Title
Diagnostics and Control of Gas Turbines Through System Identification

Author
Holcomb, Chad M.

Publication Date
2015

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Diagnostics and Control of Gas Turbines Through System Identification

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Mechanical Engineering)

by

Chad M. Holcomb

Committee in charge:

Professor Robert R. Bitmead, Co-Chair
Professor Raymond A. de Callafon, Co-Chair
Professor Philip E. Gill,
Professor Kenneth Kreutz-Delgado,
Professor Forman A. Williams

2015

The dissertation of Chad M. Holcomb is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____

____

____

____

Co-Chair

____

Co-Chair

University of California, San Diego

2015

DEDICATION

To Janel.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would not have been able to do any of this without the encouragement and unwavering support of my family, friends, and colleagues. I would like to thank my advisers, Prof. Bob Bitmead and Prof. Raymond de Callafon for their patience and support. I cannot count the number of times where I was stymied by a problem or obstacle and they offered wisdom and clarity set the wheels of progress back in motion.

Thank you to all my colleagues at Solar Turbines who supported me during these years and no doubt put in extra hours to cover the work that I should have been doing. I owe my boss and mentor, Roy Dickenson, an enormous thank you for being my champion from the first day I walked into his office and let him know I wanted to go back to school while working. Were it not for his, 'never take no for an answer' attitude, infinite creativity, and patience, I would never have started, let alone finished.

I would like to thank my parents for the strong emphasis on education and encouragement to complete my PhD, even through my nine year diversion into industry.

Finally, my most profound thanks go to my wife, Janel. Your smile and energy have lifted my spirits after many a long day. Your love and support during this time has made all the difference.

Publications: The results included in this thesis have been published in the following articles:

Chapter 3

Chad M. Holcomb, Robert R. Bitmead, Raymond A. de Callafon, "Closed-Loop Nonlinear System Identification Applied To Gas Turbine Analytics", *Proc. of the ASME Turbo Expo 2014. Turbine Technical Conference & Exposition*, Dusseldorf , GE. 2014.

Chad M. Holcomb, Robert R. Bitmead, Raymond A. de Callafon, "Closed-Loop Identification of Hammerstein Systems with Applications to Gas Turbines.", *Proc. of the 19th IFAC World Congress. Capetown*, S.A., pp. 493–498. 2014.

Chapter 4

Chad M. Holcomb, Robert R. Bitmead, "Gas Turbine Analytics using Archival Data Sets." *Accepted to ASME Turbo Expo 2015. Turbine Technical Conference & Exposition*, Montreal, CA. 2015.

Chapter 5

Chad M. Holcomb, Raymond A. de Callafon, "Subspace Identification for Disturbance Rejection Control Design in Gas Turbines." *Accepted to 19th European Control Conference*, Linz, AT. 2015.

VITA

2000	B.S. (Department of Mechanical Engineering) Rensselaer Polytechnic Institute, Troy, NY, USA
2001	B.S. (Department of Electrical Engineering) Rensselaer Polytechnic Institute, Troy, NY, USA
2012	M.S. Engineering Sciences (Mechanical Engineering) University of California, San Diego, USA
2015	Ph.D. Engineering Sciences (Mechanical Engineering) University of California, San Diego, USA

PUBLICATIONS

Chad M. Holcomb, Robert R. Bitmead, Raymond A. de Callafon, "Closed-Loop Nonlinear System Identification Applied To Gas Turbine Analytics", *Proc. of the ASME Turbo Expo 2014. Turbine Technical Conference & Exposition*, Dusseldorf , GE. 2014.

Chad M. Holcomb, Robert R. Bitmead, Raymond A. de Callafon, "Closed-Loop Identification of Hammerstein Systems with Applications to Gas Turbines.", *Proc. of the 19th IFAC World Congress. Capetown*, S.A., pp. 493–498. 2014.

Chad M. Holcomb, Raymond A. de Callafon, "Subspace Identification for Disturbance Rejection Control Design in Gas Turbines." *Accepted to 19th European Control Conference*, Linz, AT. 2015.

Chad M. Holcomb, Robert R. Bitmead, "Gas Turbine Analytics using Archival Data Sets." *Accepted to ASME Turbo Expo 2015. Turbine Technical Conference & Exposition*, Montreal, CA. 2015.

ABSTRACT OF THE DISSERTATION

Diagnostics and Control of Gas Turbines Through System Identification

by

Chad M. Holcomb
Doctor of Philosophy in Engineering Sciences (Mechanical Engineering)

University of California San Diego, 2015

Professor Robert R. Bitmead, Co-Chair

Professor Raymond A. de Callafon, Co-Chair

In this dissertation, system identification methods are applied to characterize the dynamic system behavior of gas turbines for use in diagnostics, condition monitoring, and control design. The application examples are motivated by real-world engineering challenges with industrial gas turbines.

Firstly, a method is proposed to represent the nonlinear interconnection of the fuel control system and gas turbine using a block Hammerstein model. A new iterative gradient descent prediction error minimization method is proposed to identify a low order dynamic model of the gas turbine and a parametrized static nonlinear mapping. The identified nonlinear mapping is demonstrated to capture fuel valve contamination.

Secondly, a modification to existing subspace system identification methods is formulated to permit their application to multiple concatenated, but non-contiguous data sets, which jointly can be sufficiently rich for identification while individual records are not. This is a break from traditional system identification, where one uses a single contiguous synchronous input-output record. In addition, a set of simple sufficient conditions for identifiability are derived. In tandem, these developments increase the intrinsic value of archival data as it enables the construction of a sufficiently informative data set for dynamic modeling. These conditions represent the primary theoretical contribution of this dissertation. The identified models are suitable for use in condition monitoring methods that are based on characterizations of transient behavior.

Finally, subspace system identification is applied for identification of control-relevant MIMO generalized plant models of the closed-loop gas turbine system in feedback control at a series of discrete operating points. An operating point dependent scaling method is proposed to enable a single nominal linear plant model to capture the dynamics of the closed-loop system. When combined, the procedure enables design of an outer loop controller to improve the disturbance rejection performance of the system. The method is demonstrated to significantly improve the transient load rejection of fuel and airflow control of an existing low emissions industrial gas turbine.

1

Introduction

This dissertation motivated by real-world engineering challenges associated with industrial gas turbines. In this field, there are two key areas where system identification of dynamic models will benefit the industry; the development of data analytics for condition monitoring and fault detection and for model based control design. The structure of the discussion is developed using problem statement and theoretical development pairs.

First, we investigate the contamination in fuel control valves via closed-loop identification of nonlinear systems using a Hammerstein model structure to capture the nonlinear interconnection of the fuel system and turbine plant. Second, we seek to extract linear dynamic models for use in analytics based on transient behavior by adapting subspace identification to accommodate the data and information content of archival data. Lastly, a control policy to improve the transient load rejection of a single shaft low emissions turbine is developed through closed-loop subspace system identification of generalized plant model and design of a sub-optimal $\mathcal{H}_2$ control.

1.1 Analytics Using Transient Events for Condition Monitoring

Condition monitoring is the use of data to determine the health of the engine, detection of performance degradation, estimation of remaining component life, or identification of component faults. Condition monitoring methods can range from direct physical inspections to indirect model-based methods for assessing the state of the engine systems and components. For modern machinery, and specifically industrial gas turbines, the existing data records of machines in operation "probably represent one of the greatest unexploited assets of pipeline companies" [1]. Intelligent engine condition monitoring and early fault detection is necessary for modern gas turbines to prevent unsafe operation, early detection of degradation, and identification of equipment faults. These diagnostics allow scheduled maintenance to prevent serious accidents and equipment failure [2]. The cost of unplanned service interruption is usually significantly higher than the cost of performing preventative maintenance [3]. In fact, it is quickly becoming necessary for manufacturers to develop the internal capability to perform remote diagnostics to remain competitive. Diagnostics and early detection of degradation not only enables scheduled, preventative maintenance, it may also be used to predict remaining service life. The prediction of service life for major components such as turbine discs is also generating great interest: on one hand minimizing the likelihood of catastrophic failure, and on the other maximizing time before overhaul or component replacement [4].

Some of the most commonly studied indicators of engine condition and health are the gas path parameters. In the arena of gas turbines, one of the most widespread approaches is the statistical analysis of field measurements to determine health indices, e.g. efficiencies, characteristic flows, and pressure ratios along the gas path, by comparing observed performance with a reference turbine model [5, 6]. These engine health parameters are typically functions of multiple variables such as stage pressures, temperatures, flow areas, and machine specific performance data. When trended over time or compared to a nominal "healthy" engine model, these parameters give owners, operators, and manufacturers, insight into the condition of the engine. In practice, only subsets of the measurements required to estimate gas path condition related parameters are instrumented outside of a dedicated test stand. Additionally, these measurements: may be corrupted by noise, inferred from indirect measurement with inherent uncertainty, susceptible to instrumentation and actuation faults.

Of all condition monitoring methods, the model-based approach is viewed by many as the most promising method for the real-time and trend base condition monitoring of such complex systems as gas turbine engines [7]. Often, transient measurements prove more useful than steady state measurements due to limited instrumentation on units in service. In addition, many engine faults contribute little to performance deviations at steady state conditions, but significantly in transient processes [8]. The notion of transient fault signatures has been identified in aero-turbines by Merrington et. al [9] using prediction error (PE) methods and the least squares (LS) estimator.

For the purpose of condition monitoring, archival data is typically treated as a bulk data set for; statistical analysis [10], machine learning, or neural network type model fitting [11]. The dynamic and transient information in the data is not fully exploited in these methods. While machine learning, neural network, or similar modeling techniques are capable of dealing with the many complex and nonlinear parameter interactions, the resultant models are difficult to interpret and interrogate. The model structure is typically hidden, or in layers, and the parameter interactions are not readily apparent. In addition, the tools to extract useful information from these model types is still in its infancy compared to the well-established tools for examining approximate linear dynamic model, such as bode plots fin the frequency domain and step-responses in the time domain. For example, one cannot easily extract and relate model parameters to system features or behaviors. One example where parameters can be related to system features are poles and zeros of linear dynamic models, which can behaviors both in the time and frequency domain.

Regardless of the model type, data used for identification must be sufficiently informative. In the case where the reference model identifies dynamic behavior, there is a requirement that operating conditions prevailing during the collection of this data should be sufficiently perturbing to reveal the system dynamics with suitable fidelity, otherwise known as persistence of excitation [12].

Using these observations, model-based diagnostics that consider both the dynamic response, information content of the data, and steady state operation should provide more comprehensive platform for condition monitoring. The relevant question is how to obtain these models with the available data, and what types of models are suitable for diagnostics.

1.2 Balancing Performance and Emissions with Control

Industry is continually demanding improved efficiency and reduced pollution from existing processes. A high profile example is the downward pressure on emissions from fossil fuel based energy production. As technologies such as wind and solar become economically competitive, gas turbines must follow suit. In particular, gas turbines are being pressed to comply with lower exhaust emissions of NOx and CO emission over a wider range of operating conditions [13]. In many geographic regions, failure to comply with stated limits may trigger significant economic penalties or prevent the unit from being operated [14]. New regulations may not only pertain to new equipment, but apply to a existing fleet equipment that has often been in operation from several years to over a decade.

One advantage of the gas turbine power plant is the capability to handle large transient changes on output power, termed load rejection. As emissions and efficiency performance demands become more stringent, the competing requirements for remaining compliant with emissions limits, while retaining similar transient load rejection capability pose significant control design challenges. In many cases, equipment is already installed and integrated into a larger process. Additionally, the existing control design may not have a baseline mathematical model suitable for modern control design. To address this challenge we to identify a closed-loop model of the existing turbine and feedback controller in the form of a generalized (or standard) plant model [15]. The generalized plant model is then used to design an outer-loop controller to improve the transient disturbance rejection performance in Chapter 5.

1.3 System Identification Approach for Reference Models

The primary focus of this dissertation is the development of methods to extract simple linear dynamic models of system dynamics suitable for condition monitoring and control design of industrial gas turbines. This is an inherently nonlinear application and we cannot trivialize or ignore the nonlinear characteristic of the machinery. To accommodate the nonlinear characteristics, the interconnection of the fuel system and turbine engine is approximated by a block Hammerstein system for the identification of fuel valve contamination. For control design, we apply an operating point dependent, i.e. parameter varying, scaling to enable linear control design methods to accommodate range of turbine operation.

1.3.1 Linear Dynamic Models

Linear dynamic models have been demonstrated to capture the dominant dynamics of gas turbine behavior and to be suitable for the design of diagnostic algorithms [16]. Identification of black-box linear dynamic models is applied to aircraft engines in testing and qualification [17] and reduced order dynamic models are attractive when the frequency content in the data is sparse and parsimony in the parameters is desirable [18]. In application, lumped capacitance models, characterized by first order transfer function, have been shown to be sufficiently accurate with a standard measurement suite to infer temperatures in critical locations and enable predictive component life calculations [4]. In addition, it is common to approximate a dynamic nonlinear system with a linear dynamic model around a local operating point [19]. The linear models may be obtained via linearization of a known nonlinear model or by experimentation and identification methods [12].

For example, the fuel control subsystem can be modeled using high-fidelity models as shown in Fig. 1.3, or alternatively viewed as a black-box system shown in a simple block diagram as shown in Fig. 1.1. This, or similar, simplified block structure approximations will be used throughout the discussion due to the relative simplicity and utility.

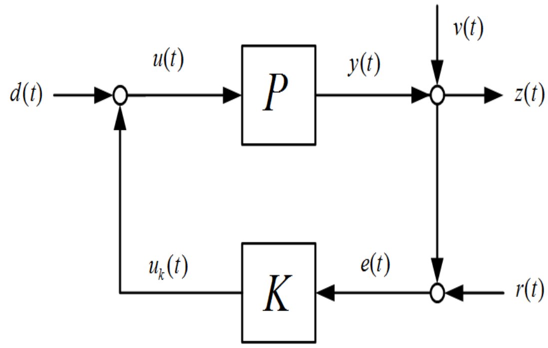

Figure 1.1: Example closed loop black-box system, where P is the plant, K the controller, with reference r, control output u_k, input noise d, plant input u, plant output y, measurement noise v, and measurement z.

1.3.2 Nonlinear Approximations with Block Nonlinear Descriptions

A Hammerstein model is a specialized model structure where a nonlinear system is approximated by a series connection of a static memoryless nonlinear function in series with a linear dynamic plant as shown in Fig. 1.2. The Hammerstein model structure is relevant to many engineering applications where nonlinear behavior is the rule rather than the exception [20]. Most physical devices have nonlinear characteristics outside a limited linear range. Control elements, such as valves and position actuators, are typically linear in a limited operating range and possess some form of saturation, hysteresis, or nonlinear characteristics in the input/output characteristic [21].

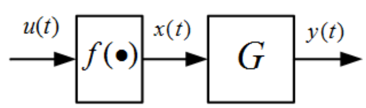

Figure 1.2: Open-loop Hammerstein system with static memoryless nonlinear map $f(\cdot)$ and linear plant model G.

In this structure, a nonlinear characteristic of the system is modeled by a static memoryless block $f(\cdot)$ and the dynamics of the plant is captured by a linear dynamic model G. The internal signal or state $x(t)$ is often unknown or estimated. The Hammerstein plant structure for the turbine plant will be used in the fault detection problem in Chapter 3.

1.4 Archival Data

Physical testing is expensive, resource intensive, and typically does cannot replicate the intended application. Conducting informative experiments to obtain control relevant models is often impractical or costly [22]. Experiments need to be disruptive to provide dynamic information, but operators and customers think this stinks. However, process disruptions do occur, and are part of the archival data record although rarely for a duration that is suitable for modeling. We do have at our disposal a large amount of archival data collected in engine rooms all over the world and transmitted to central data centers and we seek to extract "naturally occurring" appropriate disruptive events from the data archive and use them for identification.

1.5 First-Principle Models

In this thesis, to demonstrate the ideas and approaches, a high-fidelity matlab/simulink model is used in place of an actual gas turbine. This is standard industry practice and permits us introducing specific known fault conditions to test the power of the algorithms.

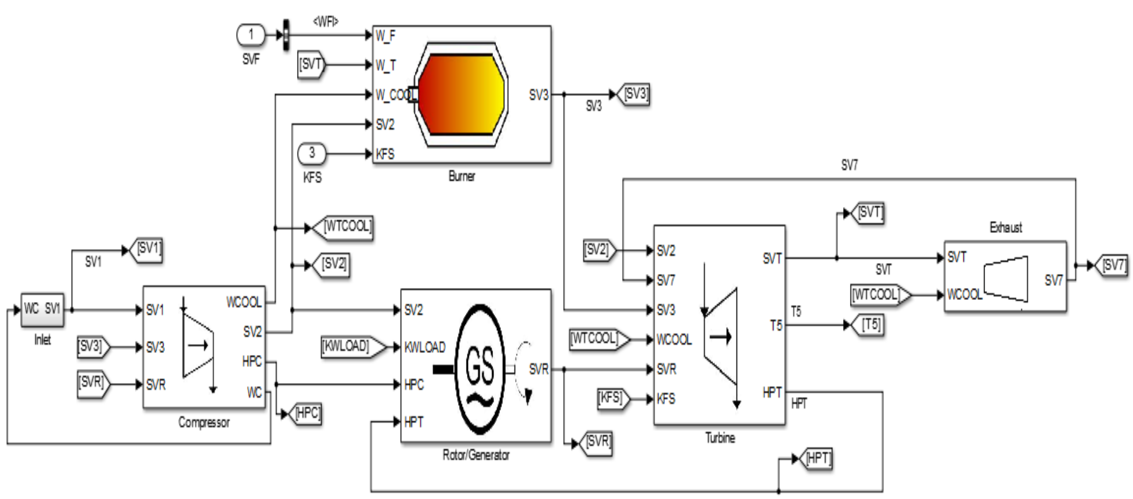

Figure 1.3: High-fidelity engine simulation implemented in Matlab / Simulink used for expereimentation and validation.

First principle models serve as a platform to conduct experiments to generate high-fidelity data for identification as well as for validation and controller design performance evaluation. They enable the simulation of experiments that may not be physically feasible in existing test facilities or may pose a risk to the machinery as well as serve to identify important relationships and measurements useful for condition monitoring. However, these types of models are extremely complex, rigid in structure and are not well suited for control design. The first-principles thermodynamic Matlab/Simulink model of a gas $Taurus^{TM}$ 60 turbine generator and feedback control system, shown in Figure 1.3, is used to generate data for use in identification and control design. The simulation contains a full representation of the Brayton cycle that includes each stage of the cycle, the full control structure for fuel and airflow including actuator and sensor models. The gas turbine model is comprised of the major engine sub-assembly models, i.e. compressor, burner, rotor, power turbine, exhaust, and their interconnections. The simulation depicted in Figure 1.3 is used as an experimentation platform for the closed-loop identification to detect contaminated fuel valves in Chapter 3. In Chapter 5, it serves as an experimentation platform to generate data as for identification of a family of linear plants for control design, validation of the identified plants, and subsequently for evaluation of a new controller for improved load rejection.

1.6 Contributions

The contributions of this dissertation can be summarized as the follows:

Closed-Loop Identification of Hammerstein Systems

- Analytic expressions a quadratic prediction error cost function for single-input single-output closed-loop Hammerstein system,
- A decoupled gradient descent method for output error identification of closed-Loop Hammerstein system,
- A method for performing a high order initialization of output error models with subsequent model order reduction of the linear plant in the Hammerstein systems
- Application to detect fuel valve contamination from dynamic experimentation

Subspace Identification with Segmented Archival Data

- A central contribution of this chapter and indeed this thesis is the extension of subspace system identification to accommodate multiple non-contiguous data sets.
- Within the framework of multiple data sets, the thesis also provides an explicit simple test for the adequacy of multiple data sets to suffice for accurate system identification.

- This specific contribution is fundamental and lies at the heart of adding value to archival data records for the purpose of developing analytics based on transient events.
- The subspace system identification method is shown to be fundamentally independent of a Hankel data structure
- A method for achieving persistency of excitation using concatenated data sets in lieu of a dedicated experiment that satisfies identifiability conditions.
- A new algorithm for subspace system identification that accommodates concatenated data sets
- The contributions in this area are fundamental and theoretical in nature. They are demonstrated with data from the gas turbines.

Closed-Loop Identification and Control Design for of a Low Emissions Gas Turbine System rejection.

- Application of covariance based subspace system identification to obtain a set of closed-loop dynamic models of a nonlinear plant in feedback control at a family of operating points
- A method for operating point dependent parameter varying input and output scaling that enables a single linear plant to approximate the dynamic response of a nonlinear system over a range of operating points
- A methodology for designing a single-input multi-output outer loop controller to improve the disturbance rejection performance of existing decentralized controllers
- Convex optimization of controller output scaling to enable a single controller to improve transient performance over a range of operating conditions
- Application of mixed $\mathcal{H}_2$ control design, with loop shaping to design a control policy for air and fuel control in a low-emissions gas turbine.

1.7 Outline

Chapter 2 gives an overview of the relevant operating principles of gas turbines, specifically the fuel and airflow control. Chapter 5 discusses the identification of nonlinear systems in closed-loop using a Hammerstein model structure and applies the algorithm to detect sulphur contamination of fuel valves. Chapter 4 discusses a novel subspace system identification method that may be applied to archival data sets to obtain linear dynamic models that are suitable for analytics of transient behavior. Chapter 5 discusses the identification and control design of closed-loop systems using a generalized plant model structure. The generalized plant model is used for control design to improve the transient load rejection performance of low-emissions gas turbine. Chapter 6 Chapter 5 concludes this thesis and provides suggestions on future research directions.

2

Gas Turbine Operation

2.1 Introduction

This chapter presents only the basic operation and terminology that is relevant to this discussion. The primary focus is to describe the function of the main components of the turbine and discussion of the gas path and the fuel/airflow control system. For a detailed treatment of gas turbine design principles, see Saravanamuttoo et. al. [23], and for combustion principles, see Williams [24].

2.2 Notation

Common notation relevant to gas turbine operation and applications.

DLE	Dry Low Emissions
GP	Gas Producer
GT	Gas Turbine Engine
N_{GP}	Speed of Gas Producer Shaft [%]
P	Pressure
PT	Power Turbine
T	Temperature [k]
Ω	Power [kW]
$\bar{\Omega}$	Load Fraction
VIGV	Variable Inlet Guide Vanes

Stage Identifiers Gas turbines have a common stage numbering to refer to locations along the gas path. The stage number is typically indicated in a subsctipt, e.g. T_5 is the temperature at stage location 5.

1	Air Inlet
2	Compressor Discharge
3_0	Primary Combustion Zone
3_1	Combustor Outlet Temperature
5	Third Stage Power Turbine Nozzle
7	Exhaust

2.3 Turbine Sections and Function

Figure 2.1 shows a cross section of a Taurus^{TM} 60 industrial gas turbine used for power generation applications. The main components of a typical single-shaft gas turbine engine, labeled in Fig. 2.1 for reference, include the compressor section, combustion chamber (combustor), turbine hot section (turbine), and exhaust collector (exhaust) are l. These components are referred to as the engine core or gas turbine engine (GT). In a single-shaft engine the compressor and turbine are connected by the central shaft and rotate at the same speed. The central shaft extends from the compressor section and transfers power through a gearbox to a mechanical load (typically an electrical generator). The flow path through in the GT is depicted in Fig 2.2.

