

where \bar{E} consists of two structures defined by $\bar{E} = \alpha_1^2$ and $\bar{E} = \alpha_2^2$, also $\alpha_1^2 > 1 > \alpha_2^2$.

A useful way to analyse the behaviour of the dynamical system is by phase plane profiles, i.e. plots of one dependent variable against another. Therefore, a plot of \dot{x} versus x gives families of ellipses for both structures (see Figure 3-1 (a) and (b)). It can be observed that neither structure is asymptotically stable, since the variables \dot{x} and x do not move towards the origin. However, asymptotic stability of the system can be obtained if the following control law is applied

$$\bar{E} = \begin{cases} \alpha_1^2, & \text{if } x\dot{x} > 0 \\ \alpha_2^2, & \text{if } x\dot{x} < 0 \end{cases} \quad (3.5)$$

As is seen in Figure 3-3 (c) the system trajectory is changing and will ultimately spiral in towards the origin. The resulting asymptotically stable system is thus attained by combining parts of trajectories of different marginally stable structures. This VSS design principle though proved difficult to generalize to systems of arbitrary order via constructive design frameworks and was only considered for second order systems

[Utkin 1984]. An alternative design approach is to create trajectories which are not inherent in any of the structures. These trajectories, the so-called sliding modes, have a particular role in the development of a general VSS theory. The next section introduces the concept of sliding modes and describes their behaviour with an illustrative example.

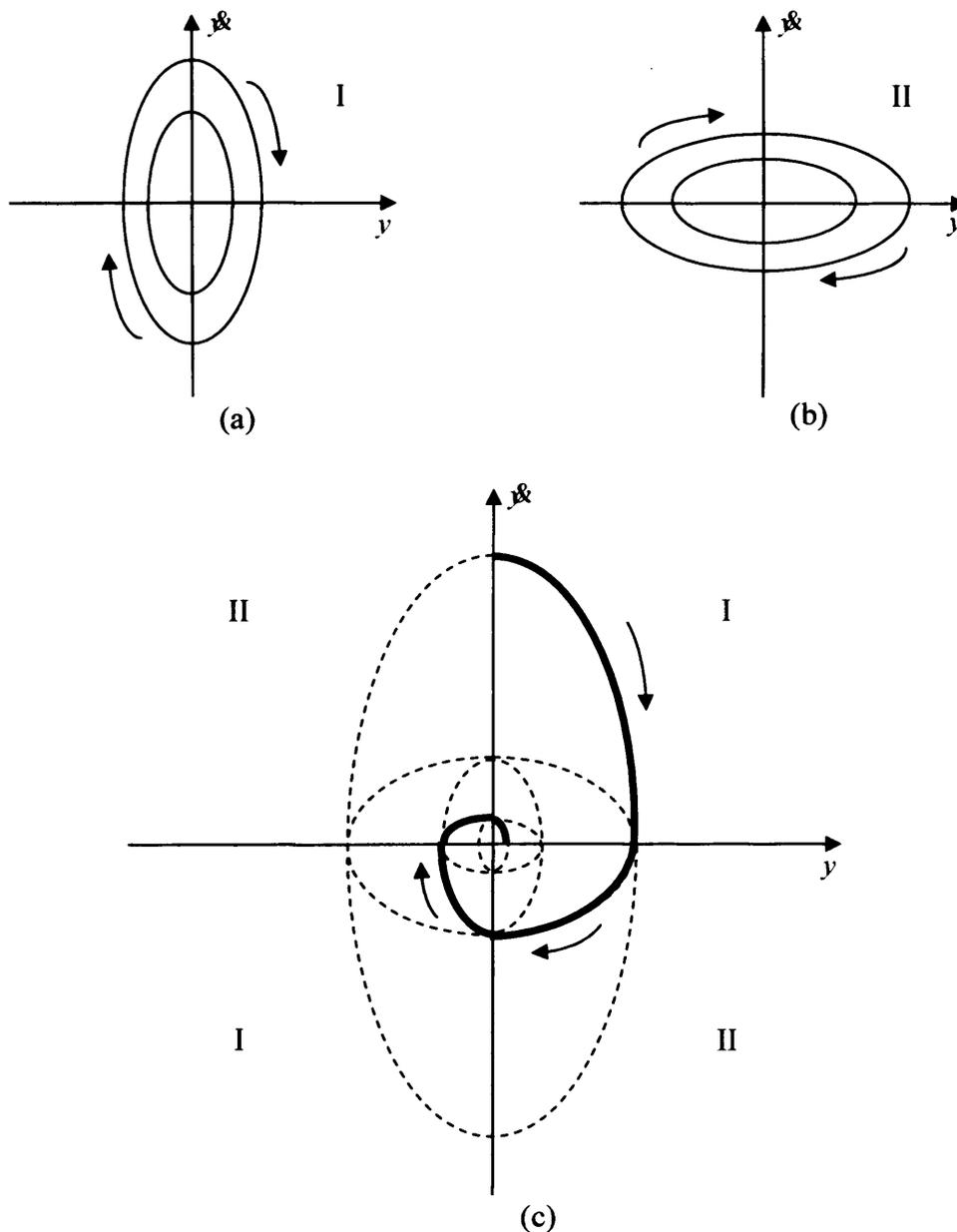


Figure 3-1: Asymptotically stable VSS consisting of two unstable structures

3.3 SLIDING MODE THEORY

In the previous section a specific structure of a variable system has been described. The design process comprised of a set of feedback control laws and a decision rule which allows the selection at any time of the right law according to the current state of the system. This section is concerned with the synthesis of a sliding mode algorithm. Essentially, this synthesis is characterised by two major tasks. The first task is to construct a sliding or switching surface in order that the system constrained to the surface demonstrates desired dynamics. The second task entails the development of a switching control law which forces the system to reach and maintain its sliding motion.

3.3.1 SWITCHING SURFACE DESIGN

To illustrate the switching surface design consider the nominal linear state space model

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (3.6)$$

in which $x(t) \in \mathfrak{R}^n$ and $u(t) \in \mathfrak{R}^m$ represent the state and control vectors respectively.

Moreover, the constant matrices $A \in \mathfrak{R}^{n \times n}$ and $B \in \mathfrak{R}^{n \times m}$ and it is assumed that B is of full rank m . In order to simplify the development of the design, the system can be transformed into a suitable canonical form. Since by assumption B is of full rank, there exists an orthogonal transformation matrix $T_r \in \mathfrak{R}^{n \times n}$ such that

$$T_r B = \begin{bmatrix} 0 \\ B_2 \end{bmatrix} \quad (3.7)$$

where $B_2 \in \mathfrak{R}^{m \times m}$ and is non-singular. The reason the transformation matrix T_r is imposed with an orthogonality restriction is that it is much easier to deal with orthogonal matrices. This is because the problem of inverting the matrix can be easily solved by straight forward transposition. By using now the coordinate transformation $x \leftrightarrow T_r x$ and by partitioning the states so that

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} \quad (3.8)$$

where $x_1 \in \mathfrak{R}^{n-m}$ and $x_2 \in \mathfrak{R}^m$, the nominal linear system (3.6) becomes

$$\dot{x}_1(t) = A_{11}x_1(t) + A_{12}x_2(t) \quad (3.9)$$

$$\dot{x}_2(t) = A_{21}x_1(t) + A_{22}x_2(t) + B_2u(t) \quad (3.10)$$

This representation is known as regular form. Further, note that the system is decomposed into two subsystems, each one describing different dynamics. Equation (3.9) describes the null space dynamics and equation (3.10) describes the range space dynamics [Edwards 1998].

Once the system is transformed, the next step is to choose a variable s as a function of the system states. Generally, this variable is selected as a linear combination of the system states, i.e.

$$s(x(t)) = Sx(t) \quad (3.11)$$

where S parameterises a manifold (the so-called switching surface) of the system. By partitioning the switching function matrix compatibly as

$$S = [S_1 \ S_2] \quad (3.12)$$

where $S_1 \in \mathfrak{R}^{m \times (n-m)}$ and $S_2 \in \mathfrak{R}^{m \times m}$ then

$$\det(SB) = \det(S_2 B_2) = \det(S_2) \det(B_2) \quad (3.13)$$

Since by construction B_2 is non-singular, a necessary and sufficient condition is that S_2 must be non-singular. Assuming this to be the case, then during the sliding mode the following sliding condition exists

$$S_1 x_1(t) + S_2 x_2(t) = 0 \quad (3.14)$$

Equation (3.14) can also be written by expressing $x_2(t)$ in terms of $x_1(t)$ as follows

$$x_2(t) = -M x_1(t) \quad (3.15)$$

where $M \in \mathfrak{R}^{m \times (n-m)}$ is defined by

$$M = S_2^{-1} S_1 \quad (3.16)$$

This indicates that during the sliding mode $x_2(t)$ is related linearly to $x_1(t)$. By substituting for $x_2(t)$ in equation (3.9) yields

$$\dot{x}_1(t) = (A_{11} - A_{12}M)x_1(t) \quad (3.17)$$

Equations (3.15) and (3.17) governed the dynamics of the sliding mode. It can be observed that the matrix S_2 has no direct effect on the dynamics of the sliding motion, but rather acts as a scaling factor for the switching function. Moreover, the design of a stable sliding mode requires the determination of the state feedback matrix M such that $(A_{11} - A_{12}M)$ has left-hand half-plane eigenvalues. Edwards and Spurgeon [Edwards 1998] state that if the pair (A_{11}, A_{12}) is controllable, then any robust linear state feedback method such as pole placement or quadratic minimisation can be used to compute M . Having found M , then the switching surface design can be completed by computing $S = [S_1 \ S_2]$.

After the design of the switching surface the next significant aspect is to guarantee the existence of a sliding mode. Details on sufficient conditions for the existence of a sliding mode are provided in the next section.

3.3.2 CONTROL LAW DESIGN

In an attempt to motivate this problem, a single-input case is first considered followed by a multivariable case. In general, the goal of the control law is to transfer the system into a prescribed manifold of the state space (i.e. sliding manifold) and to ensure the attainment within this subspace despite the presence of uncertainties. This is accomplished by setting a condition on the control law which renders the switching

surface attractive to the system. Such a condition is termed the reachability condition and the initial system trajectory under this condition is called the reaching phase.

3.3.2.1 THE SINGLE-INPUT CASE

A possible method for specifying the reachability condition for a single-input case is by Lyapunov stability analysis. By choosing the positive definite Lyapunov function candidate

$$V(s) = \frac{1}{2} s^2 \quad (3.18)$$

where s is the sliding variable and deriving its derivative as

$$\dot{V}(s) = s \dot{s} \quad (3.19)$$

then, the sliding variable will converge to zero only if

$$\dot{V} < 0 \quad (3.20)$$

This relation between the sliding variable and its first derivative is thus the reachability condition. Bhatti et al [Bhatti 1999] reported the use of a linear or a discontinuous (i.e. nonlinear) reachability condition which satisfies equation (3.20) and hence making the sliding manifold attractive. The two conditions considered are defined as follows

$$\dot{s} = -k_L s \quad (3.21a)$$

$$\dot{s} = -k_D \text{sgn}(s) \quad (3.21b)$$

where k_L and k_D are positive design constants. Moreover, sgn is the sign function which exhibits the property that $|s| = \text{sgn}(s)s$. The linear condition however, can only guarantee asymptotic convergence to the sliding surface. From equation (3.21a) it follows that

$$s(t) = s(0)e^{-k_L t} \quad (3.22)$$

Hence, in case initially the states do not lie on the sliding surface (i.e. $s(0) \neq 0$), then $s(t) \neq 0$ for all $t > 0$, which implies that the sliding surface is reached asymptotically, and $s(t) \rightarrow 0$ as $t \rightarrow \infty$. In contrast, the nonlinear condition defined in equation (3.21b) provides convergence at finite time as will be seen below. By rewriting this equation as

$$\frac{1}{2} \frac{d}{dt} s^2 = -k_D |s| \quad (3.23)$$

and integrating from 0 to t_s , it yields that

$$|s(t_s)| - |s(0)| \leq -k_D t_s \quad (3.24)$$

where t_s represents the time taken to reach the sliding surface ($s(t) = 0$) and is given by

$$t_s \leq \frac{|s(0)|}{k_D} \quad (3.25)$$

Having described the control law design procedure of a single-input case, the following subsection attempts to describe the multivariable case.

3.3.2.2 THE MULTIVARIABLE CASE

The control structure to be described in this section is based on the work of Dorling and Zinober [Dorling 1986]. The control law usually consists of two parts: a linear term u^L and a non-linear term u^N , which are added to form the control structure given below

$$u(x) = Lx + \frac{\rho}{\|Fx\|} Nx \quad (3.26)$$

Starting from the transformed state x given in section 3.3.1, a second transformation

$T_2 = \mathfrak{R}^n \rightarrow \mathfrak{R}^n$ such that

$$z = T_2 x \quad (3.27a)$$

where

$$T_2 = \begin{bmatrix} I_{n-m} & 0 \\ M & I_m \end{bmatrix} \quad (3.27b)$$

The above matrix is clearly non-singular, with inverse

$$T_2^{-1} = \begin{bmatrix} I_{n-m} & 0 \\ -M & I_m \end{bmatrix} \quad (3.28)$$

Partitioning $z^T = \begin{bmatrix} z_1^T & z_2^T \end{bmatrix}$ with $z_1 \in \mathfrak{R}^{n-m}$ and $z_2 \in \mathfrak{R}^m$

$$z_1 = x_1 \quad (3.29a)$$

$$z_2 = Mx_1 + x_2 \quad (3.29b)$$

Therefore, the transformed system becomes

$$\dot{z} = \Sigma z_1 + A_{12} z_2 \quad (3.30a)$$

$$\dot{z} = \Theta z_1 + \Phi z_2 + B_2 u \quad (3.30b)$$

where $\Sigma = A_{11} - A_{12}z_2$, $\Theta = M\Sigma - A_{22}M + A_{21}$ and $\Phi = MA_{12} + A_{22}$

In order to attain the sliding mode, it is necessary to force z_2 and \dot{z} to become identically zero. The linear part of the control is defined to be

$$u^L(z) = -B_2^{-1} \{ \Theta z_1 + (\Phi - \Phi_*) z_2 \} \quad (3.31)$$

where Φ_* is any $m \times m$ matrix with left-hand half-plane eigenvalues.

Therefore, transforming back into the original state space x gives

$$L = -B_2^{-1}[\Theta \quad \Phi - \Phi_*]T_2^T \quad (3.32)$$

However, as in the single-input case this linear control law u^L will only drive the state component z_2 to zero asymptotically. In order to attain convergence in finite time a nonlinear control term u^N is required. This nonlinear term must be discontinuous whenever $z_2 = 0$, and continuous elsewhere. Letting P_2 denote the positive-definite unique solution of the Lyapunov equation

$$P_2 \Phi_* + \Phi_*^T P_2 + I_m = 0 \quad (3.33)$$

then $P_2 z_2 = 0$ if and only if $z_2 = 0$, and the nonlinear term is given by

$$u^N(z) = \frac{-\rho}{\|P_2 z_2\|} B_2^{-1} P_2 z_2, z_2 \neq 0 \quad (3.34)$$

where $\rho > 0$ is a scalar parameter. The control law defined by (3.31) and (3.34) will drive an arbitrary initial state z^0 to the sliding manifold in a finite time

$$t \leq \frac{1}{\rho} \sqrt{\frac{\langle z_2^0, P_2 z_2^0 \rangle}{\sigma_{\min}(P_2)}} \quad (3.35)$$

where $\sigma_{\min}(P_2)$ denotes the minimum eigenvalue of P_2 and $\langle \cdot, \cdot \rangle$ is the usual Euclidean inner product on \mathfrak{R}^m . Finally, expressing the control law back into the original coordinates

$$N = -B_2^{-1}[0 \ P_2]T_2T \quad (3.36a)$$

$$F = [0 \ P_2]T_2T \quad (3.36b)$$

3.3.3 AN ILLUSTRATIVE EXAMPLE

As an example, consider the second order system defined in equation (3.3). Rewriting the equation in the regular form gives

$$\dot{x}_1(t) = x_2(t) \quad (3.37a)$$

$$\dot{x}_2(t) = u(t) \quad (3.37b)$$

$$y(t) = x_1(t) \quad (3.37c)$$

Here u is the input and y the output of the system. The overall aim of the design is to stabilise the system, i.e. drive the system states $(x_1(t), x_2(t)) = (y(t), \dot{y}(t))$ to the origin.

Recall that the design procedure consists of two phases, the determination of the sliding surface and the selection of the control action so that the sliding surface is reached. In the first phase although the sliding surface can be selected as a nonlinear function of the system states, usually it is chosen as a linear combination of the states (see equation (3.11)). Levant [Levant 2003] reports that for such a first order sliding mode algorithm, the sliding surface should have relative degree one with respect to the system input.

This means that the control input has to appear explicitly in the first time derivative of the sliding variable. Taking into account the above, consider the following sliding function

$$s(y(t), \dot{y}(t)) = cy(t) + \dot{y}(t) \quad (3.38)$$

where c is a strictly positive design parameter. It is easy to check that the sliding mode dynamics are $\dot{y}(t) = -cy(t)$, a first order dynamic wholly determined by the value of c , which can be selected by the designer. Further, this linear sliding surface function satisfies the relative degree one condition, since the first time derivative of the sliding variable is a function of the control input, i.e.

$$\dot{s}(t) = c y(t) + u(t) \quad (3.39)$$

Whilst the sliding surface has been selected, the second phase of the design relates to the construction of the control law. Initially, choose a reachability condition of the form

$$\dot{s}(t) = -k_D \operatorname{sgn}(s(t)) - k_L s(t) \quad (3.40)$$

This reachability condition clearly satisfies the inequality of equation (3.20) as follows

$$\begin{aligned} s\dot{s} &= s(t)[-k_D \operatorname{sgn}(s(t)) - k_L s(t)] \\ &= -k_D |s(t)| - k_L s(t)^2 < 0 \end{aligned} \quad (3.41)$$

provided that k_D and k_L have a positive value. The discontinuous control u can now be computed by substituting equation (3.40) into (3.39)

$$u(t) = -c y(t) - k_D \operatorname{sgn}(s(t)) - k_L s(t) \quad (3.42)$$

In this particular design let the system parameters be $c = 1$, $k_D = 1$ and $k_L = 1$. Figure 3-2 depicts the phase portrait of the system with initial conditions $x_1(0) = 1$ and

$x_2(0) = 0$. It is important to note the two distinct modes of the dynamical behaviour of the closed-loop system. The initial mode is the reaching mode in which the trajectory moves toward the sliding surface and attains the surface in finite time. Once the trajectory reaches the sliding surface, the second mode, the so-called sliding mode occurs. During the sliding mode, the trajectory tends to the origin of the phase plane which represents the equilibrium state of the system.

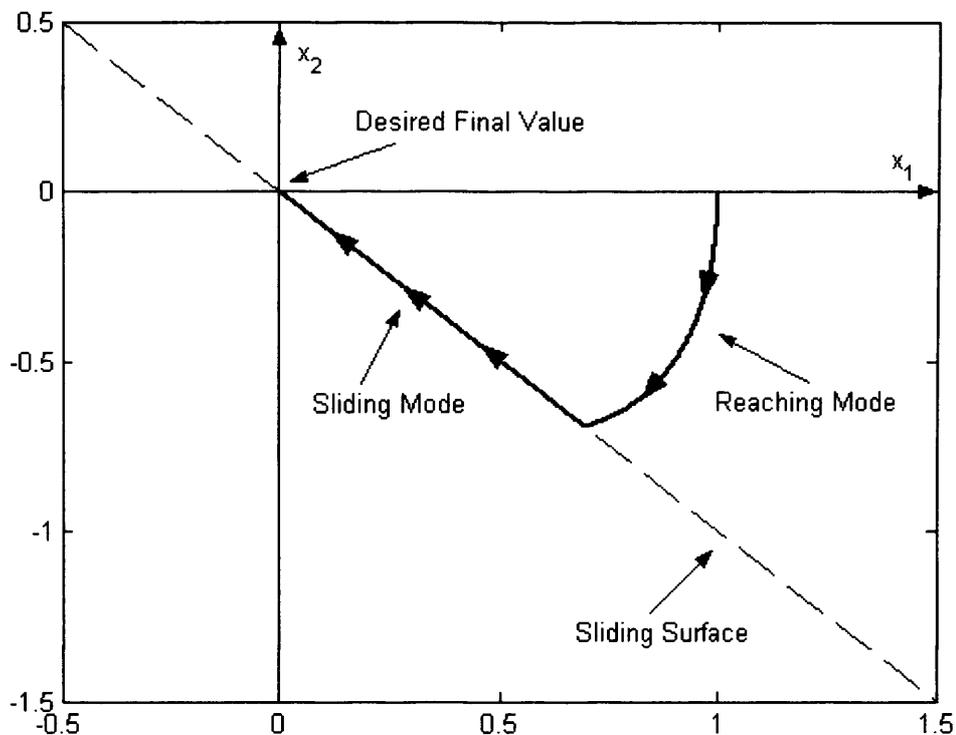


Figure 3-2: Phase portrait of the second order system

3.4 THE EQUIVALENT CONTROL METHOD

Before introducing the properties of the sliding motion, the notion of equivalent control will be considered. In essence, the equivalent control approach is a mathematical method for describing the dynamics of the sliding mode system [Utkin 1992]. Suppose

at time, t_s , the state trajectory of the second order system (see equation (3.37)) intercepts the sliding surface (see equation (3.38)) and a sliding mode is established for all $t > t_s$. As seen in section 3.3.1, during the sliding mode the switching function satisfies $s(t) = 0$, which in turn indicates that $\dot{s}(t) = 0$. Hence, from equation (3.39) it yields

$$u_{eq}(t) = -c\dot{x}(t) \quad (3.43)$$

This control law is known as the equivalent control. Note that this control signal is not the real signal which is applied to the system but it is the signal required on average to maintain a sliding motion (see Figure 3-3). A possible way to motivate the above is by passing the real control signal (see equation (3.42)) through a low-pass filter. The real control signal u is comprised of a low-frequency, or average, component and a high-frequency component. Therefore, let the average control u_{av} be the output of the low-pass filter

$$\tau \dot{u}_{av}(t) + u_{av}(t) = u(t) \quad (3.44)$$

where τ is the time constant. Figure 3-4 verifies that the average control value coincides with the equivalent control value but only when the sliding motion is established (in this example after approximately 0.7 seconds).

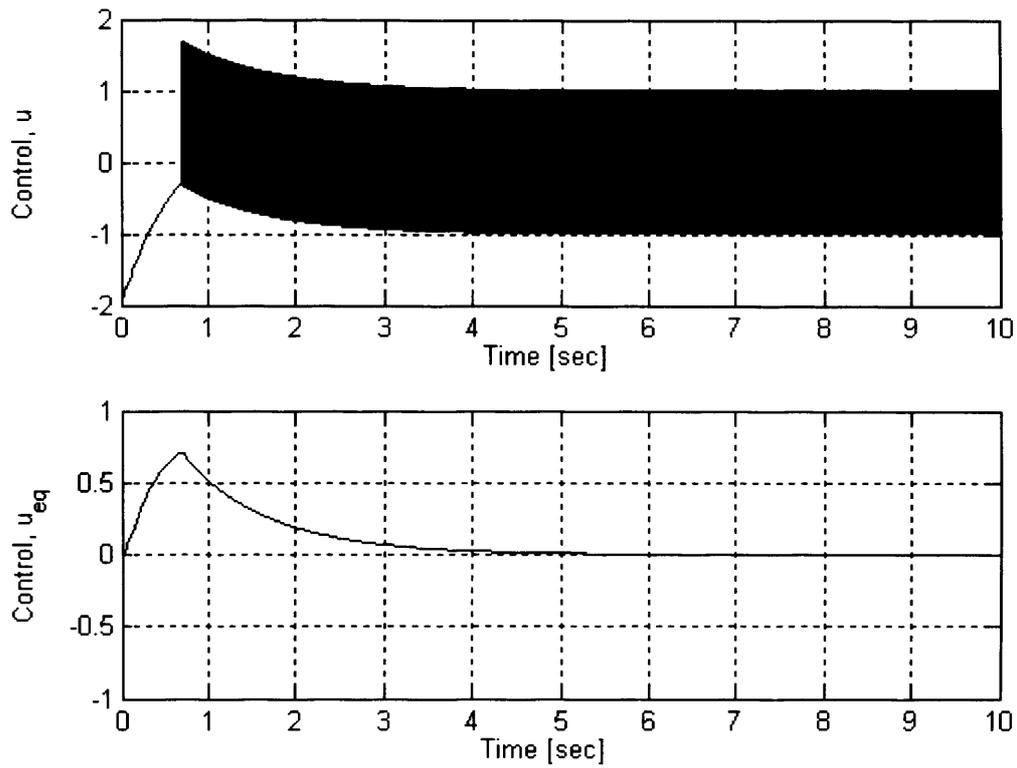


Figure 3-3: Discontinuous and equivalent control

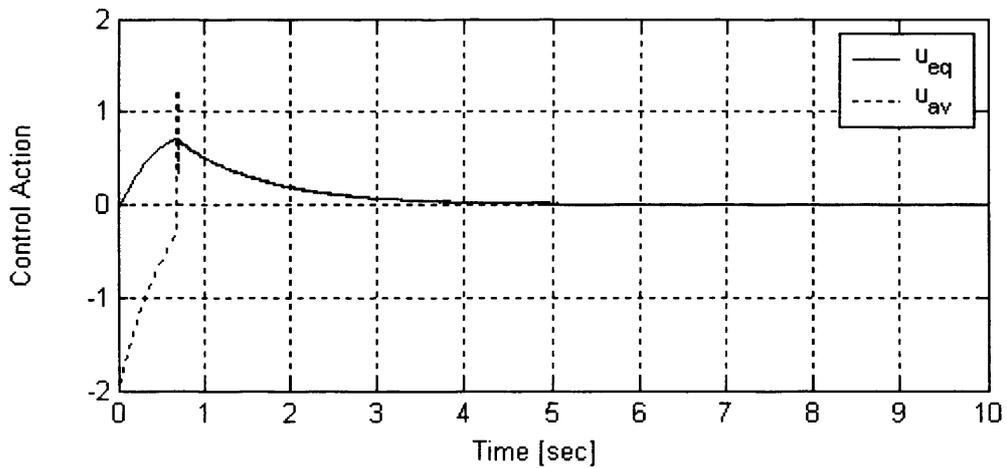


Figure 3-4: Comparison of the equivalent control with the average control

3.5 PROPERTIES OF THE SLIDING MOTION

Sliding mode control theory has attractive properties which are unique and relevant to real applications. In this section the two main properties of a sliding mode system are described.

3.5.1 REDUCTION OF ORDER

By referring back to the second order system in equation (3.37), if a sliding mode exists then the switching function defined in equation (3.38) becomes $s(t) = 0$. Hence, it follows

$$\begin{aligned} cy(t) + \dot{y}(t) &= 0 \\ \dot{y}(t) &= -cy(t) \end{aligned} \tag{3.45}$$

Observe that equation (3.45) determines the system motion on the sliding surface and is also independent of the control. Moreover, this motion on the sliding surface is governed by a first order differential equation, which the designer can select, and is one order less than the original system. In a more general case DeCarlo et al [DeCarlo 1988] reported that for a multivariable system of n th-order, the system dynamics during the sliding motion are reduced to $(n-m)$ th-order (where m is the number of inputs).

3.5.2 ROBUSTNESS PROPERTY

The most distinguished property of the sliding motion is its ability to result in very robust systems. Consider the following general linear system of the form

$$\dot{x}(t) = (A + \Delta A)x(t) + Bu(t) + f(t) \quad (3.46)$$

where ΔA and $f(t)$ represent the modelling error and external disturbance respectively.

If $\Delta \tilde{A}$ and $\tilde{f}(t)$ exist such that the matching conditions

$$\Delta A = B\Delta \tilde{A} \text{ and } f(t) = B\tilde{f}(t) \quad (3.47)$$

are satisfied, then the sliding motion is invariant to the uncertainty. In other words, all modelling uncertainties and disturbances acting in the input channels are completely rejected during the sliding motion [Edwards 1998]. Although it is difficult to imagine at this point how the effect of an unknown disturbance can be completely cancelled, it will be demonstrated later that this property results from the equivalent control.

Consider once again the second order system (see equation (3.37)) where a nonlinear term $asiny(t)$ is added in the input channel. This additional term can be regarded as a disturbance or an uncertainty. Hence, the resulting equations of the system are given by

$$\dot{x}_1(t) = x_2(t) \quad (3.48a)$$

$$\dot{x}_2(t) = u(t) + asiny(t) \quad (3.48b)$$

$$y(t) = x_1(t) \quad (3.48c)$$

where $a = -1$. Using the control law (see equation (3.42)) with initial conditions

$x_1(0) = 1$ and $x_2(0) = 0$, the phase portrait shown in Figure 3-5 is obtained. It reveals

that in finite time the sliding surface is reached and thereafter the system behaves like the ideal second

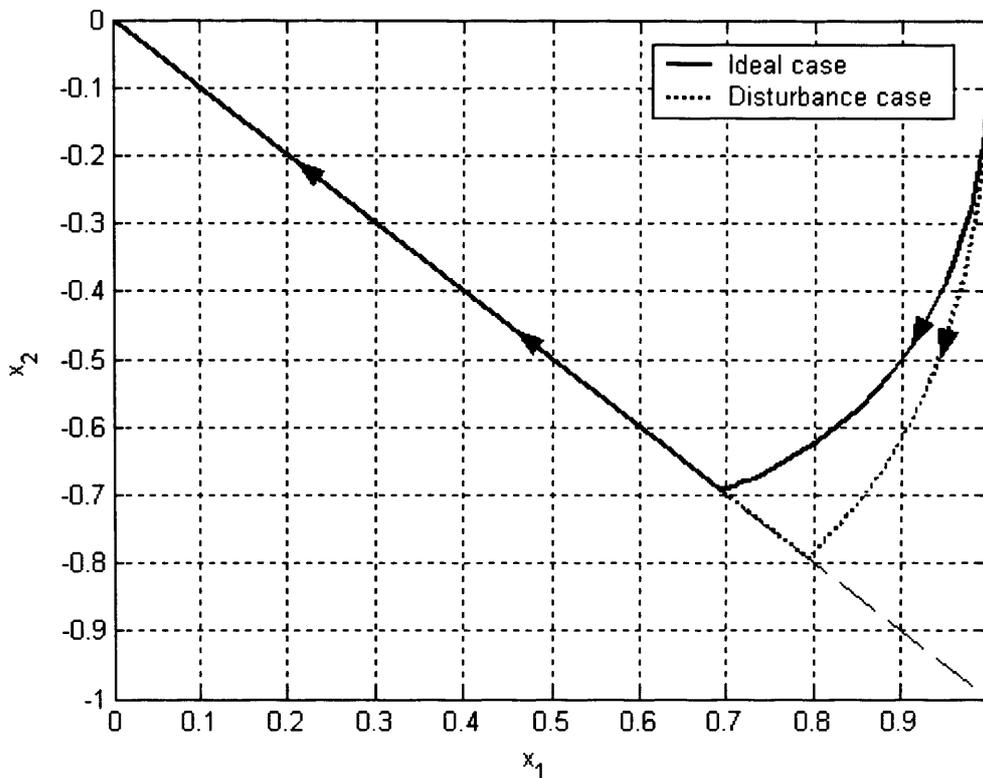


Figure 3-5: Phase portrait with disturbance case

order system (i.e. when $a = 0$). Consequently, the influence of the disturbance during the sliding motion is rejected and the system is said to be robust. This implies that the system is insensitive to mismatches between the model used for control law design and the plant on which it will be implemented. Nevertheless, the transient motion which brings the state of the system to the sliding surface (i.e. reaching mode) is different. This indicates that the disturbance term affects the system during this mode and thus preferably the transient motion should be as rapid as possible.

By using now the equivalent control method the rejection of the uncertainty can be substantiated. As argued in section 3.4, once a sliding mode is established, then $s(t) = 0$ and $\dot{s}(t) = 0$ for all subsequent time. From equations (3.38) and (3.48) follows

$$\dot{x}(t) = c\dot{y}(t) + a\sin y(t) + u(t) \quad (3.49)$$

Equating the above expression to zero, the equivalent control is obtained, i.e.

$$u_{eq}(t) = -c\dot{y}(t) - a\sin y(t) \quad (3.50)$$

Therefore, equation (3.50) shows the capability of the equivalent control to capture precisely the uncertain term ($a\sin y(t)$) in the closed-loop system and to cancel its effect.

3.6 THE CHATTERING PROBLEM

The sliding mode algorithms presented so far in this chapter are designed to drive the system state into the ideal sliding mode. Edwards and Spurgeon [Edwards 1998] stated that if infinite frequency switching of the discontinuous control was achievable, the sliding motion would be attained and maintained and the sliding variable s would be equal to zero. This type of behaviour is called an ideal sliding mode or ideal sliding motion. In practice, ideal sliding motion cannot be attained since switching is possible only at a finite frequency. This is due to the presence of switching imperfections such as time delays and limitations imposed by the physical actuators. The phenomenon of non-ideal but fast switching is known as ‘chattering’ [Burton 1986], [DeJager 1992]. The chattering behaviour may have undesirable consequences that could preclude the practical implementation of sliding mode algorithms for some classes of systems. For

example, in mechanical systems the high-frequency operation may lead to excessive wear and tear of the actuators. The same high operating frequency of the actuator can appear in the form of acoustic noise that may have harmful effects elsewhere in the system. Furthermore, the actuator is always in operation so it draws power continuously. This power is dissipated in the form of heat which may result in additional cost for its removal. Therefore, it is evident that there is a necessity to eliminate or suppress this harmful phenomenon.

Several design methods have been developed to tackle the problem of chattering. A straightforward approach is by tuning the reaching law parameters in equation (3.40). As follows from the analysis of section 3.3.3, the parameters k_D and k_L must have positive values in order to satisfy the reachability condition (see equation (3.20)). However, the amplitude of chattering is proportional to the magnitude of the discontinuous gain parameter k_D and the original setting of $k_D = 1$ is rather conservative. Thus, a lower value of k_D will reduce the amplitude of chattering. By setting $k_D = 0.1$ and $k_L = 5$ the control action illustrated in Figure 3-6 is acquired. The value for the linear gain k_L is increased in order to reduce the time taken to attain sliding. A reduction in the amplitude of the switching compared to the original setting is clearly visible.

The nature of the control signal in the above approach is still non-linear. Although the amplitude of the switching has been greatly reduced, in many cases such a signal would be unwanted. A popular design approach is to ‘soften’ the discontinuity of the control law. This is achieved by substituting a continuous approximation to the non-linear part

of the control signal. For instance, the signum function in equation (3.42) might be replaced by a

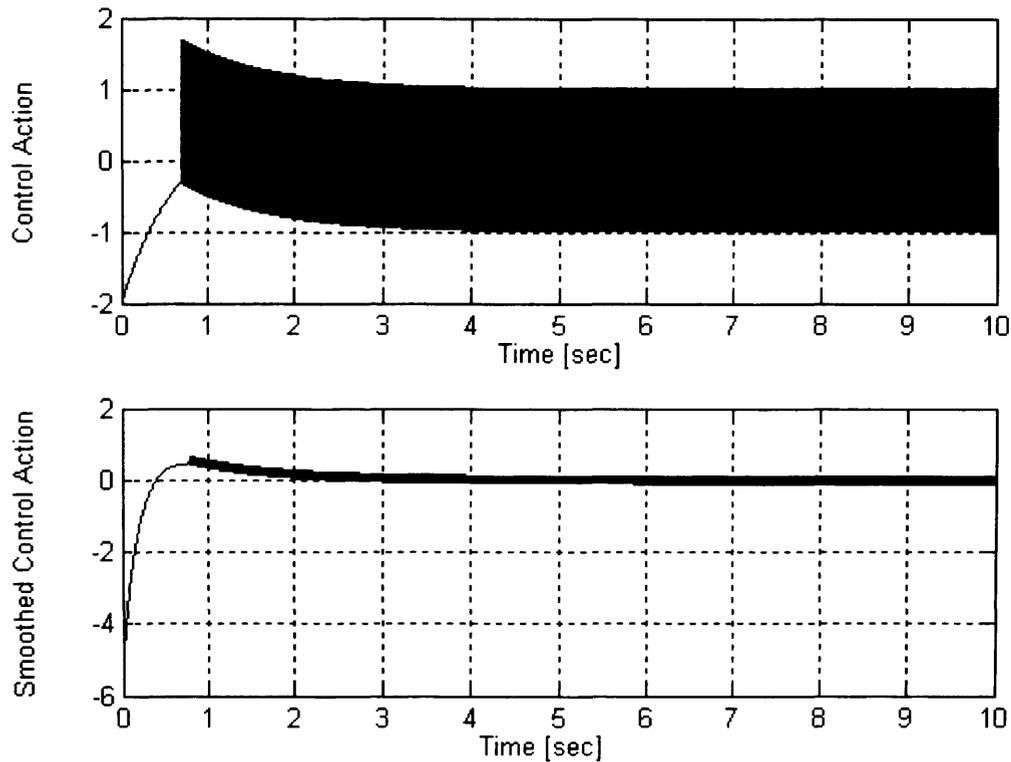


Figure 3-6: Comparison of the discontinuous and smooth control

saturation function of the form

$$\begin{aligned} \text{sat}(s) &= +1 \text{ when } s > \delta \\ &= \frac{s}{\delta} \text{ when } |s| \leq \delta \\ &= -1 \text{ when } s < -\delta \end{aligned} \quad (3.51)$$

where $\delta > 0$ and defines the size of the boundary layer. An alternative approximation is given below

$$v_{\delta}(s) = \frac{s}{(|s| + \delta)} \quad (3.52)$$

where δ is a small positive constant. As $\delta \rightarrow 0$, the function $v_{\delta}(\cdot)$ tends point-wise to the signum function, hence the name ‘signum-like’ function. Figure 3-7 depicts the signum function along with the continuous approximations described so far.

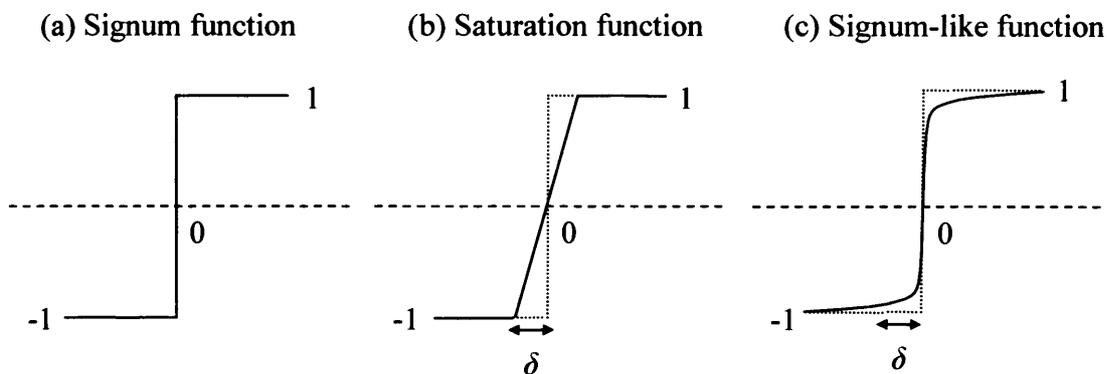


Figure 3-7: Signum, saturation and signum-like functions

At this point, it is important to note that with the linear approximations ideal sliding motion no longer takes place. This is because the continuous control action transfers the states only to within a neighbourhood of the sliding surface. As a consequence, the total invariance property with respect to matched uncertainty will be lost. Nonetheless, by choosing a small enough value for δ an arbitrarily close approximation to ideal sliding can be obtained. In the literature this is known as pseudo-sliding. Figure 3-8 shows the smooth control signal acquired by using the signum-like function with $\delta = 0.05$.