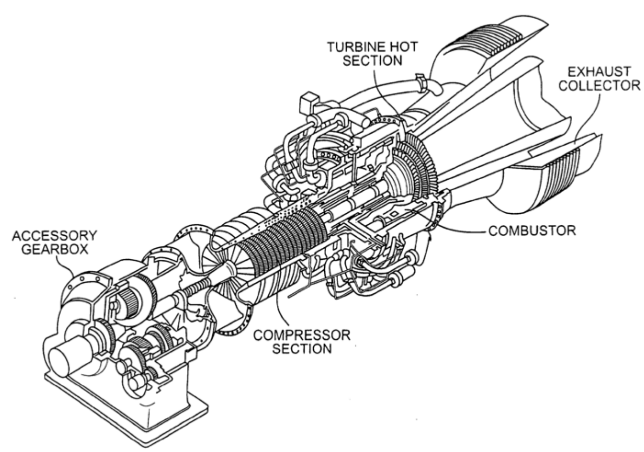

Figure 2.1: Cutaway of a typical Taurus^{TM} 60 engine used for power generation

2.4 Gas Path Overview

Ambient air enters the air inlet, where it is filtered and then enters the compressor inlet and is compressed through the axial flow compressor. The compressor section is equipped with several stages of variable inlet guide vanes (VIGVS) that control angle of the first several stages of stator vanes to regulate the airflow through the compressor. Hot compressed air exits the compressor discharge and enters the combustor, where it is mixed with fuel that enters through an array of fuel injectors. The fuel and air mixture ignites and heats the fuel/air mixture prior to entering the turbine hot section. After exiting the combustor, the fuel/air mixture passes through a series of stator nozzles and rotating vanes that are coupled mechanically to the central shaft. Each set of nozzles and vanes is considered a turbine 'stage'. As pressure and temperature drops across each of turbine stages the thermodynamic and kinetic energy of the gas is transferred into rotational mechanical energy. The temperature of the gas at the entrance to the first stage nozzle ($T_{3,0}$) is the primary mechanical limitation of the turbine and is often in excess of $1360K$ (2000^oF). The operating point is characterized but the turbine

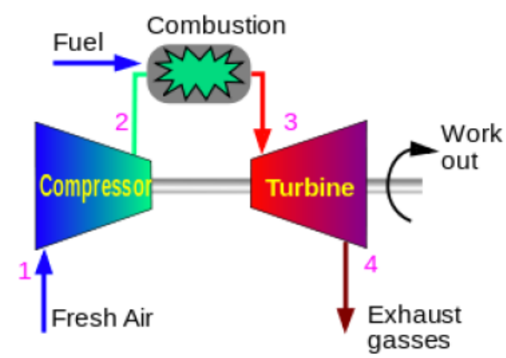

Figure 2.2: Diagram of gas path and operation of a typical single-shaft gas turbine (Wikimedia Commons 2012)

load fraction $\bar{\Omega}_t$ defined as the ratio of the instantaneous net output shaft power defined as Ω_t, over the maximum output power defined as Ω_{MAX}:

$$\bar{\Omega}(t) = \frac{\Omega(t)}{\Omega_{MAX}} \tag{2.1}$$

Figure 2.3 shows a high level block diagram of the closed-loop turbine system used to generate data identification. Here the GT resides in the *turbine plant* block. The *fuel control* and *temperature control* loops are shown as decoupled control loops, which reflects the current control architecture. In practice, the fuel and temperature control are highly coupled through

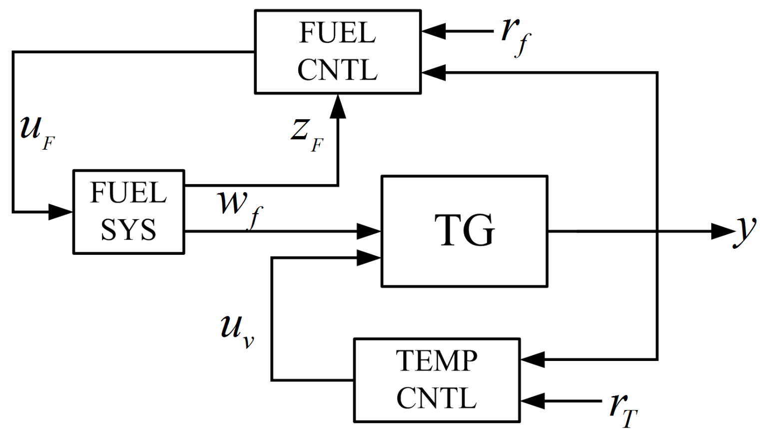

Figure 2.3: Block-diagram representation of a low-emissions closed-loop turbine plant system.

the GT, which we hope to exploit in future control design. The *fuel control* block represents the algorithms that are contained in the controller, while the *fuel system* represents models of the physical piping, actuators, instrumentation and geometry of the physical fuel system. Similarly the *temperature control* block contains the algorithms and actuators that control the VIGVs to regulate the airflow through the turbine.

2.4.1 Airflow Control

In low emissions turbines the temperature of the fuel/air mixture must be controlled in a small window for optimal emissions and combustion stability as shown in Fig 2.4 [25]. The primary zone temperature is the fluid temperature inside the combustion chamber, designed specifically to operate where the NOx and CO emissions intersect ($\sim 1750K$). This temperature is not directly measured, but inferred from a temperature measurement downstream in the turbine section where thermocouples can survive the extreme temperatures. The primary zone temperature is dictated by the fuel air ratio inside the combustion chamber. Since the fuel flow is a factor of the output power demand, the temperature is regulated by the airflow control via the VIGVS. One of the great difficulties with lean premixed systems is maintaining flame stability in the narrow flame temperature range between high NOx production and lean flame extinction. Modern low emissions combustion system designs are capable of mantling stable combustion and low emissions over a wide range of the load range, if the control policies keep the airflow and pilot fuel ratio are properly regulated. The requirement to control these two parameters has added a degree of complexity to the control system that is not required in a conventional combustion gas turbine. During start-up and low-load operation, the pilot flow rate has been optimized to maintain flame stability to handle load transients.

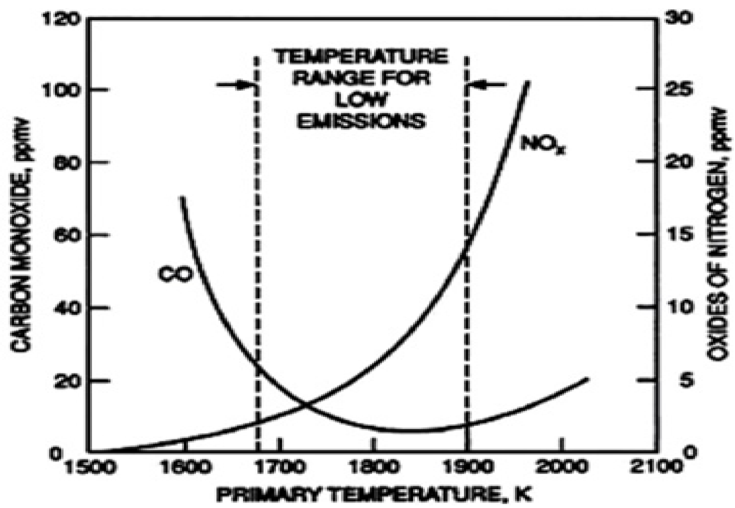

Figure 2.4: Primary-zone temperature versus NOx and CO emissions (Wikimedia Commonsd)

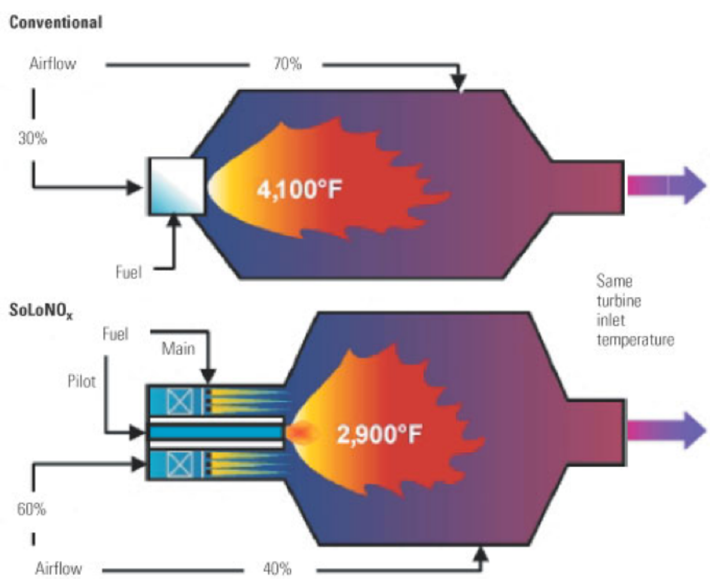

Figure 2.5: Comparison of conventional combustion and SoLoNOx^{TM} combustion System (Courtesy Solar Turbines Inc.)

Below 50% load the combustor airflow is managed in the same way as in a conventional engine. Above 50% rated load, the control transitions into the low emissions mode, modulating either the bleed valve or VIGVs to keep the combustion primary zone temperature within a specified range. Solar's dry low emissions (DLE) gas turbine use the power turbine inlet temperature T_5 as an indirect measurement of the primary zone temperature T_{PZ} to control the VIGVs as to the reference temperature that is a function of turbine load and ambient conditions.

2.4.2 Fuel Control

The fuel flow required is primarily a function of the output power required to drive the mechanical load and the efficiency of the turbine system. The compressor section alone requires between 60-70% of the gross power to compress the air entering the combustor. The main control element is a fuel control valve, or valves. These valves have a known relation between flow area and position. This relation is highly nonlinear over the range of the valve to provide for control to start the turbine up through maximum power. This flow area characteristic is used in the control algorithm to map a flow demand to a valve command. The fuel control

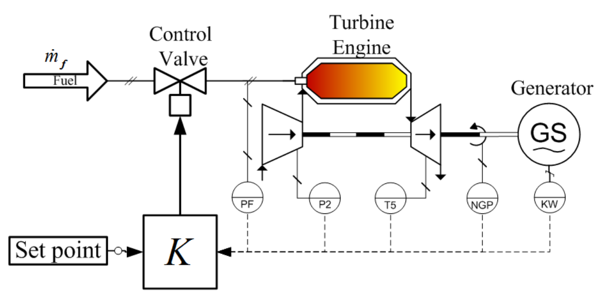

Figure 2.6: Simplified representation of fuel control system.

system in a gas turbine engine, depicted in Fig. 2.6, is perhaps the most critical system, in terms of maintaining system stability and attaining the performance targets of the turbine engine. The fuel control loop is almost universally implemented using one or several digital feedback control loops. The controller(s) use measurements of shaft speed, engine stage temperatures, and pressures to control fuel valve(s), with partially known, and often nonlinear input-output behavior, to regulate fuel flow into the combustion chamber. The fuel control must meet performance objectives subject to operational constraints for equipment protection and safety. In addition, the fuel flow reported by the fuel control system is integral in estimating condition monitoring related parameters. The critical control element in the system is the fuel metering element or valve.

Degradation of the fuel system prevents the gas turbine from meeting power demands and emissions targets. Contamination of the fuel valves, prevents the proper fuel ratio from being delivered to low emissions combustion systems as shown in Fig 2.5.

3

Closed-Loop Identification of Hammerstein Systems

3.1 Introduction

A Hammerstein model is a specialized model structure where a nonlinear system is approximated by a series connection of a static memoryless nonlinear function in series with a linear dynamic plant. The Hammerstein model structure is relevant to many engineering applications where nonlinear behavior is the rule rather than the exception. Most physical devices have nonlinear characteristics outside a limited linear range. Control elements, such as valves and position actuators, are typically linear in a limited operating range and possess some form of saturation, hysteresis, or nonlinear characteristics in the input/output characteristic. In addition, it is common to approximate a dynamic nonlinear system with a linear dynamic model around a local operating point. The linear models may be obtained via linearization of a known nonlinear model or by experimentation and identification methods.

In this chapter, the nonlinear interconnection of the gas turbine and fuel control system is approximated by a Hammerstein system. A gradient descent prediction error minimization method is developed to identify a static memoryless nonlinearity that jointly captures the nonlinear interconnection of the fuel valve in turbine along with a low order dynamic plant model. The algorithm is applied to a first principles simulation of a Taurus 60 Conventional Combustion turbine and fuel system to demonstrate that fuel valve contamination may be detected using the method.

3.2 Background on Identification of Hammerstein Systems

A rich body of research exists on open-loop Hammerstein system identification spanning several decades. The existing methods can be roughly divided into four categories: the iterative method[26, 27], the over-parametrization method [28], the correlation method [29], and and the separable least squares method. The development of methods for identification of closed-loop Hammerstein systems appears later in the literature. These include iterative prediction error minimization (PEM) methods via instrumental variables [30, 21] for SISO systems and recently for MIMO systems in a predictor based subspace system identification framework [31].

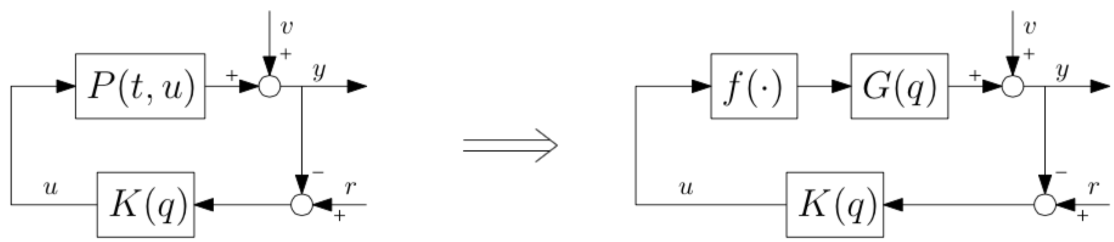

Figure 3.1: [Left]: Block diagram of nonlinear dynamic system $P(t,u)$ with known linear feedback controller $K(q)$.[Right]: Block diagram of Hammerstein approximation of $P(t,u)$ by series connection of the static memoryless nonlinearity $f(\cdot)$ and linear dynamic plant $G(q)$.

3.2.1 Parameterizations of Memoryless Nonlinearities

Typically, the static nonlinearity is approximated by a piecewise input-output map of the nonlinearity that is defined over a grid of expected input values. Piecewise linear (PWL) static maps of $f(\cdot)$ have been developed and applied for system identification of nonlinearities in open-loop and closed-loop Hammerstein frameworks [32]. When the nonlinear mapping is written as a sum of orthogonal basis functions and associated weights defined on the input grid, the identification problem may be written as a pseudo linear regression optimization [30]. Piecewise linear triangle basis functions [33], have been applied to the closed-loop identification problem [21], while other candidate basis functions, such as sigmoid, polynomial, and gaussian radial basis functions are applicable. The choice of basis function is related to the identification goal as each possess certain advantages and disadvantages. The selection of a basis to parametrize the static nonlinearity is discussed within the neural network and image processing literature. Gaussian discriminant classifiers are stated to excel at capturing localized properties as well as continuously differentiable, and are used here [34].

3.2.2 Nonlinear Feedback

In the literature, the feedback is either assumed as unity or a known linear feedback controller. In this case we also consider that the feedback is nonlinear in the sense that the controller is a known linear dynamic system $K(q)$., in series with a known static nonlinear function $f_0(\cdot)$. This may be interpreted as a gain scheduled controller with static poles and zeros with $f_0(\cdot)$ the output gain scheduling function. Alternatively, a-priori knowledge of the control actuator nonlinearity or initial estimate of the nonlinearity may be available. In either case, the system model includes a block that enables injection of a-priori knowledge into the feedback loop in the form a known, differentiable function $f_0(\cdot)$.

3.3 Problem Formulation

The system under consideration is represented by the closed-loop BIBO stable nonlinear system model shown in Fig. 3.2. The known controller $K(q)$, and nominal static mapping $f_0(\cdot)$, with output $w(t)$ is explicitly included in the model. The uncertain static memoryless nonlinearity $\delta(\cdot)$, in the series connection of $f_0(\cdot)$ with linear dynamics $G(q)$ *jointly* captures any deviation from $f(\cdot)$ from $f_0(\cdot)$ and nonlinear behavior of the plant.

The objective is to identify the unknown nonlinear map $\delta(\cdot)$, and linear dynamics $G(q)$ from the system in Fig. 3.2, from uniformly sampled experiment data consisting of input-output signals of $r(t)$, $u(t)$, and $y(t)$ with sampling time t_s over N samples. For identification, $r(t)$ is a known, additive, and persistently exciting sequence.

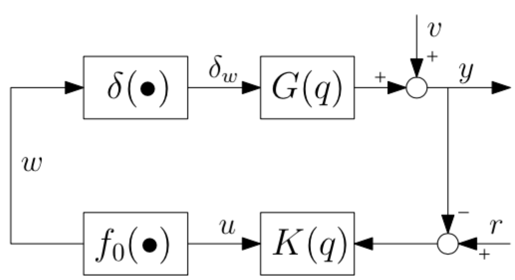

Figure 3.2: Block diagram of closed-loop Hammerstein model with known linear dynamic controller $K(q)$, known nominal $f_0(\cdot)$, unknown linear dynamic plant $G(q)$, and unknown static memoryless function $\delta(\cdot)$ to be estimated.

The signal $w(t)$, is a known function of the controller output $u(t)$ that is given by,

$$w(t) = f_0(u(t)). \tag{3.1}$$

To facilitate parameter estimation, $\delta(\cdot)$ is approximated by a set of orthogonal basis functions that allow $\delta_w(t)$ to be written

$$\delta_w(t) = \sum_{j=1}^{M} \rho_j(w(t))\mu_j. \tag{3.2}$$

The linear dynamic process $G(q)$ is the linear time invariant plant,

$$G(q) = q^{-t_d}\frac{B(q)}{A(q)}, \tag{3.3}$$

of polynomials $A(q)$ and $B(q)$, with the time shift operator q^{-1}, and input time delay t_d. Similarly, the controller may be written

$$K(q) = \frac{D(q)}{C(q)}. \tag{3.4}$$

The following assumptions apply:

Assumption 1. *The reference input $r(t)$ is known, control output $u(t)$, and noisy output $y(t)$ are measured for identification.*

Assumption 2. *The system is closed-loop BIBO stable.*

Assumption 3. *The nominal $f_0(\cdot)$ in (3.1) is a known, monotonic, continuously differentiable function.*

Assumption 4. *The input time delay t_d, to the linear plant is known.*

Assumption 5. *The reference input $r(t)$, is persistently exciting for the identification of $G(q)$.*

Assumption 6. *It is not assumed that $u(t)$ excites the full input range of the static nonlinearity $\delta(\cdot)$ and plant. The the series connection of $\delta(\cdot)$ and $G(q)$ specifically identifies $\delta(\cdot)$ over a limited range.*

3.4 Parametrization

3.4.1 Static Nonlinearity

The nonlinear mapping in (3.2) is written as a linear combination of orthogonal basis functions $\rho_j(w(t))$, with weights μ_j. The weighting vector μ, of function $\delta(\cdot)$, in this basis is an M-vector parameter to be identified. The basis uses the grid

$$m = \begin{bmatrix} m_1 & \cdots & m_M \end{bmatrix}^T, \tag{3.5}$$

to define the center locations of the basis functions and satisfies $[m_1 \le w(t) \le m_M], \forall\, t \in [1..N]$. As mentioned previously, the choice of the M-vector of basis functions $\rho(w(t))$, is closely related to the identification objective and structure of the nonlinearity. A basis that facilitates a good approximation with parsimony in the parameters is desirable. Gaussian radial basis functions are stated to excel at characterizing local properties [34]. This discussion will apply a Gaussian basis of discriminant classifiers with radial basis functions at grid node m_j, defined as

$$\rho_j(w(t)) = \exp\left(\frac{-(w(t) - m_j)^2}{\sigma_j^2}\right)$$

where σ_j controls the spread of the basis functions. The choice of σ_j is a design parameter and is taken to be constant. The center values m_j, are taken on a uniformly spaced grid. The vector of basis functions $\rho(w(t))$, and parameter μ, are written

$$\rho(w(t)) = \begin{bmatrix} \rho_1(w(t)) & \cdots & \rho_M(w(t)) \end{bmatrix}^T, \tag{3.6}$$

$$\mu = \begin{bmatrix} \mu_1 & \cdots & \mu_M \end{bmatrix}^T \tag{3.7}$$

The vector μ, specifies the weights of the basis functions at the grid points in (3.5) and $\delta(w(t))$ can be written in vector notation as

$$\delta_w(t) = \rho^T(w(t))\mu. \tag{3.8}$$

3.4.2 Closed-Loop Model Parametrization

The linear dynamic model $G(q)$ is parametrized by $G(q,\eta) = B(q,\eta)/A(q,\eta)$ where

$$\begin{aligned} A(q,\eta) &= 1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}, \\ B(q,\eta) &= b_0 + b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}. \end{aligned}$$

The stacked plant parameter vector η, is constructed from the coefficients of $A(q,\eta) \to \eta_a$ and $B(q,\eta) \to \eta_b$ as

$$\eta = \begin{bmatrix} a_1 & \cdots & a_{n_a} & b_0 & \cdots & b_{n_b} \end{bmatrix}^T = \begin{bmatrix} \eta_a & \eta_b \end{bmatrix}^T \tag{3.9}$$

The predictor of $y(t)$, using an OE model structure of the linear plant, is given by

$$\hat{y}(t) = \frac{B(q,\eta)}{A(q,\eta)} \rho^T(\hat{w}(t - t_d))\mu. \tag{3.10}$$

Substituting the expression for $w(t)$ from (3.1) allows the set of closed-loop signal relations, from Fig. 3.2 to be written:

$$\mathcal{M}(\eta,\mu) \begin{cases} \hat{y}(t) &= \dfrac{B(q,\eta)}{A(q,\eta)} \rho^T(\hat{w}(t - t_d))\mu \\ \hat{w}(t) &= f_0(\hat{u}(t)) \\ \hat{u}(t) &= K(q)\,(r(t) - \hat{y}(t)) \end{cases} \tag{3.11}$$

The prediction in (3.10) can be written as a linear combination of the parameter dependent *regressor* $\phi^T(t,\theta)$, and a non-minimal parameter θ, as

$$\hat{y}(t,\theta) = \phi^T(t,\theta)\theta. \tag{3.12}$$

The non-minimal parameter θ, and noise free data *regressor* $\phi(t,\theta)$, vectors are written

$$\theta = \begin{bmatrix}\theta_a & \theta_{b\mu}\end{bmatrix}^T = \begin{bmatrix}a_1 & \cdots & a_{n_a} & b_0\mu^T & \cdots & b_{n_b}\mu^T\end{bmatrix}^T \tag{3.13}$$

and

$$\phi(t,\theta) = \begin{bmatrix}\phi_a(t,\theta)\\ \phi_{b\mu}(t,\theta)\end{bmatrix} \tag{3.14}$$

respectively, where

$$\phi_a(t,\theta) = \begin{bmatrix}-\hat{y}(t-1) & \cdots & -\hat{y}(t-n_a)\end{bmatrix}^T, \tag{3.15}$$

$$\phi_{b\mu}(t,\theta) = \begin{bmatrix}\rho^T(\hat{w}(t-t_d))) & \cdots & \rho^T(\hat{w}(t-t_d-n_b)\end{bmatrix}^T. \tag{3.16}$$

3.5 Optimization

3.5.1 Prediction Error and Objective Function

The prediction error $\epsilon(t,\theta)$, is given by

$$\epsilon(t,\theta) = y(t) - \phi^T(t,\theta)\theta. \tag{3.17}$$

In batch estimation with N data points, the measurement vector Y, prediction vector $\hat{Y}(\theta)$, and RMS prediction error vector $E(\theta) = \frac{1}{\sqrt{N}}[Y - \hat{Y}(\theta)]$, are given by:

$$Y = \begin{bmatrix}y(1) & \cdots & y(N)\end{bmatrix}^T \tag{3.18}$$

$$\hat{Y}(\theta) = \begin{bmatrix}\hat{y}(1,\theta) & \cdots & \hat{y}(N,\theta)\end{bmatrix}^T \tag{3.19}$$

$$E(\theta) = \frac{1}{\sqrt{N}}\begin{bmatrix}\epsilon(1,\theta) & \cdots & \epsilon(N,\theta)\end{bmatrix}^T \tag{3.20}$$

We seek $\hat{\theta}$ to minimize the quadratic cost function

$$V_{OE}^N(\theta) = \frac{1}{2N}\sum_{t=1}^{N}\epsilon^2(t,\theta) = \frac{1}{2}E(\theta)^T E(\theta). \tag{3.21}$$

The minimization of (3.21) is an OE minimization where $V_{OE}^N(\theta)$ is inherently non-convex in θ due to the dependency of $\phi(t,\theta)$ in (3.17) on θ. Furthermore, $V_{OE}^N(\theta)$ is over-parameterized in θ due to the multiplicative elements of the coefficients of $B(q,\eta)$ and μ from (3.13). Despite these difficulties, we develop an iterative gradient-based algorithm that is robust to over-parameterization and ensures convergence to at least a local minima.

3.5.2 Dealing With Overparametrization

The non-minimal structure of the parameter θ is common to identification of Hammerstein systems with an arbitrary gain that must be distributed between the nonlinearity and linear plant. The common gain assignment uses the optimal rank one method for decomposing $\theta_{b\mu}$ with freedom to obtain uniqueness in the parameters by normalizing $\|b_i\|_2$ or $\|\mu_j\|$ to one and applying a SVD procedure presented by Bai [28]. While this yields a unique estimate, it might not conform to physical reality. In this analysis, we seek to identify a local range of $\delta(\cdot)$ and dynamic plant model, and will achieve uniqueness of the parameter estimates by constraining a single function value of $\delta(\cdot)$. The excitation is additive and symmetric about a constant speed reference r_0. This corresponds to a steady-state operating point $[r_0, u_0, w_0, y_0]$. Define $\bar{u}(t) = u_0$ and with (3.1), define the grid (3.5) such that it contains a central grid point $m_i = \bar{w}(t) = f_0(\bar{u}(t))$ and set $\delta(\bar{w}(t)) = 0$. This constraint allows us to use a-priori knowledge of $f_0(\cdot)$ within the model structure as it corrects for the nonlinearity as the system moves from the operating point. This allows the linear dynamic plant to capture the steady state gain as well as the dynamics of the physical system. In addition, this enables the comparison of $\delta(\cdot)$ from successive batches of data.

From the definition of the parameter vector in (3.13), we define an auxiliary parameter matrix

$$\Gamma \triangleq \begin{bmatrix} b_0\mu_1 & b_0\mu_2 & \cdots & b_0\mu_M \\ b_1\mu_1 & b_1\mu_2 & \cdots & b_1\mu_M \\ \vdots & \vdots & \ddots & \vdots \\ b_{n_b}\mu_1 & b_{n_b}\mu_2 & \cdots & b_{n_b}\mu_M \end{bmatrix}.$$

The optimal rank one estimate of the parameter vectors $\hat{\eta}_b$ and $\hat{\mu}$ from the parameter estimate $\hat{\theta}_{b\mu}^T$, are obtained by minimizing

$$[\hat{\mu}, \hat{\eta}_b] = \underset{\mu \in \Re^M,\ \eta_b \in \Re^{n_b+1}}{\arg\min} ||\Gamma - \hat{\Gamma}(\theta_{b\mu}^T)||^2, \tag{3.22}$$

where $\hat{\Gamma}(\hat{\theta}_{b\mu}^T)$ is assembled using the *blockvec* operator as

$$\begin{aligned} \hat{\Gamma}(\hat{\theta}_{b\mu}^T) =& \text{blockvec}(\hat{\theta}_{b\mu}) \\ =& \begin{bmatrix} \hat{\theta}_{b\mu}^T(1:M) \\ \hat{\theta}_{b\mu}^T(M+1:2M) \\ \vdots \\ \hat{\theta}_{b\mu}^T(n_bM+1:(nb+1)M) \end{bmatrix}. \end{aligned}$$

To assert the constraint $\delta(\bar{w}(t)) = 0$, first set an initial value of $\hat{b}_0 = 1$ and $\hat{\mu}_0 = \hat{\theta}^T_{b\mu}(1 : M)$, which gives $\hat{\delta}_0(w(t)) = \rho^T(w(t))\hat{\mu}_0$ and $\hat{\delta}_0(\bar{w}) = \rho^T(\bar{w})\hat{\mu}_0$. Define Ω as the mapping of M points in $\begin{bmatrix} w_1 \cdots w_M \end{bmatrix}$ within the range $[m_1, m_M]$ as

$$\Omega = \begin{bmatrix} \rho_1(w_1) & \cdots & \rho_M(w_1) \\ \vdots & \ddots & \vdots \\ \rho_1(w_M) & \cdots & \rho_M(w_M) \end{bmatrix} \in \Re^{(MxM)}. \tag{3.23}$$

A unique inverse of Ω is guaranteed to exist by its construction as an orthogonal set of basis functions. If the correction

$$\hat{\mu} = \hat{\mu}_0 - \Omega^{-1}\hat{\delta}_0(\bar{w}) \tag{3.24}$$

is applied to the weighting parameter vector, then $\delta(w(t)) = \rho^T(w(t))\hat{\mu}$ satisfies $\delta(\bar{w}(t)) = 0$ and $\hat{b}_i$ is updated using $\hat{\mu}$ as

$$\hat{b}_i = \frac{\hat{\mu}^T\hat{\Gamma}(i,:)}{\hat{\mu}^T\hat{\mu}}, \tag{3.25}$$

where the i^{th} row of the $\hat{\Gamma}$ matrix is $\hat{\Gamma}(i,:)$ in this notation. For an estimate $\hat{\theta}^N$, $\hat{\eta}_a$, $\hat{\eta}_b$, and $\hat{\mu}$ are created from (3.9), and (3.7) respectively.

3.5.3 Initialization

The cost function in Eq. (3.21) will likely contain multiple local minima, which stresses the importance of a good initial parameter estimate. Define n_a^* and n_b^* as such that the linear plant model $G(q, \hat{\eta})$ capture the plant dynamics. For initialization, we increase the orders of n_a and n_b to minimize the bias from the nonlinear distortions and unmodeled dynamics on the dynamic plant. Define n_e as the model order increase such that the orders of the initial estimate follow $n_{a,0} = n_a^* + n_e$ and $n_{b,0} = n_b^* + n_e$ while the parameter μ retains dimension M. A least squares initial estimate $\hat{\theta}_0^N$, is calculated generate $\hat{\eta}_0$and $\hat{\mu}_0$.

A subsequent model reduction via singular perturbation procedure applied to $G(q, \hat{\eta}_0)$ to generate a new linear plant estimate $\hat{\eta}_1$ of orders n_a^* and n_b^* [[35]]. Subsequently, $\hat{\theta}_1^N$ is constructed with $\hat{\mu}_0$ and $\hat{\eta}_1$.

The model order reduction reduces the computational expense of the iterative gradient descent algorithm due to the significant reduction in parameters. The matrix inverse in the iterative update, detailed in the next section, dominates the computational cost and is of order $O(n_{\mathcal{M}}^3)$. The model reduction enables the use a high order initial estimate, while reducing the cost of each subsequent iteration by $O((n_eM + n_e)^3)$.

3.5.4 Iterative Gradient Descent Estimation

The iterative gradient descent algorithm requires a parameter estimate θ_s, and gradient of the cost function (3.21) with respect to the parameter estimate, at each iteration step s. The gradient is commonly defined as $\Psi(\theta) = -\nabla_\theta V^N(\theta)$. The non-minimal parametrization contains multiplicative elements of η_b and μ. To relax the parametrization's constraint on the gradient direction we apply a two-step parameter update procedure. The procedure first takes an iteration step in the μ direction, with η held constant, that generates $\hat{\theta}_{\mu,s}$, and a step in the η direction, with μ held constant, that generates $\hat{\theta}_{s+1}$. The gradient descent estimation requires that the cost function gradients $F_\mu(\hat{\theta}_s)$ and $F_\eta(\hat{\theta}_s)$, to be calculated at each iteration step.

Gradient Expressions

The gradient expressions for $F_\mu(\hat{\theta}_s)$ and $F_\eta(\hat{\theta}_s)$ are developed in the following two lemmas. The gradient is simply a function of *theta*, and the evaluation of the gradient at $\hat{\theta}_s$ is found by substitution of $\theta = \hat{\theta}_s$.

Lemma 1. *The gradient of the cost function* (3.21) *with respect to μ, with η held constant, is given by*

$$F_\mu(\theta) = -\left.\nabla_\mu V^N(\theta)\right|_\eta = \Psi_\mu^T(\theta)E(\theta), \tag{3.26}$$

where $\nabla_\mu = \begin{bmatrix} \frac{\partial}{\partial \mu_1} & \cdots & \frac{\partial}{\partial \mu_M} \end{bmatrix}$ *and $\Psi_\mu(\theta)$ is the gradient of the prediction error vector $\nabla_\mu E(\theta)$. For a time sample t,*

$$\psi_\mu(t,\theta) = \begin{bmatrix} \frac{\partial \hat{y}(t)}{\partial \mu_1} & \cdots \frac{\partial \hat{y}(t)}{\partial \mu_k} & \cdots & \frac{\partial \hat{y}(t)}{\partial \mu_M} \end{bmatrix} \tag{3.27}$$

and $\Psi_\mu(\theta) \in \Re(N \times M)$ is written

$$\Psi_\mu(\theta) = \frac{1}{\sqrt{N}} \begin{bmatrix} \psi_\mu^T(1,\theta) & \cdots & \psi_\mu^T(N,\theta) \end{bmatrix}^T. \tag{3.28}$$

The prediction error gradient $\Psi_\mu(\theta)$, is calculated by filtering $\hat{w}(t)$ using the parameters μ, η, and closed-loop signal relations from (3.11).

Proof. With $[r(t), \hat{u}(t), \hat{w}(t), \hat{y}(t)]$ defined in (3.11),

$$\begin{aligned} \left.\frac{\partial \hat{y}(t)}{\partial \mu_k}\right|_\eta =& \frac{1}{A(q,\eta)}\frac{\partial \hat{y}(t)}{\partial \mu_k} + \frac{B(q,\eta)}{A(q,\eta)}\rho^T(\hat{w}(t-t_d)) \\ &+ \frac{B(q,\eta)}{A(q,\eta)}\left(\frac{\partial \rho^T(\hat{w}(t-t_d))}{\partial \hat{w}(t-t_d))}\frac{\partial \hat{w}(t-t_d)}{\partial \mu_k}\right)\mu. \end{aligned} \tag{3.29}$$

With $\hat{w}(t)$ defined in (3.1),

$$\frac{\partial \hat{w}(t)}{\partial \mu_k} = -\left(\frac{\partial f_0(\hat{u}(t))}{\partial \hat{u}(t)}\right)\left(\frac{D(q)}{C(q)}\right)\left(\frac{\partial \hat{y}(t)}{\partial \mu_k}\right), \tag{3.30}$$

and

$$\frac{\partial \rho^T(\hat{w}(t))}{\partial \hat{w}(t)} = -\frac{2}{\sigma^2}\begin{bmatrix} (\hat{w}(t)-m_1)\rho_1(\hat{w}(t)) \\ \vdots \\ (\hat{w}(t)-m_M)\rho_M(\hat{w}(t)) \end{bmatrix}^T, \tag{3.31}$$

for $[m_1 < \hat{w}(t) < m_M]$. With $\hat{w}(t)$, (3.31) can be calculated from $t = 0, 1, \cdots, N$ and $\psi_\mu(t)$ is then be calculated with (3.29) and (3.30) to assemble $\Psi_\mu(\theta)$. □

Lemma 2. *The gradient of the cost function* (3.21) *with respect to* η*, with* μ *held constant, is given by*

$$F_\eta(\theta) = -\left.\nabla_\eta V^N(\theta)\right|_\mu = \Psi_\eta^T(\theta)E(\theta) \tag{3.32}$$

where

$$\nabla_\eta = \begin{bmatrix} \frac{\partial \hat{y}(t)}{\partial a_1} & \cdots \frac{\partial \hat{y}(t)}{\partial a_{n_a}} & \frac{\partial \hat{y}(t)}{\partial b_0} & \cdots & \frac{\partial \hat{y}(t)}{\partial b_M} \end{bmatrix}$$

and $\Psi_\eta(\theta)$ *is the gradient of the prediction error vector* $\nabla_\eta E(\theta)$*. For a time sample* t*,*

$$\psi_\eta(t,\theta) = \nabla_\eta \hat{y}(t) = \begin{bmatrix} \frac{\partial \hat{y}(t)}{\partial a_1} & \cdots \frac{\partial \hat{y}(t)}{\partial a_{n_a}} & \frac{\partial \hat{y}(t)}{\partial b_0} & \cdots & \frac{\partial \hat{y}(t)}{\partial b_M} \end{bmatrix}, \tag{3.33}$$

and $\Psi_\eta(\theta) \in \Re(N \times (n_a + n_b + 1))$ *is written*

$$\Psi_\eta(\theta) = \frac{1}{\sqrt{N}}\begin{bmatrix} \psi_\eta(1) & \cdots & \psi_\eta(N) \end{bmatrix}^T. \tag{3.34}$$

The prediction error gradient $\Psi_\eta(\theta)$*, is calculated by filtering* $\hat{w}(t)$ *using the parameters* μ*,* η*, and closed-loop signal relations from* (3.11).

The proof of Lemma 2 follows the same format as Lemma 1 and is omitted.

Iterative Parameter Update

The parameter update is written in terms of an iteration step in the μ direction that generates $\hat{\theta}_{\mu,s}$, and a step in the η direction that generates $\hat{\theta}_{s+1}$, using the gradient expressions (3.28) and (3.34) and prediction error in (3.20). Using the subscript s to denote the iteration step,

$$\hat{\theta}_{\mu,s} = \hat{\theta}_s - \gamma(\hat{\theta}_s)N_\mu(\hat{\theta}_s)^{-1}F_\mu(\hat{\theta}_s), \tag{3.35}$$

$$\hat{\theta}_{s+1} = \hat{\theta}_{\mu,s} - \gamma(\hat{\theta}_{\mu,s})N_\eta(\hat{\theta}_{\mu,s})^{-1}F_\eta(\hat{\theta}_{\mu,s}), \tag{3.36}$$

where $N_\mu(\hat{\theta}_s)$ and $N_\eta(\hat{\theta}_s)$ are approximations of the Hessians given by,

$$N_\mu(\hat{\theta}_s) = \left[I + \Psi_\mu(\hat{\theta}_s)\gamma(\hat{\theta}_s)\Psi_\mu^T(\hat{\theta}_s)\right],$$
$$N_\eta(\hat{\theta}_{\mu,s}) = \left[I + \Psi_\eta(\hat{\theta}_{\mu,s})\gamma(\hat{\theta}_{\mu,s})\Psi_\eta^T(\hat{\theta}_{\mu,s})\right].$$

In the parameter updates $\gamma(\hat{\theta}_{\mu,s})$ and $\gamma(\hat{\theta}_{\eta,s})$ are the Levenberg-Marquardt step sizes [36]. $F_\mu(\hat{\theta}_s)$ and $N_\mu(\hat{\theta}_s)$ are calculated via (3.26) to update $\hat{\theta}_{\mu,s}$. Then new estimates of $\hat{u}_{\mu,s}(t)$ and $\hat{y}_{\mu,s}(t)$ are generated via (3.11). $F_\eta(\hat{\theta}_s)$ and $N_\eta(\hat{\theta}_s)$ are calculated using $\hat{u}_{\mu,s}(t)$ and $\hat{y}_{\mu,s}(t)$ and (3.32). The parameter update $\hat{\theta}_{s+1}$ is calculated via (3.35). The two-step parameter updates in (3.35) are repeated until the parameter converges using the stopping criterion, for a user defined τ by,

$$\|\hat{\theta}_s - \hat{\theta}_{s-1}\|/\|\hat{\theta}_{s-1}\| > \tau. \tag{3.37}$$

3.6 Application Example

3.6.1 Sulphur Contamination Identification

Faults due to contamination of the fuel metering valves have received particular attention with respect to elemental sulphur deposition. Contamination, referred to here as a fault, that alters the input-output characteristic of the main fuel control element negatively affects the control stability in addition to introducing measurement errors into the fuel flow regulation and limiting functions.

Figure 3.3: Image of sulphur contamination (yellow deposits) of fuel control valve. (Courtesy Solar Turbines Inc.)

An example of a common fault source is local sulphur contamination on the control surface of the fuel valve, shown in Fig. 3.3 [37]. Figure 3.4 shows histograms of the fuel valve control signal $u(t)$, and actual effective flow area $w(t)$, from the high-fidelity simulation and demonstrates how local contamination distorts the input-output relation in closed-loop control. This is due to sulphur deposits creating areas of nearly constant flow area, over a local region of input command. This effectively creates a dead band in the actuator's flow characteristic. In this section, we seek specifically to identify sites of local contamination. These typically smooth deposits are thin at the edge of the deposit and increase in height towards the center.

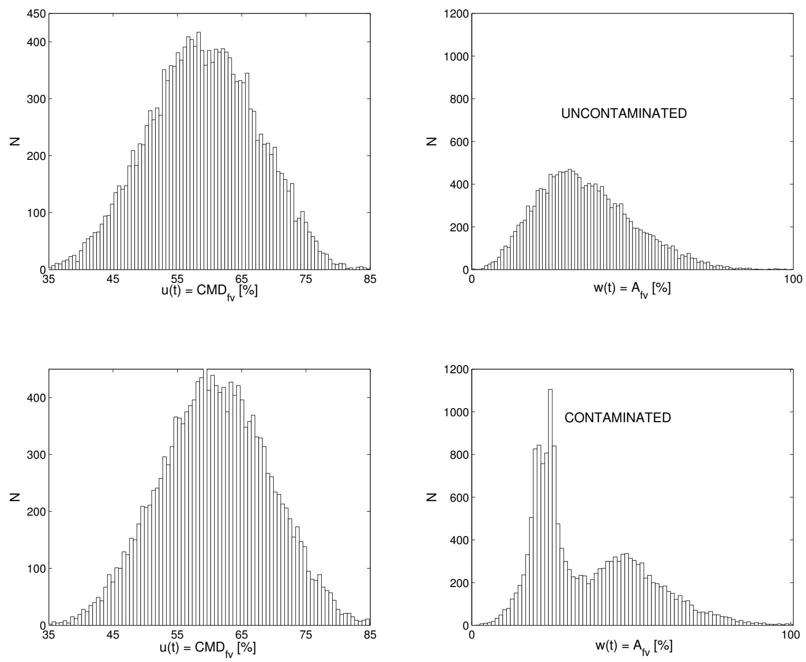

Figure 3.4: Histogram of $u(t)$ & $w(t)$. [Top] uncontaminated case [Bottom] contaminated case.

We apply the Hammerstein system identification method to a high-fidelity nonlinear simulation model of a $Taurus^{TM}$ 60 gas turbine driving an electric generator in an "island" application. The nominal reference set point, $r_0(t)$, for the fuel control is 100% of rated frequency or 60Hz. An excitation signal follows a uniform distribution is $r_e(t) = i.i.d.[-4, 4]$ is added to the reference so $r(t) = r_0(t)+r_e(t)$. The simulation is conducted such that the average output power of the generator is equivalent to 70% of the rated load and sampled at 10Hz. The experiment is conducted with a uncontaminated valve and with a contaminated valve. Each experiment used $N = 8184$ data points for identification and $N = 8184$ for validation. The contamination, Δ_f, represents approximately a 10% percent flow area reduction in a localized area of the valve. The nominal fuel valve, $f_0(\cdot)$ is a fourth order polynomial, and the disturbance to the nonlinearity is an additive gaussian function shown in Eqn. (3.38) with magnitude $\kappa_i = [0.02, 0.03]$, center locations $c_i = [55, 60]$ in percent of input span and $\sigma_i = [20, 10]$.

$$Exp : \begin{cases} f_0(u(t)) & = & \sum_{p=1}^{4} \alpha_p u(t)^p \\ \Delta_f & = & \kappa_i exp\left(\frac{-(u(t)-c_i)^2}{\sigma_i}\right) \\ f(\cdot) & = & f_0 - \Delta_f \end{cases} \tag{3.38}$$

3.6.2 Constraining the Decomposition of the Parameter Estimate

Figure 3.4 shows histograms of simulation data where the fuel valve command $u(t)$, and effective flow area $w(t)$, are shown for the case with an uncontaminated (top) and contaminated (bottom) cases. We seek to identify a local range of $\delta(\cdot)$ and the distribution of $u(t)$ from Fig. 3.4 indicates that most informative range of the data coincides with the mean value of $u(t)$, denoted by $\bar{u}$.

3.6.3 Prediction Error

Figure 3.5 demonstrates the accuracy of the Hammerstein model fits to the raw simulation data (black). The output $y(t)$, corresponds to the difference of shaft-speed from set-point in % of rated speed. The high-order plant, $[n_a = 7, n_b = 6]$, estimate $\hat{y}(t, \hat{\theta}_0^N)$ is shown in (blue) and low-order plant, $[n_a = 3, n_b = 2]$, estimate $\hat{y}(t, \hat{\theta}_1^N)$ is shown in (green). Both the high-order and low-order estimates perform with negligible difference, demonstrating that low-order Hammerstein models are a viable model structure.

3.6.4 Frequency Responses

An estimate of the linear plant dynamics $\hat{G}_N(\omega)$ for a data set containing N uniformly sampled measurements of $u(t)$ and $y(t)$ may be written as the ratio of the estimated cross-spectrum of $\hat{S}_{yu}^N(\omega)$ to the spectrum of the input $\hat{S}_{uu}^N(\omega)$ as the frequency response function (FRF) given by,

$$\hat{G}^N(\omega) = \frac{\hat{S}_{yu}^N(\omega)}{\hat{S}_{uu}^N(\omega)} \tag{3.39}$$

Figure 3.6 demonstrates the comparative fit of measured frequency response function (FRF) from the first principles model to the frequency response of third order linear plant model, identified using the closed-loop Hammerstein model framework presented here. The FRF of

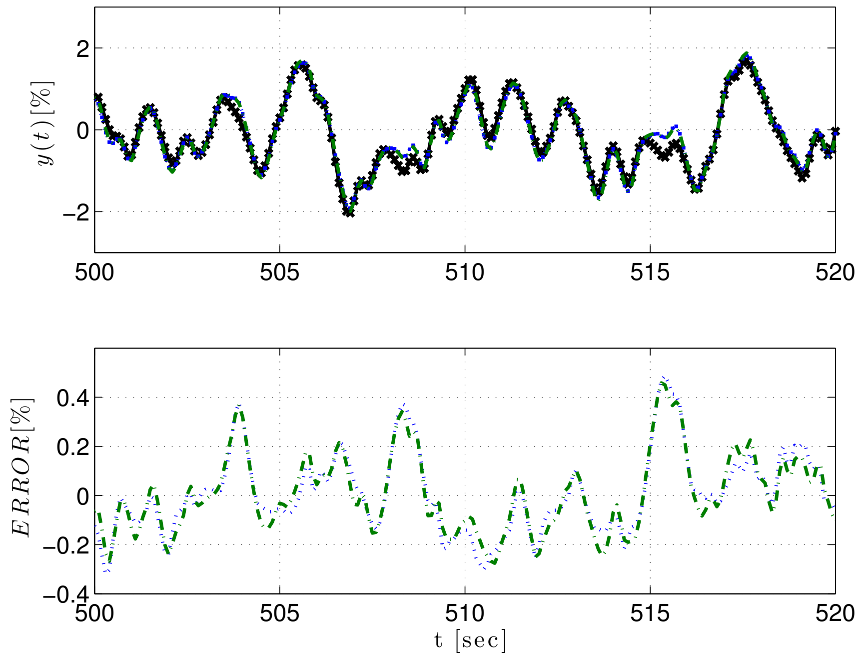

Figure 3.5: [Top] Actual and prediced values and [Bottom] prediction error. Legend: Actual speed $y(t)$ (black), initial estimate $\hat{y}(t, \hat{\theta}_0^N)$ (blue) and reduced order estimate $\hat{y}(t, \hat{\theta}_1^N)$ (green).

the transfer function from $w(t)$ to $y(t)$ from the original data is shown in black. The frequency response of the linear dynamics from $G(\hat{\theta}_0^N)$ is shown in blue and $G(\hat{\theta}_1^N)$ in green. Note that the actual flow area is only available in a simulation environment, thus demonstrating that the linear approximation of both the high and low order dynamics resemble the frequency response estimate.