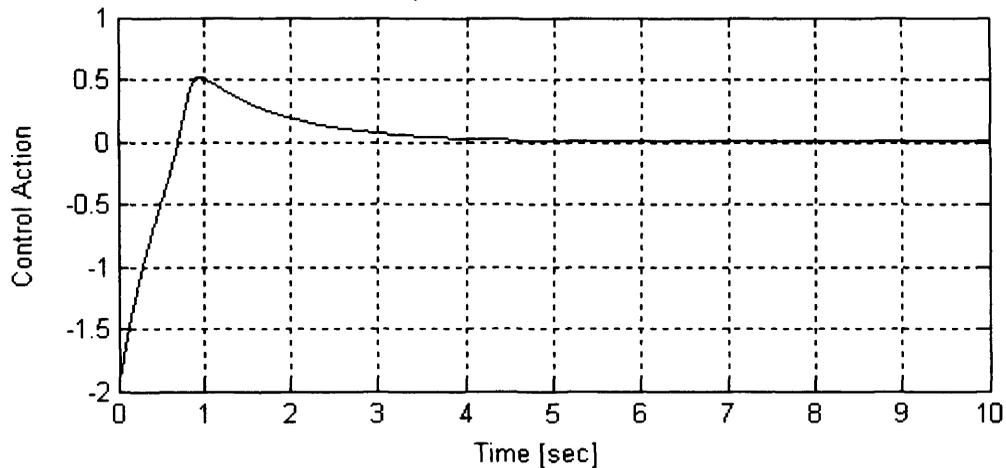


Figure 3-8: Smooth control signal using the signum-like function

3.7 SLIDING MODE OBSERVERS

In section 3.3, the design procedure to synthesise sliding mode algorithms has been presented. In this design, it is assumed that the state vector of the system is available. However, in practice the system states are not always available or sometimes are difficult to measure. A common way of overcoming this difficulty is to utilise an observer. An observer is a model-based method, specifically developed for estimation of state variables. An early example of an observer is a Luenberger observer [Luenberger 1971], as briefly presented in the following section.

3.7.1 LUENBERGER OBSERVER

Consider the observable system [Ogata 1997] represented by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{3.53}$$

where A, B and C are matrices of appropriate dimensions. The corresponding observer has the form

$$\dot{\hat{x}}(t) = (A + LC)\hat{x}(t) + Bu(t) - Ly(t) \quad (3.54)$$

The choice of the design matrix L must be such that the eigenvalues of $(A + LC)$ have negative real parts. Rewriting equation (3.54) as

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) - L(y(t) - C\hat{x}(t)) \quad (3.55)$$

it can be noticed that the observer operates like a model of the system, which is driven by the difference between the system and the observer outputs. This difference, also called the state error, is defined as

$$e(t) = x(t) - \hat{x}(t) \quad (3.56)$$

Hence, the error dynamics are governed by

$$\dot{e}(t) = (A + LC)e(t) \quad (3.57)$$

Due to the selection of the design matrix L , $(A + LC)$ is a stable matrix and therefore the error vector will converge to zero from any initial condition, just as $\hat{x}(t)$ will converge to $x(t)$.

Having described the idea of utilising a dynamical system for estimation of systems states (i.e. an observer), the next section introduces the possibility of using sliding mode techniques for robust state reconstruction.

3.7.2 OBSERVER DESIGN VIA SLIDING MODE TECHNIQUES

The main argument in favour of sliding mode techniques is the total invariance to any uncertainties which are implicit in the input channels (see section 3.5.2 for detail). This robustness property has led to the employment of sliding mode techniques to the problem of state estimation via an observer. This area of sliding mode observers is discussed in more detail by [Slotine 1987], [Walcott 1987] and the references contained within. In this section, a design framework adopted by Edwards and Spurgeon [Edwards 1998] is presented for the synthesis of a sliding mode observer for a linear system.

Consider the system given in (3.53) and assume that that the system states are unknown. Further, only the input and output signals $u(t)$ and $y(t)$ respectively are available. The aim is to generate a state estimate $\hat{x}(t)$, such that a sliding mode is attained. During the sliding motion the output error $e_y(t)$, which is defined as

$$e_y(t) = \hat{y}(t) - y(t) \quad (3.58)$$

is forced to zero in finite time. The proposed sliding mode observer has the form

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) - G_L e_y(t) + G_D v \quad (3.59)$$

where G_L and G_D are the linear and discontinuous observer gains respectively.

Moreover, v represents the discontinuous switched component given by

$$v = \begin{cases} -\rho \frac{Fe_y}{\|Fe_y\|} & \text{if } e_y \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.60)$$

From equation (3.58) it is straightforward to show that the observer error dynamics are

$$\dot{e}_y(t) = (A - G_L C)e_y(t) + G_D v \quad (3.61)$$

The linear gain G_L must be chosen such that $A_o = (A - G_L C)$ has stable eigenvalues and there exists a Lyapunov pair (P, Q) for A_o that satisfies the Lyapunov equation

$$A_o^T P + P A_o = -Q \quad (3.62)$$

where Q and P are some positive definite matrices. Also, P satisfies the following structural constraint

$$C^T F^T = P G_D \quad (3.63)$$

where F is a design matrix. To prove that the observer of equation (3.59) guarantees quadratic stability of the error system (3.61), let the expression

$$V(e) = e_y^T P e_y \quad (3.64)$$

be a candidate Lyapunov function. Differentiating the above equation and substituting equation (3.61) it follows that

$$\begin{aligned}
 \dot{V}(e) &= \dot{e}_y^T P e_y + e_y^T P \dot{e}_y \\
 &= (A_o e_y + G_D v)^T P e_y + e_y^T P (A_o e_y + G_D v) \\
 &= e_y^T A_o^T P e_y + v^T G_D^T P e_y + e_y^T P A_o e_y + e_y^T P G_D v \\
 &= e_y^T (A_o^T P + P A_o) e_y + 2 e_y^T P G_D v
 \end{aligned} \tag{3.65}$$

Therefore, by using equations (3.62) and (3.63) it follows that

$$\dot{V}(e) \leq -e_y^T Q e_y - 2\rho \|F C e_y\| \tag{3.66}$$

It is clear from the above that if $\rho > 0$, then the system is quadratically stable.

Besides using sliding mode observers in state estimation, in recent years the technique has been used successfully for condition monitoring and fault diagnosis. It will be seen in later chapters that appropriately designed sliding mode observers can be used to monitor a dry vacuum pump in order to reduce the occurrence of unplanned stoppages.

3.8 HIGHER ORDER SLIDING MODES

In previous sections the synthesis of a classical, first order sliding mode algorithm that drives the states of a system to a given manifold and keeps them within this constraint has been described. The main advantages, along with the possible presence of high frequency oscillations in the input (control) signals were presented. To avoid this ‘chattering-effect’ some approaches were proposed in section 3.6. In this section, a new

class of sliding modes is presented that has been developed to tackle the chattering phenomenon. Another advantage of this new class of sliding modes, called higher order sliding modes (HOSMs), is that they provide improved tracking (sliding) accuracy under sliding motion in sampled conditions as will be illustrated later.

HOSM techniques are a generalisation of the CSM control. In CSMs the discontinuity acts on the first order derivative of the sliding variable, whereas with HOSMs the discontinuity is acting on higher order time derivatives of the sliding variable. In general, if the total time derivatives $s, \dot{s}, \dots, s^{(r-1)}$ of the sliding variable exist, then the r th-order sliding mode (r -sliding mode) is determined by the equalities

$$s = \dot{s} = \ddot{s} = \dots = s^{(r-1)} = 0 \quad (3.67)$$

which constitutes an r -dimensional condition on the state of the dynamic system [Fridman 2002].

Therefore, the higher the order of the sliding variable derivative where the high frequency first appears, the less visible the oscillations on the variable itself will be. In other words, by moving the switching to the higher order derivatives of the control signal it is no longer hazardous, since the switching now takes place within the inner circuits of the control system (i.e. mostly in a computer) and not within the actuator.

Having briefly described how HOSMs suppress the chattering effect, their improved tracking (sliding) accuracy under sliding motion will now be presented. Levant [Levant 1993] reported that the quality of a sliding mode algorithm is related to the sliding

accuracy. This sliding accuracy can be better realized if the concepts of ideal and real sliding are introduced. Recall that every motion that takes place strictly on the constraint manifold $Sx(t) = 0$ is called an ideal sliding (see section 3.6 for detail). In reality though, ideal sliding cannot be attained due to switching imperfections such as delays. When these switching imperfections are taken into account and the constraint is kept only approximately, then real sliding is taking place. If ε is some measure of these switching imperfections then the following definition can be introduced:

Definition 1 [Levant 1993] Let $\gamma(\varepsilon)$ be a real-valued function such that $\gamma(\varepsilon) \rightarrow 0$ as $\varepsilon \rightarrow 0$. A real sliding algorithm on the constraint $s = 0$ is said to be of order r ($r > 0$) with respect to $\gamma(\varepsilon)$ if, for any compact set of initial conditions and for any time interval $[T_1, T_2]$, there exists a constant C , such that the steady-state process satisfies

$$|s| \leq C |\gamma(\varepsilon)|^r \quad (3.68)$$

for $t \in [T_1, T_2]$. In the special case when $\tau > 0$ is the minimal switching time interval, the relation can be stated as

$$|s| \leq C |\tau|^r \quad (3.69)$$

Therefore, as an example, it can be said that for a second-order sliding mode algorithm the deviation of the system from its constraint (sliding accuracy) is proportional to the square of the switching time delay.

3.9 DISCRETE TIME SLIDING MODE

Control systems that are theoretically developed for continuous time may not perform well and may even become unstable when direct digital implementation is applied. This fact is especially true for sliding mode systems because in discrete time systems the control signal is held constant between successive sampling points. The control signal cannot be altered at the very instant when the system motion crosses the sliding surface, which in turn means that it is not easy to force the system to remain on the sliding surface. Moreover, as seen in section 3.5 the invariance and robustness properties of sliding modes are only satisfied on the sliding surface. These desirable properties may therefore be lost in a discrete system. The obvious solution of decreasing the sampling period $\Delta\tau$ may not be always viable. Synthesis of sliding mode algorithms in the discrete time framework has therefore been necessary.

Koshkouei and Zinober [Koshkouei 2002] state that in discrete time systems there is a countable set of points consisting of the so-called lattice and the surface on which these sliding points lie is named the lattice-wise hyperplane. Consider the discrete time linear time invariant system

$$\begin{aligned}x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k)\end{aligned}\tag{3.70}$$

where $x(k) \in \mathfrak{R}^n$ is the state vector at the sampling instant k , $u(k) \in \mathfrak{R}^m$ is the input and $y(k) \in \mathfrak{R}^p$ is the output. A , B and C are matrices of appropriate dimensions. Define the sliding dynamical sequence as

$$s(k) = cx(k) \quad (3.71)$$

Definition 2 [Koshkouei 2002] The set of all points $x(k) \in \mathfrak{R}^n$, which lie on the hyperplane $cx = 0$ is said to be the sliding lattice-wise hyperplane or more concisely the ‘sliding lattice’. In fact, the sliding lattice-wise manifold is an infinite countable subset of the manifold $cx = 0$.

The next step in the design is to obtain a condition for the existence of the DSM, in which the sliding lattice is made attractive. The simplest condition for this existence is to rewrite equation (3.20) in the following form

$$\nabla s(k)s(k) < 0 \quad (3.72)$$

where $\nabla s(k) = s(k+1) - s(k)$. However, this condition guarantees that the states only approach the sliding lattice and do not remain on it. A stronger condition was proposed by Sarpturk et al [Sarpturk 1987]

$$|s(k+1)| < |s(k)| \quad (3.73)$$

for all k . Kim et al [Kim 2000] noted that for this condition the DSM system must be designed such that the switching function is decreased at every sampling index k and that it may be difficult to obtain a control law that satisfies this. The following condition was thus proposed

$$|s(k)| \leq \varepsilon \quad (3.74)$$

where $\varepsilon > 0$ and denotes the sliding lattice band width.

3.10 CONCLUSIONS

This chapter has described the basic concepts of sliding mode techniques. More specifically, the philosophy of the design of a sliding mode algorithm has been considered, highlighting a number of characteristics which will influence the development of the condition monitoring and fault diagnosis system in respect of the BOC Edwards' iGX dry vacuum pump.

The design procedure is based on two goals:

- First, the construction of the sliding surface to ensure the desired behaviour of the system in the sliding mode
- Second, the development of the control law that forces the system to reach and remain on the sliding surface

In discussing the properties of the sliding motion, it has been noticed that once the sliding motion is established the order of the system is reduced and it becomes totally invariant to matched uncertainties. Despite these advantages, sliding mode techniques present an important drawback that may limit their practical applicability, the chattering effect. It has been shown that these difficulties can be alleviated by smoothing the nonlinear term of the control signal. Moreover, the potential use of sliding modes for observer design has been examined and their use in fault detection schemes has also been mentioned. After this, the concept of HOSM has been discussed briefly. This

concept not only addresses the problem of chattering, but also provides an improved tracking capability compared to CSM in discrete implementation. Finally, the importance of restructuring the sliding mode design in a sampled data system framework has been presented.

In conclusion, based on the review of this chapter, it is expected that sliding mode techniques may be effectively applied for the development of on-board condition monitoring and fault diagnosis systems. The dry vacuum pump represents an excellent test bed for evaluation of sliding mode techniques because of its complexity and inherently nonlinear dynamics. In the following chapters, sliding mode observer-based fault detection systems will be developed for a number of the defined subsystems of the dry vacuum pump.

CHAPTER 4

CASE STUDY 1: THE COOLING SYSTEM

As machines become more complicated and valuable, there is a greater need to protect them, and the systems they support, from the consequences of faults. In chapter 2, a fault was defined as an unpermitted deviation of at least one characteristic property or parameter of the system from the acceptable condition. In many cases, such faults may create safety, maintenance or even environmental problems. This is also relevant for dry vacuum pumps, hence the requirement for development of a condition monitoring and fault diagnosis system.

In this chapter, the first nominated system component of the dry vacuum pump, the water cooling system, is considered. The function of the water cooling system is to dissipate excessive heat generated from the vacuum pump. This removal of heat is crucial, since by controlling the temperature of the vacuum pump the occurrence of unplanned stoppages is minimised.

The fault detection system for the water cooling system is based on sliding mode techniques. In particular, a nonlinear sliding mode observer-based approach is proposed

as a solution. The attractiveness of this approach arises from the fundamental robustness against certain kinds of parameter variations. Further, it enables faults and/or unmeasurable system parameters to be reconstructed. As a result, the concept has already gained some practical importance in the field of condition monitoring and fault detection.

Alessandri et al [Alessandri 1999] described the development of a sliding mode observer-based method for robust fault diagnosis of unmanned underwater vehicles. This system is compared with the extended version of the Kalman filter (EKF). It is concluded that the sliding mode observer performs better than the EKF, since its residuals are more stable and react faster at the appearance of faults. A specific sliding mode observer for fault detection and isolation is produced by Edwards et al [Edwards 2000]. The observer is designed to maintain the sliding motion, even in the presence of faults, and reconstructs the fault signals as a function of the so-called equivalent injection signal. This property of fault, or parameter, reconstruction is successfully employed for a ship benchmark problem by Edwards and Spurgeon [Edwards 2000]. The results obtained are compared with alternative observer-based approaches (i.e. fuzzy observers) and it is reported that the detection performance is comparable or better. Another contribution which builds on the work of Edwards et al [Edwards 2000] is presented by Tan and Edwards [Tan 2002]. The suggested method requires the design of a secondary sliding mode observer, which adds to the primary design requirements. Nonetheless, it also allows for more complex examples to be considered.

The preceding paragraphs have introduced the cooling system fault detection problem and illustrated how sliding mode techniques are being established for use in fault

diagnosis. Applying this philosophy it will be seen that parameter estimation and thus fault detection of the water cooling system can be achieved by examining the associated equivalent injection signal.

The chapter is organised as follows: In section 4.1 the cooling system fault detection problem is formulated and practical issues such as the selection of appropriate monitoring signals are considered. The development of the cooling system heat transfer models is described in section 4.2. Three different scenarios representative of a defective cooling system are introduced in section 4.3. Section 4.4 covers the development of the sliding mode observer, which has been produced for parameter estimation and hence fault detection of the cooling system. Finally, the experimental set-up and results are presented in sections 4.5 and 4.6 respectively.

4.1 PROBLEM FORMULATION

In the design of the diagnostic scheme, a number of practical considerations need to be taken into account. As pointed out in chapter 2, the success of any diagnostic scheme depends strongly on the selection of appropriate monitoring signals and transducers. The choice of which signals it is most suitable to monitor is often a trade-off for the user between cost and functionality. Therefore, in terms of the water cooling system it is proposed to monitor temperature signals. Temperature transducers are cheaply available and readily fitted to the system in a non-invasive manner.

Temperature is a key system parameter and relatively straightforward to monitor. However, monitoring temperature signals is not always sufficient to describe the state of the system. This is because a variety of faults may cause the temperature of the vacuum

pump to rise. For instance, in the water cooling system different faults such as total coolant failure or coolant flow pipe-work blockage will both lead to overheating of the vacuum pump. In addition, the time dependency of the faults may also be different. When total coolant failure occurs, the rise in the temperature is abrupt (i.e. step-wise), whereas in the case of a blockage in the pipe-work, a drift-like rise is observed.

As a solution to the above problems, certain parameters that indicate the state of the cooling system like the coolant mass flow rate and the heat transfer coefficient can be monitored. Although this may seem as an appropriate choice for the monitoring task, practically the measurement of these variables is often expensive or difficult to achieve.

The sliding mode observer based approach provides a means of estimating these system parameters with the use of the equivalent injection signal. Therefore, a major advantage of the suggested diagnostic system is that it is able to differentiate between the various faults down to the component level, something that would not be possible from the temperature measurements alone.

Once the monitoring signals have been identified, the dry vacuum pump has been instrumented in order to allow the development of the cooling system model. The details of this development are described in the following section.

4.2 HEAT-TRANSFER MODELS FOR THE DRY VACCUM PUMP COOLING SYSTEM

In the introductory section, it is stated that the function of the water cooling system is to maintain the dry vacuum pump at an appropriate working temperature and to prevent

excessive thermal expansion that can damage the pump. The cooling circuit of the iGX dry vacuum pump has evolved and three different designs have been used during the course of this research (see chapter 2 for details). However, this chapter is only concerned with the initial design, as the other two designs were later introduced into the project. The initial cooling circuit design is formed by a long stainless steel pipe that surrounds the working volume of the system in order to dissipate excess heat from the motor, bearings and pump stages.

Three explicit mathematical models that describe the water cooling system are derived from physical laws. The system can be represented by means of a block diagram, as illustrated in Figure 4-1. The primary source of heat into the coolant system is the electrical power supply. Additionally, heat is exchanged between the pump and the atmosphere, the cooling water and pumped gas.

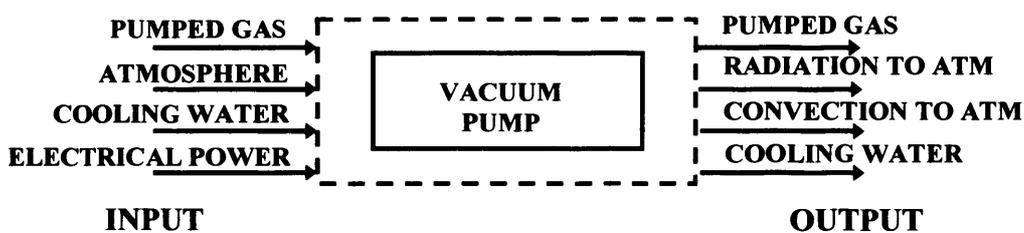


Figure 4-1: Block diagram of vacuum pump cooling system heat balance

A simple heat transfer model previously developed for a diesel engine [Bhatti 1999] is modified. The rate of change of pump body temperature is given by

$$(mc_p)_B \frac{dT_B(t)}{dt} = Q_P - Q_{CW} - Q_{CONV} - Q_{RAD} - Q_{PG} \quad (4.1)$$

where Q denotes an instantaneous heat transfer rate. The radiated heat and the net mass flow rate of pumped gas are both small, hence, Q_{RAD} and Q_{PG} can be neglected.

Moreover, the instantaneous heat transfer rate of power $Q_p = kI$ is assumed to be a linear function of the inverter current I , where k is a constant. The cooling water heat transfer is defined as

$$Q_{CW} = \dot{m}_c c_{pcw} (T_o(t) - T_i(t)) \quad (4.2)$$

and the surface heat loss to ambient through convection

$$Q_{CONV} = (hA)_B (T_B(t) - T_{atm}(t)) \quad (4.3)$$

Further, m_B is the mass of the pump body and c_{pB} is the specific heat capacity of the pump body. T_B , T_i , T_o and T_{atm} are the pump body, inlet, outlet and atmospheric temperatures respectively. Also, A_B represents the surface area of the pump body, h_B the heat transfer coefficient of the pump body, \dot{m}_c the mass flow rate of coolant through pump and c_{pcw} the specific heat capacity of the coolant.

Substituting for Q will result in the following equation

$$(mc_p)_B \frac{dT_B(t)}{dt} = kI(t) - \dot{m}_c c_{pcw} (T_o(t) - T_i(t)) - (hA)_B (T_B(t) - T_{atm}(t)) \quad (4.4)$$

Rearranging the above expression yields

$$\frac{dT_B(t)}{dt} = \frac{1}{(mc_p)_B} kI(t) - \frac{c_{pcw}}{(mc_p)_B} \dot{m}_c (T_o(t) - T_i(t)) - \frac{(hA)_B}{(mc_p)_B} (T_B(t) - T_{atm}(t)) \quad (4.5)$$

Re-labelling the coefficients for ease of exposition it follows that

$$\dot{T}_{B1}(t) = a_1 kI(t) - a_2 \dot{m}_c (T_o(t) - T_i(t)) - a_3 (T_{B1}(t) - T_{atm}(t)) \quad (4.6)$$

where a_1 , a_2 and a_3 are given by $a_1 = \frac{1}{(mc_p)_B}$, $a_2 = \frac{c_{pcw}}{(mc_p)_B}$ and $a_3 = \frac{(hA)_B}{(mc_p)_B}$.

In equation (4.6), the rate of change of pump body temperature (T_{B1}) is parameterized in terms of the mass flow rate, \dot{m}_c , but it can be also parameterized in terms of the heat transfer coefficient between the pump and the coolant h_c (T_{B2}). This will result in

$$\frac{dT_{B2}(t)}{dt} = \frac{1}{(mc_p)_B} kI(t) - \frac{A_c}{(mc_p)_B} h_c (T_{B2}(t) - T_o(t)) - \frac{(hA)_B}{(mc_p)_B} (T_{B2}(t) - T_{atm}(t)) \quad (4.7)$$

where A_c is the surface area of the surrounding pipe-work. Hence,

$$\dot{T}_{B2}(t) = a_1 kI(t) - a_4 h_c (T_{B2}(t) - T_o(t)) - a_3 (T_{B2}(t) - T_{atm}(t)) \quad (4.8)$$

where $a_4 = \frac{A_c}{(mc_p)_B}$.

The final equation is derived by considering the rate of change of the coolant temperature

$$(mc_p)_{cw} \frac{dT_o(t)}{dt} = Q_c - Q_{cw} \quad (4.9)$$

where

$$Q_c = h_c A_c (T_B(t) - T_o(t)) \quad (4.10)$$

is the pump body to coolant heat transfer and

$$Q_{cw} = \dot{m}_c c_{pcw} (T_o(t) - T_i(t)) \quad (4.11)$$

the cooling water heat transfer. In addition, m_{cw} , is the mass of the water coolant contained in the pump. Substituting the instantaneous heat transfer rates (4.10) and (4.11) into (4.9) and rearranging

$$\frac{dT_o(t)}{dt} = \frac{A_c}{(mc_p)_{cw}} h_c (T_B(t) - T_o(t)) - \frac{c_{pcw}}{(mc_p)_{cw}} \dot{m}_c (T_o(t) - T_i(t)) \quad (4.12)$$

Therefore, by renaming the coefficients

$$\dot{T}_o(t) = b_1 h_c (T_B(t) - T_o(t)) - b_2 \dot{m}_c (T_o(t) - T_i(t)) \quad (4.13)$$

where b_1 and b_2 are given by $b_1 = \frac{A_c}{(mc_p)_{cw}}$ and $b_2 = \frac{1}{m_{cw}}$.

4.3 FAULT SCENARIOS FOR A DEFECTIVE COOLING SYSTEM

Equations (4.6), (4.8) and (4.13) represent the dynamics of the cooling system. It is useful to consider the possible variation in the system parameters that may be used to indicate likely malfunction of the cooling system. For example, a variation in the coolant mass flow rate \dot{m}_c , a change in the heat transfer coefficient h_c between the pump and coolant and a change in the heat transfer k between the pump and the temperature sensor. Let

$$\dot{m}_c = \tilde{\dot{m}}_c + \Delta \dot{m}_c \quad (4.14a)$$

$$h_c = \tilde{h}_c + \Delta h_c \quad (4.14b)$$

$$k = \tilde{k} + \Delta k \quad (4.14c)$$

where $\Delta \dot{m}_c, \Delta h_c, \Delta k$ represent the deviations (i.e. the faults) and $\tilde{\dot{m}}_c, \tilde{h}_c, \tilde{k}$ the nominal parameters. Substituting the above in equations (4.6), (4.8) and (4.13) gives

$$\begin{aligned} T_{B1}^c(t) = & a_1 I(t) (\tilde{k} + \Delta k) - a_2 T_o(t) (\tilde{\dot{m}}_c + \Delta \dot{m}_c) \\ & + a_2 T_i(t) (\tilde{\dot{m}}_c + \Delta \dot{m}_c) - a_3 T_{B1}(t) + a_3 T_{atm}(t) \end{aligned} \quad (4.15)$$

$$\begin{aligned} T_{B2}^c(t) = & a_1 I(t) (\tilde{k} + \Delta k) - a_4 T_{B2}(t) (\tilde{h}_c + \Delta h_c) \\ & + a_4 T_o(t) (\tilde{h}_c + \Delta h_c) - a_3 T_{B2}(t) - a_3 T_{atm}(t) \end{aligned} \quad (4.16)$$

$$\begin{aligned} T_o^c(t) = & b_1 T_B(t) (\tilde{h}_c + \Delta h_c) - b_1 T_o(t) (\tilde{h}_c + \Delta h_c) \\ & - b_2 T_o(t) (\tilde{\dot{m}}_c + \Delta \dot{m}_c) + b_2 T_i(t) (\tilde{\dot{m}}_c + \Delta \dot{m}_c) \end{aligned} \quad (4.17)$$

It can be observed that by setting the deviations $\Delta m_c, \Delta h_c, \Delta k$ to zero in the above equations a nominal cooling system dynamics can be obtained.

4.4 SLIDING MODE OBSERVER DEVELOPMENT

An observer is a model-based method which is traditionally used for estimation of states variables [Luenberger 1971]. The sliding mode observer designed in this section is used for parameter estimation and hence fault detection of the water cooling system.

For the sliding mode observer development, it is assumed that the output of the plant is the pump body and coolant outlet temperatures. Furthermore, the sliding surface is chosen to be the error between the observer output and the plant output. It will be shown that the sliding motion will be attained even in the presence of a fault and that the sliding mode observer predicts both the temperatures and the model parameters – i.e. coolant mass flow rate, coolant heat transfer coefficient and pump body heat transfer.

The proposed observer has the following structure

$$\hat{T}_{B1}^{\&}(t) = -a_3 \hat{T}_{B1}(t) + a_1 \tilde{k} I(t) - a_2 \tilde{m}_c T_o(t) - a_2 \tilde{m}_c T_i(t) + a_3 T_{atm}(t) + v_{B1} \quad (4.18)$$

$$\hat{T}_{B2}^{\&}(t) = -(a_4 \tilde{h}_c + a_3) \hat{T}_{B2}(t) + a_1 \tilde{k} I(t) + a_4 \tilde{h}_c T_o(t) + a_3 T_{atm}(t) + v_{B2} \quad (4.19)$$

$$\hat{T}_o^{\&}(t) = -(b_1 \tilde{h}_c + b_2 \tilde{m}_c) \hat{T}_o(t) + b_1 \tilde{h}_c T_B(t) + b_2 \tilde{m}_c T_i(t) + v_o \quad (4.20)$$

where $v_i = K_i \left(\frac{\varepsilon_i}{\|\varepsilon_i\| + \delta} \right)$, $i =_{B1, B2, o}$ and K_i are the gains of the discontinuous signals v_i .

Moreover, ε_i is the observer error defined as the difference between the estimated and measured temperatures (i.e. $\varepsilon_{B1} = T_{B1} - \hat{T}_{B1}$, $\varepsilon_{B2} = T_{B2} - \hat{T}_{B2}$ and $\varepsilon_o = T_o - \hat{T}_o$). The selection of K_i must be such that the reachability problem is satisfied and the sliding motion is sustained at all times (see Appendix D for details). Finally, δ is a small positive constant used to reduce chattering (see section 3.6 for details). The following equations yield the observer error dynamics

$$\dot{\varepsilon}_{B1} = -a_3 \varepsilon_{B1} + a_1 \Delta k l(t) - a_2 \Delta r \tilde{\kappa}_c (T_o(t) - T_i(t)) - v_{B1} \quad (4.21)$$

$$\dot{\varepsilon}_{B2} = -(a_4 \tilde{h}_c + a_3) \varepsilon_{B2} + a_1 \Delta k l(t) - a_4 \Delta h_c (T_{B2}(t) - T_o(t)) - v_{B2} \quad (4.22)$$

$$\dot{\varepsilon}_o = -(b_1 \tilde{h}_c + b_2 \Delta r \tilde{\kappa}_c) \varepsilon_o - (b_1 \Delta h_c + b_2 \Delta r \tilde{\kappa}_c) T_o(t) + b_1 \Delta h_c T_B(t) + b_2 \Delta r \tilde{\kappa}_c T_i(t) - v_o \quad (4.23)$$

Assuming the K_i are chosen sufficiently large, a sliding mode will be attained and maintained. The observer errors and their derivatives will converge to zero due to the choice of the sliding surface. Thus, the sliding mode equations (4.21), (4.22) and (4.23) become

$$0 = a_1 \Delta k l(t) - a_2 \Delta r \tilde{\kappa}_c (T_o(t) - T_i(t)) - v_{B1} \quad (4.24)$$

$$0 = a_1 \Delta k l(t) - a_4 \Delta h_c (T_{B2}(t) - T_o(t)) - v_{B2} \quad (4.25)$$

$$0 = -(b_1 \Delta h_c + b_2 \Delta r \tilde{\kappa}_c) T_o(t) + b_1 \Delta h_c T_B(t) + b_2 \Delta r \tilde{\kappa}_c T_i(t) - v_o \quad (4.26)$$

The above expressions demonstrate that the observer provides a means of detecting changes in system parameters. It can also be observed that the system parameters are

interdependent and more than one discontinuous signal must be utilised to perform fault diagnosis. The idea is illustrated in Table 4-1.

Table 4-1. Fault conditions

Fault Condition	Average Value of Injection Signal		
	v_{B1}	v_{B2}	v_o
$\Delta r \& c_c$	Non-zero	0	Non-zero
Δh_c	0	Non-zero	Non-zero
Δk	Non-zero	Non-zero	0
Normal Condition	0	0	0

4.5 EXPERIMENTAL SET-UP

This section contains a description of the test equipment and instrumentation used for validating the sliding mode observer developed in the previous section. A dry vacuum pump is available in the laboratory (see chapter 2 for details). Temperature sensors are fitted on the vacuum pump in order to deliver input/output data. The motor current is captured from the system's inverter via a serial link. Similarly, the mass flow rate is recorded via a serial link connected to the dSpaceTM board.

dSpaceTM (Digital Signal Processing and Control Engineering) is the hardware interface involved in the experimental validation, since it provides all the tools required for real time data acquisition and direct data exchange with MatlabTM/SimulinkTM. Finally, a digital low pass filter generated in MatlabTM is employed to remove any unwanted high frequency components of the signal noise. Alternatively, an equivalent hardware filter

could be designed, decreasing the computational overhead but resulting in a small increase in hardware costs.

In order to replicate the changes in system parameters that may occur prior to pump failure, three different experimental scenarios are considered. A low flow or total coolant failure will result in high temperatures in the motor, stator and the bearings. These high temperatures will affect the vacuum pump and can result in total failure or an emergency stop. In order to simulate this type of fault a control valve is used to restrict the water flow. Secondly, a reduction in the rate of heat transfer from pump to coolant which can be caused by deposits on the coolant flow pipe-work is investigated. This type of fault is simulated by inserting an insulating material between the vacuum pump and the pipe-work. The final type of fault examined is a change in the heat transfer between the pump and the temperature sensor corresponding to additional internal source of heating such as bearing friction. The fault is simulated by fitting a small heater close to the location of the bearings.

Finally, to validate the heat transfer models introduced in section 4.2, a large quantity of temperature data has been acquired by using dSpaceTM. This large archive of data was then used for offline simulation and estimation of the parameters for the heat transfer models. Details of the procedure are reported in Appendix C.

4.6 RESULTS

A series of tests have been carried out to test the sliding mode observer based diagnostic system. The diagnostic scheme requires six input variables, i.e. $[T_B, T_o, T_i, T_{atm}, I, \dot{n}_c]$.

The data sampling time is selected to be equal to 1 second. In contrast, the sliding mode

observer needs a shorter sample time because the switching term of the algorithm has to keep the observer error sufficiently close to zero to maintain the sliding mode in order to maximise the accuracy of the parameter estimates.

Initially, validation data is obtained from the dry vacuum pump under fault-free conditions. Figure 4-2 illustrates the behaviour of observer T_{B1} under these conditions. The first graph represents a plot of the measured and the estimated data from the sliding mode observer. It can be noticed that the observer tracks the temperature data and that the corresponding error between them is of the order of 0-0.002 °C. The third graph represents the equivalent injection signal u_{B1} . It can be noted that it is not affected and remains close to zero under normal operating conditions.

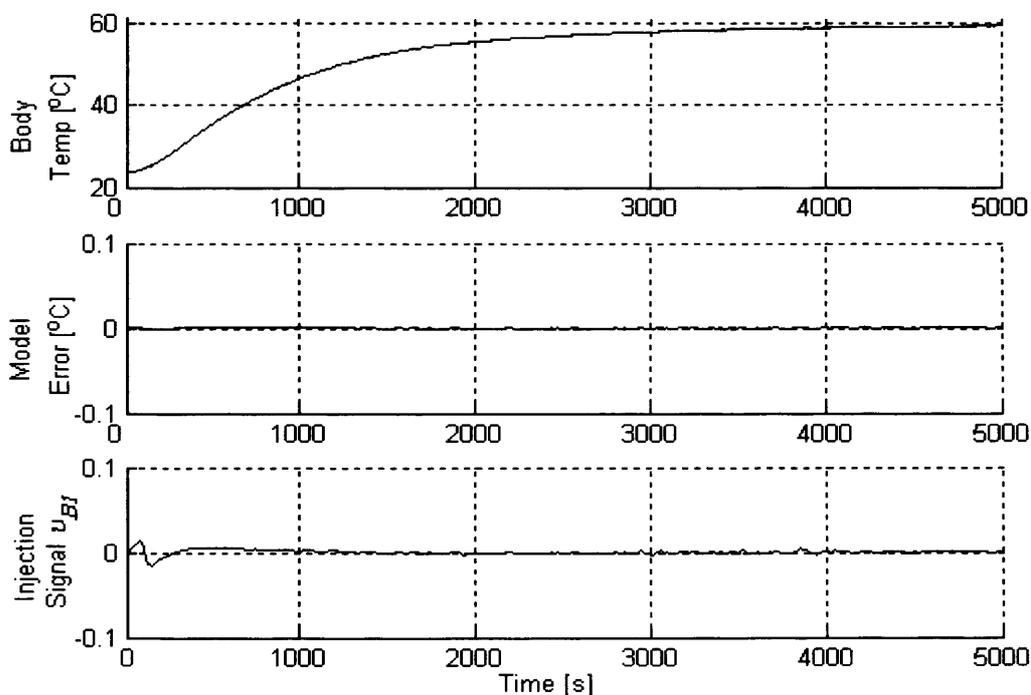


Figure 4-2: Measured and estimated body temperature T_{B1} , model error and the corresponding equivalent injection signal

Having validated the sliding mode observer under fault-free conditions, a fault is then introduced into the system. Figure 4-3 shows the measured and the estimated data under a fault condition for observers T_{B1} , T_{B2} and T_o . It can be seen that at approximately 3200 seconds a fault in the coolant mass flow rate \dot{m}_c is introduced by restricting the water flow. Nevertheless, all the three observers attain a sliding motion even in the presence of the fault. Figure 4-4 depicts the error plots between the measured and estimated data under a fault condition for observers T_{B1} , T_{B2} and T_o . It can be seen that the error magnitudes are small, which verifies that the sliding mode observers are sliding and that the estimated value tracks the measured data.