The coherence function is an excellent tool to evaluate if an linear model is appropriate as it estimates the extent to which two signals are linearly related across the input spectrum, and for signals $w(t)$ and $y(t)$, is written

$$\gamma(\omega)^2 = \frac{|\hat{S}_{yw}^N(\omega)|^2}{\hat{S}_{ww}^N(\omega)\hat{S}_{yy}^N(\omega)}. \tag{3.40}$$

Figure 3.7 shows the coherence of the FRF from $w(t)$ to $y(t)$ from the original simulation (black) and from $\hat{\delta}_w(t)$ to $\hat{y}(t)$ in the low-order Hammerstein model fit (blue). The effect of the nonlinear distortions at low frequencies has been all but removed in a single iteration, demonstrating the successful application of a Hammerstein model to a gas turbine with a high energy excitation.

In the estimation there are several design parameters available to the user, those related to identification of the linear dynamics $[n_a^*, n_b^*, n_e]$, and those related to identification

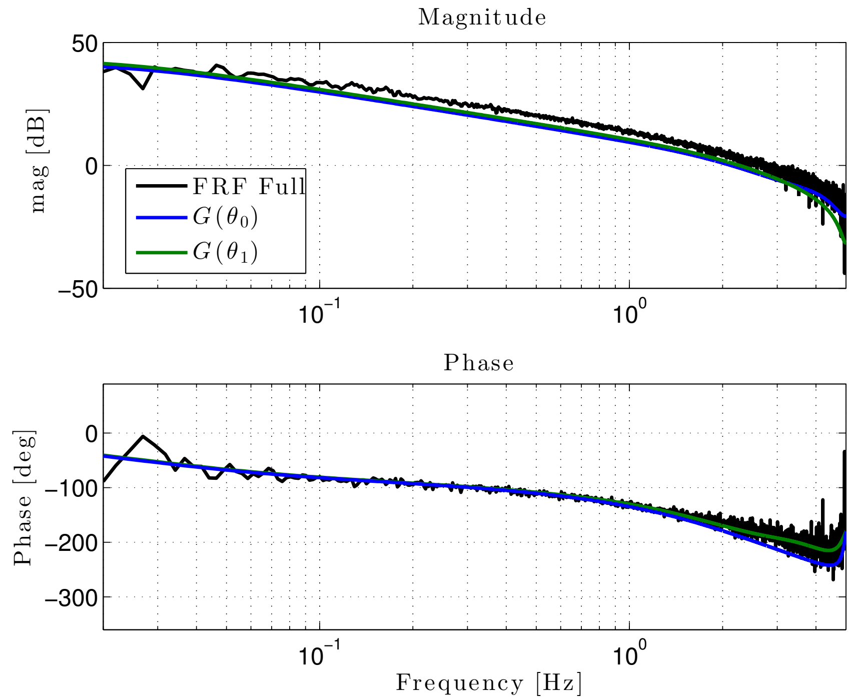

Figure 3.6: Frequency Response: FRF estimates from raw flow area $w(t)$ to shaft speed error $y(t)$ (black), bode plot plant $G(q, \hat{\theta}_0)$ from high order estimate (blue) and reduced order model $G(q, \hat{\theta}_1)$ (green).

of the nonlinearity, $[m, \sigma]$. Table 3.1 compares the performance for variation in the linear plant parametrization.

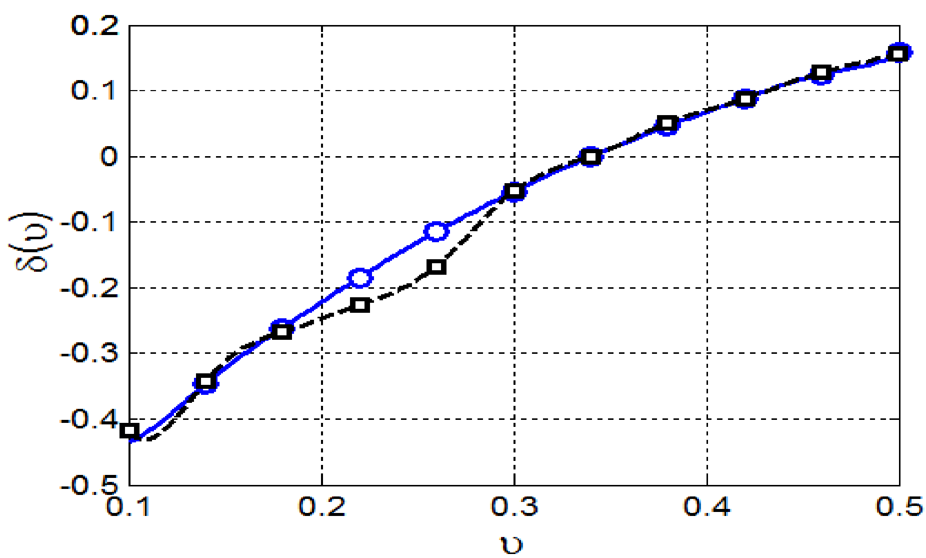

Figure 3.8: Identified $\delta(\upsilon)$: (-◯-) uncontaminated (-□-) contaminated.

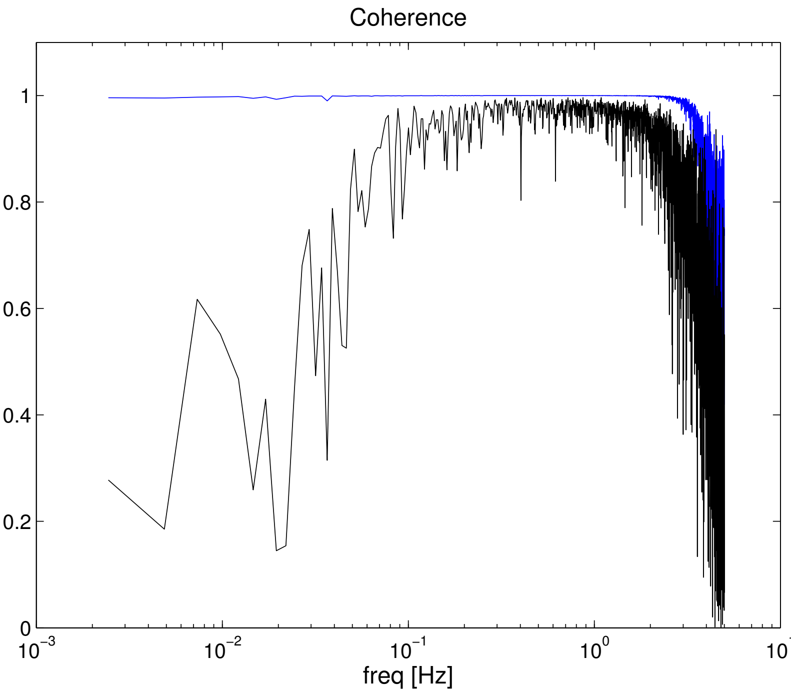

Figure 3.7: Coherence: $\gamma(\omega)^2$ of the raw FRF $w(t) \to y(t)$ from simulation data (black) & reduced order model $\hat{G}(\hat{\theta}_1^N)\ \hat{\delta}_w(t) \to \hat{y}(t)$ (blue).

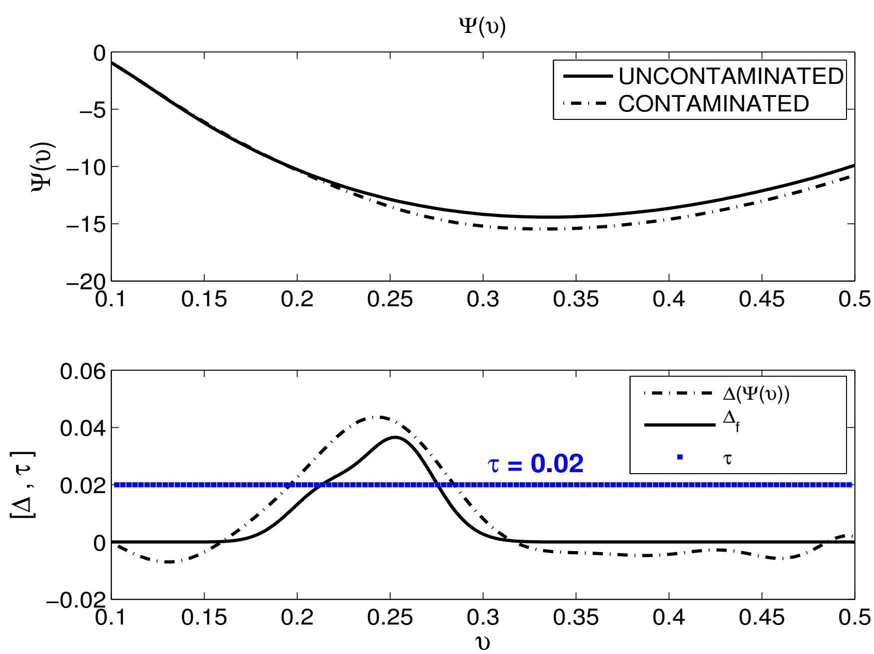

Figure 3.9: [Top] $\Psi(\upsilon)$:- integral of $\delta(\upsilon)$ [Bottom] Actual contamination function Δ_f and identifed $|\Delta\Psi(\upsilon)|$ & with a propose fault threshold $\tau = 0.02$,

Table 3.1: Experimental results illustrating the benefit of the high order initialization on $V^N(\theta)$ and the number of iterations required for convergence.

				$\sigma = 0.35$			
	N	s	n_a^*	n_b^*	n_e	M	$V^N(\theta)$
$\hat{\theta}_0^N$	8192	0	3	2	5	9	0.302
$\hat{\theta}_s^N$	8192	5	3	2	0	9	0.214
$\hat{\theta}_0^N$	8192	0	3	2	3	9	0.423
$\hat{\theta}_s^N$	8192	10	3	2	0	9	0.314
$\hat{\theta}_0^N$	8192	0	3	2	1	9	0.433
$\hat{\theta}_s^N$	8192	10	3	2	0	9	0.343

Figure 3.8 demonstrates the identification $\delta(m, \mu)$ with 8184 data points for identification and 8184 for validation for each experiment. The function is plotted against υ, corresponding to an R-vector of values spanning a range of the effective flow area, $w(t)$. Through visual inspection, the identification captures the nonlinear characteristics of the turbine plant as well as the local distortion in the contaminated area of the valve.

To extract information specific to the local contamination of the fuel valve, a fault metric based on the integral error is proposed. The fault detection for contamination relies on identifying how any deviation of $f(\cdot)$ from f_0 is detectable during the identification of $\delta(\cdot)$. The metrics shown are written using trapezoidal approximations of the local integral, $\Psi_p(\upsilon)$ for experiments $p = 1, 2$ for a normalized grid spacing. By examining the difference $\Delta\Psi = \Psi_2(\upsilon) - \Psi_1(\upsilon)$ a metric for contamination can be defined as $|\Delta\Psi| \geq \tau$ where τ is a user defined threshold indicating the maximum allowable deviation from a healthy fuel control valve.

The use of Gaussian basis functions allow the nonlinearity and its properties to be evaluated over a finer grid spacing than the system-identification, i.e $R > M$. Fig. 3.9 demonstrates that $\Delta\Psi$ recovers the local distortion associated with contamination and matches the magnitude and shape characteristic. For detection purposes, successive batch estimates may be evaluated and trended over time to maximize the reliability of the metric $max(\Delta(\psi(\upsilon))$ and provide indications that local distortions of the nonlinearity are accumulating. This allows the scheduling of maintenance and early detection that an actuator fault is imminent rather than waiting for a system shutdown or instability to interrupt operation.

3.7 Summary

A method for closed-loop identification of the nonlinear characteristics and dynamic response of gas turbine engines has been developed. The method uses a Hammerstein model framework for the identification of a nonlinear function that jointly captures the nonlinear characteristics of the series connection of a partially known actuator and nonlinear turbine plant operating in closed-loop. The direct approach is extended to the closed-loop identification of Hammerstein systems using high order linear plant models during the initialization step using a non-minimal parametrization. A model reduction step produces a low order linear dynamic plant model. The identification includes a new, data dependent gain assignment that facilitates an informed decomposition to a minimal linear plant realization and Gaussian basis function parametrization of the nonlinearity. The use of a Gaussian basis allows an arbitrarily fine grid for examining the identified nonlinearity. This provides a platform for obtaining low-order linear dynamic plant models as well as expose local distortions of the nonlinearity that correspond to actuator faults. The fault detection metric is applied to the estimation of a nonlinear map between a known nominal valve model and a Hammerstein model representation of a gas turbine engine. An example of fuel valve contamination detection is presented. Although the application example utilizes fuel valve contamination to demonstrate a candidate problem target, the algorithm can readily be applied to a wide class of physical systems, including other applications specific to gas turbine control.

3.8 Acknowledgements

This chapter, in part, is a reprint of “Closed-loop Nonlinear System Identification Applied To Gas Turbine Analytics” as it appears in *Proc. of the ASME Turbo Expo 2014. Turbine Technical Conference & Exposition*, Dusseldorf , GE. 2014. Chad M. Holcomb, Robert R. Bitmead, Raymond A. de Callafon. This chapter, in part, is a reprint of “Closed-Loop Identification of Hammerstein Systems with Applications to Gas Turbines” as it appears in that *Proc. of the 19th IFAC World Congress. Capetown*, S.A., pp. 493–498. 2014. Chad M. Holcomb, Robert R. Bitmead, Raymond A. de Callafon. The dissertation author was the primary investigator and author of this material.

4

Subspace Identification Using Archival Data

4.1 Introduction

System identification has been formulated on the basis of a single contiguous data synchronous data record. Many operational data records, including most archival operational data, are insufficiently informative to permit accurate system identification. However, multiple records taken from distinct and separate time periods might contain sufficient system information to permit the fitting of an accurate model.The methodology of system identification needs to be modified to profit from the use of these multiple data sets in fitting models. A central contribution of this chapter and indeed this thesis is the extension of subspace system identification, building on the methods of [38], to accommodate multiple non-contiguous data sets. Within the framework of multiple data sets, the thesis also provides an explicit simple test for the adequacy of multiple data sets to suffice for accurate system identification. This specific contribution is fundamental and lies at the heart of adding value to archival data records for analytics. Specifically, we focus on diagnostics that are related to transient events because many faults are not observable in steady state conditions are during in transient events [8].

Archival data presents a host of challenges to the formulation of system identification as it is not very informative by nature. It contains long periods of steady state operation with low excitation. It also has periods where equipment is shut down or data is missing [39]. Transient events and behavior are present, but not over a period that where traditional system identification methods that rely on a synchronous and contiguous data record of input and output signals to provide accurate estimates. One could imagine searching through the data record for informative transient events, cut out the segments of informative data, and then stitch several of the segments together to construct sufficiently informative experiment. Zhang and Bitmead approached channel estimation in wireless networks with the foundation for the procedure [38].

We propose an algorithm using subspace system identification to enable application to archival data. Subspace system identification is chosen because it readily handles MIMO modeling and frankly, we do not know how to handle the problem in any other way. First, we present the subspace system identification algorithm under consideration for a single data segment. We then re-examine the subspace identification formulation and observe the that it is not fundamentally connected to Hankel matrices or a contiguous data set but rather a set of data 'windows' that have the correct rank and column space [40]. Then, we formulate subspace system identification to enables identification with concatenated segments of non-contiguous data. Finally the algorithm is applied to an existing set of archival data recorded on a gas turbine for demonstration.

4.2 Subspace Identification - Single Segment

Subspace system identification (SSI), seeks to fit a linear dynamic model to measurements of input and output data using a state-space model. It readily handles MIMO models [12, 41]. In this discussion, we consider the nth-order, n_u-input n_y-output linear time invariant system model

$$\mathcal{S} := \begin{cases} x_{t+1} &= Ax_t + Bu_t \\ y_t &= Cx_y + Du_t + v_t, \end{cases} \tag{4.1}$$

where the input $u_t \in \mathbb{R}^{n_u}$, output $y_t \in \mathbb{R}^{n_y}$, and state $x_t \in \mathbb{R}^n$. The output measurements are assumed to be corrupted by noise $v_t \in \mathbb{R}^{n_y}$. The identification goal is to generate a state-space realization A, B, C, D,and the initial condition x_{s_γ}, given a set of input-output data.

4.2.1 Notation

We use the term *segment* to denote a record of input-output data taken over a contiguous time interval. For a single data segment γ, we have a starting time s_γ and a record length N_γ. Let the notation $\langle\gamma\rangle$ define the time interval associated with the segment γ segment such that

$$\langle\gamma\rangle := \{t | t \in [s_\gamma, s_\gamma + N_\gamma - 1]\}. \tag{4.2}$$

Define block-Hankel data matrices of ℓ block rows and $j = N_\gamma - \ell + 1$ columns for input signal u_t, output signal y_t, or noise signal z_t by $U_\ell\langle\gamma\rangle \in \mathbb{R}^{\ell n_u \times j}$, $Y_\ell\langle\gamma\rangle \in \mathbb{R}^{\ell n_y \times j}$, or $Z_\ell\langle\gamma\rangle \in \mathbb{R}^{\ell n_y \times j}$ respectively. Thus,

$$U_\ell\langle\gamma\rangle = \begin{bmatrix} u_{s_\gamma+1} & u_{s_\gamma+2} & \cdots & u_{s_\gamma+j} \\ u_{s_\gamma+2} & u_{s_\gamma+3} & \cdots & u_{s_\gamma+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ u_{s_\gamma+\ell} & u_{s_\gamma+\ell+1} & \cdots & u_{s_\gamma+\ell+j-1} \end{bmatrix}, \tag{4.3}$$

$$Y_\ell\langle\gamma\rangle = \begin{bmatrix} y_{s_\gamma+1} & y_{s_\gamma+2} & \cdots & y_{s_\gamma+j} \\ y_{s_\gamma+2} & y_{s_\gamma+3} & \cdots & y_{s_\gamma+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ y_{s_\gamma+\ell} & y_{s_\gamma+\ell+1} & \cdots & y_{s_\gamma+\ell+j-1} \end{bmatrix}, \tag{4.4}$$

4.2.2 SSI Data Equation

The SSI data equation using a single segment, be it N4SID, MOESP, CVA or any other variant [41], is as follows,

$$Y_\ell\langle\gamma\rangle = \Gamma_\ell X\langle\gamma\rangle + H_\ell U_\ell\langle\gamma\rangle + Z_\ell\langle\gamma\rangle, \tag{4.5}$$

where $\Gamma_\ell \in \mathbb{R}^{(\ell n_y \times n)}$ is the extended observability matrix,

$$\Gamma_\ell = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{\ell-1} \end{bmatrix}$$

$H_\ell \in \mathbb{R}^{\ell n_y \times j n_u}$ is the block-Toeplitz matrix of system impulse response parameters,

$$H_\ell = \begin{bmatrix} D & 0 & & & \cdots & 0 \\ CB & D & 0 & & \cdots & 0 \\ CAB & CB & D & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \ddots & \vdots \\ CA^{\ell-2}B & & & & \cdots & D \end{bmatrix}$$

and $X\langle\gamma\rangle \in \mathbb{R}^{n\times j}$ the following matrix of system states

$$X\langle\gamma\rangle = \begin{bmatrix} x_{s_\gamma} & x_{s_\gamma+1} & \cdots & x_{s_\gamma+j-1} \end{bmatrix}.$$

A requirement for system identifiability in this circumstance is that the number of block rows, ℓ, exceeds the maximal observability index $\boldsymbol{\ell}(\mathcal{S})$, of the system [42]. This observability index $\boldsymbol{\ell}(\mathcal{S})$, is also referred to as the *system lag* [40]. Provided $\ell > \boldsymbol{\ell}(\mathcal{S})$, the minimal number of block columns must satisfy $j \geq (\ell + n + 1)n_u + \ell + n$ as derived by Sontag [43].

4.2.3 Single Segment Algorithm

For clarity, we consider the basic 4SID algorithm for the single segment case without noise and then specialize to the instrumental variable subspace identification IV-4SID algorithm [44]. The data generating system (4.1) is considered with data U_ℓ and Y_ℓ defined in (4.3) and (4.4) and the segment identifier omitted where it is understood from the context. First, the orthogonal projection to U_ℓ

$$\Pi_{U_\ell}^{\perp} = I - U_\ell^T (U_\ell U_\ell^T)^{\dagger} U_\ell \tag{4.6}$$

multiplies Y_ℓ on the right in the data equation (4.5), taking $Z_\ell = 0$ to yield,

$$Y_\ell \Pi_{U_\ell}^{\perp} = \Gamma_\ell X_\ell \Pi_{U_\ell}^{\perp}. \tag{4.7}$$

That is, $Y_\ell \Pi_{U_\ell}^{\perp}$ contains the column space of Γ_ℓ. We next extract this column space using the singular value decomposition (SVD). With weighting matrices W_1 and W_2, the SVD

$$W_1 Y_\ell \Pi_{U_\ell}^{\perp} W_2 = \begin{bmatrix} \mathcal{U}_1 & \mathcal{U}_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \mathcal{V}_1^T \\ \mathcal{V}_2^T \end{bmatrix}, \tag{4.8}$$

yields an estimate of Γ_ℓ by assigning $\mathcal{U}_1$ as the column-space and taking $W_1 \Gamma_\ell = \mathcal{U}_1$ to give

$$\hat{\Gamma}_\ell = W_1^{-1} \mathcal{U}_1. \tag{4.9}$$

Specific choices of W_1 and W_2 are shown in [44] to correlate with the PO-MOESP [45], N4SID [46], CVA [47] and IVM [48] algorithms. The coefficients of $\hat{C}$ are taken as the first n_y rows of $\hat{\Gamma}_\ell$, written here using MATLAB style indexing as

$$\hat{C} = \hat{\Gamma}_\ell(1 : n_y, :). \tag{4.10}$$

The shift invariance property of Γ_ℓ is exploited to generate an estimate of A with $\hat{\Gamma}_\ell^{(1)}$ and $\hat{\Gamma}_\ell^{(2)}$ as the upper $(\ell - 1)$ blocks and lower $(\ell - 1)$ blocks of $\hat{\Gamma}_\ell$ respectively [44]. The shift invariance property states that $\hat{\Gamma}_\ell^{(1)} A = \hat{\Gamma}_\ell^{(2)}$ holds and the coefficients of A can be calculated via

$$\hat{A} = \left(\Gamma_\ell^{(1)}\right)^{\dagger} \hat{\Gamma}_\ell^{(2)}. \tag{4.11}$$

The estimation of B, D, and x_{s_γ} is formulated as a least squares problem in terms of the one-step ahead prediction $\hat{y}_t$, and estimates of $\hat{A}$, $\hat{C}$ [49]. The one-step ahead prediction is given by $t_{s_\gamma} = t - s_\gamma$ such that

$$\hat{y}_{t+s_\gamma} = \sum_{\tau=0}^{t-1} CA^{t-\tau+1}Bu_{s_\gamma+\tau} + Du_{t+s_\gamma} + CA^t x_{s_\gamma}. \tag{4.12}$$

Applying the *vec* operator and the Kronecker product ($\otimes$), to (4.12) gives

$$\hat{y}_{t+s_\gamma} = vec(\hat{y}_{t+s_\gamma}) = \phi_B^T(t)vec(B) + \phi_D^T(t)vec(D) + \phi_{x_{s_\gamma}}^T(t)x_{s_\gamma} \tag{4.13}$$

This is linear in the coefficients of B, D, and x_{s_γ} with multipliers

$$\phi_B^T(t+s_\gamma) = \sum_{\tau=0}^{t-1} u_{s_\gamma+\tau}^T \otimes CA^{t-\tau+1}, \tag{4.14}$$

$$\phi_D^T(t+s_\gamma) = u_{t+s_\gamma}^T \otimes I_{n_y}, \tag{4.15}$$

$$\phi_{x_{s_\gamma}}^T(t+s_\gamma) = CA^t. \tag{4.16}$$

Using (4.14)-(4.16) to assemble the full regressor for $t \in \langle\gamma\rangle$

$$\Phi^T\langle\gamma\rangle = \begin{bmatrix} \Phi_B^T\langle\gamma\rangle & \Phi_D^T\langle\gamma\rangle & \Phi_{x_0}^T\langle\gamma\rangle \end{bmatrix} \tag{4.17}$$

$$\Phi_B\langle\gamma\rangle = \begin{bmatrix} \phi_B(s_\gamma) & \cdots & \phi_B(N_\gamma - s_\gamma - 1) \end{bmatrix} \tag{4.18}$$

$$\Phi_D\langle\gamma\rangle = \begin{bmatrix} \phi_D(s_\gamma) & \cdots & \phi_D(N_\gamma - s_\gamma - 1) \end{bmatrix} \tag{4.19}$$

$$\Phi_{x_0}\langle\gamma\rangle = \begin{bmatrix} \phi_{x_{s_\gamma}}(s_\gamma) & \cdots & \phi_{x_{s_\gamma}}(N_\gamma - s_\gamma - 1) \end{bmatrix} \tag{4.20}$$

$$\theta = \begin{bmatrix} vec(B)^T & vec(D)^T & x_{s_\gamma}^T \end{bmatrix} \tag{4.21}$$

Let $\mathcal{Y}\langle\gamma\rangle = vec(\hat{y}\langle\gamma\rangle)$. Then $\mathcal{Y}\langle\gamma\rangle = \Phi^T\langle\gamma\rangle\theta$ holds and the least squares estimate θ given by

$$\hat{\theta} = \left[\Phi\langle\gamma\rangle\Phi^T\langle\gamma\rangle\right]^{-1}\left[\Phi\langle\gamma\rangle\mathcal{Y}\langle\gamma\rangle\right]. \tag{4.22}$$

and the coefficients for $\hat{B}$, $\hat{D}$, and $\hat{x}_{s_\gamma}$ assembled per the definition of θ in (4.21).

4.3 Subspace System Identification - Multiple Segments

We seek to perform identification with archival data by seeking periods of informative transient data and use multiple segments to extract a dynamic model. The data segments are non-contiguous or segmented in time and variable in length. This data may be viewed as a series of data segments as shown in Figure 4.1 with time on the bottom axis. Let the segment

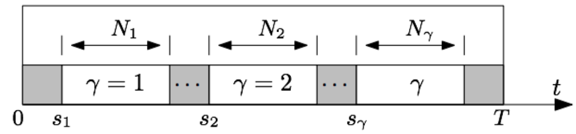

Figure 4.1: Segmented data representation

number be denoted as γ, start time index s_γ, data segment length N_γ, and number of segments γ_S. Given a selection of a data segment, let u_t and y_t be the input and output data from that segment. Using the interval notation in (4.2)

$$u\langle\gamma\rangle = \begin{bmatrix} u_{s_\gamma} & u_{s_\gamma+1} & \cdots & u_{s_\gamma+N_\gamma-1} \end{bmatrix} \tag{4.23}$$

$$y\langle\gamma\rangle = \begin{bmatrix} y_{s_\gamma} & y_{s_\gamma+1} & \cdots & y_{s_\gamma+N_\gamma-1} \end{bmatrix} \tag{4.24}$$

4.3.1 SSI Data Equation

When considering multiple data segments, the data equation (4.5) holds for each data segment $\gamma = 1, 2, \ldots, \gamma_S$. Consequently, we may immediately write the SSI data equation with concatenated data sets [38] as

$$\bar{Y}_\ell = \Gamma_\ell \bar{X} + H_\ell \bar{U}_\ell + \bar{Z}_\ell, \tag{4.25}$$

with

$$\bar{U}_\ell = \left[\begin{array}{c|c|c|c} U_\ell\langle 1\rangle & U_\ell\langle 2\rangle & \cdots & U_\ell\langle\gamma_S\rangle \end{array}\right] \tag{4.26}$$

$$\bar{Y}_\ell = \left[\begin{array}{c|c|c|c} Y_\ell\langle 1\rangle & Y_\ell\langle 2\rangle & \cdots & Y_\ell\langle\gamma_S\rangle \end{array}\right] \tag{4.27}$$

$$\bar{X} = \left[\begin{array}{c|c|c|c} X\langle 1\rangle & X\langle 2\rangle & \cdots & X\langle\gamma_S\rangle \end{array}\right] \tag{4.28}$$

$$\bar{Z}_\ell = \left[\begin{array}{c|c|c|c} Z_\ell\langle 1\rangle & Z_\ell\langle 2\rangle & \cdots & Z_\ell\langle\gamma_S\rangle \end{array}\right] \tag{4.29}$$

and the set of initial conditions

$$\mathcal{X}_0 = \begin{bmatrix} x_{s_1} & x_{s_2} & \cdots & x_{s_{\gamma_S}} \end{bmatrix}, \tag{4.30}$$

4.3.2 Segmented Data Subspace Algorithm

With the concatenated data equation (4.25), one could apply the regular subspace system identification algorithms to extract the column space of Γ_ℓ, then the coefficients of matrices A, C. The segmented data subspace algorithm generally follows the same steps as the single segment. First, the projection from (4.6) is assembled using $\bar{U}_\ell$ as

$$\Pi^{\perp}_{\bar{U}_\ell} = I - \bar{U}_\ell^T \left(\bar{U}_\ell \bar{U}_\ell^T\right)^{\dagger} \bar{U}_\ell \tag{4.31}$$

, and multiplies $\bar{Y}_\ell$ on the right in the data equation (4.25), taking $\bar{Z}_\ell = 0$, to yield

$$\bar{Y}_f \Pi^{\perp}_{\bar{U}_\ell} = \Gamma_\ell \bar{X}_\ell \Pi^{\perp}_{\bar{U}_\ell}. \tag{4.32}$$

The same SVD procedure in (4.8) provides an estimate of the column-space of Γ_ℓ. The estimate of $\hat{\Gamma}_\ell$ is again calculated via (4.9) and the matrices $\hat{A}$ and $\hat{C}$ are recovered from (4.11) and (4.10) respectively.

The formulation of the least squares problem to recover B, D, and now $\mathcal{X}_0$, requires slight modification. Each data segment used for identification introduces n coefficients associated with initial conditions into the parameter vector θ, for a total of $(n \times \gamma_S)$ parameters from $\mathcal{X}_0$. We construct a regressor matrix $\bar{\Phi}$ and parameter vector $\bar{\theta}$ using two segments $\gamma = (1, 2)$, starting indices (s_1, s_2), and data lengths (N_1, N_2), that easily generalizes to the multi-segment case. The inclusion of initial conditions $\mathcal{X}_0 = [x_{s_1}, x_{s_2}]$ from (4.30) gives the augmented parameter vector

$$\bar{\theta} = \left[\begin{array}{cccc} vec(B)^T & vec(D)^T & x_{s_1}^T & x_{s_2}^T \end{array}\right]. \tag{4.33}$$

The regressor elements for each data segment are calculated again using (4.14), (4.15) and (4.16) for the sequences $y_t\langle 1\rangle$ and $y_t\langle 2\rangle$ from (4.12) and $\hat{A}$ and $\hat{C}$. Then the augmented regressor matrix $\bar{\Phi}$ and output record $\bar{\mathcal{Y}}$ are assembled as follows.

$$\bar{\Phi}^T = \begin{bmatrix} \Phi_B^T\langle 1\rangle & \Phi_D^T\langle 1\rangle & \Phi_{\mathcal{X}_0}^T\langle 1\rangle & 0 \\ \Phi_B^T\langle 2\rangle & \Phi_D^T\langle 2\rangle & 0 & \Phi_{\mathcal{X}_0}^T\langle 2\rangle \end{bmatrix} \tag{4.34}$$

$$\bar{\mathcal{Y}} = \begin{bmatrix} vec(y_t\langle 1\rangle) \\ vec(y_t\langle 2\rangle) \end{bmatrix} \tag{4.35}$$

The least squares estimate of B, D, and $\mathcal{X}_0$ is then given by

$$\hat{\bar{\theta}} = \left[\bar{\Phi}\bar{\Phi}^T\right]^{-1} \left[\bar{\Phi}\bar{\mathcal{Y}}\right] \tag{4.36}$$

Note that $\Phi_B^T\langle\gamma\rangle$ and $\Phi_D^T\langle\gamma\rangle$ are simply stacked vertically in (4.34), as they are associated with the same parameters in $\bar{\theta}$ and share the same column dimension. The terms $\Phi_{\mathcal{X}_0}^T\langle 1\rangle$ and $\Phi_{\mathcal{X}_0}^T\langle 2\rangle$ appear in independent block columns in the regressor matrix $\bar{\Phi}^T$ as they apply only to the initial condition on that segment. In addition, the nonzero terms are located in the block rows associated with their respective data segments with zeros in all other block rows.

4.4 Identifiability

The salient question in this discussion, is not the fine detail of the identification algorithm but the conditions, under which such an algorithm uniquely identifies the system. It is these conditions which permit the systematic searching of archives for suitable data.

The ability to extract a unique estimate of the system from the data, formally identifiability, depends on both the system being identified and the data. We use the following definition of identifiability:

Definition 1. *The system*

$$\mathcal{S} = \{A, B, C, D, \mathcal{X}_0, \gamma = 1, 2, \ldots, \gamma_S\}$$

is identifiable from the data if n_u, n_y, n are fixed, $\mathcal{S}$ is linear time invariant and there is no other $\mathcal{S}' \neq \mathcal{S}$ that fits the data other than

$$\mathcal{S}' = \left\{TAT^{-1}, TB, CT^{-1}, D, T^{-1}\mathcal{X}_0, \gamma = 1, 2, \ldots, \gamma_S\right\}$$

for some similarity transformation T.

The definition of the concatenated data segments in $\bar{U}_\ell$ (4.26) and $\bar{Y}_\ell$ (4.27) are define in terms of concatenated block-Hankel matrices, but do not retain the Hankel structure as written. Let the time vector for each block column of $\bar{U}_\ell$ and $\bar{Y}_\ell$ as $t_{[i,\ell]} = t_{i,1}, t_{i,2}, \ldots, t_{i,\ell}$ where $i = 1, 2, \ldots, j$, such that each column may be written

$$u_{t_{[i,\ell]}} = \begin{bmatrix} u_{t_{i,1}} \\ u_{t_{i,2}} \\ \vdots \\ u_{t_{i,\ell}} \end{bmatrix} \quad \text{and} \quad y_{t_{[i,\ell]}} = \begin{bmatrix} y_{t_{i,1}} \\ y_{t_{i,2}} \\ \vdots \\ y_{t_{i,\ell}} \end{bmatrix} \tag{4.37}$$

We then can write an equivalent representation with no assumption of the neighboring columns being contiguous in time as

$$\bar{U}_\ell = \begin{bmatrix} u_{t_{[1,\ell]}} & u_{t_{[2,\ell]}} & \cdots & u_{t_{[j,\ell]}} \end{bmatrix} \tag{4.38}$$

$$\bar{Y}_\ell = \begin{bmatrix} y_{t_{[1,\ell]}} & y_{t_{[2,\ell]}} & \cdots & y_{t_{[j,\ell]}} \end{bmatrix} \tag{4.39}$$

and additionally define

$$\bar{W}_\ell = \begin{bmatrix} \bar{U}_\ell \\ \bar{Y}_\ell \end{bmatrix}. \tag{4.40}$$

Using these constructions, the following theorem provides sufficient conditions for identifiability of LTI systems from non-contiguous data.

Theorem 1. *Consider a linear time invariant system $\mathcal{S}$ in* (4.1) *with minimal state dimension n. If, for $\ell > n$, the input data in* (4.38)*, and stacked input/output data in* (4.40)*,*

$$rank(\bar{U}_\ell) = n_u \ell \tag{4.41}$$

$$rank(\bar{W}_\ell) = n_u \ell + n \tag{4.42}$$

then the system is uniquely identifiable. Further, the associated states at for all t in the data in $\bar{W}_\ell$ are uniquely recoverable.

Proof. In Willems and Markovsky [50], it is established that for an LTI system with known minimal state dimension n, the dimension of the length ℓ behaviors $\{w_k \in \mathscr{L}^{n_w}\ ,\ k = 1, \ldots, \ell\}$ is $m\ell + n$ where m is the input cardinality and equal to n_u in this case. Accordingly, if rank$(\bar{W}_\ell) = n_u\ell + n$, the columns of $\bar{W}_\ell$ span all possible length ℓ behaviors, denoted $\mathscr{B}|_{[1,\ell]}$ of the data generating system. Further, if $\ell > \boldsymbol{\ell}(\mathscr{B})$, where $\boldsymbol{\ell}(\mathscr{B})$ is the system lag, the leftkernel of $\bar{W}_\ell$ yields the complete minimal kernel representation of the system and thereby suffices for identifiability. The system lag is always less than the state dimension, therefore the condition that $\ell > n$ ensures existence of a complete kernel representation and does not require a-priori knowledge of $\boldsymbol{\ell}(\mathscr{B})$.

In our formulation, the inputs $\{u_k\}$ is defined as the system inputs, i.e. the inputs are not determined by the behaviors. Hence these are free variables and must satisfy rank$(\bar{U}_\ell) = n_u\ell$ in order for the subspace system identification algorithms to yield a state variable realization with $\{u_k\}$ as the input.

By construction, the extended observability matrix Γ_ℓ, has full column rank, and under the rank conditions on $\bar{U}_\ell$ and $\bar{W}_\ell$, the regression calculation to recover the coefficients of B,D, and initial conditions $\mathcal{X}_0$ yields a unique initial state for each data set. □

4.5 Applications

We consider a known noise free linear system to demonstrate the consistency of the proposed algorithm and to demonstrate that the rank conditions proposed in Theorem 1 are correct. I practice, there is always some noise on the signals or the system under investigation is being approximated by a linear system. A demonstration conducted using operational data from a gas turbine follows that characterizes the dynamic interaction between measured bearing temperatures other relevant measurements from the gas turbine.

4.5.1 Linear System

Consider the system

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t \\ y_t &= Cx_t + Du_t \end{aligned}$$

where

$$A = \begin{bmatrix} 0.9 & 0.2 \\ 0 & 0.8 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 \end{bmatrix}$$

Given a realization of the system from data, we have the estimates $\hat{A}$,$\hat{B}$,$\hat{C}$,$\hat{D}$, and initial conditions $\hat{\mathcal{X}}_0$ that are equivalent within a similarity transformation T per Definition 1. In the case of exact data, the similarity transformation can be explicitly calculated using an intermediate equivalent Jordan form for both the know system and identified state-space matrices. Define the common Jordan form of the underlying system as

$$\begin{aligned} \tilde{x}_{t+1} &= A_j \tilde{x}_t + B_j u_t \\ y_t &= C_j \tilde{x}_t + Du_t \end{aligned}$$

where

$$\begin{aligned} A_j &= T^{-1}AT &&= \hat{T}^{-1}\hat{A}\hat{T} \\ B_j &= T^{-1}B &&= \hat{T}^{-1}\hat{B} \\ C_j &= CT &&= \hat{C}\hat{T} \\ D_j &= D &&= \hat{D} \end{aligned}$$

With the estimates of $\hat{A}$, $\hat{B}$, $\hat{C}$, $\hat{D}$, the original matrices can be recovered by

$$\begin{aligned} A &= (T\hat{T}^{-1})\hat{A}(\hat{T}T^{-1}) \\ B &= (T\hat{T}^{-1})\hat{B} \\ C &= \hat{C}(\hat{T}T^{-1}) \\ D &= \hat{D} \end{aligned}$$

4.5.2 Turbine Bearing Temperature Modeling

In some locations there are multiple units operating on the same pipeline or power grid and often the model and rating of the turbines are identical, although the total operating hours may be significantly different or a particular turbine has been through an overhaul cycle.

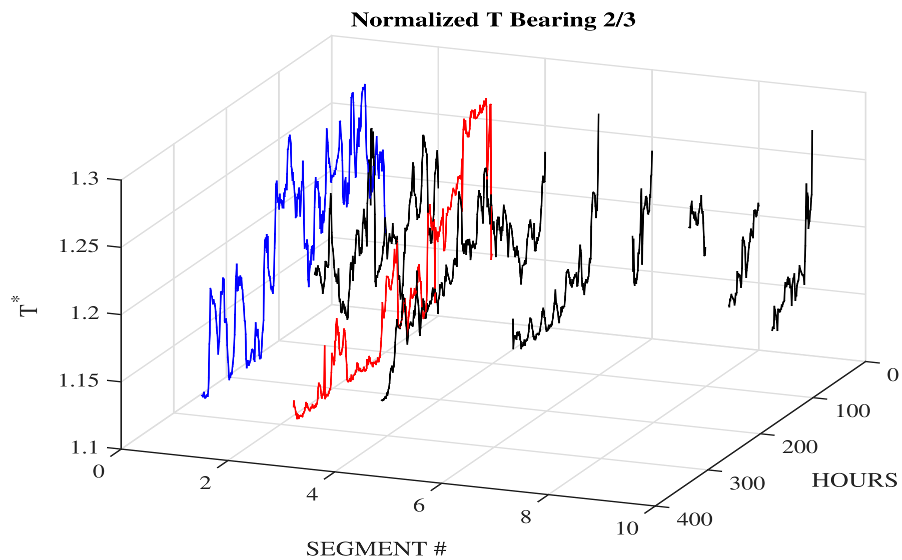

Figure 4.2: History of normalized temperatures of the GP 2/3 bearing plotted as contiguous data with two data segments (blue and red) highlighted for use in identification.

In this example, there are three 10 MW gas turbine units, (U1, U2 , U3), operating in a gas compression application. Each unit periodically experiences elevated bearing temperatures as shown in Fig. 4.2 with normalized bearing temperature plotted against the hours of continuous operation. The histogram in Fig. 4.3 shows the distribution of contiguous periods of fired operation and the duration of these periods. Analysts may be looking at this data to predict bearing life [9], isolate a root cause [8], or determine if behavior seen on a particular unit is being exhibited at another location or across the fleet. We seek to identify a dynamic model from available measurements to find relevant input output relations to enable this analysis. In this case, data is sampled hourly and aggregated remotely at a central monitoring facility. The histogram indicates that only several of the records from each unit contains a period of contiguous operation longer than 200 hours, thus limiting the model order and model fidelity if only a single data record were used.

We examine the data records for the three units to seek a suitably informative data record. With some intuition and background knowledge, we choose the gas producer shaft speed $u_1 = N_{GP}$ variable inlet guide vane command $u_2 = GV_C$, and lube oil pressure $u_3 = P_L$ as inputs with outputs chosen as the radial GP 2/3 bearing $y_1 = T_{BRG23}$ and GP thrust bearing temperature $y_2 = T_{GPTHRUST}$ temperatures as outputs. Each data set is separated into segments as shown in Fig. 4.2 and records that are longer than 200 hours are kept for identification. The data matrices $\bar{U}_l$, $\bar{Y}_l$ and regressor $\bar{\Phi}$ are assembled from (4.26), (4.27),

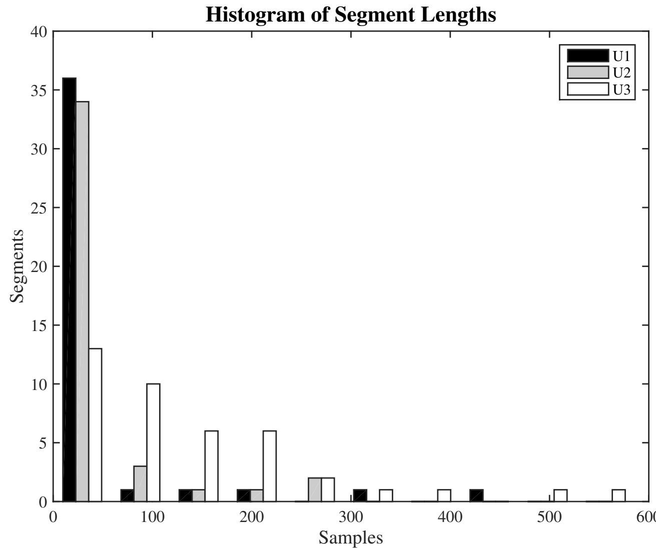

Figure 4.3: Histogram of data segment length in hours for each unit

(4.34) by adding one segment at a time. We use two segments as shown in Fig. 4.2 and developed in the concatenated SSI algorithm.

The singular value decomposition is used to approximate the rank of the matrices $\bar{U}\ell$ and $\bar{\Phi}$ by examining the singular values of the data matrices as shown in Fig. 4.4. Segment 1 failed the rank condition on $(\Phi\langle 1\rangle)$ and is not shown. This illustrates that while a single segment may not be sufficiently informative, concatenation of multiple segments may create a sufficiently informative data set to facilitate identification. Segment 1 failed the rank condition on $(\Phi\langle 1\rangle)$ and is not shown. This illustrates that while a single segment may not be sufficiently informative, the concatenation procedure facilitates identification. For identification the model order is selected as $n = 3$ and block row dimension $l = 4$.

Figure 4.5 demonstrates that linear models capture the transient dynamics of the MIMO interactions that are not easily captured through physics based modeling. In addition, the concatenation of data segments improves the information content and thus, the prediction accuracy. The RMS prediction error for T_{BRG23} and $T_{GPTHRUST}$ decreases by 26 % and 9 % respectively with the concatenated model vs. the single segment model. A comparison of the frequency response magnitudes is shown in Fig 4.6. The magnitude frequency illustrates that N_{GP} is highly relevant at low frequencies while the guide vane command and lube oil pressure show excitation at higher frequencies, in the neighborhood of 10E-4 rad/sec, with underdamped modes evident.