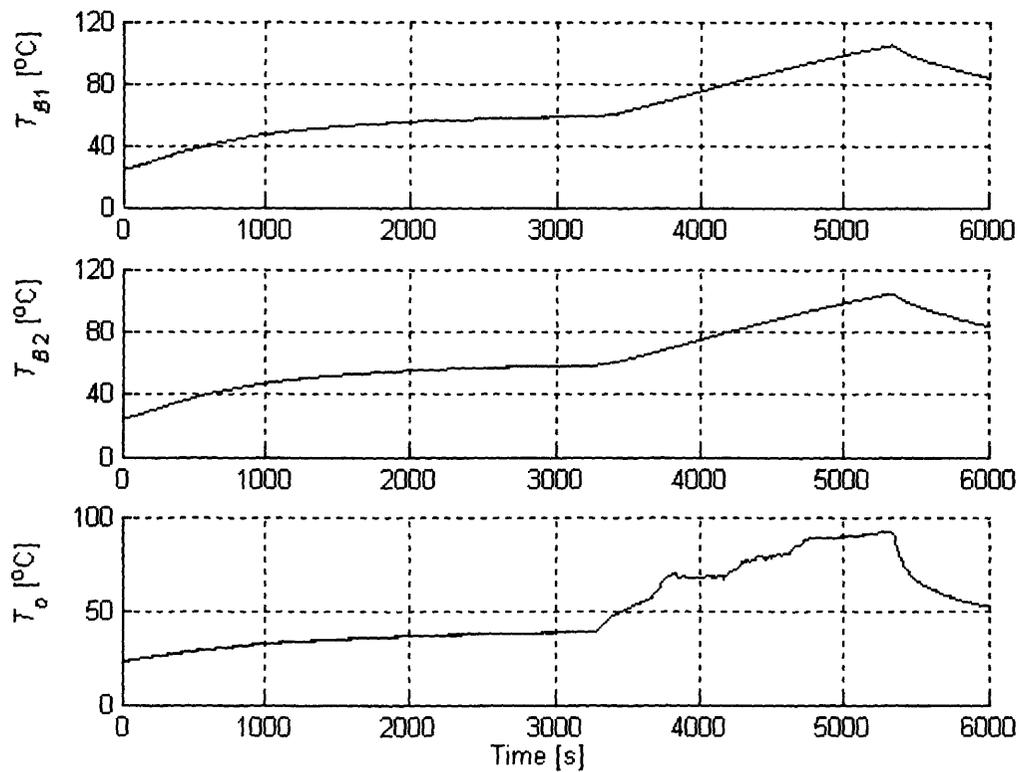


Figure 4-3: Measured and estimated data for T_{B1} , T_{B2} and T_o under a fault condition

(\dot{m}_c)

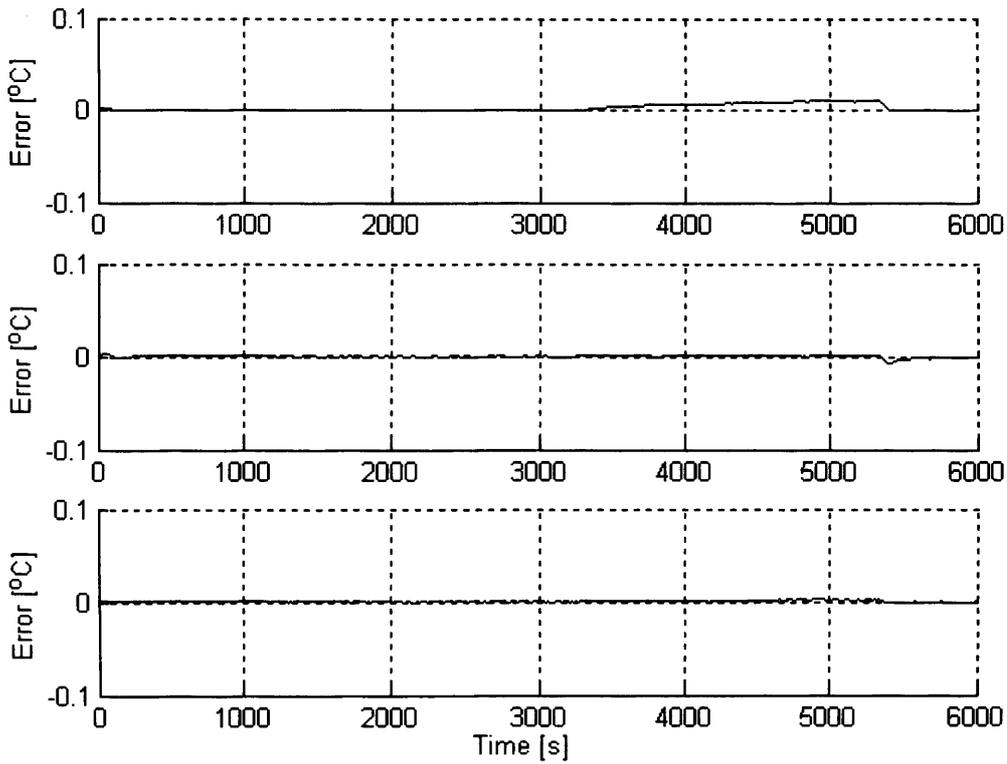


Figure 4-4: Error in temperature between the measured and estimated value for T_{B1} , T_{B2} and T_o observers under a fault condition ($\Delta n\dot{\kappa}$)

Figures 4-5 and 4-6 depict the equivalent injection signals for the three observers and the parameter estimate $\Delta n\dot{\kappa}$. The proposed diagnostic technique indicated that the parameters are interdependent so that the diagnostic system must detect non-zero values in more than one injection signal to infer a fault (see Table 4-1 for detail). As predicted, signals ν_{B1} and ν_o are affected, whereas ν_{B2} is largely unaffected. Further, good correlation between the estimates of $\Delta n\dot{\kappa}$ can be observed.

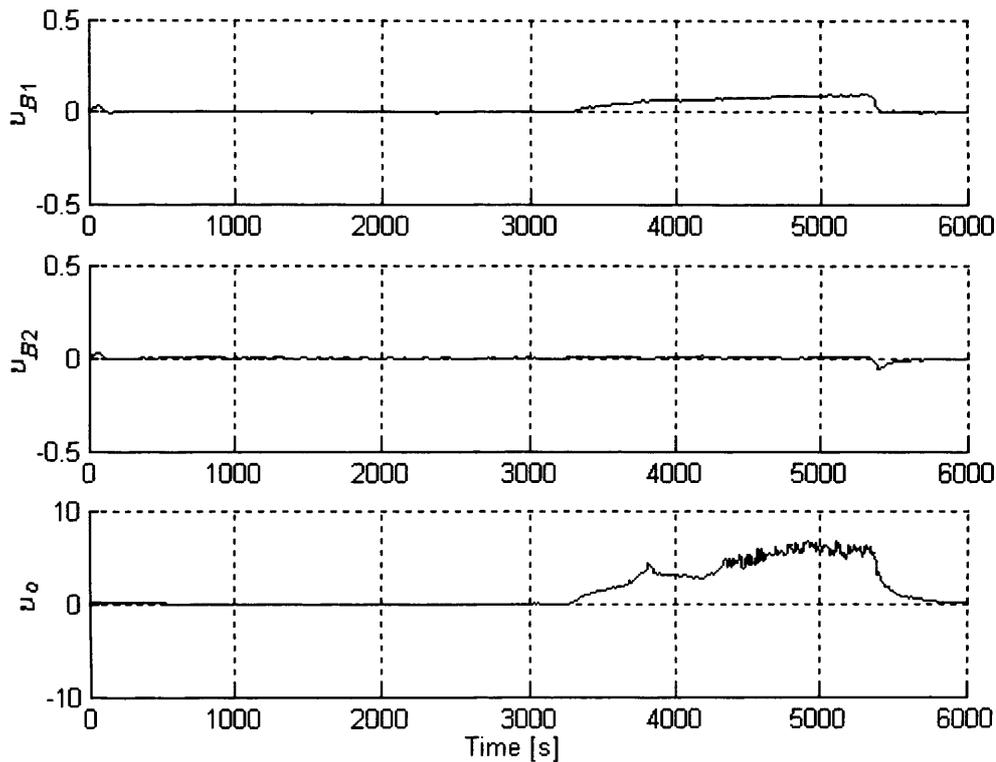


Figure 4-5: Injection signals v_{B1} , v_{B2} and v_o for coolant flow failure

Finally, Figures 4-6 and 4-7 show the non-zero parameter estimates for the remaining two fault situations. A change in the heat transfer coefficient h_c between the pump and coolant is introduced at approximately 2800 seconds. It can be seen that both observers reconstruct this change successfully. Moreover, a change in the heat transfer k between the pump and temperature sensor is introduced by the addition of approximately 120W at 1700 and 3500 seconds respectively.

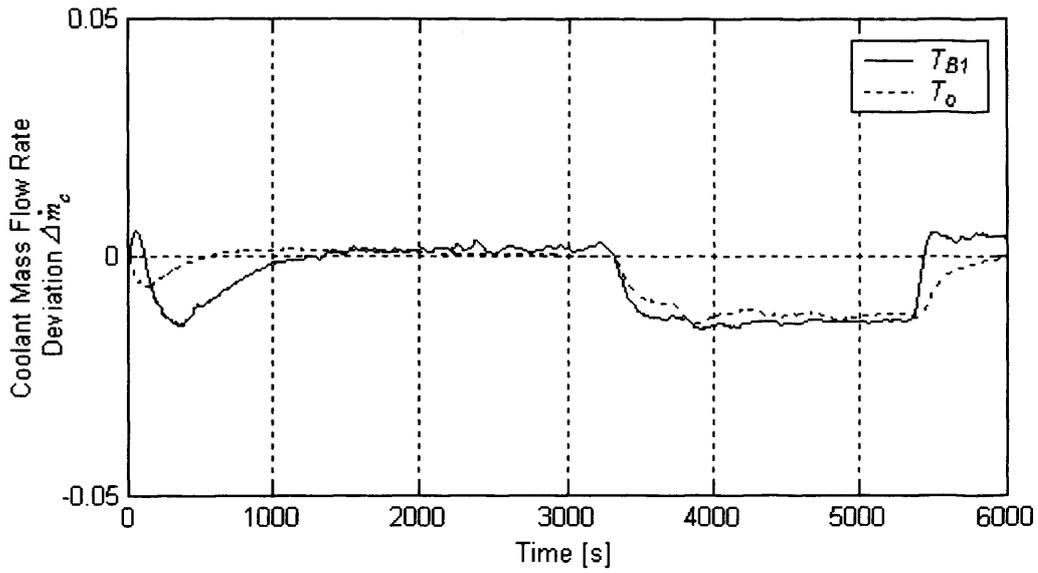


Figure 4-6: Component parameter estimate $\Delta \dot{m}_c$ for observers T_{B1} and T_o

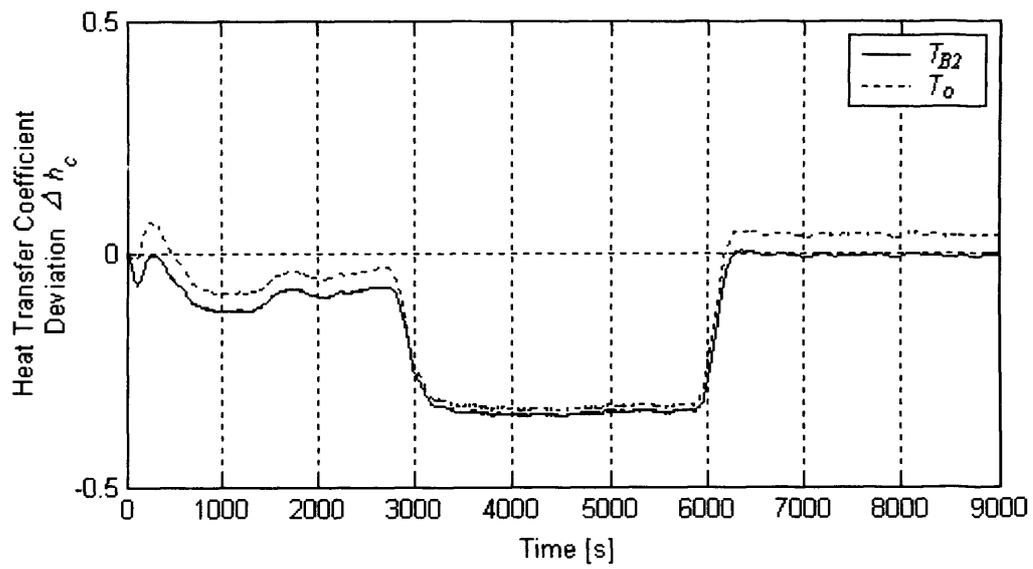


Figure 4-7: Component parameter estimate Δh_c for observers T_{B2} and T_o

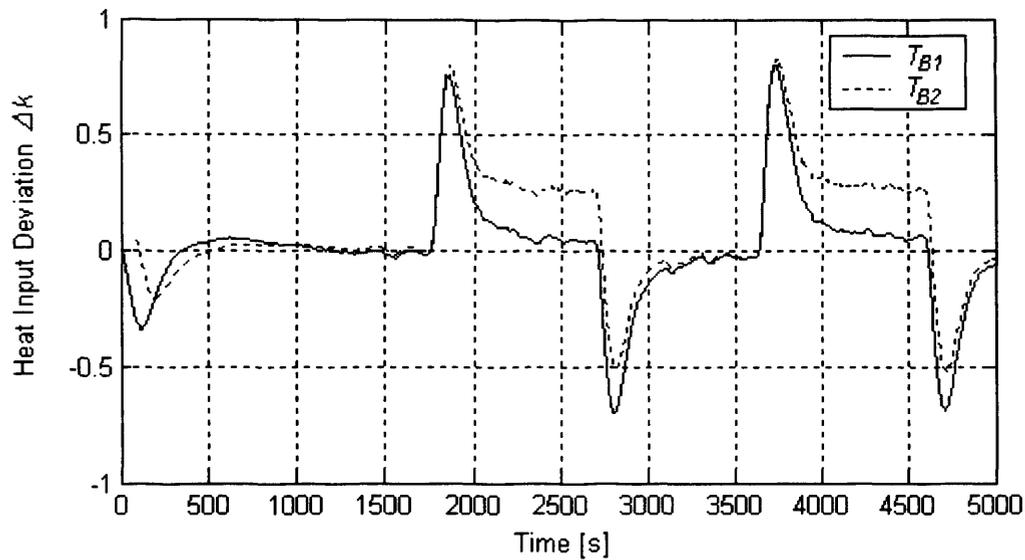


Figure 4-8: Component parameter estimate Δk for observers T_{B1} and T_{B2}

4.7 CONCLUSIONS

This chapter has presented the application of sliding mode techniques to the problem of fault detection of the cooling system for a dry vacuum pump. Specifically, a nonlinear sliding mode observer has been employed and parameter estimation, and hence fault detection has been achieved by examining the equivalent injection signal. This property of fault, or parameter, reconstruction is unique to the nonlinear sliding mode methodology and the results could not be obtained using a linear observer technique. The method also requires only low cost temperature transducers and a reading of the motor's current and coolant water flow. Further, the nature of the input signals dynamics is slow and a low sample rate of 1Hz for the heat transfer model is adequate in the implementation of the scheme.

The results from a series of experimental tests illustrate the usefulness of the approach for condition monitoring. Good correlation between the system parameter estimates is obtained from the different observers for all three targeted faults. The method provides

an earlier diagnosis and information down to a component level compared to a simple high temperature alarm, for example.

CHAPTER 5

AN ALGEBRAIC FRAMEWORK FOR IDENTIFICATION AND SLIDING MODES

Sliding mode observers are a well-developed technique with applications in condition monitoring and fault diagnosis of continuous time systems [Edwards 2000], [Edwards 2000]. Faults are detected by analysis of the so-called equivalent output injection and appropriate manipulation of this signal can effectively either reconstruct the fault signal or monitor parameters whose variation may be used to infer the health of a system. The observer is designed to maintain a sliding motion even in the presence of faults and is effective at rejecting certain kinds of parameter variations and disturbances.

Bhatti et al [Bhatti 1999] reported the use of sliding mode observers for state estimation and for reconstruction of fault signals in the parameters of a heat transfer model. A non-linear sliding mode observer was used for the detection of possible faults in a diesel engine coolant system by Goh et al [Goh 2002]. In these examples, first order differential equations are used to model heat transfer, providing a basis for the observer design.

The need for a mathematical model to formulate the observer requires that nominal values of the model parameters must be obtained. In the case of the heat transfer models referred to above, the values of some parameters can be acquired directly by measurement and some from tables of material properties. Generally however, some parameters may not be known and some means of parameter estimation for any unknown parameters must be used.

The topic of parameter estimation is broad and comprehensive coverage of all of the areas associated with it is outside the scope of this thesis. A comprehensive survey of parameter estimation techniques for continuous time models is presented by Young [Young 1981] and Unbehauen and Rao [Unbehauen 1998]. Where the model structure is known, parameter estimation techniques by such means as least squares estimation [Jiang 2004] have been used in condition monitoring and fault diagnosis schemes [Molteberg 1991].

The main problem with parameter estimation from input-output data is the presence of derivative operators in the models. The inevitable measurement of noise means that direct generation of the derivatives is not a practical option. Fliess and Sira-Ramirez [Fliess 2003] have developed a framework for identification of continuous time constant linear system parameters. Their approach exhibits good robustness properties with respect to a large variety of additive disturbances and is based on mathematical tools such as module theory, differential algebra and operational calculus. Further, the use of the algebraic framework is illustrated by means of three case studies: i.e. a first order system, a second order SISO system and a multivariable simplified version of a heat exchanger.

Adaptations of the algorithm have also been applied to the parametric identification of discrete-time linear systems [Sira-Ramirez 2002] and to fault diagnosis [Fliess 2003], [Fliess 2004], [Join 2004]. This chapter documents one of the first industrial implementations of the fast estimation methods of Fliess and Sira-Ramirez [Fliess 2003]. An algorithm for estimation of parameters of first order dynamic models used in the sliding mode observer based condition monitoring scheme is tested with simulated and captured data from the iGX dry vacuum pump.

As will be seen, the parameter estimation algorithm proves to be useful in estimating nominal parameters for the sliding mode observer. The organisation of this chapter is as follows: An algebraic framework for parametric identification is presented in section 5.1. Section 5.2 covers the mathematical development of the parameter estimation algorithm and section 5.3 presents the performance of the proposed estimation framework. Finally, section 5.4 discusses the performance of the sliding mode observer scheme.

5.1 AN ALGEBRAIC FRAMEWORK FOR IDENTIFICATION

As shown in chapter 4, first order differential equations are used to model the heat transfer through the vacuum pump, providing a basis for the sliding mode observer design developed in section 4.4. To formulate the observer, initially the nominal values of the model parameters must be obtained. These nominal values can be obtained either by direct measurement or from tables of material properties. Generally however, some parameters may not be known and thus a reliable means for parameter estimation must be employed.

In this section, an algebraic framework for identification of continuous-time constant linear system parameters developed by Fliess and Sira-Ramirez [Fliess 2003] is presented. Consider the first order constant linear system

$$\dot{y}(t) = ay(t) + u(t) + \gamma(t) \quad (5.1)$$

where $u(t)$ and $y(t)$ are, respectively, the control and output variables. The constant parameter, a , is unknown and the additive perturbation $\gamma(t)$ is assumed to be constant but of unknown amplitude. Using the notation of operational calculus as used in [Fliess 2003] and [Fliess 2004] the above equation then reads

$$sy - y(0) = ay + u + \gamma/s \quad (5.2)$$

where s is the Laplace variable. Multiplying (5.2) first by s and then deriving twice with respect to s , will eliminate the additive perturbation and the initial condition.

Therefore, multiplying by s yields

$$s^2y = asy + su + sy(0) + \gamma \quad (5.3)$$

and taking derivatives, twice, with respect to s gives

$$\left[s \frac{d^2y}{ds^2} + 2 \frac{dy}{ds} \right] a = s^2 \frac{d^2y}{ds^2} + 4s \frac{dy}{ds} + 2y - \left(s \frac{d^2u}{ds^2} + 2 \frac{du}{ds} \right) \quad (5.4)$$

Derivations with respect to time in the estimator are avoided by multiplying by s^{-2}

$$\left[s^{-1} \frac{d^2 y}{ds^2} + 2s^{-2} \frac{dy}{ds} \right] a = \frac{d^2 y}{ds^2} + 4s^{-1} \frac{dy}{ds} + 2s^{-2} y - \left(s^{-1} \frac{d^2 u}{ds^2} + 2s^{-2} \frac{du}{ds} \right) \quad (5.5)$$

Finally taking the inverse Laplace transformation of (5.5) and rearranging yields the expression for on-line estimation of a , denoted \hat{a}

$$\hat{a} = \frac{t^2 y(t) - \int_0^t (4\sigma y(\sigma) + \sigma^2 u(\sigma)) d\sigma + 2 \int_0^t \int_0^\sigma (4\lambda u(\lambda) + y(\lambda)) d\lambda d\sigma}{\int_0^t \sigma^2 y(\sigma) d\sigma - 2 \int_0^t \int_0^\sigma \lambda y(\lambda) d\lambda d\sigma} \quad (5.6)$$

The algorithm has also been adopted and applied to the problem of fault diagnosis by Fliess et al [Fliess 2004]. It has been demonstrated that the algorithm is capable of detecting faults by residual generation in simulated first and second order systems where the model parameters are unknown constants. Nonetheless, the use of integrators in the algorithm means that a time-varying parameter causes changes in output variables which are carried forward as estimation errors. This characteristic of the estimation algorithm means that in its present form (i.e. without re-initialisation of integrators) it is unsuitable for detection of intermittent faults as will be illustrated later. In the following section a mathematical development of the parameter estimation algorithm for the specific dry vacuum pump heat transfer model is provided.

5.2 MATHEMATICAL DEVELOPMENT OF THE PARAMETER ESTIMATION ALGORITHM

Consider the dry vacuum pump heat transfer models introduced in chapter 4

$$\frac{dT_B(t)}{dt} = \frac{1}{(mc_p)_B} kI(t) - \frac{c_{pcw}}{(mc_p)_B} \dot{m}_c (T_o(t) - T_i(t)) - \frac{(hA)_B}{(mc_p)_B} (T_B(t) - T_{atm}(t)) \quad (5.7)$$

$$\frac{dT_o(t)}{dt} = \frac{A_c}{(mc_p)_{cw}} h_c (T_B(t) - T_o(t)) - \frac{c_{pcw}}{(mc_p)_{cw}} \dot{m}_c (T_o(t) - T_i(t)) \quad (5.8)$$

Recall that m_B is the mass of the pump body and c_{pB} is the specific heat capacity of the pump body. T_B , T_i , T_o and T_{atm} are the pump body, inlet, outlet and atmospheric temperatures respectively. Also, A_B represents the surface area of the pump body, A_c is the coolant heat transfer area, h_B the heat transfer coefficient of the pump body, h_c is the convective heat transfer coefficient of the coolant, \dot{m}_c the mass flow rate of coolant through pump, m_{cw} is the mass of the water coolant contained in the pump and c_{pcw} the specific heat capacity of the coolant. Finally, I is the pump inverter current and k is a scalar.

Variables I , T_o , T_i , T_{atm} , T_B and parameters m_{cw} , m_B , A_B and A_c are measurable.

Parameters c_{pB} and c_{pcw} are known and assumed constant, whilst parameters h_B and k can be obtained by experimentation. The algorithm aims at estimating parameters $h_c A_c$ and \dot{m}_c , which are required for the sliding mode observer based condition monitoring scheme.

Renaming parameters

$$a_1 = k / m_B c_{pB} \qquad b_1 = 1 / m_{cw} c_{pcw}$$

$$a_2 = c_{pcw} / m_B c_{pB} \qquad b_2 = c_{pcw} / m_{cw} c_{pcw} = 1 / m_{cw}$$

$$a_3 = h_B A_B / m_B c_{pB}$$

Renaming inputs and states

$$u_1 = I \qquad x_1 = T_B$$

$$u_2 = T_i \qquad x_2 = T_o$$

$$u_3 = T_a$$

Renaming variables for estimation

$$a = \dot{m} \dot{x}$$

$$b = h_c A_c$$

Now equations (5.7) and (5.8) can be written thus

$$\dot{x}_1 = a_1 u_1 - a_2 a x_2 + a_2 a u_2 - a_3 x_1 + a_3 u_3 \qquad (5.9)$$

$$\dot{x}_2 = b_1 b x_1 - b_1 b x_2 - b_2 a x_2 + b_2 a u_2 \qquad (5.10)$$

The outputs of the model are the states, so $y = x$. Taking the Laplace transform of (5.9) and rearranging

$$sy_1 - y_1(0) = -a_3y_1 - a_2ay_2 + a_1u_1 + a_2au_2 + a_3u_3 \quad (5.11)$$

Multiply both sides by s will eliminate the initial condition

$$s^2y_1 - sy_1(0) = -a_3sy_1 - a_2asy_2 + a_1su_1 + a_2asu_2 + a_3su_3 \quad (5.12)$$

Differentiate twice with respect to s

$$\begin{aligned} & 2y_1 + 4s \frac{dy_1}{ds} + s^2 \frac{d^2y_1}{ds^2} + 2a_3 \frac{dy_1}{ds} + a_3s \frac{d^2y_1}{ds^2} \\ & - 2a_3 \frac{du_3}{ds} - a_3s \frac{d^2u_3}{ds^2} - 2a_1 \frac{du_1}{ds} - a_1s \frac{d^2u_1}{ds^2} = a \left(-2a_2 \frac{dy_2}{ds} - a_2s \frac{d^2y_2}{ds^2} + 2a_2 \frac{du_2}{ds} + a_2s \frac{d^2u_2}{ds^2} \right) \end{aligned} \quad (5.13)$$

Multiply by s^{-2}

$$\begin{aligned} & 2s^{-2}y_1 + 4s^{-1} \frac{dy_1}{ds} + \frac{d^2y_1}{ds^2} + 2a_3s^{-2} \frac{dy_1}{ds} + a_3s^{-1} \frac{d^2y_1}{ds^2} \\ & - 2a_3s^{-2} \frac{du_3}{ds} - a_3s^{-1} \frac{d^2u_3}{ds^2} - 2a_1s^{-2} \frac{du_1}{ds} - a_1s^{-1} \frac{d^2u_1}{ds^2} \\ & + a \left(2a_2s^{-2} \frac{dy_2}{ds} + a_2s^{-1} \frac{d^2y_2}{ds^2} - 2a_2s^{-2} \frac{du_2}{ds} - a_2s^{-1} \frac{d^2u_2}{ds^2} \right) = 0 \end{aligned} \quad (5.14)$$

Find the inverse Laplace transform

$$\begin{aligned}
 & t^2 y_1(t) - 4 \int_0^t \sigma y_1(\sigma) d\sigma + 2 \int_0^t \int_0^\sigma y_1(\lambda) d\lambda d\sigma - 2a_3 \int_0^t \int_0^\sigma \lambda y_1(\lambda) d\lambda d\sigma + a_3 \int_0^t \sigma^2 y_1(\sigma) d\sigma \\
 & + 2a_3 \int_0^t \int_0^\sigma \lambda u_3(\lambda) d\lambda d\sigma - a_3 \int_0^t \sigma^2 u_3(\sigma) d\sigma + 2a_1 \int_0^t \int_0^\sigma \lambda u_1(\lambda) d\lambda d\sigma - a_1 \int_0^t \sigma^2 u_1(\sigma) d\sigma \\
 & + a \left(-2a_2 \int_0^t \int_0^\sigma \lambda y_2(\lambda) d\lambda d\sigma + a_2 \int_0^t \sigma^2 y_2(\sigma) d\sigma + 2a_2 \int_0^t \int_0^\sigma \lambda u_2(\lambda) d\lambda d\sigma - a_2 \int_0^t \sigma^2 u_2(\sigma) d\sigma \right) = 0
 \end{aligned} \tag{5.15}$$

Similarly for equation (5.10)

$$\begin{aligned}
 & t^2 y_2(t) - 4 \int_0^t \sigma y_2(\lambda) d\sigma + 2 \int_0^t \int_0^\sigma y_2(\lambda) d\lambda d\sigma \\
 & + a \left(-2b_2 \int_0^t \int_0^\sigma \lambda y_2(\lambda) d\lambda d\sigma + b_2 \int_0^t \sigma^2 y_2(\sigma) d\sigma + 2b_2 \int_0^t \int_0^\sigma \lambda u_2(\lambda) d\lambda d\sigma - b_2 \int_0^t \sigma^2 u_2(\sigma) d\sigma \right) \\
 & - b \left(-2b_1 \int_0^t \int_0^\sigma \lambda y_1(\lambda) d\lambda d\sigma + b_1 \int_0^t \sigma^2 y_1(\sigma) d\sigma + 2b_1 \int_0^t \int_0^\sigma \lambda y_2(\lambda) d\lambda d\sigma - b_1 \int_0^t \sigma^2 y_2(\sigma) d\sigma \right) = 0
 \end{aligned} \tag{5.16}$$

Renaming the variables

$$\begin{aligned}
 q_1(t) = & t^2 y_1(t) - 4 \int_0^t \lambda y_1(\lambda) d\sigma + 2 \int_0^t \int_0^\sigma y_1(\lambda) d\lambda d\sigma - 2a_3 \int_0^t \int_0^\sigma \lambda y_1(\lambda) d\lambda d\sigma + a_3 \int_0^t \sigma^2 y_1(\sigma) d\sigma \\
 & + 2a_3 \int_0^t \int_0^\sigma \lambda u_3(\lambda) d\lambda d\sigma - a_3 \int_0^t \sigma^2 u_3(\sigma) d\sigma + 2a_1 \int_0^t \int_0^\sigma \lambda u_1(\lambda) d\lambda d\sigma - a_1 \int_0^t \sigma^2 u_1(\sigma) d\sigma
 \end{aligned}$$

$$q_2(t) = t^2 y_2(t) - 4 \int_0^t \lambda y_2(\lambda) d\sigma + 2 \int_0^t \int_0^\sigma y_2(\lambda) d\lambda d\sigma$$

$$\pi_1(t) = -2a_2 \int_0^t \int_0^\sigma \lambda y_2(\lambda) d\lambda d\sigma + a_2 \int_0^t \sigma^2 y_2(\sigma) d\sigma + 2a_2 \int_0^t \int_0^\sigma \lambda u_2(\lambda) d\lambda d\sigma - a_2 \int_0^t \sigma^2 u_2(\sigma) d\sigma$$

$$\pi_2(t) = -2b_2 \int_0^\sigma \int_0^\sigma \lambda y_2(\lambda) d\lambda d\sigma + b_2 \int_0^\sigma \sigma^2 y_2(\sigma) d\sigma + 2b_2 \int_0^\sigma \int_0^\sigma \lambda u_2(\lambda) d\lambda d\sigma - b_2 \int_0^\sigma \sigma^2 u_2(\sigma) d\sigma$$

$$\pi_3(t) = -2b_1 \int_0^\sigma \int_0^\sigma \lambda y_1(\lambda) d\lambda d\sigma + b_1 \int_0^\sigma \sigma^2 y_1(\sigma) d\sigma + 2b_1 \int_0^\sigma \int_0^\sigma \lambda y_2(\lambda) d\lambda d\sigma - b_1 \int_0^\sigma \sigma^2 y_2(\sigma) d\sigma$$

Therefore equations (5.15) and (5.16) are now reduced to

$$q_1(t) + \hat{a}\pi_1(t) = 0 \quad (5.17)$$

$$q_2(t) + \hat{a}\pi_2(t) - \hat{b}\pi_3(t) = 0 \quad (5.18)$$

With a small modification of notation for clarity

$$\begin{bmatrix} \pi_1 & 0 \\ \pi_2 & \pi_3 \end{bmatrix} \begin{bmatrix} \hat{a} \\ -\hat{b} \end{bmatrix} = \begin{bmatrix} -q_1 \\ -q_2 \end{bmatrix} \quad (5.19)$$

which can be written in matrix notation as

$$\Pi R = Q \quad (5.20)$$

Thus, the estimated parameters $R = [\hat{a}, -\hat{b}]^T$ are obtained from

$$R = \Pi^{-1}Q \quad (5.21)$$

5.3 PERFORMANCE OF THE PROPOSED ESTIMATION FRAMEWORK

The estimation strategy developed in the previous section will now be implemented and tested with both simulated data from a nominal model and also using measured pump data.

5.3.1 SIMULATION OF THE MODEL WITH TIME-INVARIANT PARAMETERS

A computer simulation of the heat transfer model equations (5.7) and (5.8) has been implemented using assumed parameters along with the parameter estimation algorithm (5.21). The purpose of the simulation is to test the efficacy of the parameter estimation algorithm and the correctness of its implementation. Estimation results for the test are presented in Figure 5-1. The estimated parameters converge to the actual parameters after a period of approximately 6000 seconds, demonstrating that the algorithm and the implementation are correct.

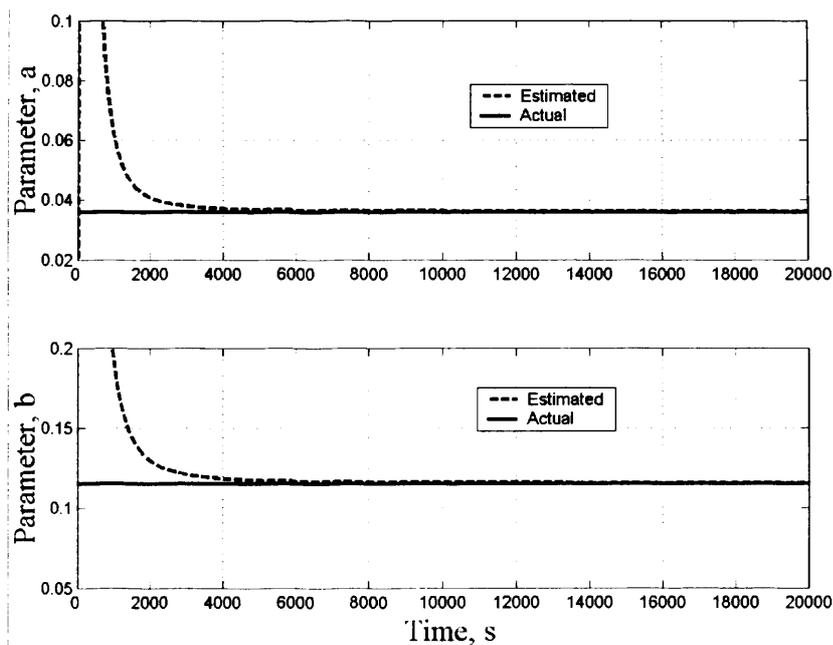


Figure 5-1: Estimates for time-invariant parameters a and b

5.3.2 SIMULATION OF THE MODEL WITH TIME-VARYING PARAMETERS

Next the algorithm is tested with a time-varying parameter as might be expected in a real pump cooling system, e.g. due to temperature variations or a sudden fault condition. In the event of a sudden 10% reduction in parameter (see Figure 5-2), the estimation results for parameter a are inaccurate even when the parameter is returned to its initial value, i.e. the estimate does not re-converge. Where the model parameters undergo a small exponential increase, the neither estimate (see Figure 5-3) converged on the correct value. These results demonstrate that the algebraic framework for parameter estimation cannot be employed for condition monitoring within this application domain.

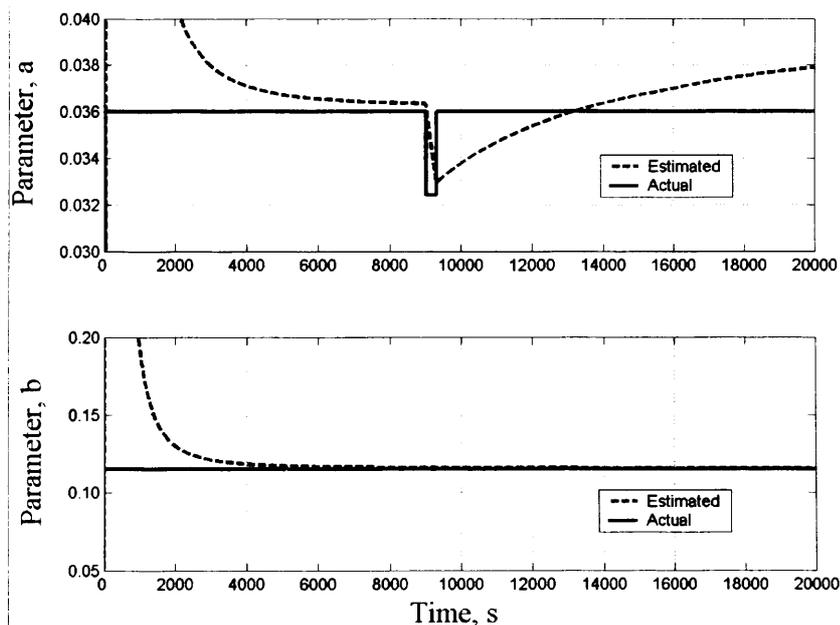


Figure 5-2: Parameter estimates in the event of a simulated 10% deviation in coolant flow rate

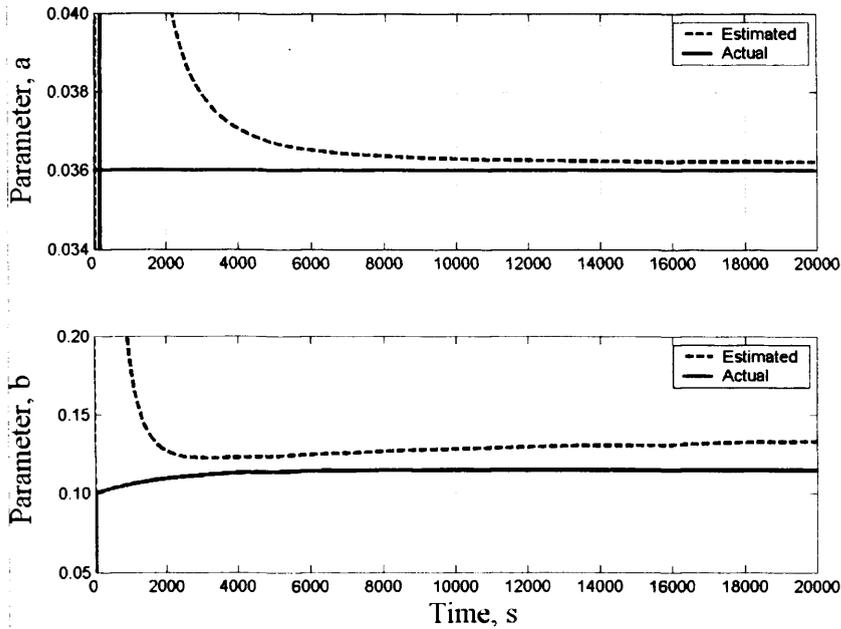


Figure 5-3: Parameter estimated in the event of a simulated temperature sensitive parameter *b*. The parameter undergoes an exponential rise of 15%

5.3.3 ESTIMATION OF PUMP MODEL PARAMETERS FROM CAPTURED PUMP DATA

Temperature data is captured at a rate of 1 Hz from the iGX dry vacuum pump running at ultimate pressure over a period of approximately 13000 seconds. The captured data is then used to estimate parameters *a* and *b*, using the algorithm as described in (5.21). Results are presented in Figure 5-4. The estimated mass flow rate is 0.014 Kg/s, 7% lower than the actual value of 0.015. The actual value of $h_c A_c$, cannot be ascertained directly although an approximation can be obtained as follows.

Consider a steady state analysis of (5.8)

$$\begin{aligned}
 0 &= A_c h_c (T_B(t) - T_o(t)) - \dot{m} c_{pcw} (T_o(t) - T_i(t)) \Rightarrow \\
 A_c h_c &= \frac{\dot{m} c_{pcw} (T_o(t) - T_i(t))}{(T_B(t) - T_o(t))}
 \end{aligned}
 \tag{5.22}$$

Inputting measured and known values gives an approximate value for $h_c A_c$

$$A_c h_c = \frac{0.015 \times 4.2 \times (20)}{(21)} = 0.06 (\text{kWK}^{-1})$$

In Figure 5-4 the parameter estimate, b , is seen to be close to this steady state approximation.

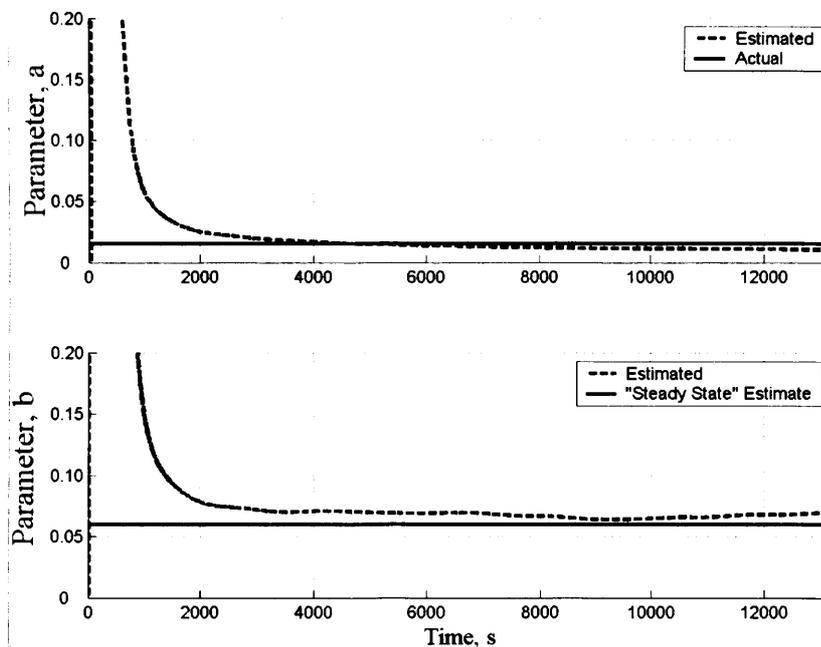


Figure 5-4: Parameter estimates from captured data

5.3.4 OVERVIEW OF PARAMETER ESTIMATION RESULTS

For the purposes of the estimation experiment the coolant flow rate is measured and maintained at a constant value. Results using captured data showed estimated parameters close to those obtained by measurement and steady state analysis of the model. The accuracy of the results suggests that the heat transfer coefficient is approximately constant during the course of the experiment. It can be seen in Figure 5-

4 that for the vacuum pump, the algorithm takes a significant time (more than 1 hour) to converge to the estimates.

Care must be taken when making assumptions about the nature of real dynamical systems. For instance, in advance of experiments it might not be known whether the parameters are indeed time invariant. Some form of parameter checking, such as the steady state estimate presented in section 5.3.3, is required to ensure that the assumptions are true and that the estimates seem reasonable.

Values for integrands calculated in the algorithm can become very large. This raises questions about the practicability of the scheme for models with large time constants, or for continuous monitoring processes, such as the pump system under consideration in this thesis. The software used to implement the algorithm must be capable of accurately processing such values which attained a magnitude of 10^{12} during the course of the experiments presented in this chapter (maximum duration 6 hours).

5.4 THE PERFORMANCE OF THE SLIDING MODE OBSERVER SCHEME

The results in this section are obtained by implementing the sliding mode observer developed in section 4.4 with the parameters estimated using the identification algorithm. Results from a fault free data set presented in Figure 5-5, clearly show that the parameter estimates are sufficiently close to the actual values for time average of the fault signal, Δn_k , to be close to zero. Figure 5-6 shows that in the event of a sudden reduction in the coolant flow rate, the fault signal, Δn_k , is perturbed from zero then returns to zero when full coolant flow is reinstated.

The above results undoubtedly demonstrate (as also seen in chapter 4) that the sliding mode observer scheme offers benefits over the estimation algorithm in the role of continuous condition monitoring in systems where intermittent faults are possible. However, the parameter estimation algorithm proved useful in estimating time invariant nominal parameters for formulation of the observer.

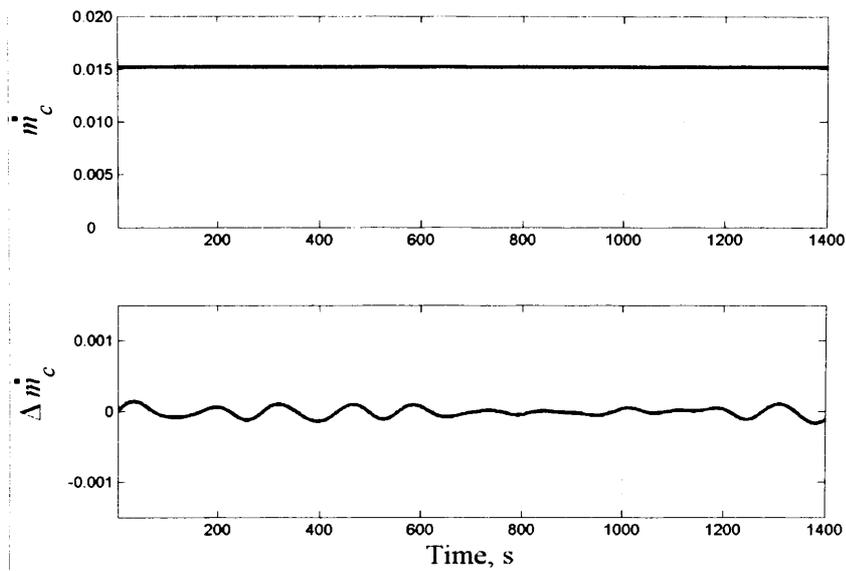


Figure 5-5: Recorder coolant flow rate, \dot{m}_c and observer fault signal, $\Delta \dot{m}_c$

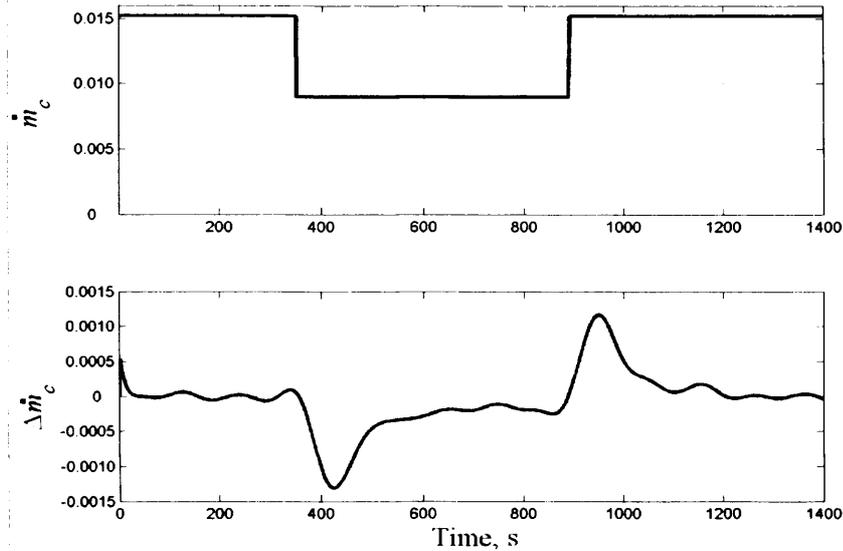


Figure 5-6: Recorded coolant flow rate, \dot{m}_c and observer fault signal, $\Delta \dot{m}_c$ in the event of a 40% reduction in coolant flow rate

5.5 CONCLUSIONS

An algebraic framework for parameter identification has been presented in this chapter. The design of a sliding mode observer requires a nominal mathematical model. If the observer methodology is to be routinely implemented in industry, automated methods are needed for parameter measurement. The developed parameter estimation algorithm has been implemented and tested on both simulated and real vacuum pump data.

The results verified the effectiveness of the algorithm in estimating time invariant nominal parameters for formulation of the sliding mode observer. Nevertheless, when the algorithm is tested with a time-varying parameter, it does not perform as well. The estimates are sensitive to even small parameter variations because the use of integrator means that the effects of past changes are carried forward into current estimates. In contrast, the sliding mode observer offered clear benefits for continuous condition monitoring in the vacuum pump system where intermittent faults are possible.

CHAPTER 6

CASE STUDY 2: THE ROLLING ELEMENT BALL BEARINGS

The present day need for increasing reliability and availability in the area of rotating machinery is now more significant than ever before and continues to grow constantly. An ordinary component for every rotating machine and of particular importance is the bearings. Healthy bearings can ensure that the machinery will run quietly and efficiently. Bearing defects occurring during manufacture or operation may lead to failure and expense to both operator and manufacturer. Continuously monitoring of the state of the bearings and scheduling maintenance as appropriate is thus desirable.

The literature on the detection of bearing defects is very extensive and mostly concentrates on analysing vibration signals. Sunnersjo [Sunnersjo 1978] reported that bearings will generate vibration during operation even if they are geometrically and elastically perfect. This is because the number of rolling elements that carry the external load is finite, which gives rise to a periodic variation of the total stiffness of the bearing assembly and consequently generates vibrations. Generally, vibration data are collected by mounting a transducer (e.g. accelerometer) externally on the machinery.

The collected vibration measurements can then be analysed in the time or frequency domain.

A simple method of time domain analysis is the investigation of the crest factor. This parameter is a measure of the ratio of the peak acceleration to the root mean square (RMS) acceleration. Mathew and Alfredson [Mathew 1984] indicated that as a bearing wears out, the peak levels of acceleration increase more rapidly than the RMS levels. Other techniques in the time domain include measurements of the probability density of acceleration or the use of a statistical parameter called kurtosis [Tandon 1999]. While time domain analysis has gained some industrial acceptance, frequency domain or spectral analysis is most widely used, since the location of defects can be detected. The basic indicator in spectral analysis is the investigation of the characteristic defect frequencies. These frequencies can be derived theoretically by considering the geometry of a bearing and its rotational frequency [Prashad 1987]. The existence of one of the characteristic defect frequencies may thus infer a fault. However, sometimes it is difficult to acquire any significant peaks at these frequencies due to noise or vibration from other sources [Kiral 2003]. An alternative and more powerful method is the high-frequency resonance technique (HFRT). The basis of the approach is that each time a defect in a bearing makes contact under load with another surface in the bearing, a short duration impulse is generated that excites a resonance at higher frequencies. The process of demodulating and analysing such signals by the HFRT is shown in McFadden and Smith [McFadden 1984]. Finally, parametric model-based spectral estimation techniques have been used with success in order to analyse vibration signals from bearings [Mechefske 1993].

In recent times, another method that attracted the attention of researchers for bearing monitoring is the analysis of acoustic signals such as noise and emission response. External microphones located adjacent to a system can be utilised to gather acoustic measurements. The analysis of acoustic noise is normally conducted by measuring the sound pressure or sound intensity, whereas for acoustic emission analysis parameters like ring-down counts and peak amplitude are measured. The reader can find a more detailed review for acoustic methods in Tandon and Choudhury [Tandon 1999].

Having identified the most common methods for bearing monitoring, an alternative approach based on sliding mode techniques is examined in this chapter. The proposed method employs first and second order non-linear sliding mode observers in order to monitor and analyse the system states. The equivalent injection signal necessary to maintain a sliding motion is utilised for parameter estimation and hence fault detection in the system. This peculiar characteristic of fault, or parameter, reconstruction along with the excellent system performance which includes insensitivity to certain kinds of parameter variations harnesses the unique capabilities of sliding mode techniques. Nonetheless, a possible drawback of the classical, first order sliding mode approach is the chattering effect. This chattering phenomenon can be perceived as high frequency oscillations in the state estimates and applied injection signal. Therefore, it is necessary to eliminate or suppress this chattering to recover the parameter or fault estimates from the applied injection signal.

In chapter 3, a number of methods were presented that alleviated chattering, though the main feature of sliding modes (i.e. robustness) is compromised. Higher order sliding modes (HOSMs) are a generalisation of classical sliding modes (CSMs) and have been

developed to tackle the problem of chattering while they preserve their main features. Further to this, HOSMs have shown an improved tracking (sliding) accuracy under sliding motion when compared to CMSs [Levant 1993]. A number of second order sliding mode algorithms are presented in Levant [Levant 1993] and Emelyanov et al [Emelyanov 1996]. It is reported that a second order algorithm needs to keep exactly two constraint conditions (i.e. $s = \dot{s} = 0$) instead of one (i.e. $s = 0$), as with CSM the discontinuity acts on the first derivative of s , whereas with a HOSM the discontinuity acts on the r^{th} derivative of s , where r is the order of the sliding mode scheme. Clearly, HOSM algorithms not only require the knowledge of the sliding variable s , but also need to know certain numbers of the derivatives of the sliding variable, which may often be unavailable. Levant [Levant 2001] introduced the design of arbitrary order sliding mode algorithms that obtain the time derivatives of the sliding variable via robust real time differentiators. These algorithms can be considered as general-purpose controllers for single-input-single-output (SISO) dynamic systems.

Although arbitrary order sliding mode algorithms are mainly still studied theoretically [Levant 2003], many of the second order sliding mode algorithms have been successfully applied to real-life problems. Levant et al [Levant 2000] demonstrated the first practical application of a second order sliding mode technique for an aircraft pitch control problem. Goh et al [Goh 2003] applied and tested the robustness of a second order sliding mode algorithm for speed control of a diesel engine generator set. The suggested method is compared with a commercial controller and a classical proportional-integral (PI) controller. The obtained results show that the second order sliding mode controller outperforms the PI controller and that it has a similar performance to the commercial controller. Finally, first and second order sliding mode

algorithms were designed by Shtessel and Poznyak [Shtessel 2004] for estimation of the parameters of a linear time varying (LTV) system. It is revealed that the second order algorithm is capable of suppressing the discontinuous nature of the applied injection signal and filtering is not required.

It will be seen in this chapter that the developed sliding mode observer based method is able to detect bearing faults by examining the equivalent injection signal. Moreover, the HOSM algorithm drastically improves the performance of the fault diagnosis scheme when compared with a CSM algorithm.

The organisation of this chapter is as follows: Section 6.1 formulates the fault prediction problem and section 6.2 considers the development of the heat transfer models for the bearings. A fault scenario for defective bearings is introduced in section 6.3. Section 6.4 presents the development of a first and second order sliding mode algorithm. The experimental set-up and the observer performance with captured data from an overheated bearing are discussed in sections 6.5 and 6.6 respectively.

6.1 PROBLEM FORMULATION

Based on the review of various methods for the detection of defects in bearings discussed in the introduction, the choice of suitable monitoring signals and transducers is now considered. Factors that affect this choice are mainly the cost and functionality. The effectiveness and the value of vibration and acoustic measurement techniques is fully acknowledged, but the complexity of developing implementable schemes can produce problems. For example, in the analysis of vibration or acoustic signals, the combined cost of transducers and high bandwidth signal processing can mean that on-

line diagnostics are not commercially viable. External environmental noise can also be a problem when implementing the methods on real systems and a noise cancellation technique has to be incorporated in the scheme.

A typical symptom of a defective bearing, in addition to the increased noise and vibration, will be a rise in the bearing temperature due to friction. Inexpensive temperature transducers can provide inputs to a heat transfer model of the iGX dry vacuum pump. Moreover, the nature of the input signal dynamics is slow and a low sample rate for the heat transfer model is adequate. Therefore, observation and analysis of temperature measurements seems an effective choice for an on board fault diagnosis system. Nevertheless, it is known that temperature signals may only provide an indication of faults and not conclusive results (see chapter 4). The condition monitoring scheme thus should also monitor symptomatic parameters in order to infer a fault. Sliding mode techniques allow this with the use of the equivalent injection signal, which estimates and detects changes in parameters that are often difficult to measure.

Once suitable monitoring signals and instrumentation has been identified, the next phase will comprise the development of the heat transfer models and sliding mode observers. Details of these developments are given in the following sections.

6.2 HEAT-TRANSFER MODELS FOR THE ROLLING ELEMENT BALL BEARINGS

The successful operation of the iGX dry vacuum pump is very much dependant on the effective heat dissipation from the motor, bearings and pump stages. A water cooling system is thus installed to maintain the vacuum pump at an appropriate working

temperature. During the second phase of this project the design of the cooling system has evolved. An extra component (i.e. temperature controlled valve) has been incorporated to regulate the temperature at more precise limits. This prevents the pump body from both over-heating and over-cooling, avoiding auto-ignition and condensation within the pump of the operational gas. Figure 6-1 illustrates the pump cooling system. At start-up, the cooling water will go through the motor but it bypasses the stator cooling plates to allow the pump to quickly attain the correct operating temperature. When the set temperature is reached, the valve will open allowing water to cool the stator.

The iGX dry vacuum pump has a set of rolling element ball bearings fitted at both the low vacuum (i.e. outlet) and high vacuum (i.e. inlet) ends (referred to as LV and HV ends respectively). Two mathematical models that describe the energy balance for the LV and HV end bearings are now described.

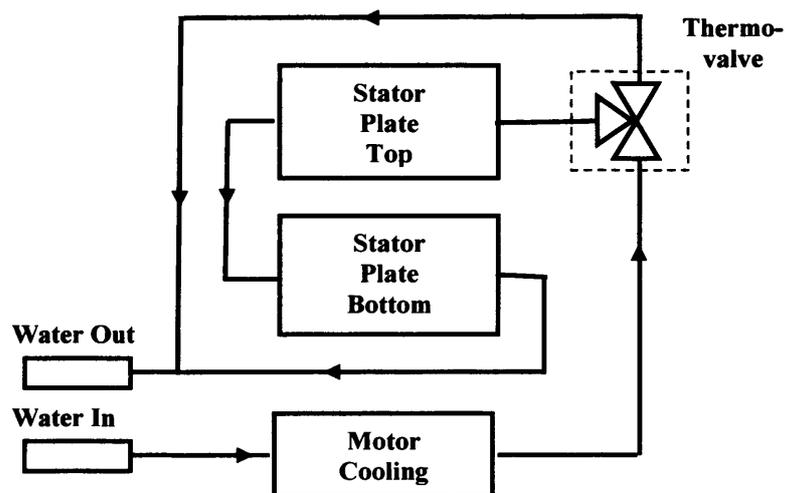


Figure 6-1: Dry vacuum pump cooling system diagram

6.2.1 LOW VACUUM BEARING TEMPERATURE MODEL

The primary source of heat to the pump is the electrical power supply Q_p . Heat is dissipated by convection Q_{CONV} and radiation Q_{RAD} from the pump body, and heat transfer to the cooling water Q_{CW} and pumped gases Q_{PG} . Let T_{LV} be the LV end bearing temperature, T_{out} the coolant outlet temperature, T_{in} the coolant inlet temperature and T_{atm} the atmospheric temperature. A mathematical expression for the rate of change of temperature for the LV bearing is given by

$$(mc_p)_{LV} \frac{dT_{LV}(t)}{dt} = Q_p - Q_{CW} - Q_{CONV} - Q_{RAD} - Q_{PG} \quad (6.1)$$

Radiated heat and the net mass flow rate of pumped gas are both small, therefore, Q_{RAD} and Q_{PG} can be neglected. Also, Q_p is assumed to be a linear function of the inverter current I and $Q_p = k_1 I$, where k_1 is a constant. Further, the cooling water heat transfer

$$Q_{CW} = m \dot{\xi} c_{pcw} (T_{out}(t) - T_{in}(t)) \quad (6.2)$$

and the surface heat loss to atmosphere through convection

$$Q_{CONV} = (hA)_{LV} (T_{LV}(t) - T_{atm}(t)) \quad (6.3)$$

It should be noted that, m_{LV} is the mass of the pump body at the LV end and c_{pLV} is the specific heat capacity. Also, A_{LV} represents the surface area of the pump body at the LV

end, h_{LV} the heat transfer coefficient, \dot{m}_c the mass flow rate of the coolant through pump and c_{pcw} the specific heat capacity of the coolant.

Substituting from (6.3), (6.2) in (6.1) and rearranging will result in the following equation

$$\frac{dT_{LV}(t)}{dt} = \frac{k_1}{(mc_p)_{LV}} I(t) - \frac{\dot{m}_c c_{pcw}}{(mc_p)_{LV}} (T_{out}(t) - T_{in}(t)) - \frac{(hA)_{LV}}{(mc_p)_{LV}} (T_{LV}(t) - T_{atm}(t)) \quad (6.4)$$

Re-labelling the coefficients for ease of exposition

$$\dot{T}_{LV}(t) = a_1 k_1 I(t) - a_2 \dot{m}_c (T_{out}(t) - T_{in}(t)) - a_3 (T_{LV}(t) - T_{atm}(t)) \quad (6.5)$$

where a_1 , a_2 and a_3 are given by $a_1 = \frac{1}{(mc_p)_{LV}}$, $a_2 = \frac{c_{pcw}}{(mc_p)_{LV}}$ and $a_3 = \frac{(hA)_{LV}}{(mc_p)_{LV}}$.

The operation of the temperature-controlled valve introduces an extra degree of non linearity to the nominal heat transfer model. Equation (6.5) is thus modified since the model parameter a_1 is a function of the temperature difference between the outlet and inlet coolant temperature (i.e. $T_{out}(t) - T_{in}(t) = \Delta T(t)$).

$$\begin{aligned} \dot{T}_{LV}(t) &= x_1(t) - a_2 \dot{m}_c x_2(t) - a_3 (T_{LV}(t) - T_{atm}(t)) \\ 250 \dot{x}_2(t) + x_2(t) &= T_{out}(t) - T_{in}(t) \\ 100 \dot{x}_1(t) + x_1(t) &= k_1 P_i(\Delta T) I(t) \end{aligned} \quad (6.6)$$

where $P_i(\Delta T)$ denotes the temperature dependent function given by

$$P_i(\Delta T) = \begin{cases} 1 \times 10^{-5} \Delta T^3(t) - 2.3 \times 10^{-4} \Delta T^2(t) \\ + 1.7 \times 10^{-3} \Delta T(t) - 5.5 \times 10^{-5} & \Delta T(t) > 2.4 \\ - 2.5 \times 10^{-5} \Delta T^3(t) + 6.9 \times 10^{-4} \Delta T^2(t) \\ - 5.2 \times 10^{-3} \Delta T(t) + 1.8 \times 10^{-2} & \Delta T(t) < 2.4 \end{cases} \quad (6.7)$$

Details of the procedure to obtain the above equation are reported in Appendix C3.

Further, $x_1(t)$ and $x_2(t)$ represent the filter dynamics introduced to the model. These dynamics are an approximation of the thermal lag (where thermal lag is defined as the delay in the heat distribution through a system) of the dry vacuum pump.

6.2.2 HIGH VACUUM BEARING TEMPERATURE MODEL

The HV bearing temperature model is formulated to provide a nominal model for the design of the sliding mode observer. The model has been formulated using the same procedure as for the LV bearing temperature model. Nonetheless, the HV end bearings are further away from the cooling circuit than the LV end bearings. Therefore, operation of the temperature-controlled valve has less effect on the HV bearing. The rate of change of temperature for the HV bearing is given by

$$\frac{dT_{HV}(t)}{dt} = \frac{k_2}{(mc_p)_{HV}} I(t) - \frac{n\dot{c}_{pcw}}{(mc_p)_{HV}} (T_{out}(t) - T_{in}(t)) - \frac{(hA)_{HV}}{(mc_p)_{HV}} (T_{HV}(t) - T_{atm}(t)) \quad (6.8)$$

where T_{HV} , T_{out} , T_{in} and T_{atm} are the HV end bearing, outlet, inlet and atmospheric temperatures respectively. Moreover, k_2 is a constant and I the inverter's current, m_{HV} is the mass of the pump body at the HV end and c_{pHV} is the specific heat capacity. Also,

A_{HV} represents the surface area of the pump body at the HV end, h_{HV} the heat transfer coefficient, \dot{m}_c the mass flow rate of the coolant through pump and c_{pcw} the specific heat capacity of the coolant. Re-labelling the coefficients and introducing $z_1(t)$ and $z_2(t)$ as the filter dynamics at the HV end will result in

$$\begin{aligned} \dot{T}_{HV}(t) &= z_1(t) - b_2 \dot{m}_c z_2(t) - b_3 (T_{HV}(t) - T_{atm}(t)) \\ 1500 \dot{z}_2(t) + z_2(t) &= T_{out}(t) - T_{in}(t) \\ 250 \dot{z}_1(t) + z_1(t) &= k_2 b_1 I(t) \end{aligned} \quad (6.9)$$

where b_1 , b_2 and b_3 are given by $b_1 = \frac{1}{(mc_p)_{HV}}$, $b_2 = \frac{c_{pcw}}{(mc_p)_{HV}}$ and $b_3 = \frac{(hA)_{HV}}{(mc_p)_{HV}}$.

6.3 FAULT SCENARIO FOR DEFECTIVE BEARINGS

A worn bearing will cause a rise in temperature due to friction. This fault scenario can be simulated by attaching a small heating element external to the bearing case at both ends. Equations (6.6) and (6.9) represent the dynamics of the LV and HV end bearings. A deviation in the system parameters is defined to indicate malfunction of the bearings.

Let

$$k_1 = \tilde{k}_1 + \Delta k_1 \quad (6.10a)$$

$$k_2 = \tilde{k}_2 + \Delta k_2 \quad (6.10b)$$

where Δk_1 , Δk_2 represent the deviations (i.e. the faults) and \tilde{k}_1 , \tilde{k}_2 the nominal parameters. Substituting the above equations in equations (6.6) and (6.9) gives

$$\begin{aligned}
\dot{T}_{LV}(t) &= x_1(t) - a_2 m_{\xi} x_2(t) - a_3 (T_{LV}(t) - T_{am}(t)) \\
250 \dot{x}_2(t) + x_2(t) &= T_{out}(t) - T_{in}(t) \\
100 \dot{x}_1(t) + x_1(t) &= (\tilde{k}_1 + \Delta k_1) P_i(\Delta T) I(t)
\end{aligned} \tag{6.11}$$

$$\begin{aligned}
\dot{T}_{HV}(t) &= z_1(t) - b_2 m_{\xi} z_2(t) - b_3 (T_{HV}(t) - T_{am}(t)) \\
1500 \dot{z}_2(t) + z_2(t) &= T_{out}(t) - T_{in}(t) \\
250 \dot{z}_1(t) + z_1(t) &= (\tilde{k}_2 + \Delta k_2) b_1 I(t)
\end{aligned} \tag{6.12}$$

Clearly, by setting the deviations Δk_1 and Δk_2 to zero in equations (6.11) and (6.12), the nominal dynamics can be obtained.

6.4 SLIDING MODE OBSERVER FOR PARAMETER ESTIMATION AND FAULT PREDICTION

This section is concerned with the development of sliding mode observers which are used for parameter estimation and fault prediction of overheated bearings. Assuming that the output of the plant is the LV and HV end bearing temperatures, then the sliding surface is defined as the error between the observer output and the plant output. The discontinuous injection signal is chosen to force the system to reach and remain on the sliding surface. It will be seen that the use of this discontinuous signal can provide the information needed for predicting defective bearings. The proposed observers for both LV and HV ends have the following structure

$$\begin{aligned}
\hat{\dot{T}}_{LV}(t) &= -a_3 \hat{T}_{LV}(t) + \hat{x}_1(t) - a_2 m_{\xi} x_2(t) + a_3 T_{am}(t) + v_{LV} \\
250 \hat{\dot{x}}_2(t) + \hat{x}_2(t) &= T_{out}(t) - T_{in}(t) \\
100 \hat{\dot{x}}_1(t) + \hat{x}_1(t) &= \tilde{k}_1 P_i(\Delta T) I(t)
\end{aligned} \tag{6.13}$$

$$\begin{aligned}
\hat{T}_{HV}(t) &= -b_3 \hat{T}_{HV}(t) + \hat{z}_1(t) - b_2 \tilde{z}_2(t) + b_3 T_{am}(t) + \nu_{HV} \\
1500 \hat{z}_2(t) + \hat{z}_2(t) &= T_{out}(t) - T_{in}(t) \\
250 \hat{z}_1(t) + \hat{z}_1(t) &= \tilde{k}_2 b_1 I(t)
\end{aligned} \tag{6.14}$$

where ν_i is the observer injection signal.

6.4.1 A FIRST ORDER SLIDING MODE ALGORITHM

For the first order algorithm define $\nu_i = K_i (\varepsilon_i / \|\varepsilon_i\| + \delta)$, $i = LV, HV$ where K_i are the gains of the discontinuous signals ν_i . The observer error ε_i is defined as the difference between the estimated and measured temperatures (i.e. $\varepsilon_{LV} = T_{LV} - \hat{T}_{LV}$ and $\varepsilon_{HV} = T_{HV} - \hat{T}_{HV}$). The two gains K_{LV} and K_{HV} must be chosen so that the reachability problem is satisfied and the sliding motion is sustained even in the presence of faults. Finally, a small positive constant $\delta = 0.05$ is used for suppression of chattering.

In the sliding mode the following two equations give the observer error dynamics

$$\begin{aligned}
\mathcal{E}_{LV} &= -a_3 \varepsilon_{LV} + \tilde{x}_1(t) - \nu_{LV} \\
100 \tilde{x}_1(t) + \tilde{x}_1(t) &= \Delta k_1 P_i (\Delta T) I(t)
\end{aligned} \tag{6.15}$$

$$\begin{aligned}
\mathcal{E}_{HV} &= -b_3 \varepsilon_{HV} + \tilde{z}_1(t) - \nu_{HV} \\
250 \tilde{z}_1(t) + \tilde{z}_1(t) &= \Delta k_2 b_1 I(t)
\end{aligned} \tag{6.16}$$

where $\tilde{x}_1 = x_1 - \hat{x}_1$ and $\tilde{z}_1 = z_1 - \hat{z}_1$ represent the filter dynamics introduced to the LV and HV bearing temperature models.

Similarly to chapter 4, Lyapunov stability analysis can be utilised to prove stability of the observer system. Assuming sufficiently large gains K_i are chosen, a sliding motion will be attained and maintained. In this case the observer error and its derivatives will converge to zero due to the choice of sliding surface. Hence, the above equations become

$$\begin{aligned} 0 &= \tilde{x}_1(t) - v_{LV} \\ 100\tilde{x}_1(t) + \tilde{x}_1(t) &= \Delta k_1 P_i(\Delta T) I(t) \end{aligned} \quad (6.17)$$

$$\begin{aligned} 0 &= \tilde{z}_1(t) - v_{HV} \\ 250\tilde{z}_1(t) + \tilde{z}_1(t) &= \Delta k_2 b_1 I(t) \end{aligned} \quad (6.18)$$

Thus the discontinuous injection signal provides a means of detecting parameter changes in the heat transfer model symptomatic of fault conditions.

6.4.2 A SECOND ORDER SLIDING MODE ALGORITHM

A general review of the most popular second order sliding mode algorithms is presented in Levant [Levant 1993]. Based on this review, it can be noticed that the majority of the algorithms require the knowledge of the time derivative of the sliding variable. This is often a problem in implementing such algorithms due to the increased information demand. The only second order sliding mode algorithm that does not explicitly use the derivative of the sliding variable is the so-called ‘super-twisting’ algorithm. In the ‘super-twisting’ algorithm the discontinuous injection signal is constituted by two terms, an integral term and a term given by a discontinuous function of the sliding variable. Therefore, the ‘super-twisting’ algorithm is defined as follows

$$v = v_1 + v_2 \quad (6.19a)$$

$$v_1 = \begin{cases} -v, & |v| > 1 \\ -W \operatorname{sgn}(s), & |v| \leq 1 \end{cases} \quad (6.19b)$$

$$v_2 = \begin{cases} -\lambda |s_0|^\rho \operatorname{sgn}(s), & |s| > s_0 \\ -\lambda |s|^\rho \operatorname{sgn}(s), & |s| \leq s_0 \end{cases} \quad (6.19c)$$

where $W, \lambda > 0, \rho \in (0,1)$, and sufficient conditions for finite time convergence are

$$W > \frac{\Phi}{\Gamma_m} > 0 \quad (6.20a)$$

$$\lambda^2 \geq \frac{4\Phi\Gamma_m(W + \Phi)}{\Gamma_m^3(W - \Phi)} \quad (6.20b)$$

$$0 < \rho \leq 0.5 \quad (6.20c)$$

The main advantages of the ‘super-twisting’ algorithm are its robustness and that it does not require any information on the time derivative of s . In contrast, the discontinuous injection signal it produces is not Lipschitz when s is small, which may result in noise on the injection signal.

A simplified version of the ‘super-twisting’ algorithm has been adopted from Shtessel and Poznyak [Shtessel 2004] and employed here for the particular problem of parameter estimation and fault prediction.

$$\begin{cases} v_i = \lambda |\varepsilon_i|^{1/2} \operatorname{sgn}(\varepsilon_i) + W v \\ \dot{\varepsilon}_i = \operatorname{sgn}(\varepsilon_i) \end{cases} \quad (6.21)$$

where ε_i is the sliding variable, $i =_{LV, HV}$, λ and W the gain parameters. Equation (6.21) will also ensure $\varepsilon_i \rightarrow 0$ and thus the same argument employed in 6.4.1 can be used to show that the corresponding v_i will provide an estimate of any parameter changes.

6.5 EXPERIMENTAL SET-UP

To examine the effectiveness of the sliding mode observer based fault detection scheme the iGX dry vacuum pump should be first carefully instrumented. Five integrated-circuit sensors are used to monitor temperature signals along with a flow meter that allowed the measurement of the coolant flow rate. In addition, the motor current signal is captured from the general purpose inverter which is part of the system.

The raw signals from the vacuum pump are acquired using a dSPACETM DS1103 real-time processor board that is interfaced to a personal computer. This hardware set-up provides a facility for data collection and data transmission to MatlabTM/SimulinkTM.

6.6 RESULTS

The parameter estimation and fault prediction system is realized by using the algorithms described in section 6.4, and their performance is investigated by the following experiments.

6.6.1 FAULT DETECTION

Various tests are conducted on the iGX dry vacuum pump in order to simulate a defective bearing on both the LV and HV ends. The pump is initially operated normally to obtain validation data for the nominal model.

The data sampling time is selected to be 1 second. Conversely, it is noticed that in order to maximise the accuracy of the parameter estimates and keep the observer error sufficiently close to zero, the first order sliding mode algorithm requires a shorter sample time than 1 second.

From equations (6.17) and (6.18), it is evident that the discontinuous injection signals v_{LV} and v_{HV} described in 6.4.1 provide an estimate of the parameter changes. Because of the discontinuity, the signals must first be ‘filtered’ when the first order scheme is employed as reported by Edwards and Spurgeon [Edwards 1998]. The observer gains K_{LV} and K_{HV} are both set to 1.

Validation data are obtained from the pump under fault-free conditions. The dry vacuum pump is started from cold and run at constant gas load for approximately 30 minutes. After this point the gas load is varied using the flow control valve in order to vary the pump inlet pressure in the range 0-100 mbar. It is observed that extra heat is

generated at higher gas loads causing the temperature-controlled valve to switch on and off frequently, representing a challenging case for model validation.

Figure 6-2 depicts the error between the measured and estimated data from the first order sliding mode observer of the LV and HV end respectively. It can be seen that the error magnitudes are small ($0.0015\text{ }^{\circ}\text{C}$ or approximately 0.01%). This proves that the sliding mode observers are sliding and that the estimated value tracks the measured data.

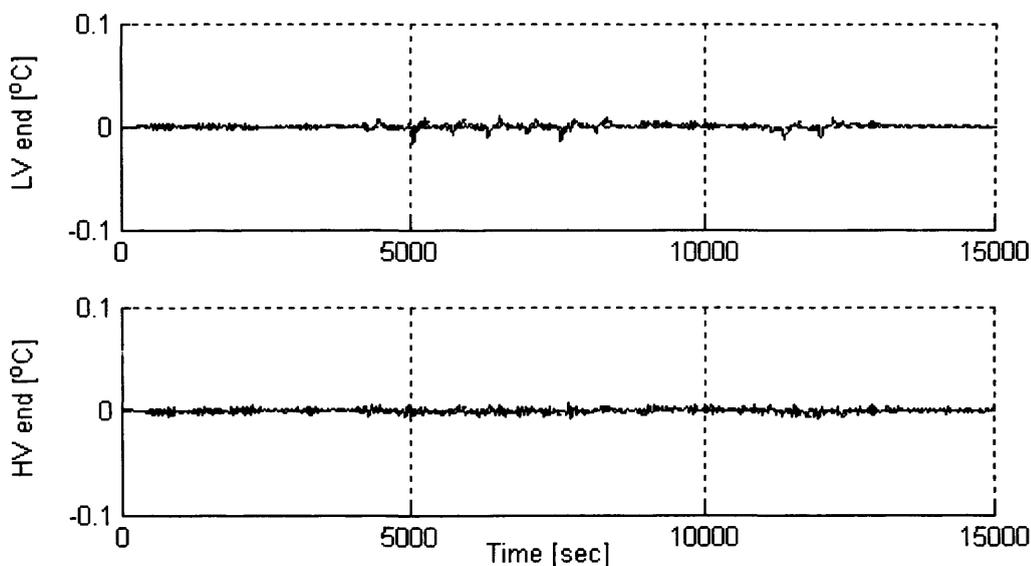


Figure 6-2: Error in temperature between LV and HV end models and corresponding measured experimental data under fault-free conditions

Having validated the observers, a series of tests to simulate a defective, overheated bearing are performed. Some representative results for both the LV and HV end are illustrated in Figure 6-3. In both cases, the dry vacuum pump is operated fault-free for a period of 4100 seconds. At that point, a heater attached close to the bearings, is switched on at an RMS power level of 110W. Heat is applied for a period of time

before switching off the heater. The procedure is repeated once again and a second rise on the bearing temperature is recorded. It should be noted that during the fault scenarios, the pump is operating over a range of gas loads. Figure 6-3 also shows that the two sliding mode observers remain in the sliding mode under all system conditions, demonstrating the robustness of the approach. Finally, Figure 6-4 shows the deviation in parameters, k_i , against time. It can be observed that the parameter estimates remain unchanged for a period of 4100 seconds. At that point a simulated fault is introduced and clearly some changes in the parameter changes can be noticed. The changes in these parameters indicate that the sliding mode observers can detect the induced fault in both cases. It can also be noticed that the LV end is more responsive than the HV end. This is due to the LV end bearing being closer to the coolant pipe-work.

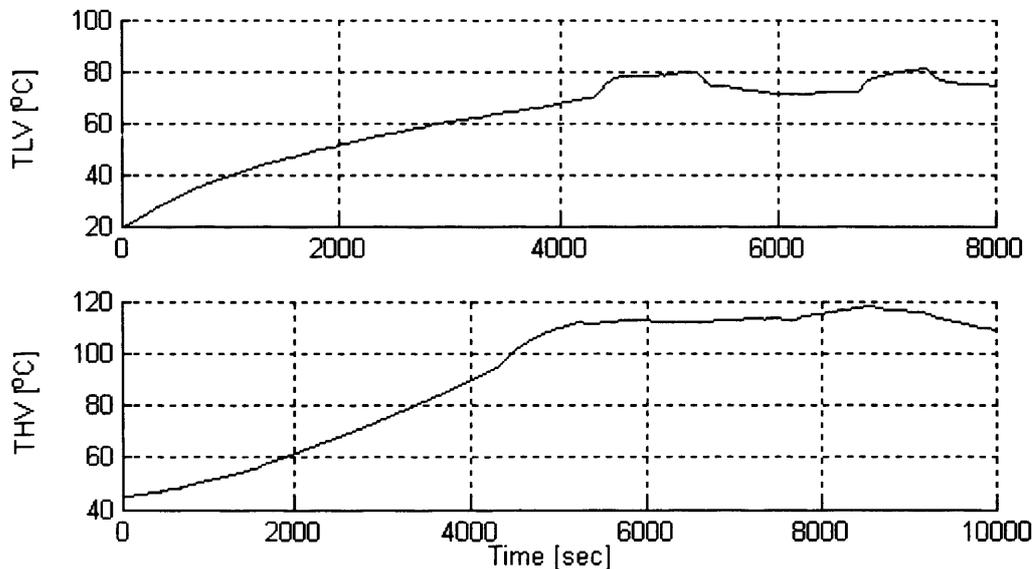


Figure 6-3: Measured and estimated data for the T_{LV} and T_{HV} under fault conditions

6.6.2 A COMPARISON OF FIRST AND SECOND ORDER SLIDING MODE OBSERVER PERFORMANCE

A new series of tests are now performed to investigate whether the HOSM algorithm is capable of suppressing chattering as would be expected from the theoretical considerations, and to assess if it provides an improved tracking (sliding) accuracy.