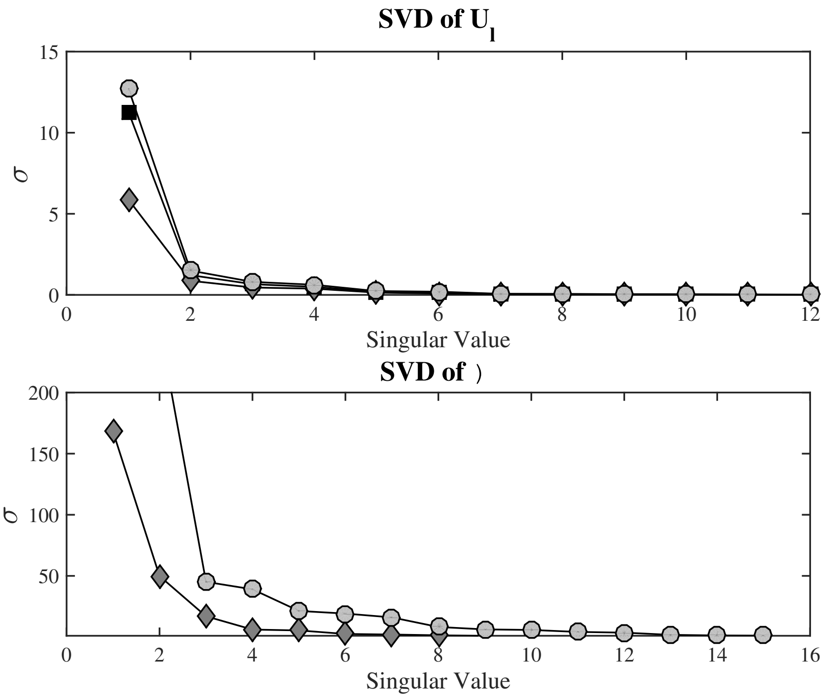

Figure 4.4: [Top] Singular values from SVD($U_\ell\langle 1\rangle$) ($\square$), SVD($U_\ell\langle 2\rangle$) ($\diamond$), and SVD($\bar{U}_\ell$) ($\bigcirc$) from concatenated segments demonstrating rank improvement via data concatenation. [Bottom] Singular values from SVD($\Phi\langle 2\rangle$) ($\diamond$) and SVD($\bar{\Phi}$) ($\bigcirc$) from concatenated segments, $\ell = 1, 2$.

4.6 Summary

The methods developed here are directed to trawling archival data records to reconstruct a sufficiently informative experiment in lieu of dedicated testing. Archival data records, by nature may contain long periods of steady state operation with little to no excitation to expose system dynamics. In fact, discarding periods of uninformative data has been shown to improve the accuracy of transfer function estimates [51]. It is often the case that archival data is missing measurements, there are periods when the process is shutdown or the data is corrupted. Markovsky recently addressed both the exact and approximate identification question with missing data. In the exact identification case, complete sub-matrices are extracted from an incomplete mosaic-Hankel matrix to compute their left-kernels to subsequently identify a non-minimal representation of the data generating system. In the approximate framework zero/infinite weights are assigned to the missing/exact data points in an ARMAX model structure [39].

Here, we propose a modification to subspace system identification methods to permit their application to multiple concatenated, but non-contiguous data sets, which jointly can

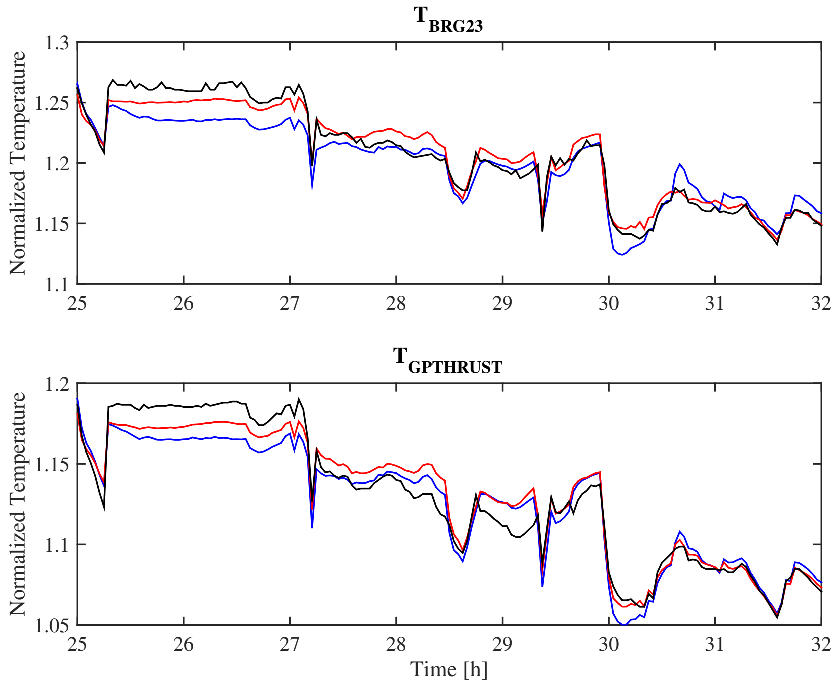

Figure 4.5: [Top] Time plot of actual T_{BRG23} (black), prediction from segment 2 model (blue) and concatenated data model (red). [Bottom] Time plot of actual $T_{GPTHRUST}$ (black), prediction from segment 2 model (blue) and concatenated data model (red).

be sufficiently rich for identification while individual records are not. This is a break from traditional system identification, where one uses a single contiguous synchronous input-output record. In addition, sufficient identifiability conditions are provided based on measurable rank conditions on the data matrices independent of any specific structure, i.e. not the block-Hankel data matrices that are a cornerstone of much of the literature on subspace identification.

Theorem 1 effectively removes the enforced Hankel matrix structure requirement on data used for subspace system identification, thereby enabling any set of sufficiently informative data to be candidates for subspace identification. This is not a trivial result as Hankel data structures are a cornerstone of subspace system identification algorithms. This result will will undoubtedly spark some lively debate for both the theoretical and application oriented community that employ subspace system identification.

It also provides the parameters for a user to search for informative sections of data within a time record by definition the minimum length of the record and enables the rank of the matrices $\bar{U}_\ell$, $\bar{W}_\ell$, to determine if a data set or selection of data sets is suitable. The sufficiency conditions are only applicable to exact identification, but the stochastic noisy setting is out of

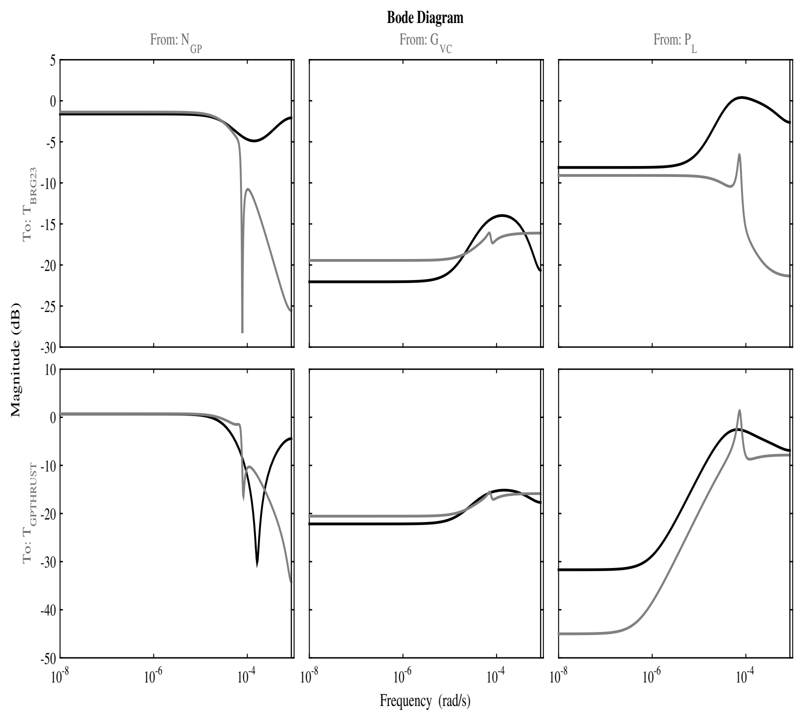

Figure 4.6: Magnitude plot from each input channel for segment 2 (black) and the concatenated data (gray). The columns are associated with inputs N_{GP}, GV_C, and P_L. The top row is for output T_{BRG23} and bottom row for output $T_{GPTHRUST}$.

the scope of this discussion. Regardless, the sufficient conditions serve as a useful guide when seeking to apply the proposed algorithm to archival data.

The multiple segment algorithm is demonstrated on a set of archival data for a gas turbine to model and replicate the transient behavior observed with a potential bearing problem. The dynamic model and associated frequency responses provide valuable information and new tools to analysts and engineers for isolating the root cause. In addition, prognostic component life prediction have been developed that use simplified linear models as a foundation [4]. This type of algorithm is well suited to these applications.

As with any engineering problem, there are trade offs when moving from theory to practice. While it is attractive to think that any selection of data windows of input and output data that satisfy the rank conditions, the for each new data segment and the additional n coefficients for the initial conditions, make the computational expense of the numerical implementation increase rapidly. In addition, the matrices $\bar{\Phi}$ used to recover B and D become ill-conditioned the dimension of the regressors grow with additional initial conditions to estimate.

4.7 Acknowledgements

This chapter, in part, is a reprint of "Gas Turbine Analytics using Archival Data Sets" that has been accepted to the *ASME Turbo Expo 2015. Turbine Technical Conference & Exposition*, Montreal, CA. 2015. Chad M. Holcomb, Robert R. Bitmead. The dissertation author was the primary investigator and author of this material.

5

Disturbance Rejection - Identification to Control

5.1 Introduction

Industry is continually demanding improved efficiency and reduced pollution from existing processes. A high profile example is the downward pressure on emissions from fossil fuel based energy production. As technologies such as wind and solar become economically competitive, gas turbines must follow suit to remain competitive with developing renewable energy sources. In particular, gas turbines are being pressed to comply with lower exhaust emissions of NOx and CO emission over a wider range of operating conditions [13]. In many geographic regions, failure to comply with stated limits may trigger significant economic penalties or prevent the unit from being operated [14]. New regulations may not only pertain to new equipment, but apply to a existing fleet equipment that has often been in operation from several years to over a decade.

One advantage of the gas turbine power plant is the capability to handle large transient changes on output power, termed load rejection. As emissions and efficiency performance demands become more stringent, the competing requirements for remaining compliant with emissions limits, while retaining similar transient load rejection pose significant control design challenges. In many cases, equipment is already installed and integrated into a larger process. Additionally the existing control design may not have a baseline mathematical model suitable for modern control design. To address this challenge we to identify a closed-loop model of the existing turbine and feedback control in the form of a generalized plant model [15].

One of the central considerations when applying system identification is to keep the

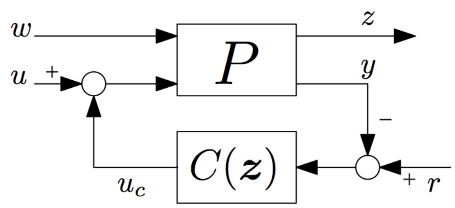

Figure 5.1: Feedback control of nonlinear turbine plant

intended use of the model in mind, such that identification produces a model suitable for the design task. In this case, the objective is a control relevant model of the existing system that facilitates control design that improves the disturbance rejection of the temperature and fuel control to disturbances on the turbine output shaft power. The model must capture the interaction of the disturbance input and control signals to system performance channels and existing instrumentation.

In practice, experiments for identification are only practical in closed loop and physical testing is costly. Secondly the power disturbance we wish to design for may potentially damage the physical equipment without a proven control policy. A first principles dynamic turbine simulation is available and provides an excellent and inexpensive experimentation and validation platform

In this setting, the nonlinear turbine plant and feedback control are identified in closed-loop around a nominal operating point and system identification produces a generalized plant model that is a suitable framework for control design and robustness analysis.

The generalized plant model is then used to design an outer-loop controller to improve the transient disturbance rejection performance.

5.2 Nominal System

Figure 5.1 depicts the closed-loop interconnection of the existing turbine plant P with the known linear controller $C(z)$. The controller output $u_c \in \mathbb{R}^{n_u}$ contains the fuel flow demand and the VIGV command. The measurements $y(t) \in \mathbb{R}^{n_y}$ available for control are the shaft speed y_1, combustion temperature as measured at the power turbine y_2, and and compressor discharge pressure y_3. The control references $r(t) \in \mathbb{R}^{n_r}$ are the shaft speed set-point r_1, and combustion temperature set-point r_2. The input signal w represents an internal plant or *middle* disturbance. The existing controller $C(z)$ is a known linear decentralized controller with strictly proper stable transfer functions that maps shaft speed to fuel flow demand and combustion temperature to

VIGV command.

$$\begin{bmatrix} u_{c_1}(t) \\ u_{c_2}(t) \end{bmatrix} = \begin{bmatrix} C_{11}(z) & 0 \\ 0 & C_{22}(z) \end{bmatrix} \begin{bmatrix} r_{11}(t) - y_1(t) \\ r_{12}(t) - y_2(t) \end{bmatrix}$$

For identification, we introduce the additive excitation signals u on both output channels of $C(z)$ as well as perturb w.

5.2.1 Model Structure

We consider that closed-loop interconnection of P and existing $C(z)$ in Fig 5.1 can be described around a local operating point by a system model as shown in Fig 5.2. The operating point is characterized but the turbine load fraction $\bar{\Omega}(t)$ defined as the ratio of the instantaneous net output shaft power defined as $\Omega(t)$, over the maximum output power defined as Ω_{MAX}:

$$\bar{\Omega}(t) = \frac{\Omega(t)}{\Omega_{MAX}} \tag{5.1}$$

Let $G_\ell(z)$ be a linear discrete time MIMO generalized plant model that describes dynamics at a local operating point ℓ where $\ell = 1, 2, \ldots, L$ defines the index of the local operating point, i.e. $\bar{\Omega}_\ell$ from (5.1), for $G_\ell(z)$. The relation from input to output may be written

$$\begin{bmatrix} z(t) - \bar{z}_\ell \\ y(t) - \bar{y}_\ell \end{bmatrix} = W_{O,\ell} G_\ell(\boldsymbol{z}) W_{I,\ell} \begin{bmatrix} w(t) - \bar{w}_\ell \\ u(t) - \bar{u}_\ell \end{bmatrix}, \tag{5.2}$$

where

$$W_{I,\ell} = \begin{bmatrix} W_{w,\ell} & 0 \\ 0 & W_{u,\ell} \end{bmatrix} \tag{5.3}$$

$$W_{O,\ell} = \begin{bmatrix} W_{z,\ell} & 0 \\ 0 & W_{y,\ell} \end{bmatrix}, \tag{5.4}$$

are memoryless static functions of the load fraction that operate as scaling matrices that account for changes in system gain over the load range. The parameters and $\bar{u}_\ell, \bar{w}_\ell, \bar{z}_\ell, \bar{y}_\ell$ denote the known expected value of each signal at operating point. The dynamics of the system are captured in a linear block $G_\ell(\boldsymbol{z})$, which is a linear parameter varying (LPV) dynamic system. Define the scaled signals; disturbance input $w_g(t)$, excitation inputs $u_g(t)$, performance channels $z_g(t)$, and outputs $y_g(t)$ as

$$\begin{aligned} w_g(t) &= W_{w,\ell}(w(t) - \bar{w}_\ell) && \in \mathbb{R}^{n_w} \\ u_g(t) &= W_{u,\ell}(u(t) - \bar{u}_\ell) && \in \mathbb{R}^{n_u} \\ y_g(t) &= W_{y,\ell}^{-1}(y(t) - \bar{y}_\ell) && \in \mathbb{R}^{n_y} \\ z_g(t) &= W_{z,\ell}^{-1}(z(t) - \bar{z}_\ell) && \in \mathbb{R}^{n_z} \end{aligned} \tag{5.5}$$

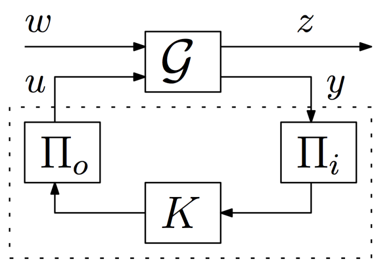

Figure 5.2: Overall control and identified generalized plant structure where the input and output weightings are absorbed into the control policy defined by the dotted box.

such that the linear linear dynamics are written as the state-space model

$$G_\ell(\boldsymbol{z}) := \begin{cases} x_g(t+1) & = Ax_g(t) + B_1 w_g(t) + B_2 u_g(t) \\ z_g(t) & = C_1 x_g(t) + D_{11} w_g(t) + D_{12} u_g(t) \\ y_g(t) & = C_2 x_g(t) + D_{21} w_g(t) + D_{22} u_g(t) \end{cases} . \tag{5.6}$$

In so-called 'packed notation' we have

$$G : \left\{ \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array} \right] \right. . \tag{5.7}$$

We assume that $W_{I,\ell}$, $W_{O,\ell}$ and $\bar{u}_\ell, \bar{w}_\ell, \bar{z}_\ell, \bar{y}_\ell$ are known such that for any ℓ we can define the affine mappings

$$\begin{array}{ccc} \Pi_I(\ell) & | & u_g = \Pi_I(\ell) u_k \\ \Pi_O(\ell) & | & y_k = \Pi_O(\ell) y_g \end{array} \tag{5.8}$$

such that we can seek to design a single controller $K(\boldsymbol{z})$, for some $\ell = n$ from the family of linear models $G_\ell(\boldsymbol{z})$ and implement the control over the range of $\ell = 1, \ldots, L$ such that the closed loop system is well posed and internally stable as proposed in Fig 5.2.

5.2.2 COBRA Identification

For identification, let the inputs and outputs in the generalized plant in (5.5) be denoted defined as $\eta(t)$ and the output as $\nu(t)$

$$\begin{aligned} \eta(t) &:= [w_g(t), u_g(t)]^T \quad \in \mathbb{R}^{n_\eta}, \\ \nu(t) &:= [z_g(t), y_g(t)]^T \quad \in \mathbb{R}^{n_\nu}, \end{aligned} \tag{5.9}$$

where $n_\eta = n_w + n_u$ and $n_\nu = n_z + n_y$. This allows (5.6) to be written compactly as

$$\nu(t) = G_\ell(z)\eta(t) \tag{5.10}$$

Similar to subspace identification methods that employ instrumental variables, let $\xi(t) \in \mathbb{R}^{n_\xi}$ be a quasi-stationary signal that is correlated with the inputs $\eta(t)$ [44] and uncorrelated with any noise that may be on the signals. The possible candidates for $\xi(t)$ include the usual suspects that are used in the instrumental variable (IV) methods. If experiments are performed in open loop a typical choice is past values of the input or when experiments are performed closed-loop a typical choice is a reference or excitation signal [48]. Using $\mathtt{E}$ to denote the expectation operator the N-point estimate of the cross-covariances of the input and output data $\xi(t)$ over some domain defined by $\tau \in [\tau_q, \tau_r]$ be defined:

$$\hat{R}_{\eta\xi}(\tau) = \frac{1}{N}\sum_{t=0}^{N-\tau-1} \mathtt{E}\left[\eta(t+\tau)\xi(t)^T\right] \tag{5.11}$$

$$\hat{R}_{\nu\xi}(\tau) = \frac{1}{N}\sum_{t=0}^{N-\tau-1} \mathtt{E}\left[\nu(t+\tau)\xi(t)^T\right] \tag{5.12}$$

The covariance estimates have dimensions $\hat{R}_{\eta\xi}(\tau) \in \mathbb{R}^{n_\eta \times n_\xi}$ and $\hat{R}_{\nu\xi}(\tau) \in \mathbb{R}^{n_\nu \times n_\xi}$. Defining the cross-covariance of the state $x(t)$ and $\xi(t)$ as

$$\hat{R}_{x\xi}(\tau) = \frac{1}{N}\sum_{t=0}^{N-\tau-1} \mathtt{E}\left[x(t+\tau)\xi(t)^T\right], \tag{5.13}$$

then enables the relations between the estimates of the covariance functions to be expressed in terms of the state-space matrices (A, B, C, D) as

$$\begin{aligned} \hat{R}_{x\xi}(\tau+1) = &\quad A\hat{R}_{x\xi}(\tau) + B\hat{R}_{\eta\xi}(\tau) \\ \hat{R}_{\nu\xi}(\tau) = &\quad C\hat{R}_{x\xi}(\tau) + D\hat{R}_{\nu\xi}(\tau) + \hat{R}_{\gamma\xi}(\tau). \end{aligned} \tag{5.14}$$

Define $\boldsymbol{R}_{\nu\xi} \in \mathbb{R}^{(n_\nu i \times n_\xi j)}$ as the block Hankel matrix assembled as j block columns and i block rows of covariance function estimates from (5.12) as

$$\boldsymbol{R}_{\nu\xi} = \begin{bmatrix} \hat{R}_{\nu\xi}(\tau_q) & \hat{R}_{\nu\xi}(\tau_q+1) & \cdots & \hat{R}_{\nu\xi}(\tau_q+j-1) \\ \hat{R}_{\nu\xi}(\tau_q+1) & \hat{R}_{\nu\xi}(\tau_q+2) & \cdots & \hat{R}_{\nu\xi}(\tau_q+j) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{R}_{\nu\xi}(\tau_q+i-1) & \hat{R}_{\nu\xi}(\tau_q+i) & \cdots & \hat{R}_{\nu\xi}(\tau_q+i+j-2) \end{bmatrix} \tag{5.15}$$

The block Hankel matrix $\boldsymbol{R}_{\eta\xi} \in \mathbb{R}^{(n_\eta i \times n_\xi j)}$ is defined similarly with covariance estimates from (5.11). The of cross-covariance of the state $x(t)$ and $\xi(t)$ is defined as $\boldsymbol{R}_{x\xi} \in \mathbb{R}^{(ni \times n_\xi j)}$ as

$$\boldsymbol{R}_{x\xi} = \begin{bmatrix} \hat{R}_{x\xi}(\tau_q) & \hat{R}_{x\xi}(\tau_q+1) & \cdots & \hat{R}_{x\xi}(\tau_q+j-1) \end{bmatrix} \tag{5.16}$$

The covariance based subspace data equation is written

$$\boldsymbol{R}_{\nu\xi} = \Gamma_i \boldsymbol{R}_{x\xi} + T_{0|i-1}\boldsymbol{R}_{\eta\xi} + \boldsymbol{R}_{\upsilon\xi} \tag{5.17}$$

where

$$\Gamma_i = \begin{bmatrix} C^T & (CA)^T & \cdots & (CA^{i-1}))^T \end{bmatrix}^T, \tag{5.18}$$

is the extended observability matrix and

$$T_{0|i-1} = \begin{bmatrix} G(0) & & & \\ G(1) & G(0) & & \\ \vdots & \vdots & \ddots & \\ G(i-1) & G(i-2) & \cdots & G(0) \end{bmatrix} \tag{5.19}$$

is the block-lower-triangular Toeplitz matrix of impulse response parameters

$$G(k) \begin{cases} 0 & k < 0, \\ D & k = 0, \\ CA^{k-1}B & k > 0. \end{cases}$$

Define $\boldsymbol{R}^{+}_{\eta\xi}$ as the matrix $\boldsymbol{R}_{\eta\xi}$ with an additional block row appended to the bottom corresponding to $[(\tau_q + i) \cdots (\tau_q + j + i - 1)]$. Additionally, define ($\bar{\boldsymbol{R}}_{\nu\xi}$ and $\bar{\boldsymbol{R}}_{v\xi}$) as ($\boldsymbol{R}_{\nu\xi}$ and $\boldsymbol{R}_{v\xi}$) with the lag index for each block element is shifted forward by one respectively. The lag shifted data equations may be written as

$$\bar{\boldsymbol{R}}_{\nu\xi} = \Gamma_i A \boldsymbol{R}_{x\xi} + T_{0|i} \boldsymbol{R}^{+}_{\eta\xi} + \bar{\boldsymbol{R}}_{v\xi} \tag{5.20}$$

The estimates of the matrices A and C are obtained by first estimating the signal subspace by recovering the range space of Γ_i. Introduce the projection matrix $\Pi_{R^{\perp}_{\eta\xi}}$, that projects the row space of a matrix onto the orthogonal compliment of the row space (i.e null-space) of the matrix $\boldsymbol{R}^{+}_{\eta\xi}$:

$$\Pi_{R^{\perp}_{u\xi}} = I - (\boldsymbol{R}^{+}_{\eta\xi})^T \left((\boldsymbol{R}^{+}_{\eta\xi})(\boldsymbol{R}^{+}_{\eta\xi})^T) \right)^{\dagger} (\boldsymbol{R}^{+}_{\eta\xi}) \tag{5.21}$$

Post multiplication of (5.17) and (5.20) by (5.21) gives

$$\begin{aligned} \boldsymbol{R}_{\nu\xi} \Pi_{R^{\perp}_{u\xi}} &= \Gamma_i \boldsymbol{R}_{x\xi} \Pi_{R^{\perp}_{u\xi}} + \boldsymbol{R}_{v\xi} \Pi_{R^{\perp}_{u\xi}} \\ \bar{\boldsymbol{R}}_{\nu\xi} \Pi_{R^{\perp}_{u\xi}} &= \Gamma_i A \boldsymbol{R}^{x\xi} \Pi_{R^{\perp}_{u\xi}} + \bar{\boldsymbol{R}}_{v\xi} \Pi_{R^{\perp}_{u\xi}} \end{aligned}$$

An estimate of Γ_i is obtained from an SVD of $\boldsymbol{R}_{\nu\xi} \Pi_{R^{\perp}_{u\xi}}$

$$\boldsymbol{R}_{\nu\xi} \Pi_{R^{\perp}_{u\xi}} = \begin{bmatrix} \mathcal{U}_1 & \mathcal{U}_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} \mathcal{V}_1^T \\ \mathcal{V}_2^T \end{bmatrix} \tag{5.22}$$

An estimate

$$\hat{\Gamma}_i = \mathcal{U}_1 \Sigma_1^{1/2} \tag{5.23}$$

The least squares solution for an estimate of $\hat{A}$ from (5.23) has been shown to be given by

$$\hat{A} = \Sigma_1^{-1/2} \mathcal{U}_1^T \bar{\boldsymbol{R}}_{\nu\xi} \Pi_{R_{u\xi}^{\perp}} \mathcal{V}_1 \Sigma_1^{-1/2} \tag{5.24}$$

The algorithm for extracting the signal subspace and recovering the realization matrices A, B, C, and D, is detailed in [52]. The identified system matrices can be partitioned via (5.7) for control design and robust analysis.