Prior to these tests, the gain parameters λ and W are tuned by trial and error. Hence, the gains λ and W are set to 0.1 and 0.001 respectively.

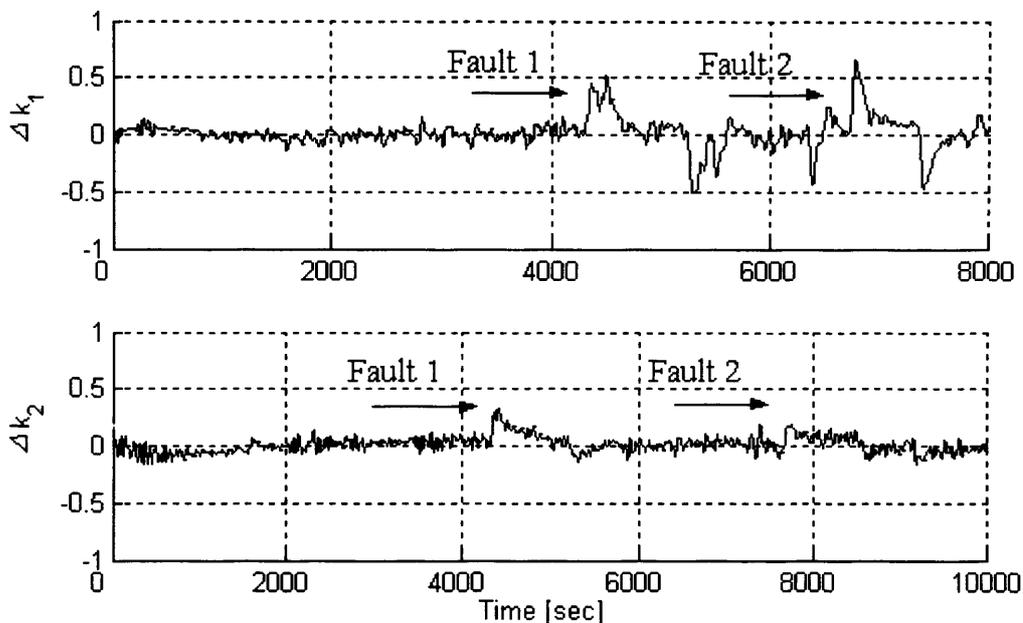


Figure 6-4: Component parameter estimate Δk_i under fault conditions

The capability of the second order sliding mode algorithm for suppressing chattering is first investigated. As discussed in chapter 3, in HOSM systems the discontinuities act on the higher order time derivatives of the sliding variable. Therefore, the higher the order of the sliding variable where the high frequency discontinuity first appears, the less the oscillations of the variable are evident. Figure 6-5 represents the discontinuous injection signal from the CSM and HOSM approach. In both cases the vacuum pump is

running fault-free with varying gas load. Clearly, it can be observed that the second order algorithm provides a smoother injection signal than the CSM.

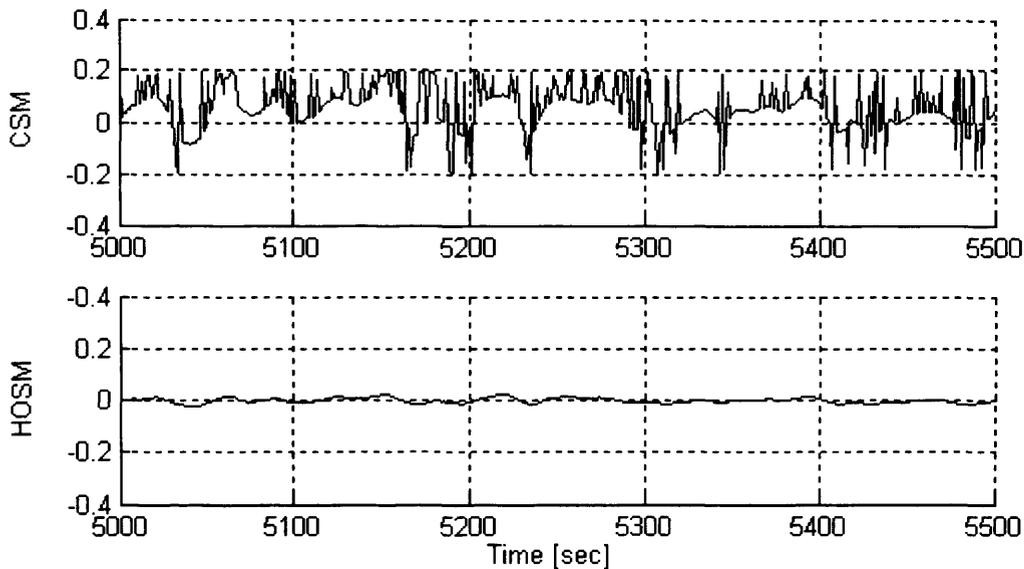


Figure 6-5: Observer injection signal for the classical (CSM) and higher order sliding mode (HOSM) schemes

The next tests investigated the tracking accuracy and reduced sampling rate of the HOSM when compared with the CSM scheme. Levant [Levant 1993] stated that the sliding accuracy of a first order sliding mode algorithm is proportional to the switching time delay. In contrast, the second order sliding mode algorithm provides an increased accuracy proportional to the square of the switching time delay. This switching time delay is represented as the sampling rate in the simulations of both CSM and HOSM algorithms. Therefore, in these particular tests various measurements of the sliding accuracy (i.e. observer error) are taken with respect to sampling rate.

Tests show that the CSM observer needs a sampling frequency faster than 1Hz in order to keep the observer error small. Here the first and second order algorithms are tested

with sampling times faster than 1 second and the RMS of the observer error is used as a measure of the performance. Table 6-1 shows that the RMS value of the HOSM observer error is considerably smaller for all sampling times. Representative error plot at a sampling time of 0.1 seconds are illustrated in Figure 6-6.

Table 6-1. Observer error statistics for CSM and HOSM

Sampling Time [sec]	Observer Error RMS [$^{\circ}$ C]	
	CSM	HOSM
0.01	1.339×10^{-4}	1.085×10^{-5}
0.1	0.1672	0.0129
1	2.3412	0.0184

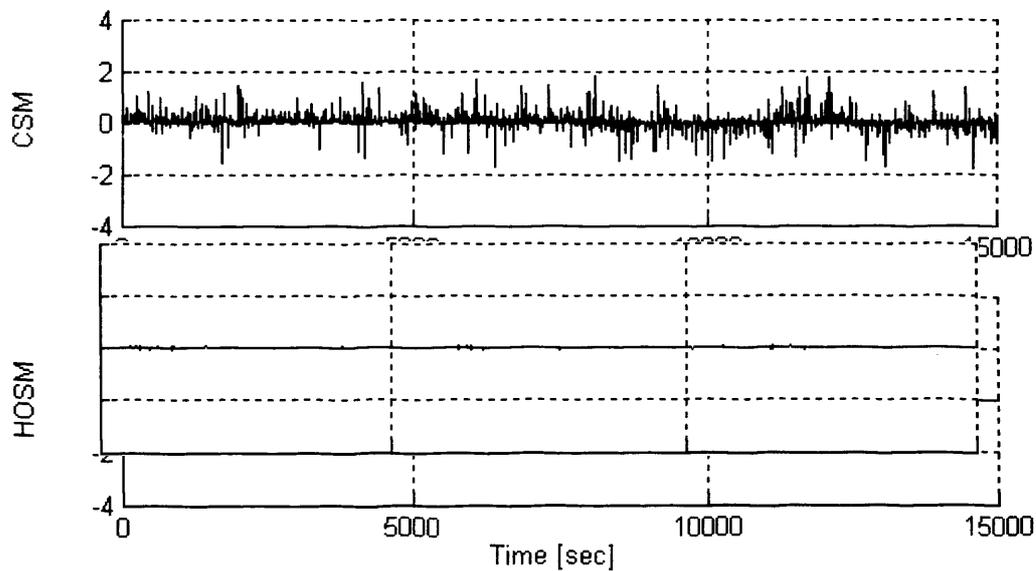


Figure 6-6: Observer error for CSM and HOSM at $T_s = 0.1s$

6.7 CONCLUSIONS

The development of a sliding mode observer-based scheme for the detection of overheated bearings has been demonstrated by application to a dry vacuum pump. In

general on such a system a defective bearing would result in either a shut-down or catastrophic failure. This chapter has shown that by examining the equivalent injection signal parameter estimation and hence fault prediction is possible. Additionally, the scheme is developed using signals from low-cost transducers sampling at a rate of 1Hz.

The experimental results have also been extended to a comparison between CSM and HOSM algorithms. Apart from the suppression of chattering, the second order sliding mode algorithm has shown an improved tracking capability and to be capable of running at lower sample rates for a given root mean square state error when compared with a first order sliding mode scheme.

CHAPTER 7

CASE STUDY 3: THE EXHAUST SYSTEM

It is widely recognised among plant operators that well-planned plant maintenance is of great benefit in reducing costs. Scheduled maintenance allows time for spares to be purchased and if necessary a dedicated service team, or the manufacturer's service engineers to be present. Scheduled maintenance stops tend to be shorter because the problem has already been identified and the solution is ready to implement.

The aim of this chapter is to develop a condition monitoring scheme for on-line fault prediction for the final nominated system component of the iGX dry vacuum pump, the exhaust system. Konishi and Yamasawa [Konishi 1999] reported that vacuum pumps are at risk of seizure caused by deposits of solid by-products of gas processes.

Similarly, waste gases in low-pressure chemical vapour deposition semiconductor processes can react strongly with residual air in the exhaust system of the dry vacuum pump to form solid deposits, obstructing the exhaust [Abreu 1994]. A characteristic of the dry vacuum pump operation is a low net mass flow rate of gas, meaning that deposits can drastically reduce the effective diameter of the flow path through the exhaust system without significantly affecting the average exhaust pressure. When the

diameter of the flow path is small, the risk of a small additional particle completely blocking the flow path and causing sudden failure of the pump system is increased. Therefore, the diagnostic scheme should be able to provide timely detection of a reduction in the free volume of the exhaust so that total blockage of the system can be prevented.

The proposed diagnostic scheme considers the use of a sliding mode observer as a discrete time estimator for an auto-regressive (AR) model where a second order sliding mode is enforced on the output error. AR modelling is a technique that can be used to predict future samples of a discrete time series as a linear function of previous samples. A regression is performed on a set of samples from the discrete time series in order to obtain a vector of model coefficients. Moreover, provided that the AR model is a good approximation of the signal's observed behaviour, it can be utilised to estimate the power spectrum of the signal. A thorough review of different modelling techniques for spectral analysis of discrete time series is presented by Kay and Marple [Kay 1981].

Up to now only the design of sliding mode systems for the continuous time domain have been considered. These systems have been shown to be effective at rejecting certain kinds of parameter variations and disturbances. However, their direct digital implementation may result in the loss of these desirable features due to the finite sampling time. Discrete time sliding modes (DSMs) have been developed to address the problems associated with digital implementation. Unlike their continuous time counterpart, they have not been studied to the same extent.

Early developments in the DSM theory are given in [Sarpturk 1987], [Futura 1990], [Spurgeon 1992]. Bartolini et al [Bartolini 2001] considered a discrete time implementation of a second order sliding mode algorithm, deriving sufficient conditions under which the sliding accuracy is improved. A practical implementation of a similar second order digital sliding mode algorithm for robust speed and torque estimation in electrical drives is presented by Bartolini et al [Bartolini 2003]. Experimental results of the proposed method have shown good performance, robustness against measurement errors and small computational demand. Edelbaheer et al [Edelbaheer 2006] developed a discrete time controller-observer pair for low speed sensor-less control of an induction machine. It is reported that in the derived discrete time expression the discontinuous term of the controller is replaced by a continuous one, thus eliminating the chattering of the control input signal. A sliding mode observer for discrete linear systems is developed by Koshkouei and Zinober [Koshkouei 2002]. Properties of DSMs and conditions for their existence are among the problems that are studied.

It will be seen in this chapter that the reduction in the free volume of the exhaust can be detected by monitoring the injection signal of a discrete sliding mode observer. The chapter starts by formulating the fault prediction problem where practical issues like selection of appropriate monitoring signals are considered. In section 7.2 a discrete sliding mode observer is formulated for monitoring changes in the parameters of an AR model. Section 7.3 discusses the case study in detail. The experimental arrangements and the pre-processing of the signals are some of the problems presented. Finally, the experimental results are discussed in section 7.4 and conclusions are presented in section 7.5.

7.1 PROBLEM FORMULATION

As already noted in chapter 2, the function of the exhaust system is to direct the pumped gases away to the atmosphere. The system consists of two components: a flexible pipe-work and a set of absorptive silencers that minimise the noise levels.

The primary task in the case study is to select appropriate monitoring signals and transducers for the exhaust system. Abreu et al [Abreu 1994] indicated that all dry vacuum pumps generate approximately sinusoidal pressure fluctuations in their exhaust pipes. Moreover, exhaust pressure sensors are already part of the typical sensor package for some pump processes. Therefore, it is clear that the study of exhaust pressure signals can be an effective solution for the condition monitoring problem.

A very important aspect of the exhaust pressure signal is its natural periodicity, since it is associated with the periodic motion of the rotors and opening and closure of the exhaust ports from the low vacuum stage. This periodicity can be examined by extracting a moving average of the signal. If the signal is averaged over a large number of rotations, all signal components except those at frequencies harmonically related to the rotation are removed. A deterministic periodic signal is produced which can then be efficiently analysed using conventional fast fourier transform (FFT) or auto-regressive (AR) methods.

From the signal frequency spectra, it is also noticed that under steady state gas load the relative power content of the exhaust pressure signal at the fundamental frequency and its second harmonic are related to the free volume in the exhaust system. The ratio between the power content at these frequencies is referred to in this thesis as the power

ratio (PR). If the PR changes, then the location of the AR model poles move to reflect the change in relative power magnitudes.

Another characteristic of the iGX dry vacuum pump is that it is driven by an asynchronous induction motor and therefore the motor slip frequency tends to increase with the gas load. The motor slip frequency is the difference between the reference frequency of the motor and the actual one. In order to compensate for these small changes in pump rotational frequency a phase-locked loop (PLL) can be used to exploit the natural periodicity of the exhaust pressure signal generating a trigger signal for data sampling in-phase with the rotor angle. A PLL [Stensby 1997] is a closed-loop feedback control signal consisting of a phase comparator, loop filter and voltage controlled oscillator (VCO). The PLL can be designed to reduce to zero or 'pull-in' the phase difference between the reference signal and the VCO output signal. Thus a well-designed PLL can generate an output signal which is phase coherent with a reference input signal and at the same frequency.

In the following section a novel sliding mode observer is designed and implemented to monitor changes in a nominal set of AR model parameters.

7.2 SLIDING MODE OBSERVER DESIGN

The exhaust pressure time series signal investigated in this case study can be described in an AR format as

$$y(k) = -\sum_{n=1}^p a_n y(k-n) + e(k) \quad (7.1)$$

where $y(k)$ denotes the exhaust pressure signal, $e(k)$ is white noise and a_n the AR model coefficients.

The above AR model can then be converted to a general discrete time model in the state space form. Shtessel and Poznyak [Shtessel 2004] represented a linear time varying (LTV) system in the state space observable canonical format. In this study the discrete time model is reformulated such that the state vector is presented in terms of the signal, $y(k)$. The state vector is defined here by

$$\begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_n(k) \end{bmatrix} = \begin{bmatrix} y(k) \\ y(k-1) \\ \vdots \\ y(k-n+1) \end{bmatrix} \quad (7.2)$$

resulting in the description;

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \mathbf{M} \\ x_n(k+1) \end{bmatrix} = \begin{bmatrix} -a_1 - a_2 \Lambda - a_n \\ 0 - a_1 \Lambda - a_{n-1} \\ \mathbf{M} \quad \mathbf{M} \quad \mathbf{M} \\ 0 \quad 0 \quad \Lambda - a_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \mathbf{M} \\ x_n(k) \end{bmatrix} + \begin{bmatrix} 0 & 0 & \Lambda & 0 \\ -a_n & 0 & \Lambda & 0 \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} \\ -a_2 - a_3 \Lambda - a_n & & & \end{bmatrix} \begin{bmatrix} y(k-n) \\ y(k-n-1) \\ \mathbf{M} \\ y(k-2n+1) \end{bmatrix} \quad (7.3a)$$

$$y(k) = [1 \ 0 \ \Lambda \ 0 \ 0]^T x(k) \quad (7.3b)$$

The objective is to design an observer to generate a state estimate $\hat{x}(k)$ such that a sliding mode can be attained in which the state errors defined by

$$\varepsilon(k) = x(k) - \hat{x}(k) \quad (7.4)$$

are forced to zero in finite time. Thus the observer is designed to detect deviation of the system behaviour from its nominal, fault-free operation by monitoring variations in the average value of the applied injection signal required to maintain sliding. A second order 'super-twisting' sliding mode observer can be written, with some abuse of notation in the following form

$$\hat{x}(k+1) = A_n \hat{x}(k) + B_n \begin{bmatrix} y(k-n) \\ y(k-n-1) \\ \mathbf{M} \\ y(k-2n+1) \end{bmatrix} + v(k) \quad (7.5)$$

where $A_n \in \mathcal{R}^{n \times n}$, $B_n \in \mathcal{R}^{n \times m}$ are the nominal matrices in equation (7.3a). The components of the vector $v(k)$ are defined by v_i , $i = 1, 2, \dots, n$;

$$\begin{cases} v_i(k) = \alpha |\varepsilon_i(k)|^{1/2} \text{sgn}(\varepsilon_i(k)) + \beta v_i(k) \\ v_i(k+1) = \tau \text{sgn}(\varepsilon_i(k)) + v_i(k) \end{cases} \quad (7.6)$$

where α and β are the gains of the discontinuous signals and τ is the observer sampling interval. The observer system (7.5), (7.6) is in fact updated at a much faster rate, $f_o = 1/\tau$, than that at which the plant data samples are captured, f_s , i.e. $f_s \ll f_o$. From the work of Levant [Levant 1998] it is known that if measurements are available at successive measurement instants an appropriately parameterized second order sliding mode will converge. Equation (7.6) is formulated in this discrete fashion due to the implementation requirements of the experimental hardware. Provided the noise component of the exhaust pressure signal is zero and the state errors are due to parameter uncertainty, Δ_i , then equation (7.3a) can be written as follows

$$\begin{aligned} \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ \mathbf{M} \\ x_n(k+1) \end{bmatrix} &= \begin{bmatrix} -a_1 + \Delta_1 & -a_2 + \Delta_2 \Lambda & -a_n + \Delta_n \\ 0 & -a_1 + \Delta_1 \Lambda & -a_{n-1} + \Delta_{n-1} \\ \mathbf{M} & \mathbf{M} & \mathbf{M} \\ 0 & 0 & \Lambda & -a_1 + \Delta_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ \mathbf{M} \\ x_n(k) \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & \Lambda & 0 \\ -a_n + \Delta_n & 0 & \Lambda & 0 \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} \\ -a_2 + \Delta_2 & -a_3 + \Delta_3 \Lambda & -a_n + \Delta_n \end{bmatrix} \begin{bmatrix} y(k-n) \\ y(k-n-1) \\ \mathbf{M} \\ y(k-2n+1) \end{bmatrix} \end{aligned} \quad (7.7)$$

If the observer error dynamics are defined as $\varepsilon(k+1) = x(k+1) - \hat{x}(k+1)$, then it is straightforward to show that

$$\begin{aligned}
\begin{bmatrix} \varepsilon_1(k+1) \\ \varepsilon_2(k+1) \\ \mathbf{M} \\ \varepsilon_n(k+1) \end{bmatrix} &= \begin{bmatrix} -a_1 & -a_2\Lambda & -a_n \\ 0 & -a_1\Lambda & -a_{n-1} \\ \mathbf{M} & \mathbf{M} & \mathbf{M} \\ 0 & 0 & \Lambda & -a_1 \end{bmatrix} \begin{bmatrix} \varepsilon_1(k) \\ \varepsilon_2(k) \\ \mathbf{M} \\ \varepsilon_n(k) \end{bmatrix} \\
&+ \begin{bmatrix} x_1 & x_2 & \Lambda & x_n \\ x_2 & x_3 & \Lambda & y(k-n-1) \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ x_n & y(k-n-1) & \Lambda & y(k-2n+1) \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \mathbf{M} \\ \Delta_n \end{bmatrix} - \begin{bmatrix} v_1 \\ v_2 \\ \mathbf{M} \\ v_n \end{bmatrix}
\end{aligned} \tag{7.8}$$

Assuming a sliding motion will be attained and maintained, $\varepsilon(k+1)$ and $\varepsilon(k)$ will converge to zero. Hence equation (7.8) becomes

$$\begin{bmatrix} v_1 \\ v_2 \\ \mathbf{M} \\ v_n \end{bmatrix} = \begin{bmatrix} y(k-1) & y(k-2) & \Lambda & y(k-n) \\ y(k-2) & y(k-3) & \Lambda & y(k-n-1) \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ y(k-n) & y(k-n-1) & \Lambda & y(k-2n+1) \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \mathbf{M} \\ \Delta_n \end{bmatrix} \tag{7.9}$$

Re-writing equation (7.9) as

$$v_i = Y_d \Delta_i \tag{7.10}$$

It is seen that the discontinuous signals v_i provide a means to estimate the parameter uncertainties Δ_i for an AR model, i.e.

$$\Delta_i = Y_d^{-1} v_i \tag{7.11}$$

These parameter uncertainties can then be used to compute the values for the AR model parameters $\hat{\alpha}_i$

$$\hat{\alpha}_i = \alpha_{ni} + \Delta_i \quad (7.12)$$

where α_{ni} are the nominal coefficients of the AR nominal model. Once the estimated values for the AR model parameters, $\hat{\alpha}_i$, are obtained, the power spectrum of the exhaust pressure signal can be estimated. Moreover, the time-averaged magnitude of the injection signal is related to the difference between the location of the poles of the nominal AR model and those of the sliding mode observer estimated model.

7.3 EXPERIMENTAL SET-UP

The sliding mode observer developed in the previous section is tested on the iGX dry vacuum pump test bed. This section details the description of the hardware set-up of the system and discusses issues such as pre-processing of signals and observer implementation.

7.3.1 HARDWARE SET-UP

The hardware interface for the exhaust system consists of a dSPACE™ data acquisition system and a BOC-Edwards active strain gauge (ASG) pressure transducer. The ASG is mounted on the ballast gas port on the final (low-vacuum) stage of the pump. The pumped gas is air and the gas load is set using a flow control valve to maintain constant pump inlet pressure. Furthermore, a barocel gauge is fitted to the vacuum vessel in order to measure the inlet pressure of the vacuum pump. It should be noted that this

transducer is not part of the condition monitoring system, but rather offers an indication of the inlet pressure for the experimental tests.

The exhaust system includes a 'cigar-box'-shaped silencer of volume 675 ml. Silencer obstruction is effected by introducing small metal objects which could be easily manipulated and presented no risk of potentially damaging back-ingestion into the pump mechanism. Finally, in the data collection experiments obstruction volumes are measured using a liquid displacement method.

7.3.2 SIGNAL PRE-PROCESSING

A phase-locked loop (PLL) is initially designed to trigger data sampling in synchronicity with the rotor angle. Figure 7-1 illustrates a simple block diagram of a PLL. The basic idea of a PLL is to inject a sinusoidal signal into the reference input. Then an internal oscillator locks to the reference signal and produces a signal that is phase coherent with the supplied reference. Since all dry vacuum pumps generate approximately sinusoidal pressure fluctuations, a normalised band pass filtered version of the exhaust pressure signal can be used as a phase reference signal, thus eliminating the need for a rotor angle encoder. The phase comparator is modelled as a simple multiplier. The scalar, A_v , can be defined to be 1, whilst the reference signal is normalised so that $A_i = 1$. The voltage controlled oscillator (VCO) is modelled as an integrator and a cosine function. The integrator is reset to zero when the output reaches $\pm 2\pi$ radians. A first order Butterworth low pass filter with cut-off frequency 5 Hz is implemented as the loop filter.

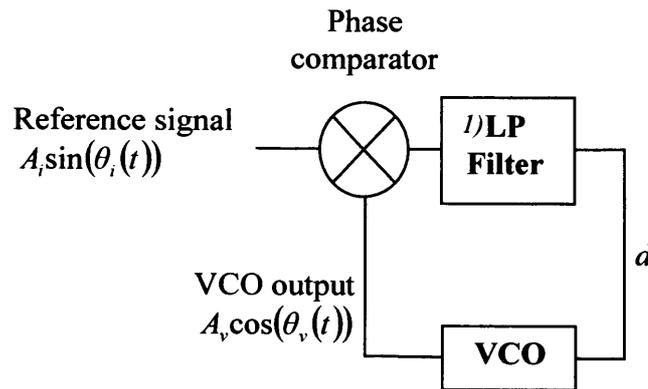


Figure 7-1: Block diagram of a PLL

When the phase is locked; the VCO output, $\cos\theta_v(t) = \cos(\omega_v t + \phi_v(t))$ is phase coherent with the phase reference signal, $\sin(\theta_i(t)) = \sin(\omega_i t + \phi_i(t))$, therefore $\omega_v = \omega_i$.

Thus the output of the integrator is an approximation of the pump's rotor angle,

$$\theta = \omega_v t + \phi_v(t).$$

It should be pointed out that for the spectrum and sliding mode observer analysis the rotor angle does not require a physical datum point, so the exact phase difference between the VCO output and the reference signal is not important.

The rotor angle approximation can be used to trigger the re-sampling of the exhaust pressure signal at 8 samples per revolution (800 Hz approximately), essentially an interpolation process.

Ensemble averaging of the re-sampled signal is also employed in order to attenuate random noise. The signal averaging process resulted in a series of values

$$\bar{p}_{\theta_m}(k) = \frac{1}{N} \sum_{n=0}^{N-1} p_{\theta_m}(k-n) \quad (7.13)$$

where, N is the number of revolutions over which the signal is averaged, θ_m is the sampling point in terms of rotor angle. $m = 1, 2, \dots, M$ where M is the number of samples per revolution. In the results that follow $M = 8$ and $N = 100$.

7.3.3 IMPLEMENTATION OF THE OBSERVER

Captured data is analysed using a simulation of the discrete time sliding mode observer. The choice of the observer sample rate is decided by analysis of the root mean square (RMS) of the state errors, given by;

$$\text{RMS state errors} = (\bar{\varepsilon}_{\text{RMS}} / y_{\text{RMS}}) \times 100\% \quad (7.14)$$

Plotting the RMS errors against the ratio of the observer sample rate, f_{s2} , and the AR model sampling rate f_{s1} (see Figure 7-2) shows that the RMS state errors are minimised when $f_{s2} \geq 1000 f_{s1}$. However, satisfactory results for the observer are obtained with $f_{s2} = 200 f_{s1}$. Therefore at a rotational pump speed of approximately 100 Hz and 8 samples per revolution, the observer sample rate is required to be approximately 160 kHz. Finally, the observer gains, α and β , are given the values 10,000 and 2,500 respectively.

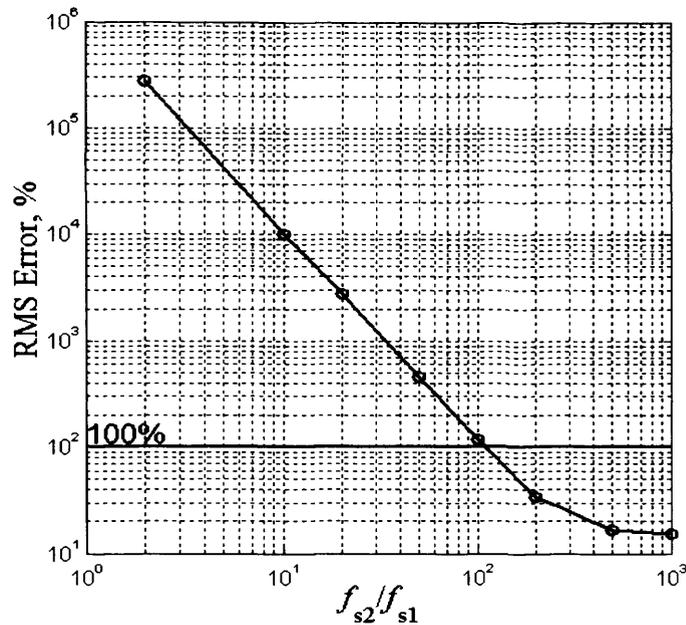


Figure 7-2: RMS state error as a function of sampling frequency

7.4 RESULTS

In order to demonstrate the efficacy of the condition monitoring scheme, a number of experiments are carried out on the test bed. Data collection is performed with an unobstructed silencer and three obstruction volumes; 50 ml, 115 ml and 170 ml. The tests are performed at constant gas loads with pump inlet pressure set to approximately 10 mbar in each case.

Conventional AR techniques can be used to analyse the spectrum of the synchronously sampled data and calculate the PR of the exhaust pressure signal. An 8th order AR model is fitted to a set of obstruction-free data using the Yule-Walker algorithm. An example of the spectrum from 160 point (20 pump revolutions) frame of data obtained is presented in Figure 7-3. The PR is simply calculated as the magnitude of the peak at the fundamental frequency (1 cycle per rev) divided by that of the second harmonic (3 cycles per rev).

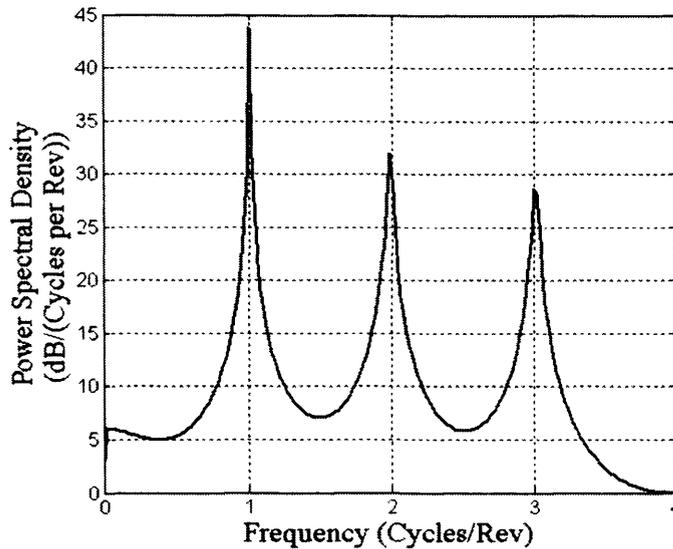


Figure 7-3: Yule-Walker AR model power spectrum

Figure 7-4 depicts the ensemble-averaged data series in the presence of different degrees of obstruction ('fill volume'). It can be noticed that the changes in amplitudes of the exhaust pressure signal fluctuations due to the obstruction are quite small and thus it would be difficult to infer accurately blockage of the exhaust from this information. The time-averaged values of the PR, calculated by means of the 8th order AR model, increase with the obstruction volume as illustrated in Figure 7-6. However, the instantaneous values of the PR shown in Figure 7-5 are not constant, indicating the need for some additional filtering. Further, the coincidental peaks at 300 revolutions for the 50 ml and 115 ml data sets are not thought to have a common cause.

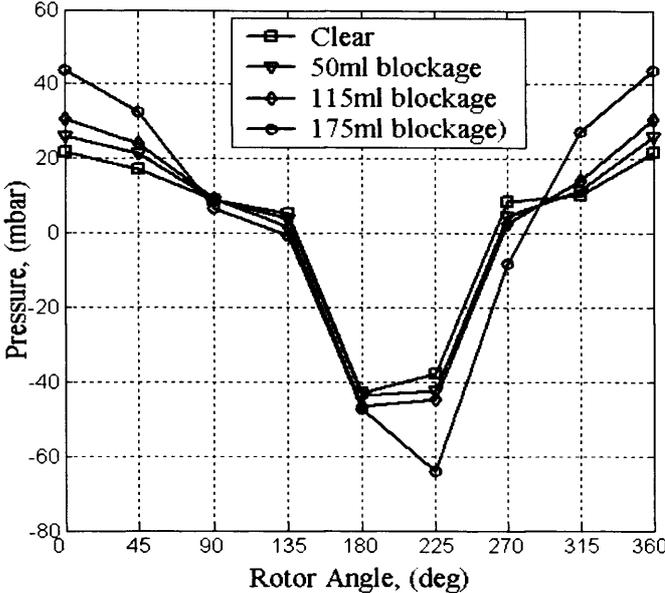


Figure 7-4: Ensemble-averaged exhaust pressure amplitude vs. rotor angle, showing the variation in amplitudes with increasing volume of fill

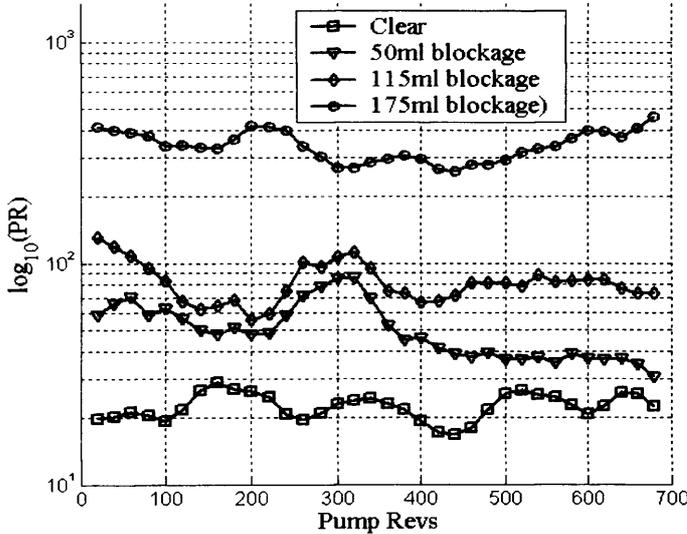


Figure 7-5: AR estimates of PR vs. pump revolutions for a series of silencer blockages.

Note that spectra are calculated every 20 revolutions

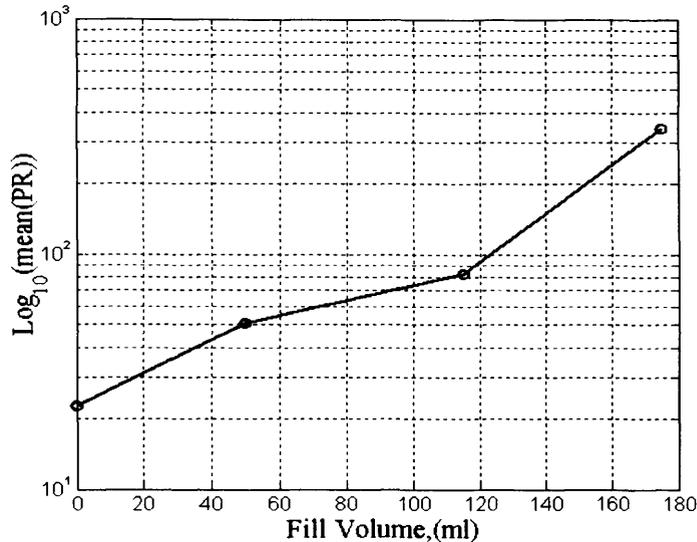


Figure 7-6: The PR increases with the volume of silencer obstruction. These average results are the means of the four data series in Figure 7-5

A change in the frequency content of the exhaust pressure signal implies a change in the AR model coefficients with respect to the nominal sliding mode observer model. The RMS value of the sliding mode observer injection signal rises to match the uncertainty in the AR model parameters. Hence, a change in the signal content caused by the exhaust obstruction causes the RMS injection signal to rise (see Figure 7-7). The observer produces a new estimate of the AR model parameters based on the nominal model and the injection signal. This AR model can be used to produce a spectrum. Nonetheless, even with ensemble averaging of the signal over 100 revolutions, the short data set used to generate the parameters ($2P-1$ samples, where P is the AR model order) resulted in high sample-to-sample variations for the PR estimate. Figure 7-8 shows a comparison of PR values calculated using FFT, AR and the sliding mode observer method.

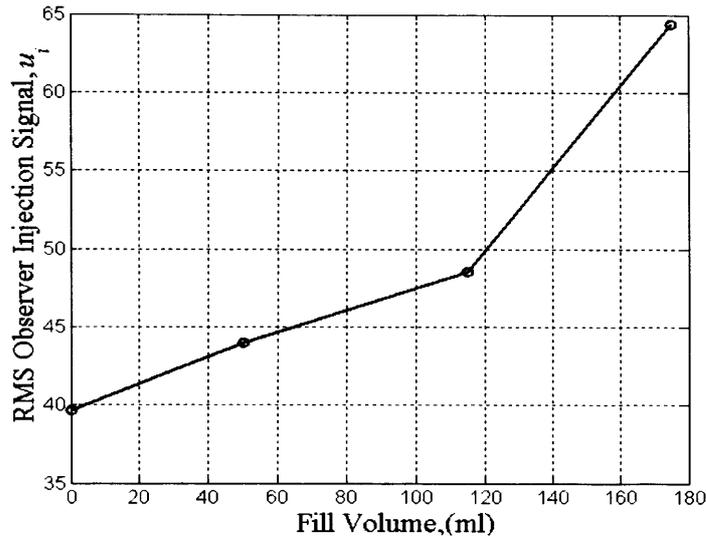


Figure 7-7: RMS of the observer injection signal increases with the volume of silencer obstruction

Therefore, further work is required to establish the reason for these sample-to-sample variations of the sliding mode approach that provide an unreliable estimate for the PR. Furthermore, whilst the sliding mode approach successfully indicates the increasing obstruction volume, it also requires a sampling rate greater than 100 times faster than that of the AR modelling approach for example, with obvious implications for computational costs. Finally, a sequence of spectra calculated using the sliding mode observer technique and by using the same data sequence is presented in Figure 7-9.