5.3 Control Synthesis

For control design, the system given by the generalized plant (5.6), and pencil (5.7) is used with estimates of the matrices given by the COBRA identification algorithm above. We apply H_2-optimal control in tandem some loop-shaping, to compute a single input, two output controller, i.e. a SIMO control, at a single operating condition, using the identified standard plant. In addition, an output scaling optimization is applied to the calculated H_2-controller to improve off-design performance.

5.3.1 Discrete-Time H_2-Optimal Control

The H_2-optimal controller is the modern incarnation of the LQG controller placed in the systematic framework developed to synthesize the robust H_∞-optimal controller as developed in [53] and [54]. The H_2-optimal controller employed here will mainly follow [55], [56], [57], and [58] (in the unconstrained case). A more rigorous treatment can be found in [59] . The assumptions on (5.6) are as follows:

Assumption 7. *$D_{11} = 0$ and $D_{22} = 0$,*

Assumption 8. *(A, B_2) is stabilizable and (A, C_2) is detectable,*

Assumption 9. *D_{12} has full column rank and D_{21} has full row rank,*

Assumption 10. *For all $\mathbf{z}^{-1}$ on the complex unit circle*

$$\begin{bmatrix} A - \mathbf{z}^{-1}I & B_2 \\ C_1 & D_{12} \end{bmatrix} \text{ has full column rank and } \begin{bmatrix} A - \mathbf{z}^{-1}I & B_1 \\ C_2 & D_{21} \end{bmatrix} \text{ has full row rank },$$

where $\mathbf{z}$ is the z-transform variable.

Assumption 11.

$$\begin{bmatrix} C_1^T C_1 & C_1^T D_{12} \\ D_{12}^T C_1 & D_{12}^T D_{12} \end{bmatrix} \geq 0$$

Let the generalized plant have a frequency domain representation and partition of

$$\begin{bmatrix} z \\ y \end{bmatrix} = G(\mathbf{z}) \begin{bmatrix} w \\ u \end{bmatrix} \iff \begin{bmatrix} z \\ y \end{bmatrix} = \left[\begin{array}{c|c} G_{11}(\mathbf{z}) & G_{12}(\mathbf{z}) \\ \hline G_{21}(\mathbf{z}) & G_{22}(\mathbf{z}) \end{array}\right] \begin{bmatrix} w \\ u \end{bmatrix}. \tag{5.25}$$

The criterion to optimize is given as

$$J = \|F_\ell(G, K)\|_2^2, \tag{5.26}$$

where $F_\ell(G, K)$ is the lower fractional transformation given by

$$F_l(G\ell, K) = G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21},$$

which is also the transfer function from w to z. Let,

$$T_{zw}(\boldsymbol{z})_\ell = F_l(G_\ell, K) = \begin{bmatrix} T_{z_1 w}(\boldsymbol{z})_\ell \\ T_{z_2 w}(\boldsymbol{z})_\ell \end{bmatrix},$$

such that the mapping from from w to z_1 and z_2 may be written

$$z_1(t) = T_{z_1 w}(\boldsymbol{z})_\ell w(t) \tag{5.27}$$

$$z_2(t) = T_{z_2 w}(\boldsymbol{z})_\ell w(t). \tag{5.28}$$

The controller K_p is computed for a single plant $G_\ell(\boldsymbol{z})|_{\ell=p}$ and applied to all plant realizations across the operating range. The nominal model $G_p(\boldsymbol{z})$ is an identified closed-loop model and is taken as the reference using the native *dh22lqg* Matlab function.

5.3.2 Output Scaling

Although the control formation is an H_2 control design, the goal is to reduce the effect of a step input disturbance to the performance channels. In the frequency domain, this is more of a measure of the resultant $\mathcal{H}_\infty$ norm and in the time domain, the ratio of the peak excursion of the performance channels z to a step input on w. To improve the off-design performance the output scaling for this control policy is computed via,

$$\begin{bmatrix} \sigma_{1,\ell}, \sigma_{2,\ell} \end{bmatrix} = \underset{\sigma_1, \sigma_2, \ell}{\arg\min} \, ||F_l(G_\ell, \begin{bmatrix} \sigma_{1,\ell} & 0 \\ 0 & \sigma_{2,\ell} \end{bmatrix} K)||_\infty \tag{5.29}$$

The controller from the nominal plant K_p is applied to each of the off-design plants and the output weighting from (5.29) is computed and absorbed into the scaling matrices Π_I and Π_O and treated as a function of load.

5.4 Application to a Low Emissions Gas Turbine

A first principles model of a 5MW dry-low emissions turbine generator operating in closed loop with fuel and temperature control with data sampled with $Ts = 0.25s$. The external disturbance $w(t)$ is an excitation to the load applied to the turbine shaft $\Omega_d(t)$ where the disturbance signal is uniformly distributed signal with an amplitude equal to approximately 5% of the maximum output power turbine load. The control excitation signals are the uniformly distributed random additive a fuel control excitation $u_{c1}(t)$, and airflow control excitation $u_{c2}(t)$ that are also set to perturb the steady state control signals by approximately 5% of the steady state value at each operating point. The measured outputs are the shaft speed $y_1 = N_{GP}$, turbine inlet temperature $y_2 = T_5$ and compressor outlet pressure $y_3 = P_2$. The performance channels are considered as the shaft speed error $z_1 =$ and combustion temperature error z_2 which gives:

$$\begin{aligned} u(t) &= \begin{bmatrix} u_{c1}(t) & u_{c2}(t) \end{bmatrix}^T \\ w(t) &= \begin{bmatrix} \Omega_d(t) \end{bmatrix}^T \\ y(t) &= \begin{bmatrix} y_1(t) & y_2(t) & y_3(t) \end{bmatrix}^T \\ z(t) &= \begin{bmatrix} z_1(t) & z_2(t) \end{bmatrix}^T \end{aligned} \tag{5.30}$$

5.4.1 Identification Results and Model Order Selection

The identification procedure uses a first principles high-fidelity turbine model and controllers operating in closed loop, from $L(t) = 0.9, 0.8, \cdots 0.4$. At each operating load, 3000 samples were logged with 1500 of the samples used for identification and validation. Linear models of various orders are fit to the data set. During the identification procedure, the selection of the lags τ_q and τ_r proved critical to obtaining suitable performance from the estimated plant. Figure (5.3) shows the cross-covariance estimate of the data obtained at an operating point $\bar{\Omega} = 0.4$. The information content falls off at lag values above 100 samples, shown in the dashed black window. The windows with lag values set to 400 samples, shown in dashed blue spans the entire range of lags where the cross-covariances are non-zero. Although the window corresponding to $\tau \in [-400, 400]$ captures the informative range of the cross-covariance estimate the estimated models often displayed significant bias as well as produced unstable plant realizations while the the window $\tau \in [-100, 100]$ consistently yielded stable plant realizations. For identification, the model orders $n = 4, 5, 6$ were considered. There was not an appreciable difference between the $n = 4$ and $n = 5$ models, so the $n = 5$ model results are not presented. The the speed error signal z_1 and temperature error channel z_2 are used as the benchmark signals

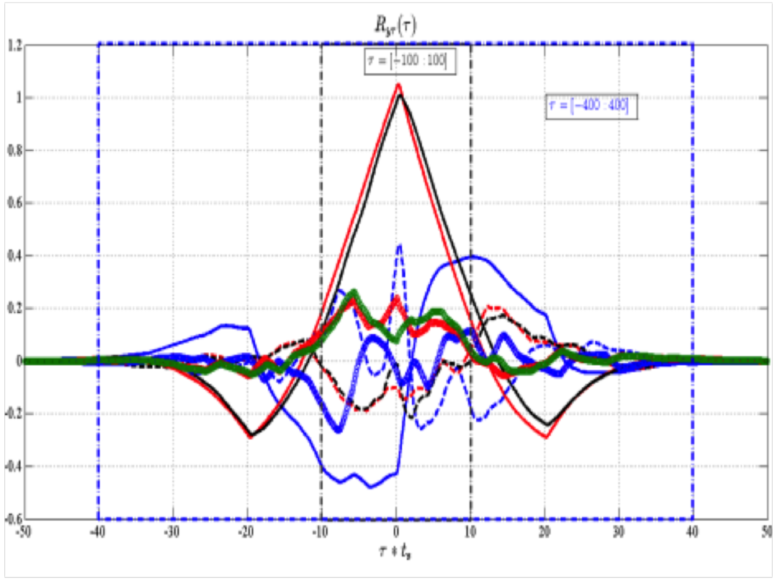

Figure 5.3: Cross-covariance estimate at a normalized load of $\bar{\Omega} = 0.4$ of max power with τ at 100 (black dashed) and 400 (blue dashed box)

for evaluation of the identification accuracy in the time domain. Figures 5.4 and 5.5 compare the time domain signals from the first principle simulation and identified model predictions at the upper range of the load range $\bar{\Omega} = 0.9$ on the left two plots and for $\bar{\Omega} = 0.4$ on the right. For each set of models, the speed error z_1 is well modeled over the load range and resemble the simulation signals. The prediction of temperature error z_2 however degrades in the lower loads, particularly in the model set where $n = 4$. The summary of the prediction error properties in Fig. 5.6 illustrate that the root mean square prediction error and variance are significantly better using the model set with $n = 6$ across the load range and a sixth order model should be considered for control design.

One of the central assumptions was that the closed-loop dynamics are consistent throughout the operating regime except for variations in the gains from input to output channels and steady state values. If the scaling in Π_I and Π_O is done properly, then the family of identified plants should demonstrated comparable frequency responses and step responses. Figure 5.7 demonstrates that the step from the disturbance to the performance channels are similar in time scale and shape for the identified plants $G_\ell(z)$, shown at four load conditions spanning 50% of the total operating range. Figure 5.8 depicts the estimated magnitude of the sensitivity functions from the disturbance w to both performance channels of z and nominal plant with output scaling at $L = [0.9, 0.7.0.5, 0.4]$. In addition, the absolute value of the plant eigenvalues are summarized in Table 5.1. The approximation of linear dynamics appear similar in both time and frequency domain, which lends support to proceed with the choice of a single nominal plant model for control design.

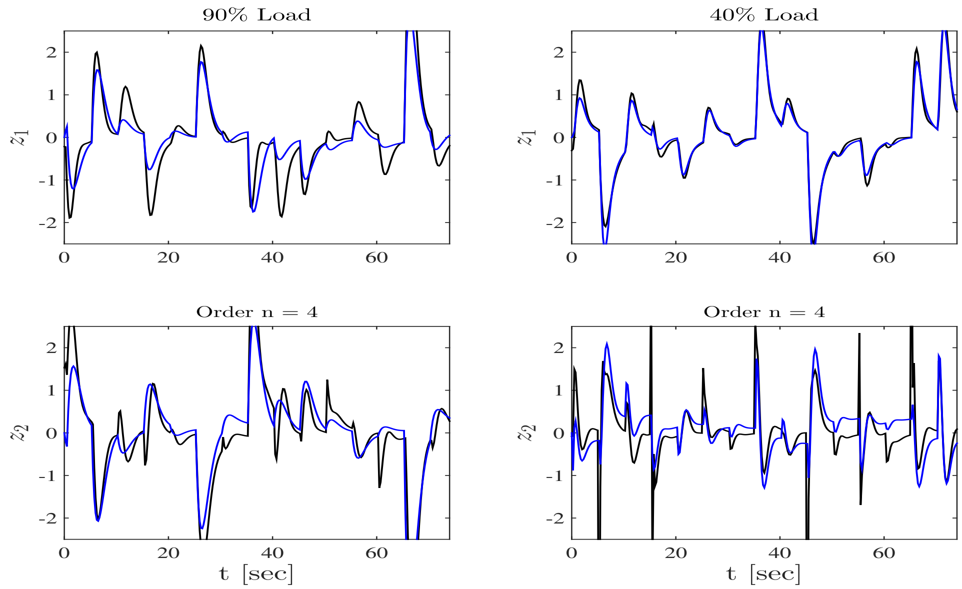

Figure 5.4: Comparison of z_1, z_2 from first principles simulation signals (black) and identified plant (blue) for $n = 4$ at $\bar{\Omega} = 0.9$ [left] and $\bar{\Omega} = 0.4$ [right]

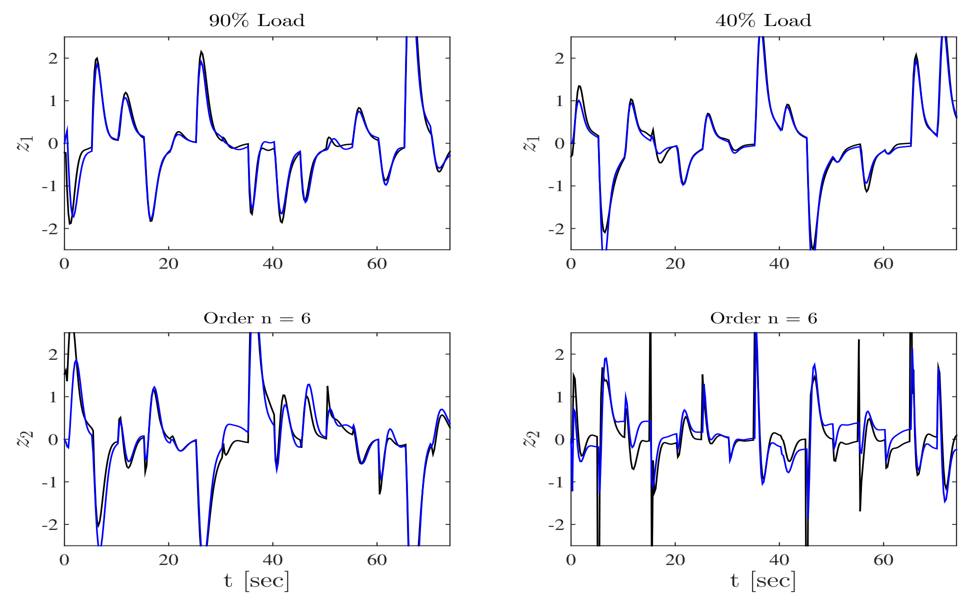

Figure 5.5: Comparison z_1, z_2 from first principles simulation signals (black) and identified plant (blue) for $n = 6$ at $\bar{\Omega} = 0.9$ [Left] and $\bar{\Omega} = 0.4$ [Right]

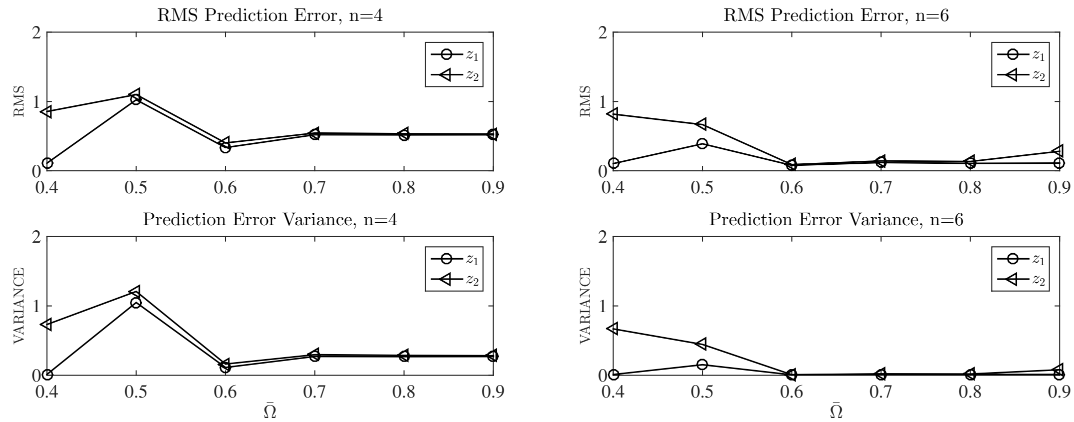

Figure 5.6: [Left] Prediction error and variance for $n = 4$ and [Right] Prediction error and variance for $n = 6$.

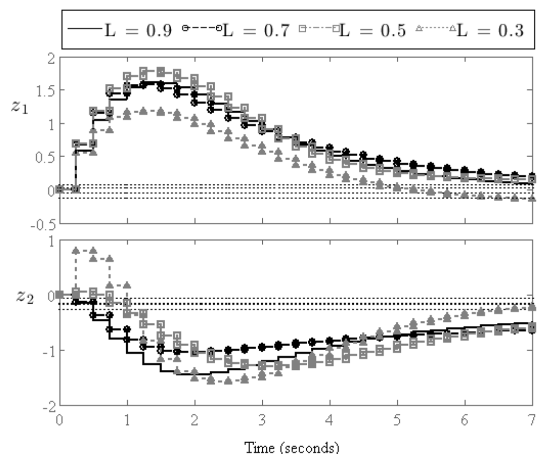

Figure 5.7: Step response of from $w(t)$ to z_1 and z_2 at $\bar{\Omega} = 0.9$ (-), 0.7 ($\bigcirc$), 0.5 ($\square$), and 0.4 ($\triangle$) demonstrating the dynamic similarity in the transient response of the performance channels.

Table 5.1: Magnitude of eigenvalues of A over identified load range for $n = 4$ and $n = 6$.

$\bar{\Omega}$	Magnitude of $\lvert\lambda\rvert$ for $n = 4$					
0.9			0.8224	0.8224	0.9259	0.9756
0.8			0.8475	0.8475	0.9038	0.9038
0.7			0.8425	0.8425	0.8987	0.9566
0.6			0.7856	0.7856	0.8809	0.8809
0.5			0.8327	0.8327	0.9198	0.9841
0.4			0.6560	0.6560	0.8642	0.8642
$\bar{\Omega}$	Magnitude of $\lvert\lambda\rvert$ for $n = 6$					
0.9	0.7590	0.7590	0.8943	0.8943	0.9740	0.9740
0.8	0.7172	0.7172	0.7227	0.8674	0.9094	0.9862
0.7	0.7085	0.7085	0.7752	0.8602	0.9416	0.9800
0.6	0.6342	0.7347	0.7347	0.8824	0.9038	0.9742
0.5	0.7671	0.7671	0.8501	0.8501	0.9865	0.9865
0.4	0.1483	0.7463	0.7463	0.8441	0.8764	0.9386

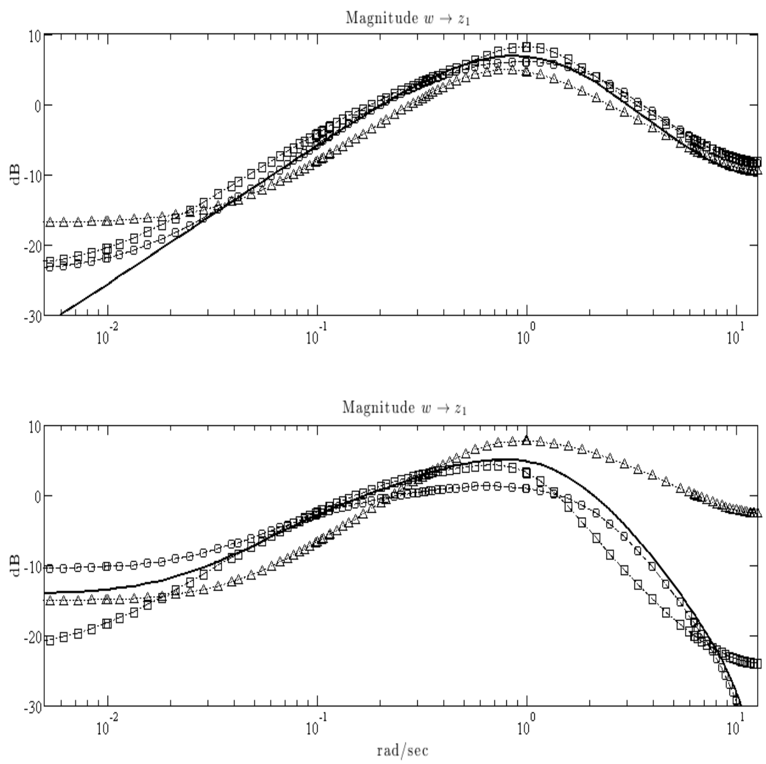

Figure 5.8: Magnitude of the sensitivity function of the disturbance w to both performance channel z at $\bar{\Omega} = 0.9$ (-), 0.7 ($\bigcirc$), 0.5 ($\square$), and 0.4 ($\triangle$)

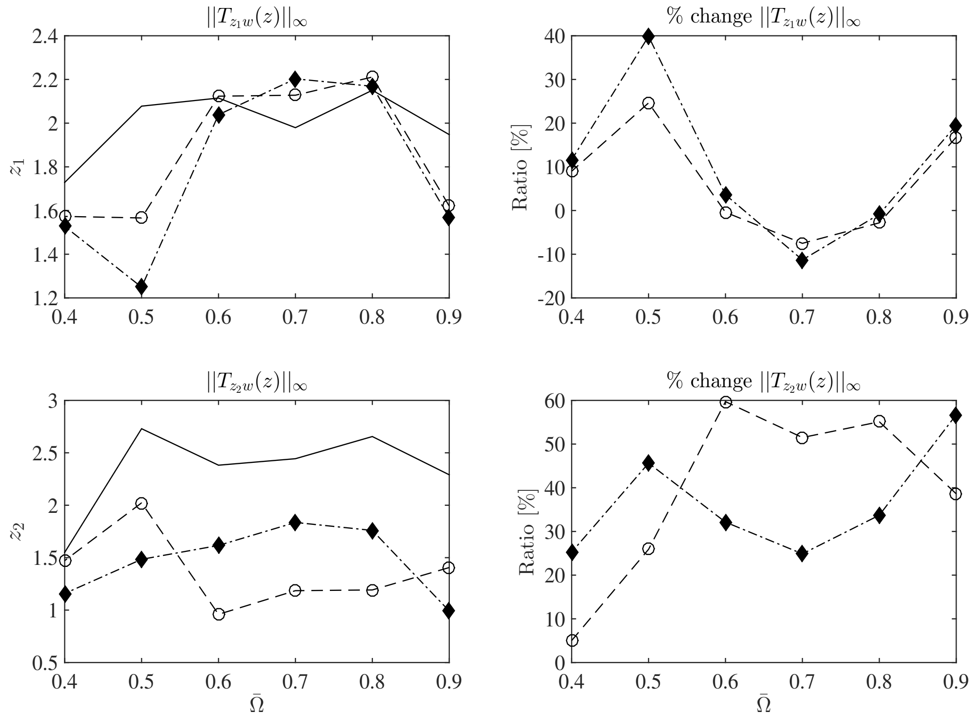

Figure 5.9: [Left] $||T_{zw}(\boldsymbol{z})||_\infty$ for each performance channel, [Right] percent change of the infinity norm with both control policies with respect to the identified model. Legend: Identified plant (-), initial K (◯) and K with output scaling (⧫).