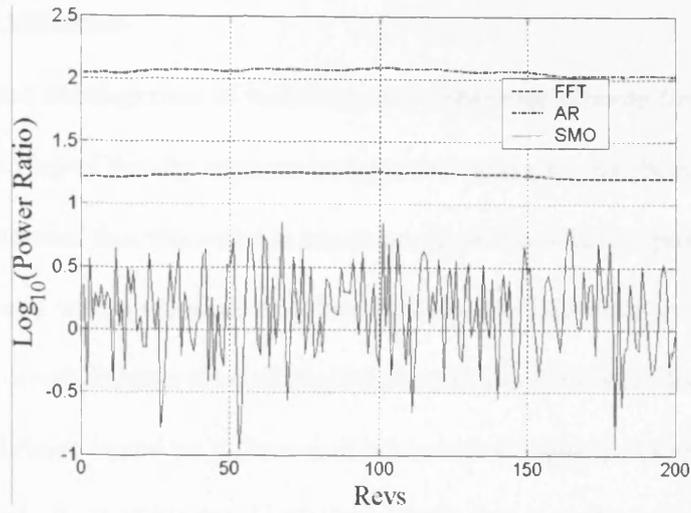


Figure 7-8: Comparison of PR values calculated using FFT, AR and sliding mode observer methods

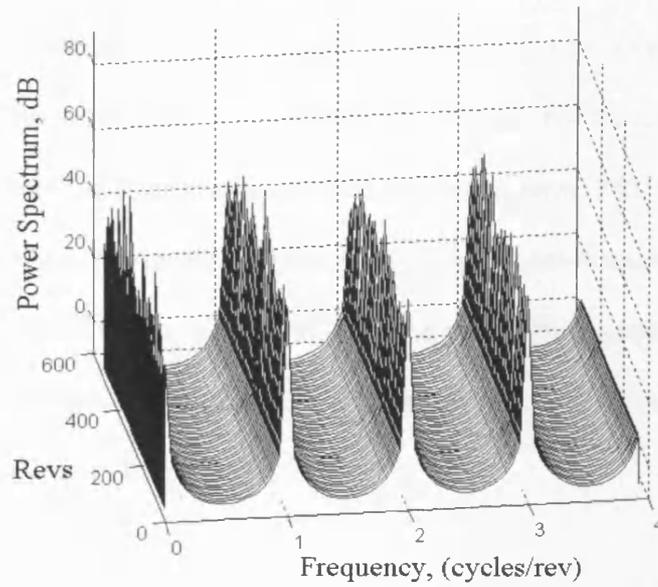


Figure 7-9: Spectra calculated from sliding mode observer generated AR model coefficients for clear silencer

7.5 CONCLUSIONS

In this chapter, the development of a sliding mode observer scheme for on-line condition monitoring of the dry vacuum pump exhaust system has been considered. It has been demonstrated that the exhaust pressure signal from such a pump can be practically acquired with a standard transducer, and described with an auto-regressive (AR) model. A novel discrete time sliding mode observer has been designed to estimate AR model coefficients based on a short data set sampled from the exhaust pressure signal, and a nominal set of model coefficients estimated from fault-free data.

Two approaches have been used to monitor the state of the exhaust system. Firstly, by monitoring the injection signal of the sliding mode observer, since it provides indication of a deviation of the signal spectrum away from that of the nominal model and can detect changes in the exhaust signal which indicate exhaust blockages. A second approach is to monitor an empirical parameter, the power ratio (PR). The PR is defined as the ratio of the signal power magnitudes of the fundamental frequency and the second harmonic of the exhaust signal. Therefore, to monitor the PR directly by periodically estimating the signal spectrum and detecting the increasing trend symptomatic of decreasing free volume in the exhaust.

In the context of this thesis, the sliding mode observer scheme can be concluded to perform the same condition monitoring function but at higher computational cost than the conventional AR approach for example. Moreover, the value for PR varies with gas load and thus in this simple form the method is only valid for pump at constant gas load. Further work is necessary to map the PR with respect to gas load or motor power.

The results presented in the case study establish the general principles of the monitoring scheme but further testing under process conditions with more realistic fill material is required. Model structures other than AR may be a more appropriate basis for the observer, particularly for detection of possible transient fault conditions. Currently the phase locked loop is implemented as a simulation however a hardware implementation may have cost benefits. Finally, the potential for using aspects of the scheme in other applications is still to be explored.

CHAPTER 8

CONCLUDING REMARKS AND FUTURE WORK

8.1 CONCLUDING REMARKS

This thesis has concentrated on predicting malfunction in quasi steady-state rotating machines. In particular, a practical condition monitoring and fault diagnosis scheme for dry vacuum pumps has been presented.

The thesis commenced by introducing the significance of vacuum systems in industry and the benefits of ‘dry’ pumps. The need for condition monitoring and fault diagnosis has been clearly established. After reviewing a number of condition monitoring and fault diagnosis techniques, the application of the sliding mode methodology to this area has been explored. The subject of this study, a BOC-Edwards iGX dry vacuum pump and the particular project related subsystems have been described. With many typical fault symptoms for each pump subsystem, a selection of suitable monitoring signals and sensors has been discussed. In deference to factors such as cost, functionality and robustness, the best potential solution in terms of sensor suite to the problem of condition monitoring and fault diagnosis in vacuum systems has been considered.

Sliding mode techniques, presented in chapter 3, have been shown to be an effective solution for the development of on-board condition monitoring and fault diagnosis systems due to their fundamental properties of robustness and the parameter reconstruction properties they provide. Simulation results have demonstrated that once a sliding motion has been established, the order of a system is reduced and it becomes totally invariant to certain kinds of uncertainties. Further, a disadvantage of sliding modes, the chattering effect, which may limit their practical applicability has been discussed. It has been shown by the application of simulation results that these difficulties can be alleviated by smoothing the nonlinear term of the control signal or employing higher order sliding mode techniques. Apart from these properties, the potential use of sliding modes for observer design has been examined and it has been concluded that appropriately designed sliding mode observers can be used to monitor the iGX dry vacuum pump.

To this end, a sliding mode observer based diagnostic scheme has been designed and implemented in chapter 4 to detect possible faults of the cooling system. Parameter estimation and hence fault detection has been achieved by analysis of the so-called equivalent injection signal. This property of fault, or parameter, reconstruction is unique to the nonlinear sliding mode methodology and of great importance since it allows signals or parameters that may be expensive or difficult to measure to be estimated. The results from a series of experimental tests have illustrated the usefulness of the approach and good correlation between the system parameters and experimentally determined values has been obtained for all three targeted coolant faults. Most importantly, the method is economically attractive because it only requires low cost temperature transducers that can be readily fitted to the system in a non-invasive

manner and a reading of the motor's current and coolant water flow. It also provides an earlier diagnosis and information down to a component level compared to a simple high temperature alarm.

Without the exact nominal parameter values of the mathematical model required to formulate a sliding mode observer, some means of parameter estimation is desirable. A parameter estimation algorithm has been developed, implemented and tested on both simulated and real vacuum pump data. The algorithm has been shown to estimate the time invariant nominal parameter values to an acceptable accuracy. However, an issue of concern with regard to the algorithm is that it did not perform well when time-varying parameters were tested. Results obtained from the sliding mode observer scheme have shown that it offers clear benefits over the estimation algorithm in the role of continuous condition monitoring and fault diagnosis in systems where intermittent faults are possible.

Chapter 6 has considered the development of a condition monitoring scheme for the detection of overheated bearings. The proposed sliding mode observer-based scheme utilises a heat transfer model to estimate parameters of the system with parameter changes being symptomatic of faults. This has been achieved with an analysis of the so-called equivalent injection signal. Whilst in this case study the nominal heat transfer model has been more complex to derive, the sliding mode scheme still offers benefits in terms of robustness and cost. The experimental results have been extended to effect comparison between a first order and a second order sliding mode algorithm. The practical importance of the second order algorithm has been demonstrated, since the results have shown that the chattering effects have been suppressed. The second order

algorithm has also shown to be capable of running at lower sample rates for a given root mean square error when compared with the first order sliding mode algorithm.

The problem of detecting exhaust blockages in the dry vacuum pump by using spectral estimation and sliding mode observer techniques has been addressed in chapter 7. Two independent techniques have been proposed which are capable of providing timely detection of a reduction in the free volume of the exhaust. Firstly, aspects of pump performance have been monitored via an empirical relationship, called the power ratio, observed in periodic features in the measured exhaust pressure signal. These features can be characterised by the magnitudes at particular frequencies in the power spectrum of the signal. Experimental results have shown that at a constant gas load the method can be used to provide an early indication of exhaust blockage. Secondly, a novel discrete time sliding mode observer has been designed to estimate auto-regressive model coefficients based on a short data set sampled from the exhaust pressure signal, and a nominal set of model coefficients estimated from fault-free data. The results have demonstrated that the reduction in free volume of the exhaust can be detected by monitoring the injection signal of the sliding mode observer. This is because the magnitude of the injection is related to the difference in location between the poles of the nominal auto-regressive model and those of the estimated model.

8.2 RECOMMENDATIONS FOR FUTURE WORK

The produced diagnostic system has shown to be capable of reducing the occurrence of unplanned pump stoppages, by monitoring appropriate subsystems and parameters.

Development of a fault classification method in order to determine the size of a fault is a natural progression of the work presented in this thesis. The sensitivity of the

diagnostic system will be improved and unnecessary alarms will no longer be a problem. One possibility is the use of fuzzy systems and more specifically fuzzy membership functions to classify fault symptoms into diagnostic conditions such as normal, low and high. Membership functions may be expressed using any appropriate function like a Gaussian function or in a piecewise linear manner such as a triangular function. One advantage of using this fuzzy-based approach rather than a fixed threshold alarm is that a smooth transfer between the diagnostic conditions is achievable.

An effective method for parameter estimation based on an algebraic framework has been presented in chapter 5. However, the approach did not perform well when time-varying parameters were tested. Thus, there is a need to extend the method in order to investigate estimation of time-varying parameters. Furthermore, a discrete implementation of the algebraic framework for identification should be developed. This could overcome some of the problems associated with the low sample rates used for some of the pump measurements.

The work undertaken during the second case study has considered the second design of the cooling circuit. This second design has a single temperature-controlled valve that maintained the pump stator temperature within some set limits. The development of the heat transfer models presented in chapter 6 is based on this design. However, towards the end of this thesis a third and final cooling circuit design has been installed to the iGX dry vacuum pump. The final design consists of two temperature-controlled valves located at either end of the vacuum pump. These valves have the ability to regulate the temperature of the system more accurately. It is obvious that the heat flow through the

pump will behave differently under this design and that it is desirable to re-evaluate the heat transfer models. Early heat transfer model developments provide a sound basis for future research as may be seen in Appendix E.

Although a novel condition monitoring and fault diagnosis tool for the vacuum pump exhaust system has been successfully developed, additional testing under process conditions with more realistic fill material is needed. Moreover, model structures other than auto-regressive may be a more appropriate basis for the sliding mode observer, particularly for detection of possible transient fault conditions. The phase locked loop has been implemented as a simulation however a hardware implementation may have cost benefits. Finally, the diagnostic scheme is versatile and various aspects of it can be used in other systems or applications.

The diagnostic scheme has been implemented using the Matlab™/Simulink™/dSPACE™ environment. A typical process for this environment would be first to build a model in Simulink™, followed by an auto-generation of C-code using the Matlab™ real time workshop. This code is finally downloaded into the dSPACE™ board. Obviously, it would not be feasible to incorporate the scheme into a BOC-Edwards or other industrial vacuum pump control system architecture in this form. Hence, there is a need to explore an alternative way for generation of implementable code for the diagnostic system.

REFERENCES

[Abreu 1994] Abreu, R.A.; Troup, A.P.; Sahm, M.K., *Causes of anomalous solid formation in the exhaust systems of low-pressure chemical vapor deposition and plasma enhanced chemical vapor deposition semiconductor processes*, Journal of Vacuum Science and Technology B: Microelectronics and Nanometer Structures, Vol. 12(4), pp. 2763-2767, 1994.

[Alcorta-Garcia 1997] Alcorta-Garcia, E.; Frank, P.M., *Deterministic nonlinear observer-based approaches to fault diagnosis: A survey*, Control Engineering Practice, Vol. 5(5), pp. 663-670, 1997.

[Alessandri 1999] Alessandri, A.; Hawkinson, T.; Healey, A.J.; Veruggio, G., *Robust model-based fault diagnosis for unmanned underwater vehicles using sliding mode-observers*, 11th International Symposium on Unmanned Untethered Submersible Technology, August 22-25, 1999.

[Bachmann 1990] Bachmann, P.; Kuhn, M., *Evaluation of dry pumps vs. rotary vane pumps in aluminum etching*, Vacuum, Vol. 41(7-9), pp. 1825-1827, 1990.

- [Bartolini 1998]** Bartolini, G.; Ferrara, A.; Usai, E., *Chattering avoidance by second-order sliding mode control*, IEEE Transactions on Automatic control, Vol. 43(2), pp. 241-246, 1998.
- [Bartolini 2000]** Bartolini, G.; Punta, E., *Chattering elimination with second-order sliding modes robust to coulomb friction*, Journal of Dynamic Systems Measurement and Control, Vol. 122, pp. 679-686, 2000.
- [Bartolini 2001]** Bartolini, G.; Pisano, A.; Usai, E., *Digital second-order sliding mode control for uncertain nonlinear systems*, Automatica, Vol. 37, pp. 1371-1377, 2001.
- [Bartolini 2003]** Bartolini, G.; Damiano, A.; Gatto, G.; Marongiu, I.; Pisano, A.; Usai, E., *Robust speed and torque estimation in electrical drives by second-order sliding modes*, IEEE Transactions on Control Systems Technology, Vol. 11(1), pp. 84-90, January 2003.
- [Basseville 1985]** Basseville, M.; Benveniste, A., *Detection of abrupt changes in signals and dynamical systems*, Lecture Notes in Control and Information Sciences, Vol. 77, Springer, Berlin, 1985.
- [Beard 1971]** Beard, R.V., *Failure accommodation in linear systems through self-reorganization*, Dept. MVT-71-1, Man Vehicle Laboratory, Cambridge, MA, 1971.

REFERENCES

[Bhatti 1999] Bhatti, A.I.; Spurgeon, S.K.; Dorey, R.; Edwards, C., *Sliding mode configurations for automotive engine control*, International Journal of Adaptive Control and Signal Processing, Vol. 13, pp. 49-69, 1999.

[Bhatti 1999] Bhatti, A.I.; Twiddle, J.A.; Spurgeon, S.K.; Jones, N.B., *Engine coolant system fault diagnostics with sliding mode observers and fuzzy analyser*, IASTED International Conference on Modelling, Identification and Control, Innsbruck, Austria, 1999.

[Burton 1986] Burton, J.A.; Zinober, A.S.I., *Continuous approximation of variable structure control*, International Journal of Systems Science, Vol. 17, pp. 876-885, 1986.

[Chen 1996] Chen, J.; Patton, R.J.; Zhang, H., *Design of unknown input observers and robust fault detection filters*, International Journal of Control, Vol. 63, pp. 85-105, 1996.

[Chen 1999] Chen, J.; Patton, R.J., *Robust model based fault diagnosis for dynamic systems*, Kluwer Academic Publishers, 1999.

[Chow 1984] Chow, E.Y.; Willsky, A.S., *Analytical redundancy and the design of robust failure detection systems*, IEEE Transactions on Automatic Control, Vol. 29(7), pp. 603-614, July 1984.

REFERENCES

- [Davis 2000]** Davis, R.P.; Abreu, R.A.; Chew, A.D., *Dry vacuum pumps: A method for the evaluation of the degree of dry*, Journal of Vacuum Science and Technology A, Vol. 18(4), pp. 1782-1788, July/August 2000.
- [DeCarlo 1988]** DeCarlo, R.A.; Zak, S.H.; Matthews, G.P., *Variable structure control of nonlinear multivariable systems: a tutorial*, Proceedings of the IEEE, Vol. 76(3), pp. 212-232, 1988.
- [DeJager 1992]** DeJager, B., *Comparison of methods to eliminate chattering and avoid steady state errors in sliding mode digital control*, Proceedings of the IEEE VSC and Lyapunov Workshop, Sheffield, pp. 37-42, 1992.
- [Ding 1994]** Ding, X.; Guo, L.; Frank, P., *Parameterization of linear observers and its application to observer design*, IEEE Transactions on Automatic Control, AC-39, pp. 1648- 1652, 1994.
- [Dorling 1986]** Dorling, C.M.; Zinober, A.S.I., *Two approaches to hyperplane design in multivariable variable structure control systems*, International Journal of Control, Vol. 44, pp. 65-82, 1986.
- [Duval 1987]** Duval, P., *Will tomorrow's high-vacuum pumps be universal or highly specialized?* Journal of Vacuum Science and Technology A, Vol. 5(4), pp. 2546-2551, July/August 1987.

REFERENCES

- [Duval 1989] Duval, P., *Selection criteria for oil-free vacuum pumps*, Journal of Vacuum Science and Technology A, Vol. 7(3), pp. 2369-2372, May/June 1989.
- [Edelbaher 2006] Edelbaher, G.; Jezernik, K.; Urlep, E., *Low-speed sensorless control of induction machine*, IEEE Transactions on Industrial Electronics, Vol. 53(1), pp. 120-129, February 2006.
- [Edwards 1994] Edwards, C.; Spurgeon, S.K., *On the development of discontinuous observers*, International Journal of Control, Vol. 59(5), pp. 1211-1229, 1994.
- [Edwards 1998] Edwards, C.; Spurgeon, S.K., *Sliding mode control: theory and applications*, UK: Taylor & Francis, 1998.
- [Edwards 2000] Edwards, C.; Spurgeon, S.K., *A sliding mode observer based FDI scheme for the ship benchmark*, European Journal of Control, Vol. 6(4), pp. 341-355, 2000.
- [Edwards 2000] Edwards, C.; Spurgeon, S.K.; Patton, R.J., *Sliding mode observers for fault detection and isolation*, Automatica, Vol. 36, pp. 541-553, 2000.
- [Emelyanov 1996] Emelyanov, S.V.; Korovin, S.K.; Levant, A., *High order sliding model in control systems*, Computational Mathematics and Modelling, Vol. 7(3), pp. 294-318, 1996.

[Fliess 2003] Fliess, M.; Sira-Ramirez, H., *An algebraic framework for linear identification*, ESAIM Control Optimisation Calculus Variations, Vol. 9, pp. 151-168, 2003.

[Fliess 2003] Fliess, M.; Join, C., *An algebraic approach to fault diagnosis for linear systems: An algebraic setting with examples*, Proceedings of the Computer Engineering System Applications, 2003.

[Fliess 2004] Fliess, M.; Join, C.; Sira-Ramirez, H., *Robust residual generation for linear fault diagnosis: An algebraic setting with examples*, International Journal of Control, Vol. 77(14), pp. 1223-1242, 2004.

[Frank 1990] Frank, P.M., *Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy – A survey and some new results*. Automatica, Vol. 26(3), pp. 459-474, 1990.

[Frank 1994] Frank, P.M., *Enhancement of robustness in observer based fault detection*, International Journal of Control, Vol. 59, pp. 955-981, 1994.

[Fridman 2002] Fridman, L.; Levant, A., *Higher order sliding modes*, In Perruquetti, W. and J.P. Barbot, *Sliding mode control in engineering*, Marcel Dekker, Inc, pp. 53-101, 2002.

[Furuta 1990] Furuta, K., *Sliding mode control of a discrete system*, Systems and Control Letters, Vol. 14, pp. 145-152, 1990.

REFERENCES

[Gao 1995] Gao, W.; Wang, Y.; Homaifa, A., *Discrete-time variable structure control systems*, IEEE Transactions on Industrial Electronics, Vol. 42(2), pp. 117-122, 1995.

[Gertler 1988] Gertler, J.J., *Survey of model-based failure detection and isolation in complex plants*, IEEE Control System Magazine, pp. 3-11, December 1988.

[Gertler 1998] Gertler, J.J., *Fault detection and diagnosis in engineering systems*, New York: Marcel Dekker, 1998.

[Goh 2002] Goh, K.B.; Spurgeon, S.K.; Jones, N.B., *Fault diagnostics using sliding mode techniques*, Control Engineering practice, Vol. 10(2), pp. 207-217, 2002.

[Goh 2003] Goh, K.B.; Spurgeon, S.K.; Jones, N.B., *Higher-order sliding mode control of a diesel generator set*, Proceedings of the IMechE Part I-Journal of Systems and Control Engineering, Vol. 217(3), pp. 229-241, 2003

[Hablanian 1984] Hablanian, M.H., *Comments on the history of vacuum pumps*, Journal of Vacuum Science and Technology A, Vol. 2(2), pp. 118-125, April/June 1984.

[Hablanian 1988] Hablanian, M.H., *The emerging technologies of oil-free vacuum pumps*, Journal of Vacuum Science and Technology A, Vol. 6(3), pp. 1177-1182, May/June 1988.

REFERENCES

- [Hablanian 1990]** Hablanian, M.H., *Design and performance of oil-free pumps*, Vacuum, Vol. 41(7-9), pp. 1814-1818, 1990.
- [Hablanian 2003]** Hablanian, M.H., *Major advances in transfer pumps: 1953-2003*, Journal of Vacuum Science and Technology A, Vol. 21(5), pp. S15-S18, September/October 2003.
- [Harris 2001]** Harris, N., *Modern vacuum practice*, 2nd Edition, Bell and Bain, 2001.
- [Hermans 1996]** Hermans, F.J.J.; Zarrop, M.B., *Sliding mode observer for robust sensor monitoring*, Proceedings of the 13th IFAC World Congress, pp. 211-216, 1996.
- [Herrmann 2003]** Herrmann, G.; Spurgeon, S.K.; Edwards, C., *A model-based sliding mode control methodology applied to the HAD-plant*, Journal of Process Control, Vol. 13, pp. 129-138, 2003.
- [Hung 1993]** Hung, J.Y.; Gao, W.; Hung, J.C., *Variable structure control: a survey*, IEEE Transactions on industrial electronics, Vol. 40(1), pp. 2-22, 1993.
- [Isermann 1984]** Isermann, R., *Process fault detection based on modelling and estimation methods – A survey*, Automatica, Vol. 20(4), pp. 387-404, 1984.
- [Isermann 1997]** Isermann, R., *Supervision, fault-detection and fault-diagnosis methods-An introduction*, Control Engineering Practice, Vol. 5(5), pp. 639-652, 1997.

REFERENCES

[Isermann 1997] Isermann, R.; Balle, P., *Trends in the application of model-based fault detection and diagnosis of technical processes*, Control Engineering Practice, Vol. 5(5), pp. 709-719, 1997.

[Isermann 2005] Isermann, R., *Model-based fault-detection and diagnosis – Status and applications*, Annual Reviews in Control, Vol. 29, pp. 71-85, 2005.

[Isermann 2005] Isermann, R., *Fault-diagnosis systems: An introduction from fault detection to fault tolerance*, Springer, Berlin, 2005.

[Jiang 2004] Jiang, J.; Zhang, Y., *A revisit to block and recursive least squares for parameter estimation*, Computers and Electrical Engineering, Vol. 30, pp. 403-416, 2004.

[Join 2004] Join, C.; Fliess, M.; Sira-Ramirez, H., *Fault diagnosis of closed loop linear systems with parametric uncertainties*, Proceedings of the 15th International Workshop on Principles Diagnosis, pp. 111-116, 2004.

[Jones 1973] Jones, H.L., *Thesis: Failure detection in linear systems*, Cambridge, MA, 1973.

[Jones 2000] Jones, N.B.; Spurgeon, S.K.; Twiddle, J.A.; Lim, C.L.; Parikh, C.R.; Goh, K.B., *Aspects of diagnostic schemes for biomedical and engineering systems*, IEE Proceedings Science, Measurement and Technology, Vol. 147(6), pp. 357-362, 2000.

REFERENCES

[Kalman 1960] Kalman, R.E., *A new approach to linear filtering and prediction problems*, Transactions of the ASME: Journal of Basic Engineering, Vol. 82, pp. 35-45, 1960.

[Kay 1981] Kay, S.M.; Marple, S.L., *Spectrum analysis-a modern perspective*, Proceedings of the IEEE, Vol. 69(11), pp. 1380-1419, 1981.

[Kim 2000] Kim, J.; Oh, S.; Cho, D.; Hedrick, J.K., *Robust discrete-time variable structure control methods*, Transactions of the ASME, Vol. 122, pp. 766-775, 2000.

[Kiral 2003] Kiral, Z.; Karagulle, H., *Simulation and analysis of vibration signals generated by rolling element bearing with defects*, Tribology International, Vol. 36, pp. 667-678, 2003.

[Konishi 1999] Konishi, S.; Yamasawa, K., *Diagnostic system to determine the in-service life of dry vacuum pumps*, IEE Proceedings Science, Measurement and Technology, Vol. 146(6), pp. 270-276, November 1999.

[Koshkouei 2000] Koshkouei, A.J.; Zinober, A.S.I., *Sliding mode control of discrete-time systems*, Journal of Dynamic Systems Measurement and Control, Vol. 122, pp. 793-802, 2000.

[Koshkouei 2002] Koshkouei, A.J.; Zinober, A.S.I., *Sliding mode state observers for discrete-time linear systems*, International Journal of Systems Science, Vol. 33(9), pp. 751-758, 2002.

REFERENCES

[Krishnaswami 1994] Krishnaswami, V.; Luh, G-C.; Rizzoni, G., *Fault detection in IC engines using nonlinear parity equations*, Proceedings of the American Control Conference, American Automatic Control Council, Vol. 2, pp. 2001-2005, 1994.

[Lessard 2000] Lessard, P.A., *Dry vacuum pumps for semiconductor processes: Guidelines for primary pump selection*, Journal of Vacuum Science and Technology A, Vol. 18(4), pp. 1777-1781, July/August 2000.

[Levant 1993] Levant, A., *Sliding order and sliding accuracy in sliding mode control*, International Journal of Control, Vol. 58(6), pp. 1247-1263, 1993.

[Levant 1998] Levant, A., *Robust exact differentiation via sliding mode technique*, Automatica, Vol. 34(3), pp. 379-384, 1998.

[Levant 2000] Levant, A.; Pridor, A.; Ben-Asher, J.Z.; Gitizadeh, R.; Yaesh, I., *2-sliding mode implementation in aircraft pitch control*, Journal of Guidance Control and Dynamics, Vol. 23(4), pp. 586-594, 2000.

[Levant 2001] Levant, A., *Universal single-input-single-output (SISO) sliding-mode controllers with finite-time convergence*, IEEE Transactions on Automatic Control, Vol. 46(9), pp. 1447-1451, 2001.

[Levant 2003] Levant, A., *Higher-order sliding modes, differentiation and output-feedback control*, International Journal of Control, Vol. 76(9-10), pp. 924-941, 2003.

REFERENCES

[Lim 2003] Lim, C.L.; Jones, N.B.; Spurgeon, S.K.; Scott, J.J.A., *Reconstruction of human neuromuscular control signals using a sliding mode control technique*, Simulation Modelling Practice and Theory, Vol. 11, pp. 223-235, 2003.

[Luenberger 1971] Luenberger, D.G., *An introduction to observers*, IEEE Transactions on Automatic Control, Vol. A.C.16, pp. 596-602, 1971.

[Mathew 1984] Mathew, J.; Alfredson, R.J., *The condition monitoring of rolling element bearings using vibration analysis*, Journal of Vibration, Acoustics, Stress, and Reliability in Design, Vol. 106, pp. 447-453, 1984.

[McFadden 1984] McFadden, P.D.; Smith, J.D., *Vibration monitoring of rolling element bearings by the high-frequency resonance technique-A review*, Tribology International, Vol. 17(1), pp. 3-10, February 1984.

[Mechefske 1993] Mechefske, C.K., *Machine condition monitoring: Part 1-Optimum vibration signal lengths*, British Journal of NDT, Vol. 35(9), pp. 503-507, September 1993.

[Mechefske 1993] Mechefske, C.K., *Machine condition monitoring: Part 2-The effects of noise in vibration signal*, British Journal of NDT, Vol. 35(10), pp. 574-579, October 1993.

[Mehra 1971] Mehra, R.K.; Peshon, I., *An innovations approach to fault detection and diagnosis in dynamic systems*, Automatica, Vol. 7, pp. 637-640, 1971.

[Miguel 1996] Miguel, L.J.; Mediavilla, M.; Peran, J.R., *Fault diagnosis system based on sensitivity analysis and fuzzy logic*, 26th International Symposium on Multiple-valued Logic, Santiago De Compostela, Spain, pp. 50-55, 1996.

[Molteberg 1991] Molteberg, G.A., *Thesis: The application of system identification techniques to performance monitoring of four stroke turbocharged diesel engines*, Division of Marine Engineering, Norwegian Institute of Technology, University of Trondheim, 1991.

[Muenchhof 2004] Muenchhof, M.; Isermann, R., *Fault detection and diagnosis of a linear hydraulic servo axis*, Proceedings of the 16th IFAC Symposium on Automatic Control in Aerospace (ACA), International Federation of Automatic Control, 2004.

[Muenchhof 2006] Muenchhof, M., *Model based fault management for a hydraulic servo axis*, Proceedings of the 5th International IFAS, Aachen, 2006.

[Ogata 1997] Ogata, K., *Modern control engineering*, Prentice-Hall International (UK) Limited, London, 3rd Edition, 1997.

[Parikh 2001] Parikh, C.R., *Thesis: The design of classifiers for condition monitoring and fault diagnosis applications*, University of Leicester, Engineering Department, Leicester, LE1 7RH, England, 2001.

REFERENCES

[Patton 1989] Patton, R.; Frank, P.; Clark, R., *Fault diagnosis in dynamic systems: Theory and application*, Prentice Hall International Series in Systems and Control Engineering, 1989.

[Patton 1994] Patton, R.; Chen, J., *A review of parity space approaches to fault diagnosis for aerospace systems*, Journal of Guidance Control and Dynamics, Vol. 17, pp. 278-285.

[Pouliezos 1994] Pouliezos, A.; Stavrakakis, G.S., *Real time fault monitoring of industrial processes*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1994.

[Prashad 1987] Prashad, H., *The effect of cage and roller slip on the measured defect frequency response of rolling-element bearings*, ASLE Transactions, Vol. 30(3), pp. 360-367, 1987.

[Ryans 2001] Ryans, J.; Bays, J., *Run clean with dry vacuum pumps*, Eastman Chemical Co, October 2001.

[Sarpturk 1987] Sarpturk, S.Z.; Istefanopoulos, Y.; Kaynak, O., *On the stability of discrete-time sliding mode control systems*, IEEE Transactions on Automatic Control, Vol. 32(10), pp. 930-932, 1987.

REFERENCES

[Shtessel 2004] Shtessel, Y.; Poznyak, A., *Parameter identification of linear time varying systems via traditional and high order sliding modes*, Proceedings of the International Workshop on Variable Structure Systems, Villanova, Spain, September 6-8, 2004.

[Simonson 1998] Simonson, J.R., *Engineering heat transfer*, 2nd Edition, MacMillan Education Ltd, 1988.

[Sira-Ramirez 2002] Sira-Ramirez, H.; Fliess, M., *On the discrete-time uncertain visual based control of planar manipulators: An on-line algebraic identification approach*, Proceedings 41st IEEE Conf. Decision Control, 2002.

[Slotine 1987] Slotine, J.-J.E.; Hedrick, J.K.; Misawa, E.A., *On sliding observers for nonlinear systems*, Journal of Dynamic Systems, Measurement, and Control, Vol. 109, pp. 245-252, September 1987.

[Spurgeon 1992] Spurgeon, S.K., *Hyperplane design techniques for discrete-time variable structure control systems*, International Journal of Control, Vol. 55(2), pp. 445-456, 1992.

[Stensby 1997] Stensby, J.L., *Phase-locked loops: Theory and application*, CRC Press, Boca Raton, 1997.

[Sunnarsjo 1978] Sunnarsjo, C.S., *Varying compliance vibrations of rolling bearings*, Journal of Sound and Vibration, Vol. 58(3), pp. 363-373, 1978.

REFERENCES

- [Tan 2002] Tan, C.P.; Edwards, C., *Sliding mode observers for detection and reconstruction of sensor faults*, *Automatica*, Vol. 38, pp. 1815-1821, 2002.
- [Tandon 1999] Tandon, N.; Choudhury, A., *A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings*, *Tribology International*, Vol. 32, pp. 469-480, 1999.
- [Thanagasundram 2007] Thanagasundram, S., *Thesis: Fault prediction in quasi steady state rotating machines*, University of Leicester, Engineering Department, Leicester, LE1 7RH, England, 2007.
- [Troup 1989] Troup, A.P.; Turrell, D., *Dry pumps operating under harsh conditions in the semiconductor industry*, *Journal of Vacuum Science and Technology A*, Vol. 7(3), pp. 2381-2386, 1989.
- [Troup 1991] Troup, A.P.; Dennis, N.T.M., *Six years of "dry pumping": A review of experience and issues*, *Journal of Vacuum Science and Technology A*, Vol. 9(3), pp. 2048-2052, May/June 1991.
- [Twiddle 2001] Twiddle, J.A., *Thesis: Fuzzy systems in real-time condition monitoring and fault diagnosis, with a diesel engine case study*, University of Leicester, Engineering Department, Leicester, LE1 7RH, England, 2001.
- [Unbehauen 1998] Unbehauen, H.; Rao, G.P., *A review of identification in continuous-time systems*, *Annual Reviews in Control*, Vol. 22, pp. 145-171, 1998.

REFERENCES

- [Utkin 1977] Utkin, V.I., *Variable structure systems with sliding modes*, IEEE Transactions on Automatic Control, Vol. AC22(2), pp. 212-222, 1977.
- [Utkin 1984] Utkin, V.I., *Variable structure systems: present and future*, Automation and Remote Control, Vol. 44(1) (9), pp. 1105-1120, 1984.
- [Utkin 1992] Utkin, V.I., *Sliding modes in control and optimization*, London: Springer-Verlag, 1992.
- [Utkin 1993] Utkin, V.I., *Sliding mode control design principles and applications to electric drives*, IEEE Transactions of Industrial Electronics, Vol. 40(1), pp. 23-36, February 1993.
- [Viswanadham 1987] Viswanadham, N.; Sarma, V.V.S.; Singh, M.G., *Reliability of computer and control systems*, North Holland Systems and Control Series, Vol. 8, North Holland, Amsterdam, 1987.
- [Walcott 1987] Walcott, B.L.; Zak, S.H., *State observation of nonlinear uncertain dynamical systems*, IEEE Transaction on Automatic Control, Vol. 32, pp. 166-170, 1987.
- [Willsky 1976] Willsky, A.S., *A survey of design methods for failure detection in dynamic systems*, Automatica, Vol. 12, pp. 601-611, 1976.

REFERENCES

[Wycliffe 1987] Wycliffe, H., *Mechanical high-vacuum pumps with an oil-free swept volume*, Journal of Vacuum Science and Technology A, Vol. 5(4), pp. 2608-2610, July/August 1987.

[Yang 1995] Yang, H.; Saif, M., *Nonlinear adaptive observer design for fault detection*, Proceedings of the American Control Conference, pp. 1136-1139, 1995.

[Young 1981] Young, P., *Parameter estimation for continuous time model. A survey*, Automatica, Vol. 17(1), pp. 23-39, 1981.

[Young 1999] Young, K.D.; Utkin, V.I.; Ozguner, U., *A control Engineer's guide to sliding mode control*, IEEE Transactions on Control Systems Technology, Vol. 7(3), pp. 328-342, 1999.

[Zakrzewski 1988] Zakrzewski, E.; May, P.L.; Emslie, B.S., *Developments in vacuum pumping systems based on mechanical pumps with an oil free swept volume*, Vacuum, Vol. 38(8-10), pp. 757-760, 1998.

APPENDICES

APPENDIX A

DETAILS OF DRY VACUUM PUMPS

A.1 INTRODUCTION

Appendix A gives greater details on the principles of dry vacuum pumps. Most of the material discussed is relevant to the existing literature on rotating machinery. Section A.2 covers the theory behind basic vacuum technology, whereas section A.3 provides a brief review of different dry vacuum pump designs.

A.2 BASIC VACUUM THEORY

The need and benefits of performing different processes under a vacuum has led to the firm establishment of vacuum pumps systems in industry. In order to choose correctly a vacuum system the requirement to understand the basic principles of vacuum theory is essential.

Vacuum can be defined as a space from which all air and other gases have been removed. However, this is an ideal condition since to remove all gaseous material from a space is practically not possible. Harris [Harris 2001] stated that vacuum is a pressure below that of the local atmospheric pressure (approx. 1013 mbar at sea level). Pressure

is defined as the force divided by the area and force as mass multiplied by acceleration. The smaller the mass or force at a given area it will result to a lower pressure. Thus, vacuum can be created in a vessel by a vacuum system by simply removing mass from it. The more mass removed, the lower the pressure that exists inside the vessel.

Given that the term vacuum is used to describe pressures below the local atmospheric, then it is desirable to separate this wide vacuum spectrum into ranges [Harris 2001]:

Low or rough vacuum:	1013 mbar to a few mbar
Medium vacuum:	a few mbar to 10^{-3} mbar
High vacuum:	10^{-3} mbar to 10^{-7} mbar
Ultra-high vacuum:	below 10^{-7} mbar

Knowing the type of application someone may indicate in which range it belongs. For instance, chemical process industries (CPI) operate mostly in the rough vacuum range, while semiconductor applications span from rough to ultra-high vacuum.

Once the importance of vacuum is understood concepts such as pump speed, mass flow rate and conductance can be studied in order to choose the best type of system for a particular application. The following section contains a brief review of various dry vacuum pump designs.

A.3 A REVIEW OF DRY VACUUM PUMP DESIGNS

A.3.1 PISTON DESIGN

The piston design resembles a car engine with opposed pistons and it is one of the oldest technologies in the area. Designers have a long experience with this type of pump and the most popular design has four pistons adjusted in a parallel and series arrangement (see Figure A-1).

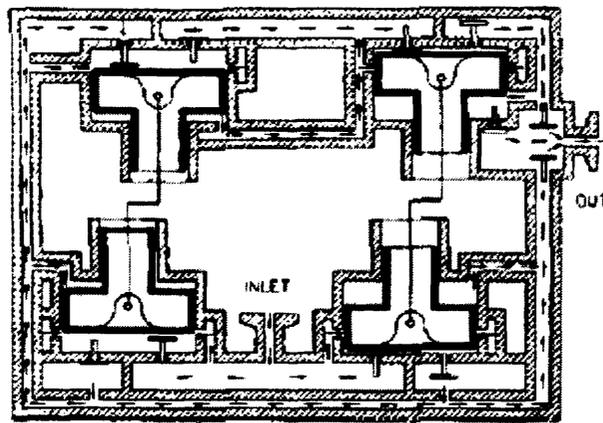


Figure A-1: Schematic arrangement of a four-stage piston pump (adapted from Hablanian [Hablanian 1990])

Although the piston design provides poor pumping speed at low pressures it also has important benefits. The most important feature is the ability to remain cool while in operation. Other designs develop high temperatures and often require a cooling system. In contrast, on piston pumps the large surface areas and the fact that the entire drive mechanism is in atmospheric air, allows the friction and compression generated heat to be distributed easily. Therefore, at the time of its operation a piston pump is only a few degrees above atmospheric temperature.

A.3.2 SCREW COMPRESSOR DESIGN

Even though the design of screw vacuum pumps was known since the mid-1950s, only recently have they become noticed due to their simple structure. The design adapts the principle of Archimedean Screws. Two meshing screws rotate in opposite direction in a case or a stator. The gas moves along the length of the screws and gas compression occurs in the final half-turn of the screws. The mechanism is illustrated in Figure A-2.

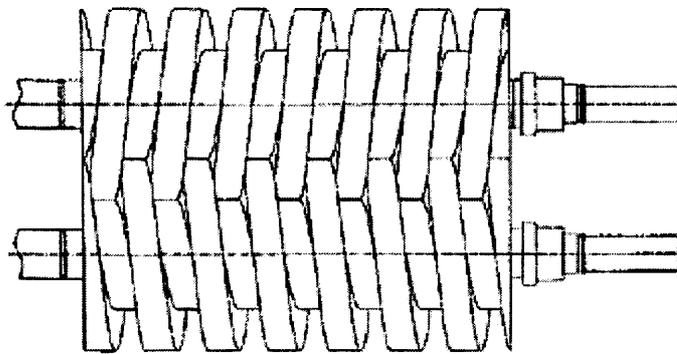


Figure A-2: The screw type mechanism (adapted from Lessard [Lessard 2000])

Screw compressors run at high speeds and give a very high compression ratio at high pressure. However, to achieve such a performance very tight clearances between the rotors are necessary. Further, high rotational speeds result in high noise levels and high temperatures.

A.3.3 SCROLL DESIGN

The scroll design consists of a pair of helical scrolls, one fixed and one orbiting (see Figure A-3). This mechanism traps the volume of gas and progressively forces it to the exhaust, producing a vacuum. The design provides good compression ratios at low pressures but at very high production cost.



Figure A-3: Scroll type design (adapted from Lessard [Lessard 2000])

A.3.4 ROOTS LOBE DESIGN

Two different types of roots design exist: a two-lobe and a three-lobe rotor pump. The latter provides a higher compression ration but runs at lower pumping speed when compared with the two-lobe design. Usually, the design is comprised by several sets or stages of roots lobes to improve the compression ratio. The set of rotors are mounted on two parallel shafts and held in phase by timing gears. A roots lobe vacuum pump is depicted in Figure A-4. An important issue that affects the design is the extreme high temperatures generated while the system is running. This overheating may cause thermal expansion, which in turn may result in a contact between the rotors and stator. An inter-stage cooler is a popular solution to this problem, although this results in non-compact designs.

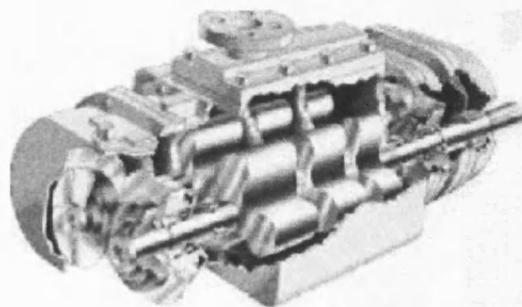


Figure A-4: Schematic of a three stage roots dry pump (adapted from Ryans and Bays [Ryans 2001])

A.4 CONCLUSION

This appendix provided details of the basic vacuum theory in order to appreciate some concepts that have been studied throughout this research. It also presents a number of vacuum pump designs used in the industry today.

APPENDIX B

INSTRUMENTATION AND DATA ACQUISITION

B.1 INTRODUCTION

Appendix B provides a list of the instrumentation fitted on the iGX dry vacuum pump test-bed along with a number of tests performed throughout this research.

B.2 iGX DRY VACUUM PUMP INSTRUMENTATION LIST

Table B-1. Instrumentation list (see Figure B-1 for instrument arrangement)

Signal		Sensor		Notes
		Type	Location	
1.	Vibration	Accelerometer Bruel and Kjaer type 4370-V S	As required	Calibration Cert. Available Signal Amplifier 2692-A-011
2.	Exhaust pressure	ASM gauge	Thermal break purge gas port	Installed and tested BOC-E Calibration certificates supplied
3.	Vacuum	CG16k dial gauge	Vacuum vessel inlet	Installed and tested BOC-E Calibration certificates supplied
4.		Pirani gauge		
5.		Barocel gauge		

APPENDIX B

6.	Air mass flow transducer	Lucas 4AM	Vacuum vessel inlet	Applicable only for Flow rates > 5slpm
7.	Gear tooth proximity	Magnetic pick-up (RS part no. 304-172)	Rotor drive gear	Installed and tested
8.	Current DC Voltage Power Torque Voltage	Yasakawa varispeed 606V7 inverter		Connecting cable from inverter to dSPACE is installed, allowing one signal at a time to be recorded
9.	Temperature 1	IC LM35 (Farnell part no. 409-080)	Cooling water inlet	Installed with signal conditioning
10.	Temperature 2	IC LM35	Cooling water outlet	Installed with signal conditioning
11.	Temperature 3	IC LM35	Pump body	Installed with signal conditioning
12.	Temperature 4	IC LM35	Pump low vacuum bearing section (external)	Installed with signal conditioning
13.	Temperature 5	IC LM35	Pump high vacuum bearing section (external)	Installed with signal conditioning
14.	Coolant mass flow	Pulsed output single jet water meter (RS part no. 399-5018)	Cooling water inlet	Output = 1 pulse per litre

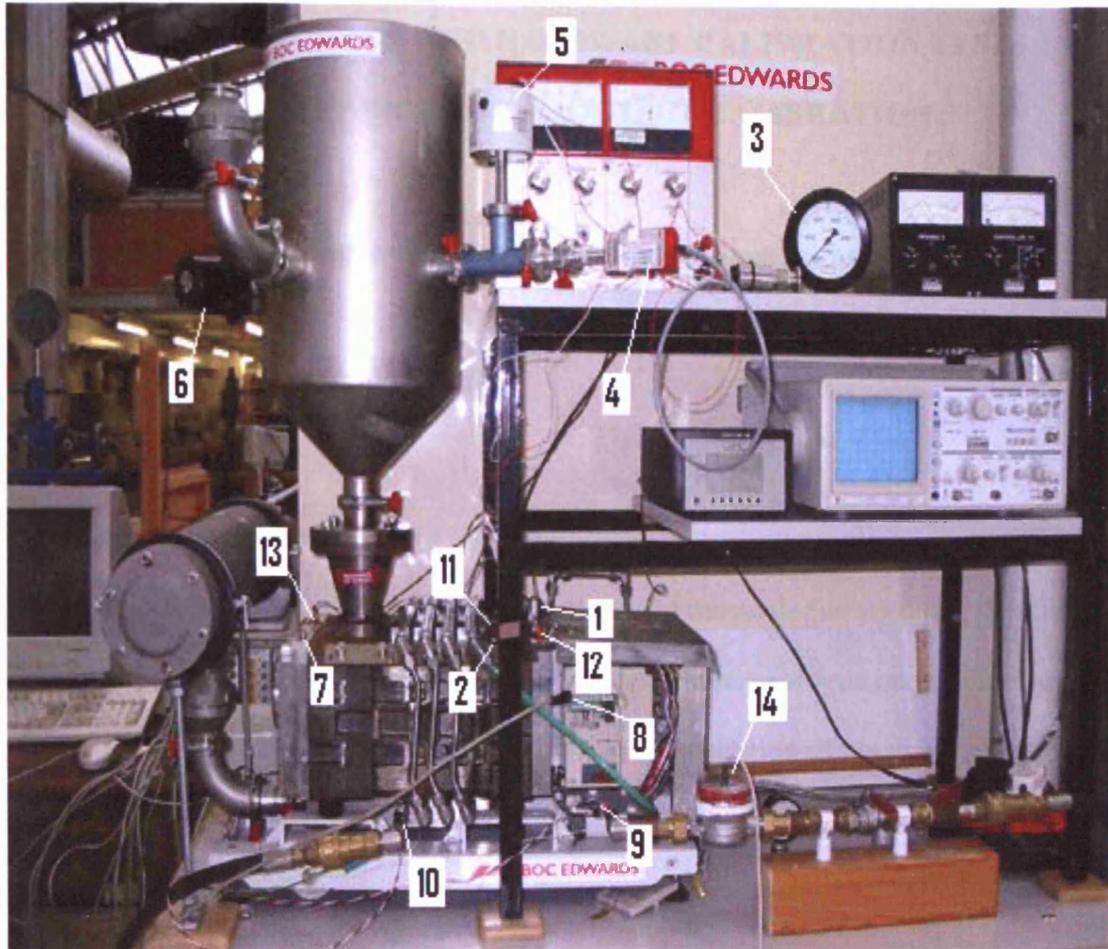


Figure B-1: Dry vacuum pump and Instrumentation. See Table B-1 for key to instruments. Note that the exhaust pressure sensor and inlet mass flow meter were not fitted at the time of the photograph

Table B-1 provides a full list of the sensors fitted on the iGX dry vacuum pump. Many of these sensors such as the accelerometer and air mass flow meter are not part of the condition monitoring and fault detection system presented in this thesis. However, they have been used by fellow researchers in the University or in order to investigate the performance of the vacuum.

B.3 DATA ACQUISITION AND HARDWARE CALIBRATION TESTS

B.3.1 INVERTER CURRENT SIGNAL OUTPUT CALIBRATION

B.3.1.1 OBJECTIVE

To calibrate the voltage output from the inverter with the current indication on the inverter display.

B.3.1.2 METHOD

Set the analogue monitor output (AMO) required constant value setting (i.e. n066) to 1 in order to select the current signal as the output from the analogue terminal. Set the monitored parameter on the display to U-03 and record the analogue output voltages for a series of inverter current values which can be set by variation of the pump inlet control valve.

B.3.1.3 RESULTS

The results are shown in Figure B-2 with a first order least squares best-fit line for the voltage/indicated current relationship.

B.3.1.4 CONCLUSION

A first order linear relationship between the AMO voltage signals V_{AMO} and the indicated current, I_{ind}

$$I_{ind} = 1.0893V_{AMO} - 0.02$$

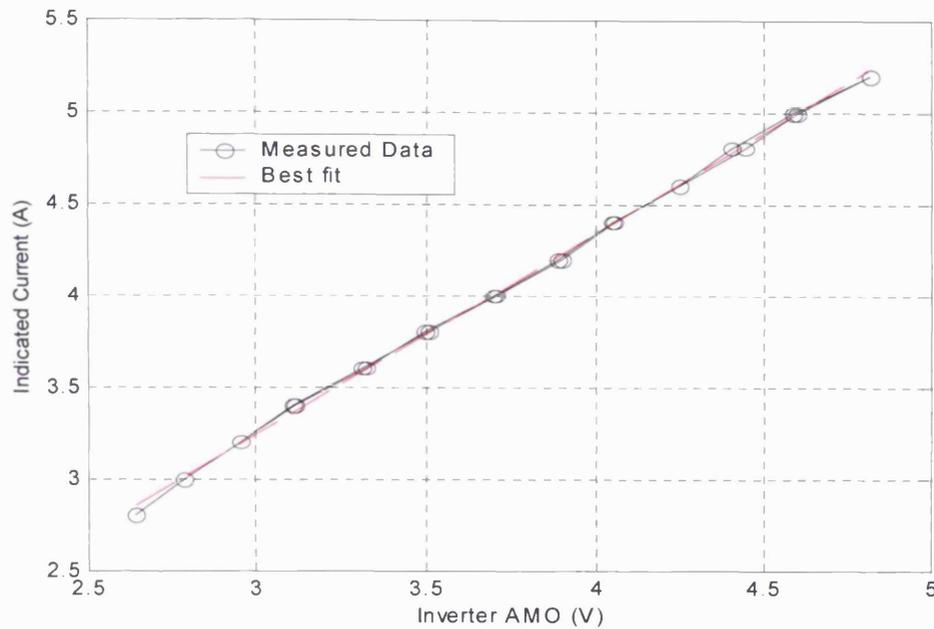


Figure B-2: Inverter signal output calibration data

B.3.2 SIMULTANEOUS RECORDING OF TEMPERATURE AND INVERTER SIGNALS FROM THE iGX DRY VACUUM PUMP

B.3.2.1 OBJECTIVE

Temperature, cooling water flow and inverter current data is required for the identification and validation of the temperature models for the iGX vacuum pump. In order to acquire inputs and outputs for the dynamic models it is desirable to simultaneously acquire the both temperature and pump current signals.

B.3.2.2 PROBLEM FORMULATION

Pump current is obtained as an analogue voltage (Figure B-3) ref [Yasakawa inverter manual pp. 64-65].

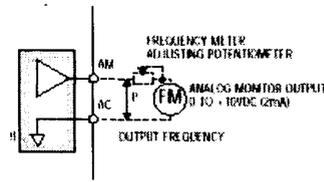


Figure B-3: Analogue output connections

Temperature is measured using a set of three k-type thermocouples (t/cs). The thermocouples are ‘single ended’ BNC connections to the dSPACE™ ADC board. The three thermocouples are electrically insulated from the pump body with heat resistant polyamide tape. The dSPACE™ ADC board PC and t/cs signal amplifiers are earthed to 240 Vmains. The pump is earthed to 415 V 3-phase supply.

If the inverter output signal is connected to the oscilloscope (there is no direct connection between the oscilloscope and the dSPACE™ ADC board or thermocouples) and the pump is switched on then the thermocouple reading is seen to jump (see Figure B-4). If the inverter output is connected directly to the dSPACE™ board via coaxial cable and a BNC connector then the jump in the thermocouple reading is of a much higher magnitude. If the inverter parameter number n080 [Yasakawa Inverter manual pp. 67] is changed to adjust the balance between motor noise and ‘motor current leakage’, then the magnitude of this jump is reduced.

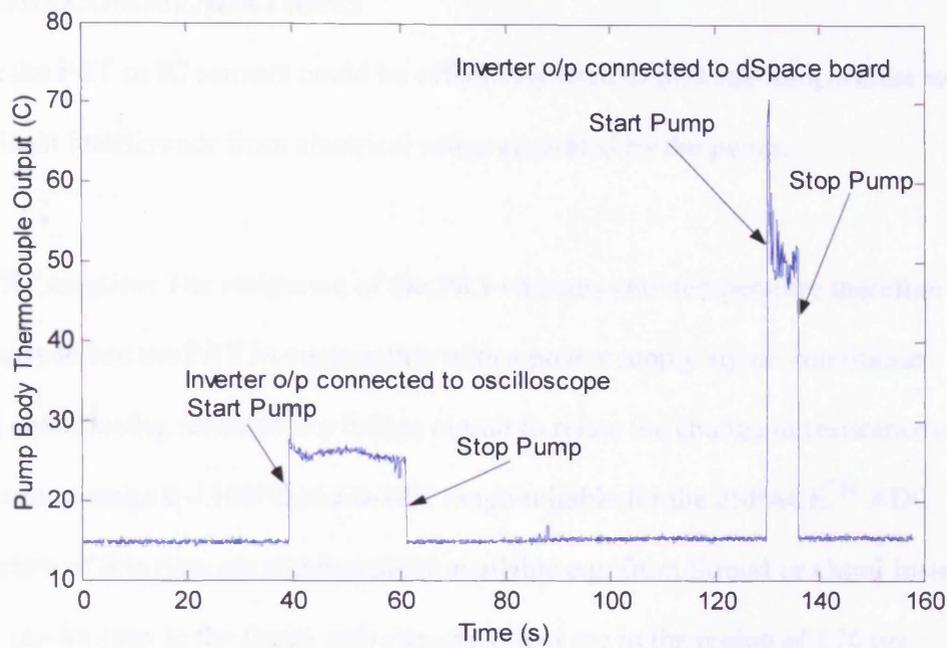


Figure B-4: Effect of inverter output connections on thermocouples readings

Therefore it seems that there are two effects occurring to disturb the thermocouple readings, an effect due to the different ground voltages of the pump and data acquisition system and also due to the effect of the motor leakage current.

As an alternative to the thermocouples two other types of transducer have been tested. A platinum resistance thermometer (PRT) previously used on a diesel engine has been fitted and continued to work well when the inverter current monitoring was connected to the dSPACE™ ADC. Also an integrated circuit (IC) temperature sensor has been tested. This sensor has been found to be highly susceptible to high frequency noise from the pump, however unlike the thermocouple noise, there was no significant offset, or proportionality with the pump current level.

B.3.2.3 RECOMMENDATIONS

Either the PRT or IC sensors could be effectively used to measure temperature without significant interference from electrical noise generated by the pump.

The PRT solution: The resistance of the PRT changes with temperature therefore it is necessary to use the PRT in conjunction with a power supply/signal conditioner. The signal conditioning required is a bridge circuit to relate the change in resistance over the temperature range 0 – 100°C to a 0-10V range suitable for the dSPACE™ ADC. Signal converters of this type are commercially available e.g. from Stroud or Omni Instruments but as can be seen in the Omni web site, the prices are in the region of £70 per instrument channel. An alternative to this would be to design and build a multi-channel signal-conditioning box in-house. Components should be fairly cheap and available off-the-shelf, but the cost would lie in terms of man-hours for design and building of the circuit. The price of a thin film PRT element suitable for surface temperature measurement is £15 for 5 at Farnell (item no. 721-8850)

The IC solution: The IC (R.S. item no. 317-960, @ £1.43 ea.) is a small integrated circuit with a voltage output proportional to temperature over a range of -55 – 150°C. The sensor is contained in a small (5mm) ‘metal can’ housing, not ideal for surface temperature measurements. The sensor is affected by high frequency noise and capacitance of ‘long’ cable connections. This can be countered by fixing a suitable resistor and capacitor in series between the output and ground connections.

The best solution in terms of temperature measurement from a surface would be the thin film PRT, but the IC approach is less costly.

experienced with the thermocouples has improved with only an insignificant change in the temperature signal ($<1^{\circ}\text{C}$) occurring on start-up of the motor. Therefore the acquisition of simultaneous temperature and pump current signals is possible.

Finally, Figure B-7 shows a comparison of the signals obtained from the pump body IC and the pump body thermocouple. Although a difference on the readings can be observed, the outcome is acceptable and within the manufacturers stated tolerances for IC and thermocouple temperature measurements of $\pm 1.5^{\circ}\text{C}$ for the thermocouples and $\pm 0.75^{\circ}\text{C}$ over the full temperature range of -55 to 150°C for the IC sensor.

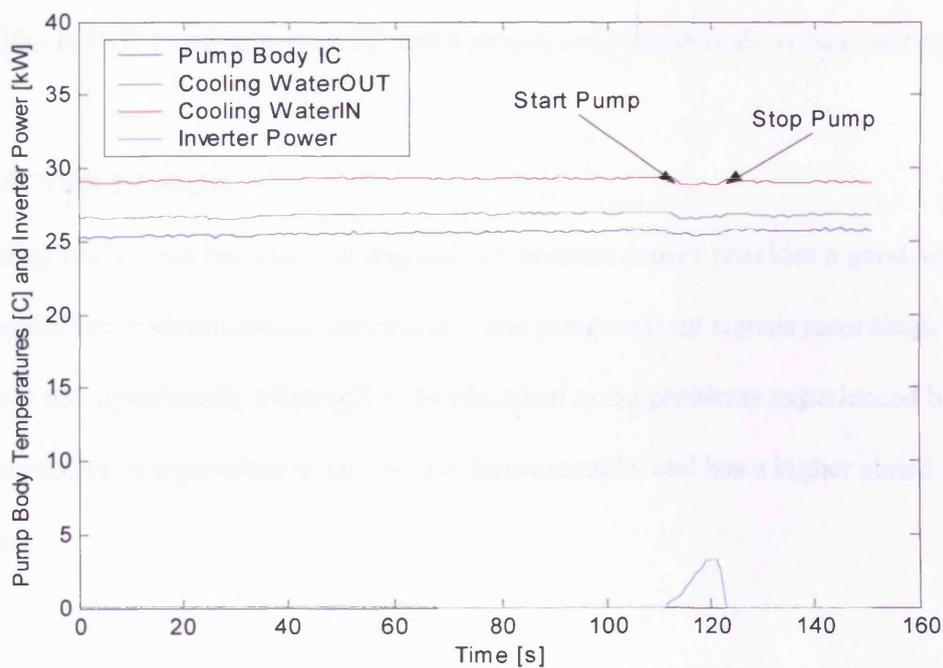


Figure B-6: Effect of inverter output connection on IC readings

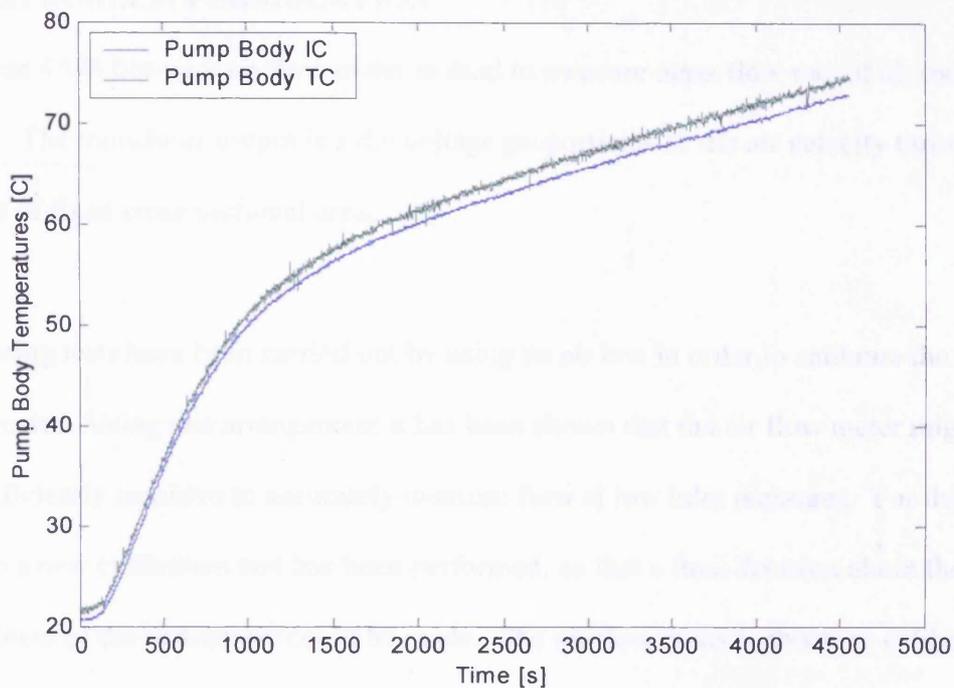


Figure B-7: Comparison on IC and thermocouple pump body sensor readings

B.3.2.5 CONCLUSION

The integrated circuit precision centigrade temperature sensor provides a good solution for acquisition of simultaneous temperature and pump current signals recordings. The sensor is not significantly affected by the electrical noise problems experienced by the thermocouples, is equivalent in cost to the thermocouple, and has a higher stated accuracy.

B.3.3 CALIBRATION OF LUCAS 4AM AIR FLOW METER

B.3.3.1 OBJECTIVE

To calibrate an air flow meter for the measurement of air mass flow rate of the iGX dry vacuum pump.

B.3.3.2 PROBLEM FORMULATION

A Lucas 4AM hot-wire air flow meter is used to measure mass flow rate of air into the pump. The transducer output is a d.c voltage proportional to the air velocity through an orifice of fixed cross sectional area.

Foregoing tests have been carried out by using an air box in order to calibrate the air flow meter. Using this arrangement it has been shown that the air flow meter might not be sufficiently sensitive to accurately measure flow at low inlet pressures. For this reason a new calibration test has been performed, so that a final decision about the usefulness of the instrument could be made. The air flow meter is this time calibrated by using an instrument provided by the University of Leicester. The type 55D41/42 calibration equipment provides a turbulence-free variable-velocity air source for static calibration of hot-wire and hot-film probes. A schematic of the experimental arrangement for air flow meter calibration and the equation used to calculate the flow velocity are given in the following subsection.

B.3.3.3 TYPE 55D41/42 CALIBRATION EQUIPMENT INFORMATION

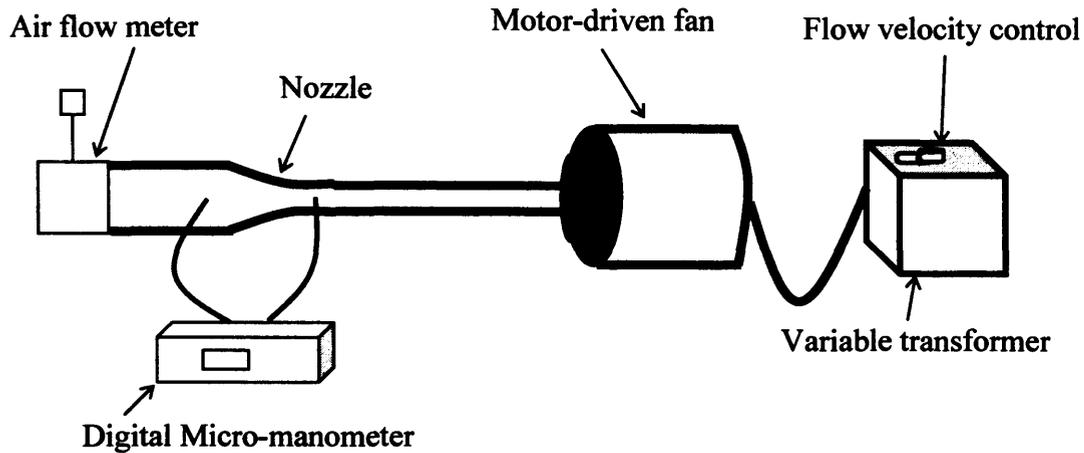


Figure B-8: Schematic showing type 55D41/42 calibration equipment

A digital micro-manometer is used to indicate the pressure difference in the nozzle. The flow velocity, v , in the nozzle measuring-section therefore may be calculated by means of the formula

$$v = \frac{2\kappa}{\kappa - 1} \times R \times T_o \times \left[1 - \left(1 - \frac{\Delta p}{P_o} \right)^{\frac{\kappa - 1}{\kappa}} \right] \quad (\text{B.1})$$

where k is the isentropic exponent of the gas in question, and R is its gas constant (for air, $k = 1.4$ and $R = 287.1 \text{ J/kgK}$). T_o and P_o are ambient temperature and ambient pressure in the nozzle reservoir, in other words: at the nozzle inlet. Finally, Δp is the pressure difference measured across the nozzle.

B.3.3.4 TEST PROCEDURE

The equipment has been set-up as shown in Figure B-8. The air flow meter is connected before the type 55D41/42 calibration equipment and readings are recorded from the micro-manometer and air flow meter at various airflow velocities.

B.3.3.5 SIGNALS

Signals to be measured:

- a) Air flow meter reading (voltage, V)
- b) Micro-manometer reading (pressure, Pa)
- c) Ambient temperature ($^{\circ}\text{C}$)
- d) Ambient pressure (mmHg)
- e) Flow velocity (m/s)

B.3.3.6 RESULTS

Atmospheric pressure at start, $P_o = 757.9$ mmHg

Ambient temperature at start, $T_o = 291.2$ K

The output voltages from the air flow meter are tabulated with the pressure values recorded from the micro-manometer. Additionally the calculated flow velocity from equation (B.1) is also given. The recorded data are shown in figure B-9.

Table B-2. Air flow meter calibration test data

Pressure [Pa]	AFM Voltage [V]	Flow Velocity [m/s]
0	0.580	0
0.31	0.582	0.0716
0.56	0.585	0.0963

APPENDIX B

1.19	0.603	0.1403
2.06	0.639	0.1846
2.27	0.654	0.1938
4.68	0.730	0.2783
5.30	0.754	0.2962
6.55	0.783	0.3292
7.70	0.800	0.3570
9.85	0.837	0.4038
14.8	0.900	0.4949
18.7	0.945	0.5563
29.3	1.050	0.6964
35.5	1.080	0.7665
49.5	1.180	0.9079
70.5	1.265	1.0802
85.2	1.310	1.1875
98.1	1.350	1.2742
117	1.390	1.3915
133	1.420	1.3915

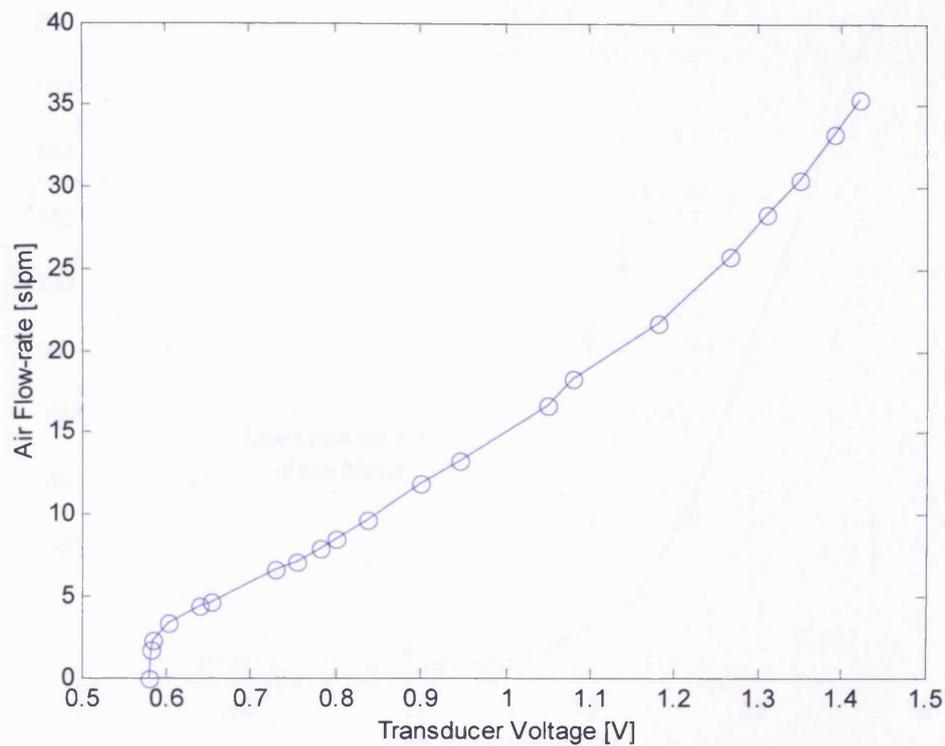


Figure B-9: Plot of air flow meter output voltage vs. air flow rate

B.3.3.7 DISCUSSION

From Figure B-9 the relationship of the transducer output voltage to air flow rate is presented. As with the previous tests, the results show that the air flow meter provides a non-linear but repeatable output, which is related to mass flow rate. Moreover, the plot of Figure B-10 shows that the air flow meter is not sensitive enough to measure flow at low vacuum pump inlet pressures (<5 mbar).

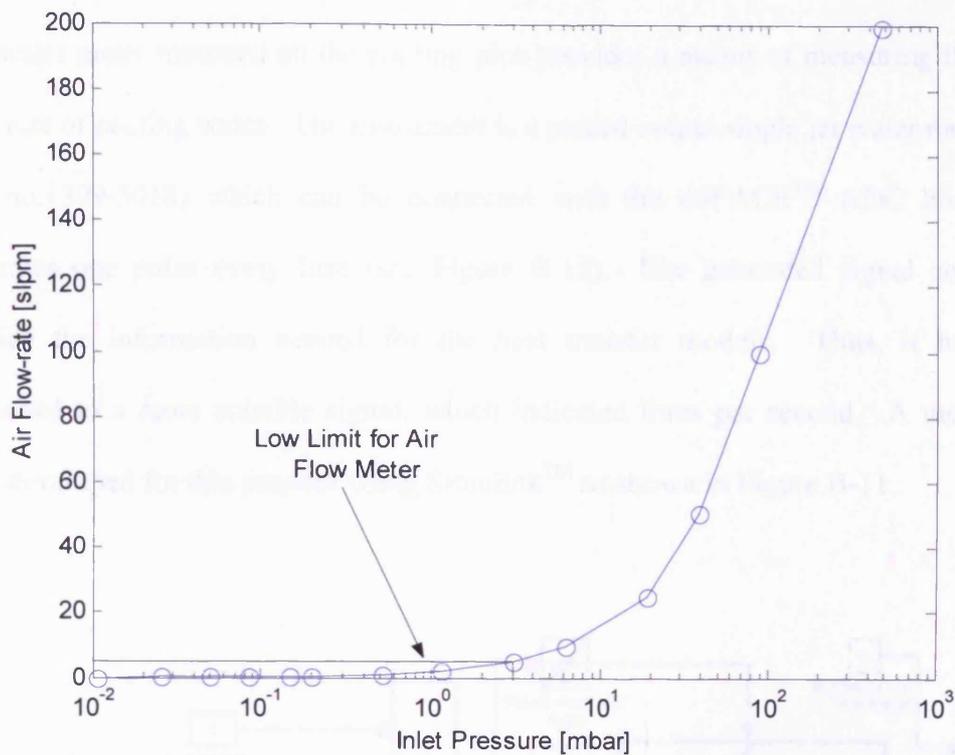


Figure B-10: Plot of the inlet pressure vs. air flow rate

B.3.3.8 CONCLUSION

The Lucas 4AM hot-wire air flow meter has been shown to provide an indication of air-flow rates above 5 slpm. Comparing this with the pump performance curves shows that the meter can be used to indicate pump mass flow-rates at typical pump inlet pressures (5 mbar-100 mbar).

B.3.4 COOLING WATER FLOW METER

B.3.4.1 OBJECTIVE

To convert the coolant flow pulse from water meter to a measurable signal in order to be used in the condition monitoring and fault detection scheme.

B.3.4.2 PROBLEM FORMULATION

The water meter mounted on the cooling pipe provides a means of measuring the mass flow rate of cooling water. The instrument is a pulsed output single jet water meter (RS part no. 399-5018) which can be connected with the dSPACE™ ADC board and generates one pulse every litre (see Figure B-12). The generated signal could not provide the information needed for the heat transfer models. Thus, it has been converted to a more suitable signal, which indicated litres per second. A model has been developed for this purpose using Simulink™ as shown in Figure B-11.

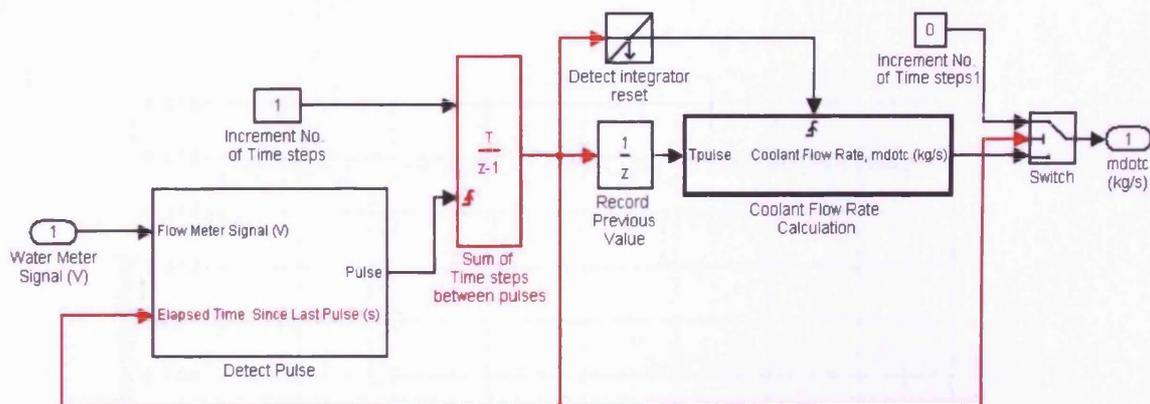


Figure B-11: Block diagram of cooling water pulse converter

B.3.4.3 RESULTS

The resulting signal from the water meter for this test is shown in Figure B-13, indicating a cooling water flow rate of approximately 0.015 l/sec when switched on.

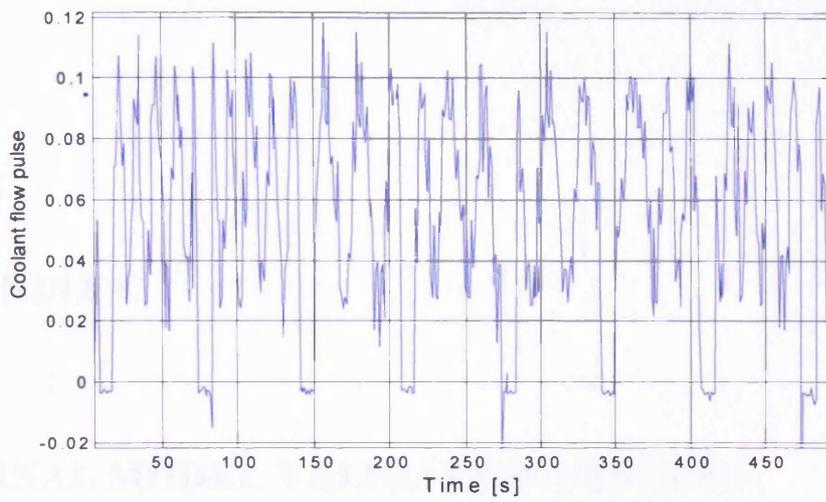


Figure B-12: Graph of the cooling water pulse

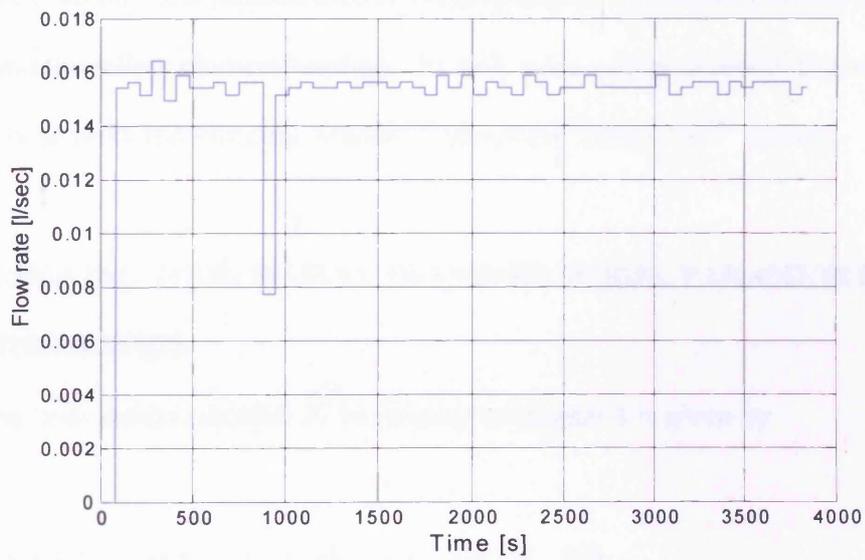


Figure B-13: Graph of the cooling water flow rate

APPENDIX C

NOMINAL MODEL VALIDATION RESULTS

C.1 INTRODUCTION

Appendix C outlines the method used to validate the nominal models for the cooling system and the rolling element bearings. In both cases a large quantity of temperature data has been recorded using the MatlabTM/SimulinkTM/dSPACETM system.

C.2 COOLING SYSTEM HEAT TRANSFER MODEL PARAMETER ESTIMATION

The pump body model equation as formulated in chapter 4 is given by

$$\dot{T}_{B1}(t) = a_1 k l(t) - a_2 m \dot{\kappa} (T_o(t) - T_i(t)) - a_3 (T_{B1}(t) - T_{am}(t)) \quad (C.1)$$

where a_1 , a_2 and a_3 are given by $a_1 = \frac{1}{(mc_p)_B}$, $a_2 = \frac{c_{pcw}}{(mc_p)_B}$ and $a_3 = \frac{(hA)_B}{(mc_p)_B}$.

Let $a_1 k = a$, $a_2 m \dot{\kappa} = b$ and $a_3 = c$. In order to estimate the three parameters a superposition method has been adopted. This is because it has been noticed that some

parts of the mathematical equation can be eliminated. Therefore, initially if the pump is cooling down with the power and the cooling water both disconnected then equation (C.1) can be rewritten as

$$\dot{T}_{B1}(t) = -c(T_{B1}(t) - T_{atm}(t)) \quad (C.2)$$

Hence,

$$\frac{1}{(T_{B1}(t) - T_{atm}(t))} \dot{T}_{B1}(t) = -c \quad (C.3)$$

Integrating equation (C.3) with respect to time, t , gives

$$\log_e(T_{B1}(t) - T_{atm}(t)) = -ct + C \quad (C.4)$$

where C is the constant of integration. At time $t = 0$, $T_{B1} = T_{B0}$ therefore

$$\log_e(T_{B0}(t) - T_{atm}(t)) = C \quad (C.5)$$

and

$$\begin{aligned} \log_e(T_{B1}(t) - T_{atm}(t)) &= -ct + \log_e(T_{B0}(t) - T_{atm}(t)) \Rightarrow \\ \log_e\left(\frac{T_{B1}(t) - T_{atm}(t)}{T_{B0}(t) - T_{atm}(t)}\right) &= -ct \end{aligned} \quad (C.6)$$

Taking the inverse logarithm results

$$\frac{T_{B1}(t) - T_{atm}(t)}{T_{B0}(t) - T_{atm}(t)} = e^{-ct} \quad (C.7)$$

The model parameter c can be estimated from the gradient of a graph of

$\log_e(T_{B1}(t) - T_{atm}(t) / T_{B0}(t) - T_{atm}(t))$ against time (see Figure C-1). Using a least

squares approach the parameter, c has been found to be equal to 6.1×10^{-5} .

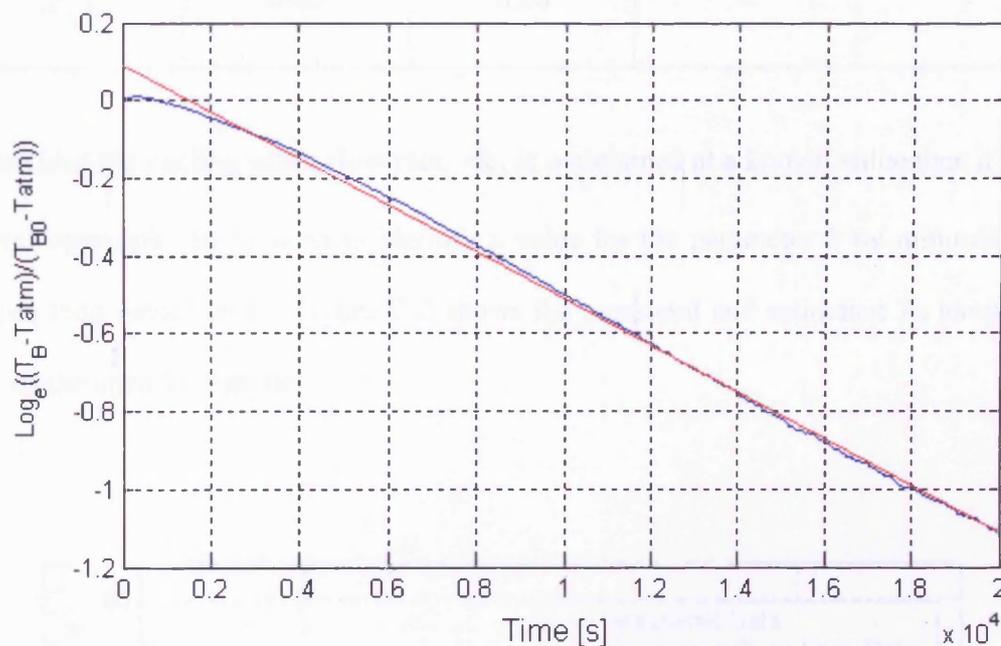


Figure C-1: Graph of regression data and best-fit line for parameter c

After estimating c a second data set has been collected. Once again the cooling down temperature is observed, but this time the valve that controls the coolant has been left on with the power disconnected. Therefore, equation (C.1) becomes

$$\dot{T}_{B1}(t) = -b(T_o(t) - T_i(t)) - c(T_{B1}(t) - T_{atm}(t)) \quad (C.8)$$

Parameter b is given by $b = a_2 \dot{m}_c$ and $a_2 = \frac{c_{pcw}}{(mc_p)_B}$. Some of these parameters are

known for instance the mass of the pump body (80 Kg) and others can be found from tables [Simonson 1988] (see Table C-1 below).

Table C-1. Specific heat capacity of water and cast iron

	0°C	20°C	40°C	60°C
$(c_p)_{cw}$	4.218	4.182	4.178	4.184
$(c_p)_B$	0.42	0.46	-	-

Provided the cooling water flow rate, \dot{m}_c , is maintained at a known value then a trial and error approach can be used to identify a value for the parameter b by minimising the open loop model error. Figure C-2 shows the measured and estimated T_B temperature from the open loop model.

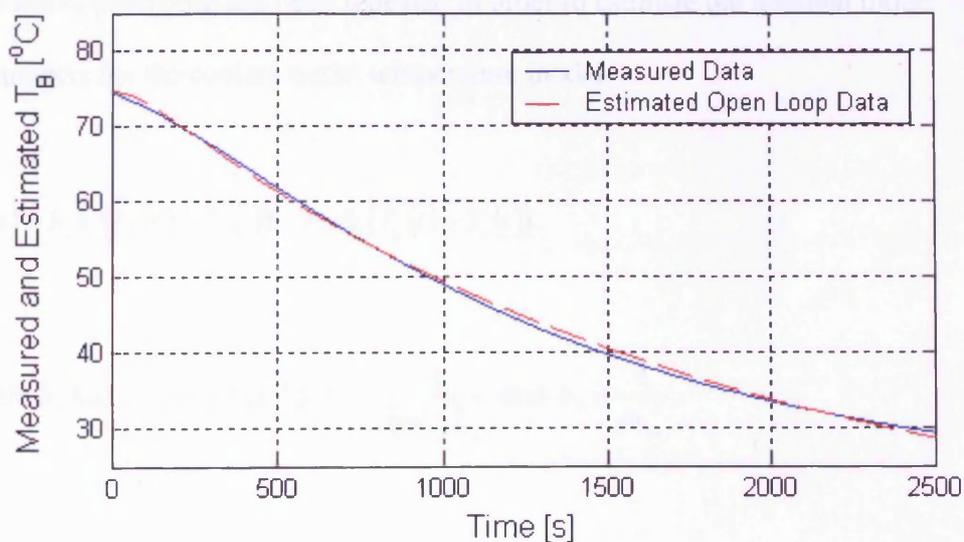


Figure C-2: Plot of measured and estimated T_B open loop data with coolant water on

Having estimated parameters b and c then a third and final data set can be collected with current and flow rate maintained at a known value. This final set has been used in order to estimate parameter a , again by minimising the open loop error with a trial and error approach. Figure C-3 depicts the measured and estimated pump body temperature T_B .

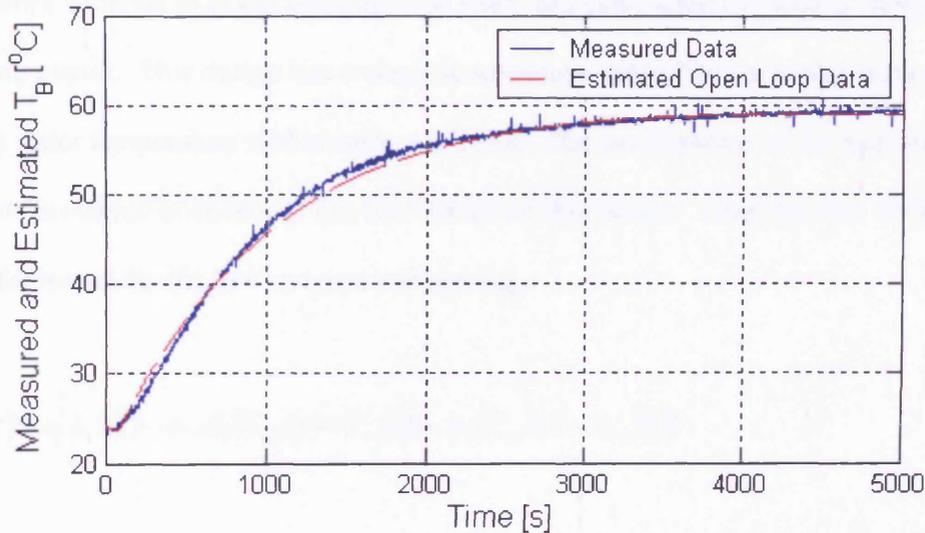


Figure C-3: Plot of measured and estimated T_B

The above procedure has been repeated in order to estimate the nominal model parameters for the coolant outlet temperature model

$$\dot{T}_o(t) = b_1 h_c (T_B(t) - T_o(t)) - b_2 \dot{T}_i(t) \quad (C.9)$$

where b_1 and b_2 are given by $b_1 = \frac{A_c}{(mc_p)_{cw}}$ and $b_2 = \frac{1}{m_{cw}}$.

Parameter b_2 can be easily found by calculating the mass of the coolant contained in the coolant system pipe-work. Substituting for this value then the remaining nominal parameter b_1 can be estimated.

C.3 BEARINGS HEAT TRANSFER MODEL PARAMETER ESTIMATION

The work undertaken in the bearings case study has considered the second design of the cooling circuit. This design has a single temperature-controlled valve that maintains the pump stator temperature within some set limits. The development of the heat transfer models presented in chapter 5 has been based on this design. Consider first the heat transfer model for the low vacuum end bearing

$$\dot{T}_{LV}(t) = a_1 k_1 I(t) - a_2 \dot{m}_c (T_{out}(t) - T_{in}(t)) - a_3 (T_{LV}(t) - T_{am}(t)) \quad (C.10)$$

where a_1 , a_2 and a_3 are given by $a_1 = \frac{1}{(mc_p)_{LV}}$, $a_2 = \frac{c_{pcw}}{(mc_p)_{LV}}$ and $a_3 = \frac{(hA)_{LV}}{(mc_p)_{LV}}$.

Similarly to section C-2, three sets of data have been obtained by eliminating parts of the mathematical equation. However, when the final term of equation (C-10) has been included and a data set has been acquired, it was noticed that the parameter a_1 was induced by the non-linear thermostatic valve. In order to investigate how the parameter changed a sliding mode observer has been used.

From equation (C.10) the following sliding mode observer can be formulated

$$\hat{\dot{T}}_{LV}(t) = \hat{a}_1 k_1 I(t) - a_2 \dot{m}_c (T_{out}(t) - T_{in}(t)) - a_3 (\hat{T}_{LV}(t) - T_{am}(t)) + v_1 \quad (C.11)$$

where $v_1 = K_1(\varepsilon_{LV} / \|\varepsilon_{LV}\| + \delta)$ is the discontinuous signal and K_1 its gain. Moreover, ε_{LV} is defined as the difference between the estimated and measured T_{LV} temperatures.

Since $e_1 = T_{LV} - \hat{T}_{LV}$ and $\dot{e}_1 = \dot{T}_{LV} - \hat{\dot{T}}_{LV}$, substituting it yields

$$\dot{e}_1 = I(a_1 - \hat{a}_1) - a_3 e_1 - v_1 \quad (C.12)$$

During the sliding motion the observer error and its derivative will converge to zero.

Hence,

$$0 = I(a_1 - \hat{a}_1) - v_1 \quad (C.13)$$

It is therefore possible to use the discontinuous signal to estimate the parameter a_1 .

Using this method a plot of the estimated parameter has been drawn as shown in Figure C-4.

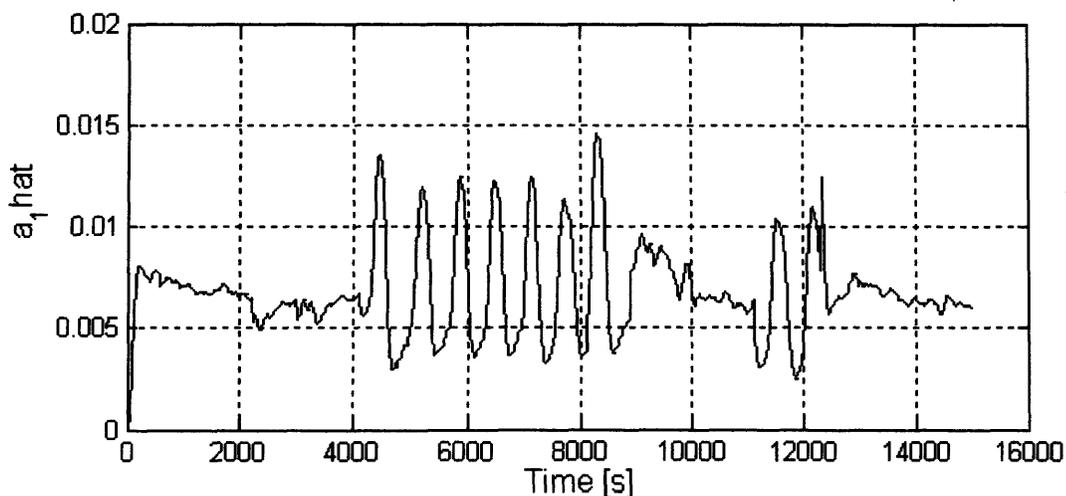


Figure C-4: Plot of the estimated parameter a_1 from the sliding mode observer

The changes noticed in Figure C-4 on the parameter a_1 are occurring when the thermostatic valve switches on and off. Hence, it can be assumed that a relationship between the parameter and the temperature difference $\Delta T(t) = T_{out}(t) - T_{in}(t)$ exists. The graph between the estimated parameter a_1 and the temperature difference is plotted in Figure C-5.

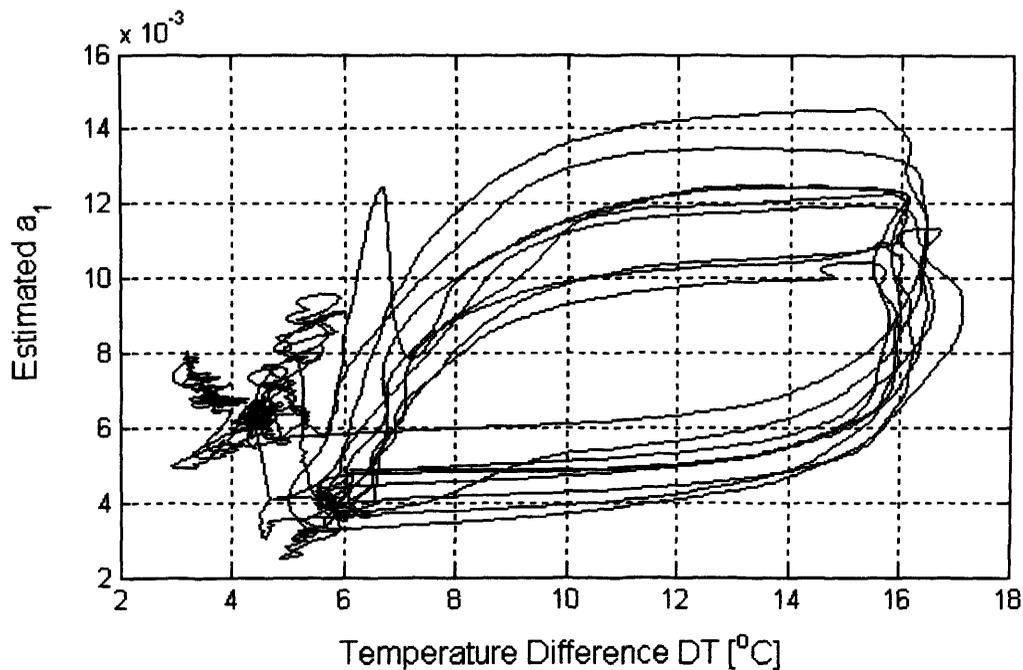


Figure C-5: Relationship between a_1 and ΔT for open/close valve

With the use of MatlabTM, a function can be found which fits in a least squares the measured data. The function coefficients have been first estimated and then the function has been evaluated so that a comparison between the measured and estimated a_1 can take place. The function estimated parameter a_1 and the observed estimated one can then be compared as shown in Figures C-6 and C-7. Finally, substituting this function for a_1 and introducing a second term $x_1(t)$ (since it takes some time for the heat to be generated) yields

$$\begin{aligned}
 \dot{x}_1(t) &= x_1(t) - a_2 m x_2(t) - a_3 (T_{LV}(t) - T_{atm}(t)) \\
 250 \dot{x}_2(t) + x_2(t) &= T_{out}(t) - T_{in}(t) \\
 100 \dot{x}_1(t) + x_1(t) &= k_1 P_i(\Delta T) I(t)
 \end{aligned}
 \tag{C.14}$$

where $P_i(\Delta T)$ denotes the temperature dependent function given by

$$P_i(\Delta T) = \begin{cases} 1 \times 10^{-5} \Delta T^3(t) - 2.3 \times 10^{-4} \Delta T^2(t) \\ + 1.7 \times 10^{-3} \Delta T(t) - 5.5 \times 10^{-5} & \Delta T(t) > 2.4 \\ - 2.5 \times 10^{-5} \Delta T^3(t) + 6.9 \times 10^{-4} \Delta T^2(t) \\ - 5.2 \times 10^{-3} \Delta T(t) + 1.8 \times 10^{-2} & \Delta T(t) < 2.4 \end{cases}
 \tag{C.15}$$

The above threshold temperature value has been calculated after observing the difference between the inlet and outlet coolant temperature, which indicated that the thermostatic valve has been switching on and off. The estimated T_{LV} temperature for the low vacuum end bearing obtained from the nominal model is shown in Figure C-8.

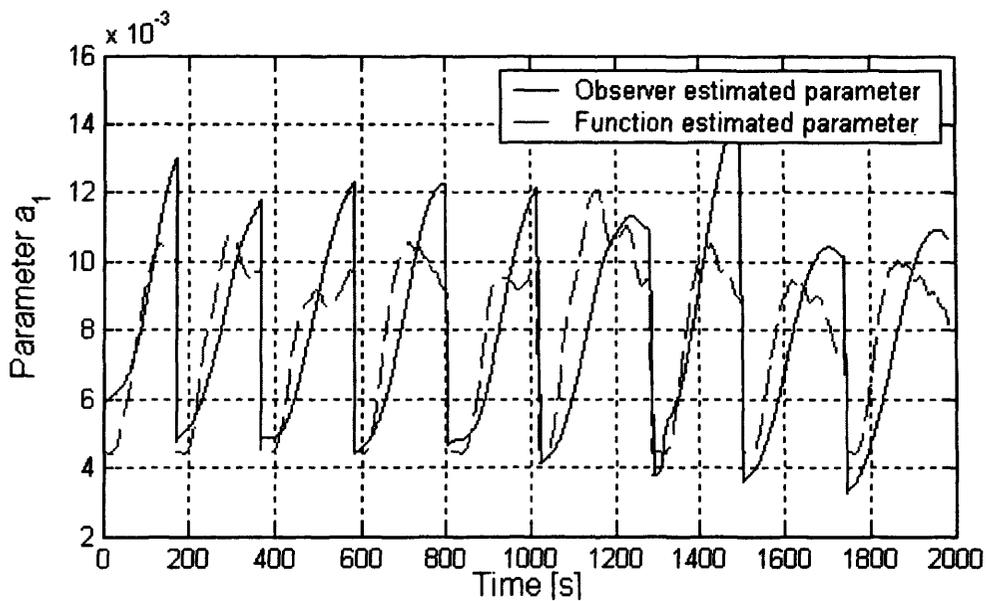


Figure C-6: Comparison between observer and function data for open valve

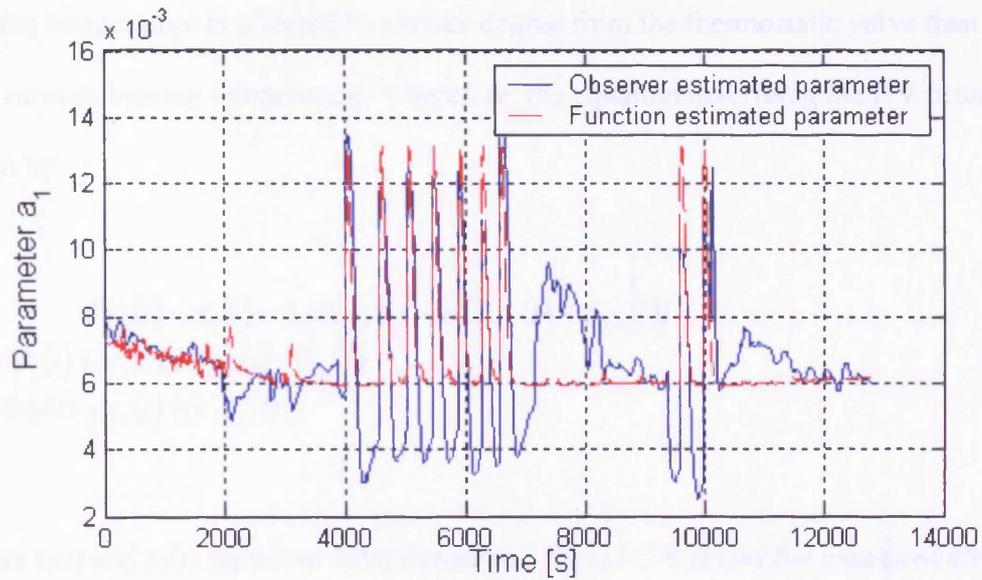


Figure C-7: Comparison between observer and function data for close valve

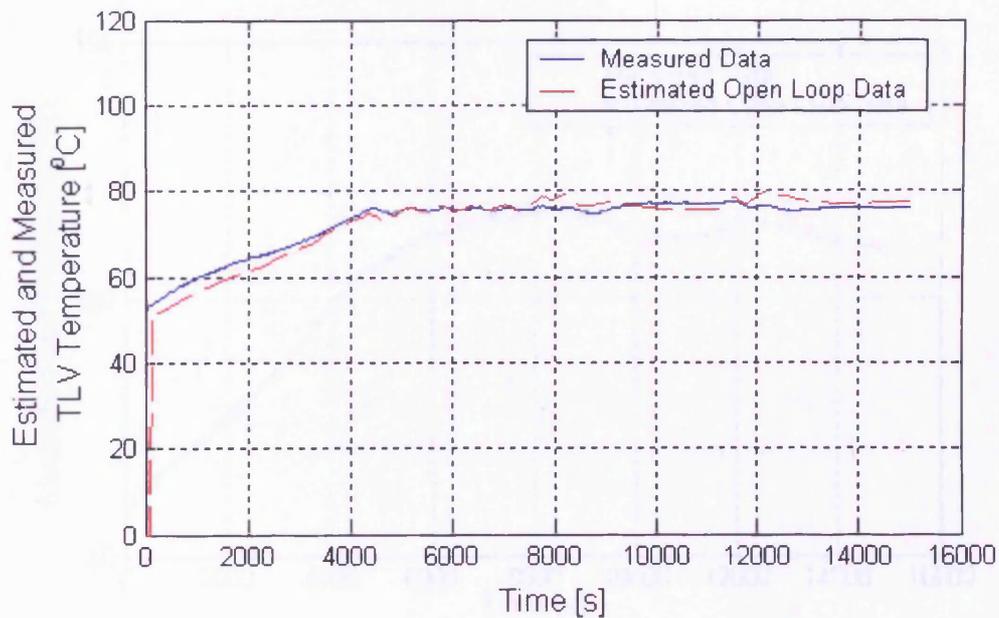


Figure C-8: Measured and estimated T_{LV} for low vacuum bearing temperature model

Having estimated the nominal parameters for the low vacuum end bearing model, a similar superposition method has been followed for the high vacuum end bearing model. Nonetheless, it has been noticed that the rate of change of the high vacuum

bearing temperature is affected to a lesser degree from the thermostatic valve than the low vacuum bearing temperature. Therefore, the equation describing the HV bearing is given by

$$\begin{aligned} \dot{T}_{HV}(t) &= z_1(t) - b_2 \dot{z}_2(t) - b_3 (T_{HV}(t) - T_{atm}(t)) \\ 1500 \dot{z}_2(t) + z_2(t) &= T_{out}(t) - T_{in}(t) \\ 250 \dot{z}_1(t) + z_1(t) &= k_2 b_1 I(t) \end{aligned} \quad (C.16)$$

where $z_1(t)$ and $z_2(t)$ represent filter dynamics. Figure C-9 shows the measured and estimated T_{HV} temperature for the high vacuum bearing nominal model.

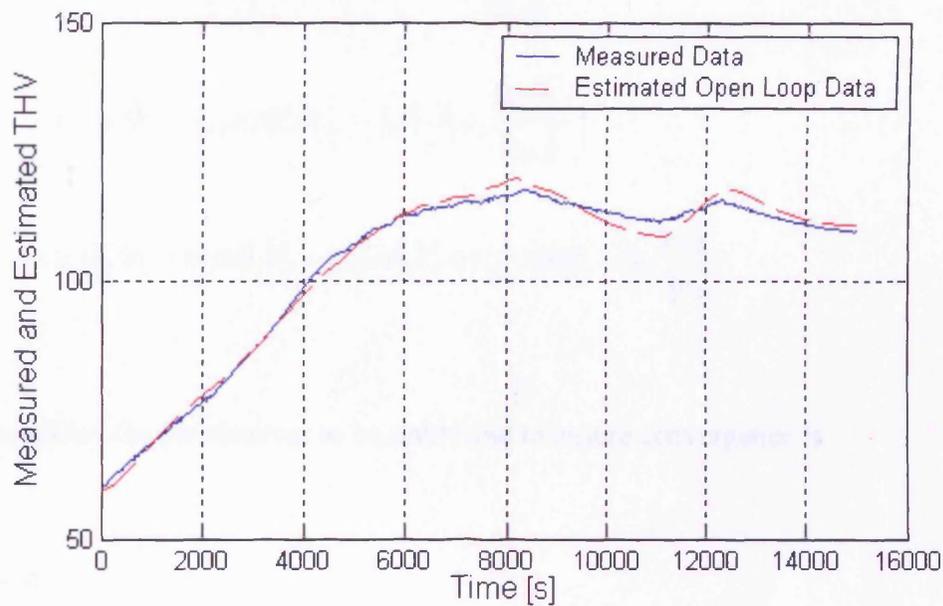


Figure C-9: Measured and estimated T_{HV} for high vacuum bearing temperature model

$$\begin{aligned}
\mathcal{L}(\varepsilon) &= \varepsilon_{B1} \left[(-a_3 \varepsilon_{B1} + a_1 \Delta k l - a_2 \Delta n \tilde{\kappa} (T_o - T_i) - \nu_{B1}) \right] \\
&\quad + \varepsilon_{B2} \left[\left(- (a_4 \tilde{h}_c + a_3) \varepsilon_{B2} + a_1 \Delta k l - a_4 \Delta h_c T_{B2} + a_4 \Delta h_c T_o - \nu_{B2} \right) \right] \\
&\quad + \varepsilon_o \left[\left(- (b_1 \tilde{h}_c + b_2 \tilde{m} \tilde{\kappa}) \varepsilon_o - (b_1 \Delta h_c + b_2 \Delta n \tilde{\kappa}) T_o + b_1 \Delta h_c T_B + b_2 \Delta n \tilde{\kappa} T_i - \nu_o \right) \right] \quad (D.2)
\end{aligned}$$

Substituting for the injection signals $\nu_i = K_i \frac{\varepsilon_i}{\|\varepsilon_i\|}$, where $i =_{B1, B2, o}$ and rearranging will

yield the following

$$\begin{aligned}
\mathcal{L}(\varepsilon) &= -a_3 \varepsilon_{B1}^2 - a_4 \tilde{h}_c \varepsilon_{B2}^2 - a_3 \varepsilon_{B2}^2 - (b_1 \tilde{h}_c + b_2 \tilde{m} \tilde{\kappa}) \varepsilon_o^2 \\
&\quad + \varepsilon_{B1} a_1 \Delta k l - \varepsilon_{B1} a_2 \Delta n \tilde{\kappa} (T_o - T_i) - K_{B1} \frac{\|\varepsilon_{B1}\|^2}{\|\varepsilon_{B1}\|} \\
&\quad + \varepsilon_{B2} a_1 \Delta k l - \varepsilon_{B2} a_4 \Delta h_c (T_{B2} - T_o) - K_{B2} \frac{\|\varepsilon_{B2}\|^2}{\|\varepsilon_{B2}\|} \\
&\quad - \varepsilon_o (b_1 \Delta h_c + b_2 \Delta n \tilde{\kappa}) T_o + \varepsilon_o b_1 \Delta h_c T_B + \varepsilon_o b_2 \Delta n \tilde{\kappa} T_i - K_o \frac{\|\varepsilon_o\|^2}{\|\varepsilon_o\|} \quad (D.3)
\end{aligned}$$

The condition for the observer to be stable and to ensure convergence is

$$\mathcal{L}(\varepsilon) < 0 \quad (D.4)$$

Now $\varepsilon_i^2 = \|\varepsilon_i\|^2$, for $i =_{B1, B2, o}$ hence,

$$\begin{aligned}
\mathcal{L}(\varepsilon) &= -a_3 \varepsilon_{B1}^2 - a_4 \tilde{h}_c \varepsilon_{B2}^2 - a_3 \varepsilon_{B2}^2 - (b_1 \tilde{h}_c + b_2 \tilde{m} \tilde{\kappa}) \varepsilon_o^2 \\
&\quad + \varepsilon_{B1} a_1 \Delta k l - \varepsilon_{B1} a_2 \Delta n \tilde{\kappa} (T_o - T_i) - K_{B1} \|\varepsilon_{B1}\|
\end{aligned}$$

$$\begin{aligned}
& + \varepsilon_{B2} a_1 \Delta k l - \varepsilon_{B1} a_4 \Delta h_c (T_{B2} - T_o) - K_{B2} \|\varepsilon_{B2}\| \\
& - \varepsilon_o (b_1 \Delta h_c + b_2 \Delta n_{\xi}) T_o + \varepsilon_o b_1 \Delta h_c T_B + \varepsilon_o b_2 \Delta n_{\xi} T_i - K_o \|\varepsilon_o\| \\
& \leq -a_3 \varepsilon_{B1}^2 - a_4 \tilde{h}_c \varepsilon_{B2}^2 - a_3 \varepsilon_{B2}^2 - (b_1 \tilde{h}_c + b_2 \tilde{n}_{\xi}) \varepsilon_o^2 \\
& - \|\varepsilon_{B1}\| (K_{B1} - a_1 \Delta k l + \|T_o - T_i\| \Delta n_{\xi} a_2) \\
& - \|\varepsilon_{B2}\| (K_{B2} - a_1 \Delta k l + \|T_{B2} - T_o\| \Delta h_c a_4) \\
& - \|\varepsilon_o\| (K_o + \|T_o - T_B\| \Delta h_c b_1 + \|T_o - T_i\| \Delta n_{\xi} b_2)
\end{aligned} \tag{D.5}$$

Therefore to ensure $\mathcal{V}(\varepsilon) < 0$

$$K_{B1} \geq \|\Delta n_{\xi} a_2 (T_o - T_i) - a_1 \Delta k l\| \tag{D.7}$$

$$K_{B2} \geq \|\Delta h_c a_4 (T_{B2} - T_o) - a_1 \Delta k l\| \tag{D.8}$$

$$K_o \geq \|\Delta h_c b_1 (T_o - T_B) + \Delta n_{\xi} b_2 (T_o - T_i)\| \tag{D.9}$$

APPENDIX E

HEAT TRANSFER MODEL DEVELOPMENTS

E.1 INTRODUCTION

Appendix E describes early heat transfer model developments for the commercial version of the iGX dry vacuum pump. The results presented in this Appendix provide a sound basis for future research.

E.2 THE iGX DRY VACUUM PUMP COOLING CIRCUIT DESIGN

The third and final design of the cooling circuit consists of two temperature-controlled valves, which monitor the temperature of the system. One of the valves is located on the motor housing while the second is positioned on the pump 4th stage stator. Figure E-1 depicts a diagram of this design. Each water valve's state is controlled by a single temperature set-point. Once the monitored temperature goes above the set-point, then the valve is switched on. Clearly, once the temperature decreases beneath the set-point minus a hysteresis value set by the designers, then the valve is switched off. These hysteresis values are provided by BOC Edwards and are set to 5 °C for the motor and 1 °C for the stator. Table F-1 illustrates the four possible coolant flow conditions as given by BOC Edwards.

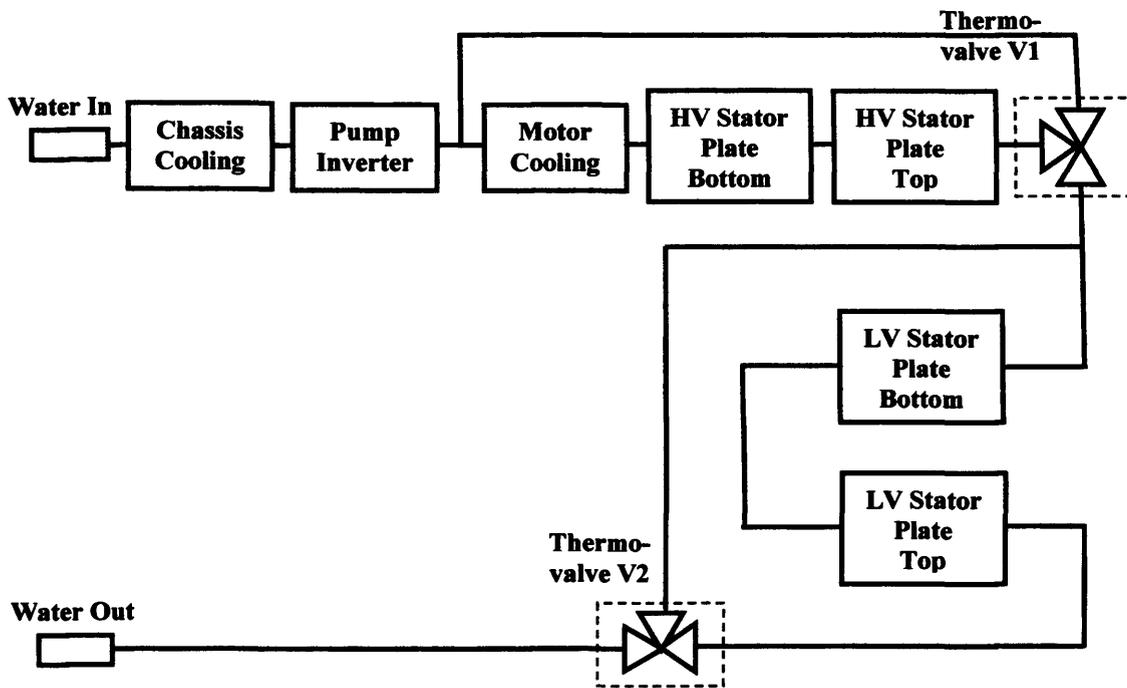


Figure E-1: Dry vacuum pump cooling circuit diagram, design No. 3

Table E-1. Coolant flow conditions

Valve State		Temperature [°C]		Flow Condition
V1	V2	Motor	Stator	
OFF	OFF	<40	<120	No cooling, but there will still be flow to and from valve manifold
ON	OFF	>40	<120	Inverter and motor cooling
OFF	ON	<40	>120	Inverter and stator cooling
ON	ON	>40	>120	Inverter, motor and stator cooling

E.3 PRELIMINARY RESULTS

Based on the information presented in the previous section, a Matlab™/Simulink™ heat transfer model has been developed. The model considers the heat flow through the various stages of the vacuum pump and employs non-linear switches to simulate the temperature-controlled valves. Figure E-2 and Figure E-3 show pump temperatures at different stages of the system. These early model developments provide a sound basis for future research, since a clear relationship between the temperature and the valves can be noticed.

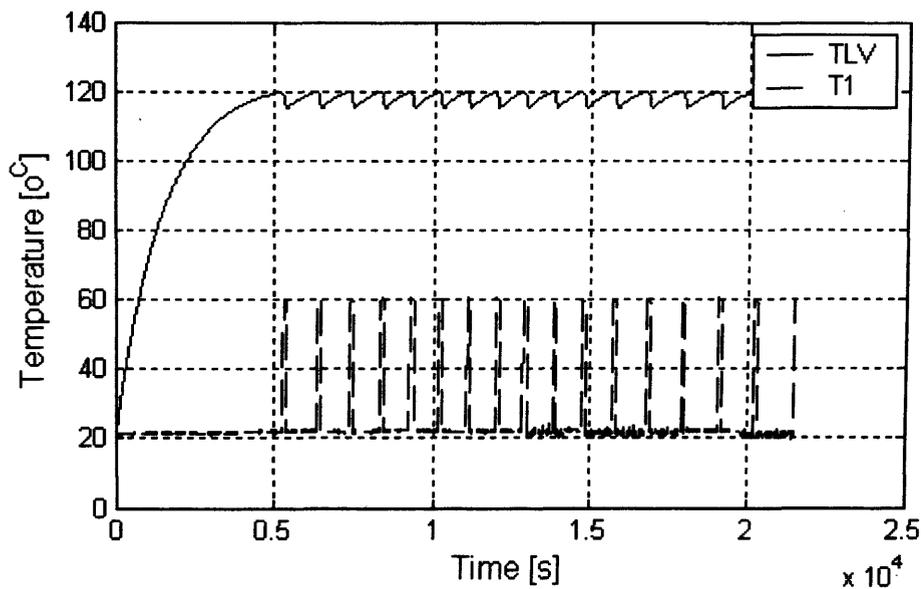


Figure E-2: Vacuum pump and coolant temperature at stage 1

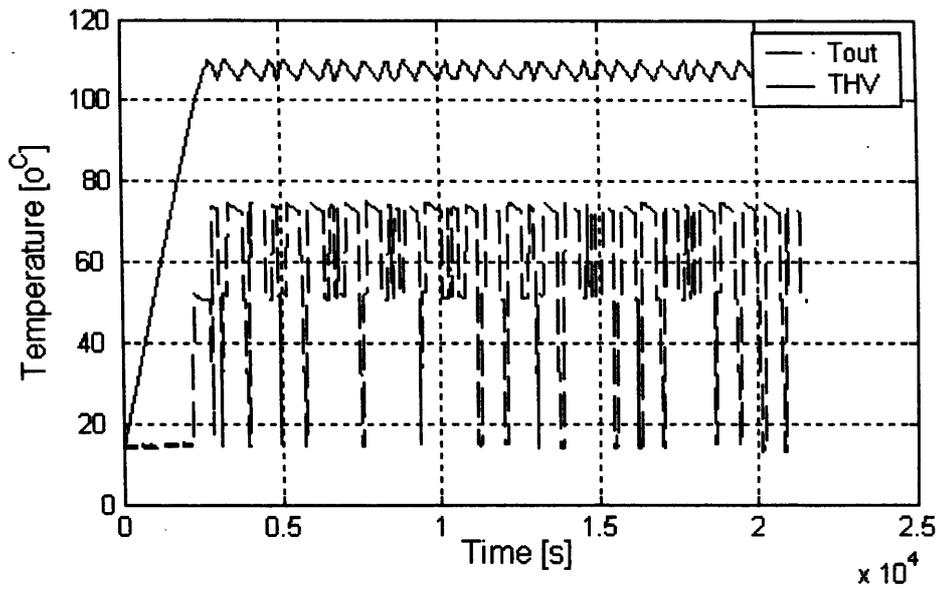


Figure E-3: Vacuum pump and coolant temperature at stage 2