5.4.2 Controller Performance

For control design, the shaft speed output $y_1 = N_{GP}$ is chosen to design a single-input multi-output controller to improve the transient disturbance rejection. The primary objective is to reduce the overshoot of the temperature channel during a transient load change. The temperature must be maintained within a tight window to avoid exceeding the mechanical upper limit during a increasing transient load condition and avoid going below a combustion stability limit during a increasing transient load condition. The latter forces the fuel control to make the fuel air ratio in the combustion chamber richer to remain stable. In this condition, the emissions rapidly exceed the nominal values. The worst case is assumed as a step input to the disturbance channel w. The plant associated with $\bar{\Omega} = 0.6$, i.e. is chosen as the design plant and an H_2-optimal controller is designed for this load condition.

Figure 5.9 summarizes the control performance in terms of the infinity norms of the transfer functions from the disturbance input w to the performance channels z_1 and z_2 for a controller designed for the plant identified at $\bar{\Omega} = 0.6$ using a sixth order $n = 6$. Overall, the controller calculated for this case demonstrates an appreciable improvement for both performance channels across the load range . The most significant improvement is the temperature control channel with an average of a 39% and 36% reduction in $||T_{z_2w}(\boldsymbol{z})||_\infty$ for the case of the controller without output scaling and with output scaling respectively. In addition, the speed demonstrates an average of a 6% and 10% reduction in $||T_{z_1w}(\boldsymbol{z})||_\infty$ for the case of the controller without output scaling and with output scaling respectively.

The best demonstration of control performance in this case is to compare the step responses from the disturbance input to the performance channels, since that is stated as the objective of the exercise. Figure 5.10 illustrates the step responses for the off-design conditions at $\bar{\Omega} = [0.9, 0.4]$ for the controller K_p designed at $\bar{\Omega} = [0.6]$. At $\bar{\Omega} = 0.9$ both the speed error and temperature error responses are improved significantly with the output scaling version (♦) demonstrating superior performance as compared to the nominal controller and existing plant. At $\bar{\Omega} = 0.4$, there is little change to the speed response for each controller although the temperature response is improved and the scaled controller again outperforms the nominal policy.

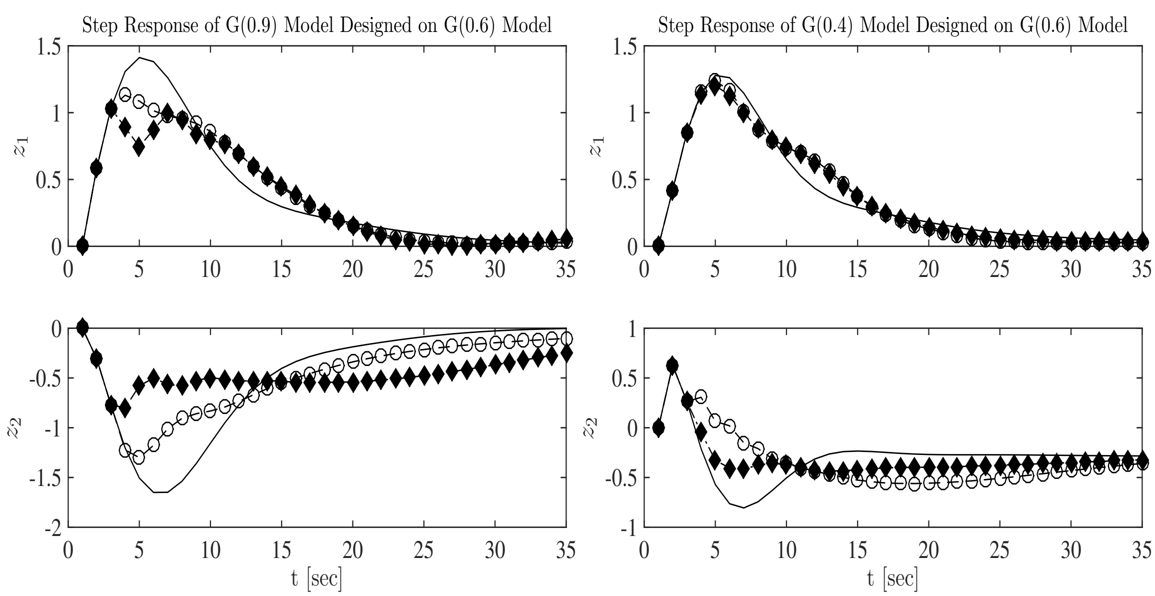

Figure 5.10: [Left] Step responses for $\bar{\Omega} = 0.4$ and [Right] $\bar{\Omega} = 0.9$ with controller designed on identified plant for $\bar{\Omega} = 0.6$. Identified plant (-), initial $K(\bigcirc)$ and K with output scaling (♦).

Figure 5.11 shows the results for control policy applied across the load range for the nominal controller K_p ($\bigcirc$) and K with output scaling (♦) as compared to the existing controller. For the nominal controller K_p, the mean improvement in peak response is 14% for the speed channel and 31% for the temperature channel. For the nominal controller with output scaling, the mean improvement in peak response is 20% for the speed channel and 31% for the temperature channel which suggest that the output scaling proposed is a tractable method for applying a single controller. The plot does indicate that at a given performance condition, the output scaling does not always outperform the nominal controller.

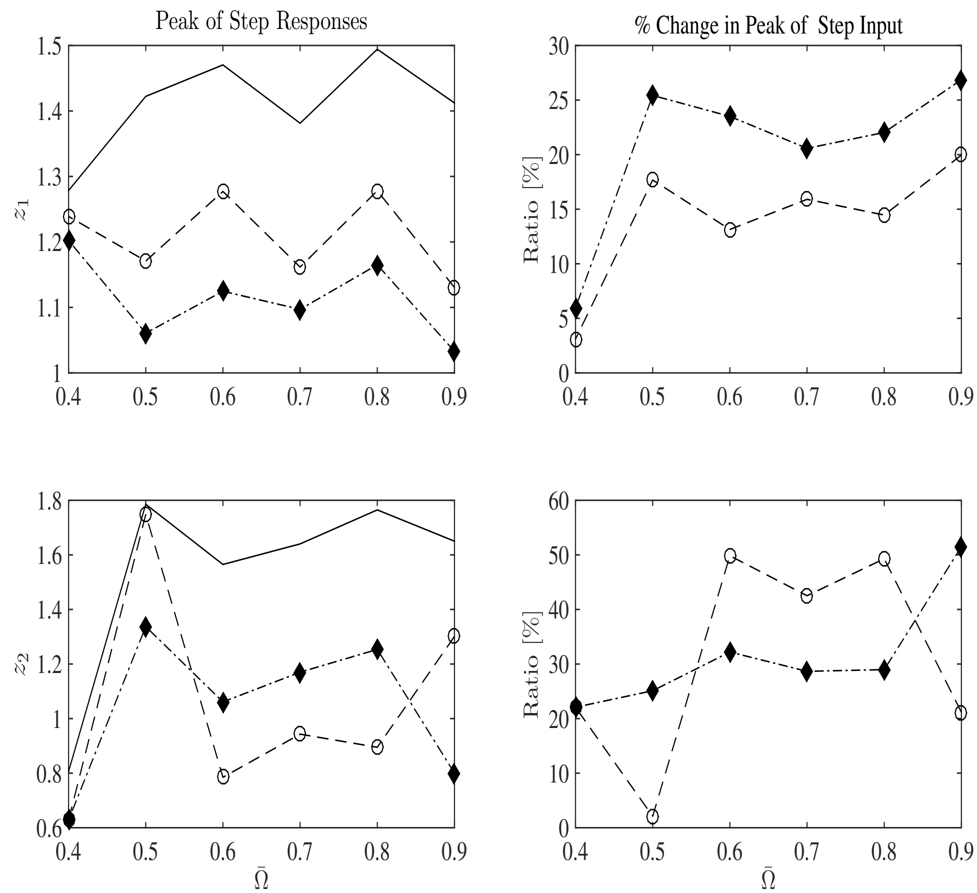

Figure 5.11: [Left] Peak of values of z_1, z_2 to step input of w. [Right] Percent change in peak in % w.r.t identified model. The top and bottom row correspond to z_1 and z_2 respectively. Legend: Identified plant (-), initial K (○) and K with output scaling (♦).

5.5 Summary

In this example, the process for identification of control-relevant multi-input multi-output (MIMO) generalized plant models of the closed-loop gas turbine system, is demonstrated. First a set of models is identified using data from a high fidelity simulation of the closed-loop gas turbine and control system at a series of operating points. A nominal dynamic model is selected from the set of identified models to serve as the design model. The variation of the system gain due to nonlinear properties of the closed-loop system is accommodated with an operating point dependent static scaling is proposed. When combined, the procedure offers a path to leave all or a subset of the existing control system in place while also facilitating design and performance analysis of an outer control loop that is able to leverage the benefits of modern multivariable control design. The procedure is demonstrated on data acquired from a first principles simulation of a gas turbine operating in closed-loop.

The identification of a standard generalized plant works well to capture the time and frequency characteristics of the closed-loop turbine plant and existing controller. A sixth order closed-loop plant model is demonstrated to be sufficiently complex to capture the dynamics of the non-linear closed-loop system to enable control design. The method enables a controller to be designed at each load point to improves the transient response of the air and fuel control to reject step load disturbances. The scaling of the input and output does allow a family of controllers to be designed at each design point, however, there is no guaranteed stability or similar performance in neighboring operating points. It many cases, a stable control policy designed for a plant realization for one operating point will prove unstable in neighboring points. The current formulation does not take advantage of knowledge of all the plant models at the various operating points, nor does it contain a method to describe the differences in the family of identified plants in a manner that enables robust control design.

For future work, a gain scheduling control such as a linear parameter varying gain $\mathcal{H}_\infty$ approach using a system of linear matrix inequalities may prove suitable to the task [60]. Regardless, creating the bridge between system identification and robust control design, has and remains a topic for future discussion.

5.6 Acknowledgements

This chapter, in part, is a reprint of “Subspace Identification for Disturbance Rejection Control Design in Gas Turbines” that has been accepted to the *19th European Control Conference*, Linz, AT. 2015. Chad M. Holcomb, Raymond A. de Callafon. The dissertation author was the primary investigator and author of this material.

6

Conclusions and Future Work

In this dissertation, system identification methods are applied to characterize the dynamic system behavior for use in diagnostics, condition monitoring, and control design of industrial gas turbines. In each of the scenarios, the target system is considered in a closed-loop setting.

The identification of contamination of fuel control valves is formulated as a prediction error minimization problem of closed-loop Hammerstein system. A new decoupled iterative gradient descent minimization is proposed based on the Levenberg-Marquardt formulation. The method successfully replicated the relationship between the fuel flow and shaft speed in a first principles simulation demonstration. The excitation requirements to span the entire range of the nonlinearity to be identified are significantly higher than what is normally achievable in practice. The use of Gaussian radial basis functions is sensitive to the choice of grid locations, particularly at the boundaries. A basis with compact support the choice of more robust numerical minimization may improve the performance in the future.

A modification to existing subspace system identification methods is formulated to permit their application to multiple concatenated, but non-contiguous data sets, which jointly can be sufficiently rich for identification while individual records are not. This is a break from traditional system identification, where one uses a single contiguous synchronous input-output record. In addition, a set of simple sufficient conditions for identifiability are derived. In tandem, these developments increase the intrinsic value of archival data as it enables the construction of a sufficiently informative data set for dynamic modeling. The identifiability conditions are derived in the noise-free data case. A treatment of the consistency of the proposed SSI algorithms in the presence of input and measurement noise is needed to complete the discussion.

The generalized plant model is introduced to model the gas-turbine system to link the identification problem formulation with the control design problem. Though, the system identification and control design consistently worked for a single design choice, the treatment of the parameter varying nature of the problem should be investigated using a formulation that uses all the available plant models for the control-synthesis rather than the somewhat ad-hoc search that was employed in this discussion. Regardless, the use of the generalized plant in the literature for control design in turbo-machinery is largely absent from the literature.

Appendix A

The Gory Details of Subspace Identification

Subspace methods are the primary tool of the identification for MIMO systems and there are a well-established family of algorithms exist for obtaining MIMO state-space system realizations. The literature focused a great deal of attention on the data-projections and novel modifications to recover estimates to the signal subspace through the extended observability matrix to obtain estimates of the coefficients of A and C. The remaining task of obtaining the estimates of B, D, and initial conditions is typically a statement that they may be obtained through a least squares procedure and direct the reader to a previous work, often to McKelvey [49]. In the segmented data treatment, proper handling of the initial conditions for each of the segments when solving for the B and D matrices is central to obtaining an unbiased estimate and one of the primary contributions of this dissertation.

A.1 System Model

We again consider a discrete time nth-order, n_u-input n_y-output linear time invariant system model with an allowance for noise on both the state w_t and output z_t as

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t \\ y_t &= Cx_t + Du_t + v_t, \end{aligned} \tag{A.1}$$

where the input $u_t \in \mathbb{R}^{n_u}$, output $y_t \in \mathbb{R}^{n_y}$, and state $x_t \in \mathbb{R}^n$. Additionally noise may be present on the state $w_t \in \mathbb{R}^n$ and output $v_t \in \mathbb{R}^{n_y}$. where the input noise and measurement noise.

A.2 Signal Recursions

If the data is available from time $t = 0$, the evolution of the state x_t and output y_t from (A.1) may be rewritten to include the initial condition x_0, and illustrate how state and measurement noise propagate through to the output.

$$x_t = \sum_{\tau=0}^{t-1} A^{t-\tau-1} B u_\tau + \sum_{\tau=0}^{t-1} A^{t-\tau-1} w_\tau + A^t x_0 \tag{A.2}$$

$$y_t = \sum_{\tau=0}^{t-1} C A^{t-\tau-1} B u_\tau + D u_t + C A^t x_0 + \eta_t \tag{A.3}$$

$$\eta_t = \sum_{\tau=0}^{t-1} C A^{t-\tau-1} w_\tau + v_t \tag{A.4}$$

We define η_t in (A.4) as the combined contribution of the state and measurement noise terms and assume it is zero-mean and independent for η_t.

It is unreasonable to assume the a data set from equipment or a process in operation will always contain measurements available from $t = 0$ with information available for the initial conditions. In order to understand the connection between the subspace identification formulation, it is important to understand the propagation of input and noise through the system dynamics over time the state (A.2) and the output (A.3). With no loss of generality, take $D = 0$ and in this case (A.1) reduces to

$$\begin{aligned} x_{t+1} &= A x_t + B u_t + w_t \\ y_t &= C x_t + v_t, \end{aligned} \tag{A.5}$$

Taking s as the starting index of the data record such that $0 \leq s \leq t$, the output $y_t = y_{t|s}$ may be written in $t - s$ equations of the form:

$$y_{t|s} \triangleq \sum_{\tau=s}^{t-1} C A^{t-\tau-1} B u_\tau + D u_t + C A^{t-s} x_s + \eta_{t|s} \tag{A.6}$$

$$\eta_{t|s} \triangleq \sum_{\tau=s}^{t-1} C A^{t-\tau-1} w_\tau + v_t \tag{A.7}$$

We next note three distinct recursions that may be written from (A.6) and (A.7).

Fixed Point (Horizon) Recursion: For the first recursion, given a starting index s and stepping s forward in time, the measurement y_t may be written as:

$$y_t = \begin{cases} y_{t|s} & = \sum_{\tau=s}^{t-1} CA^{t-\tau-1}Bu_\tau + Du_t + CA^{t-s}x_s + \eta_{t|s} \\ y_{t|s+1} & = \sum_{\tau=s+1}^{t-1} CA^{t-\tau-1}Bu_\tau + Du_t + CA^{t-s-1}x_{s+1} + \eta_{t|s+1} \\ & \vdots \\ y_{t|t-1} & = CBu_{t-1} + Du_t + Cx_{t-1} + \eta_{t|t-1} \\ y_{t|t} & = Cx_t + Du_t + \eta_{t|t} \end{cases} \tag{A.8}$$

If s is chosen such that for every t in the data record $max(t-s+1) = l$, the expressions in (A.8) are equivalent to the elements along the anti-diagonals of the block-Hankel SSI data equation. In general there are l anti-diagonal recursions for each measurement y_t except near the end points of the data sequence.

Fixed Lag (Row) Recursion For the second recursion, consider the case where $t-s$ is held constant at Δ_{ts}. Given a measurement y_t, starting index s, and stepping both t and s forward in time gives $(t-s+1)$ recursions of the form:

$$\begin{aligned} y_{t|s} & = \sum_{\tau=s}^{t-1} CA^{t-\tau-1}Bu_\tau + Du_t + CA^{t-s}x_s + \eta_{t|s} \\ y_{t+1|s+1} & = \sum_{\tau=s+1}^{t} CA^{t-\tau}Bu_\tau + Du_{t+1} + CA^{t-s}x_{s+1} + \eta_{t+1|s+1} \\ & \vdots \\ y_{t+l|s+l} & = \sum_{\tau=s+l}^{t+l-1} CA^{t+l-\tau-1}Bu_\tau + Du_{t+l} + CA^{t-s}x_{s+l} + \eta_{t+l|s+l} \end{aligned} \tag{A.9}$$

With a starting time index $s = 0$ and the lag defined as $\Delta_{ts} = t - s$, the fixed lag recursion in (A.9) is equivalent to marching along the $\Delta_{ts} + 1$ block row of the block-Hankel SSI data equation, e.g. $\Delta_{ts} = 0$ is equivalent to the j equations along the first block row.

Fixed Initial Condition (Column) Recursion For the third recursion, given a measurement y_t, starting index $t = s$, and advancing in time from $y_s \rightarrow y_{s+l-1}$ and gives l recursions of the form for each index $\left\{ \; t = s \quad |0 \leq s \leq t-1 \; \right\}$:

$$\begin{aligned} y_{s|s} & = Cx_s + Du_s + \eta_{s|s} \\ y_{s+1|s} & = CBu_s + Du_{s+1} + CAx_s + \eta_{s+1|s} \\ & \vdots \\ y_{s+l-1|s} & = \sum_{\tau=s}^{s+l-2} CA^{s-l-\tau}Bu_\tau + Du_{s+l-1} + CA^{l-1}x_s + \eta_{s+l-1|s} \end{aligned} \tag{A.10}$$

With a starting time index $s = 0$ and the initial condition fixed at $t = s$, the recursion in (A.10) is equivalent to marching along the $s + 1$ block column of a block-Hankel SSI data equation.

Bibliography

[1] I. Advanced Engine Technology, "Assessment of the benefits of on board diagnostics and diagnostic monitoring when applied to pipeline engines," *Gas Machinery Journal*, no. 3rd Quarter, 2013.

[2] R. Isermann, "Fault diagnosis of machines via parameter estimation and knowledge processing, a tutorial paper," *Automatica*, vol. 29, no. 4, pp. 815–835, 1993.

[3] G. J. Orme and M. Venturini, "Property risk assessment for power plants: Methodology, validation and application," *Energy*, vol. 36, pp. 3189–3203, 2011.

[4] A. Paridon, M. Bacic, P. Ireland, C. Barnes, and L. Lewis, "Reduced order transient disc temperature models for online health monitoring," in *Proceedings of ASME Turbo Expo 2014*. Dusseldorf, GE: ASME, June 2014.

[5] M. Venturini and D. Therkorn, "Application of a statistical methodology for gas turbine degradation prognostics to Alstom field data," *J. Eng. Gas Turbines Power*, vol. 135, no. 9, pp. 091 603–091 603, Aug. 2013.

[6] M. Venturini and N. Puggina, "Prediction reliability of a statistical methodology for gas turbine prognostics," *J. Eng. Gas Turbines Power*, vol. 134, no. 10, pp. 101 601–101 601, 2012.

[7] T. Breikin, G. Kulikov, V. Arkov, and P. Fleming, "Dynamic modelling for condition monitoring of gas turbines: Genetic algorithms approach," in *Proc. 16th IFAC World Congress*, 2005.

[8] Y. G. Li, "A gas turbine diagnostic approach with transient measurements," *Proc. of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, vol. 217, no. 2, pp. 169–177, 2003.

[9] G. L. Merrington, "Fault diagnosis of gas turbine engines from transient data," *J. Eng. Gas Turbines Power*, vol. 111, no. 2, pp. 237–243, Apr. 1989.

[10] D. Scheianu, "Methods and results in remote monitoring and diagnosing a fleet of industrial gas turbines," in *ASME Turbo Expo 2014: Turbine Technical Conference and Exposition*, vol. 6, Dusseldorf, GE, June 2014, p. V006T06A014.

[11] A. J. Volponi, H. DePold, R. Ganguli, and C. Daguang, "The use of kalman filter and neural network methodologies in gas turbine performance diagnostics: A comparative study," *J. Eng. Gas Turbines Power*, vol. 125, no. 4, pp. 917–924, Nov. 2003.

[12] L. Ljung, *System Identification: Theory for the User.* Upper Saddle River, New Jersey: Prentice-Hall, Inc, 1998.

[13] M. J. Moore, "Nox emission control in gas turbines for combined cycle gas turbine plant," *Proc. of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, vol. 211, no. 1, pp. 43–52, 1997.

[14] B. Ummels, M. Gibescu, E. Pelgrum, W. Kling, and A. Brand, "Impacts of wind power on thermal generation unit commitment and dispatch," *IEEE Transactions on Energy Conversion*, vol. 22, no. 1, pp. 44–51, 2007.

[15] K. Zhou and J. C. Doyle, *Essentials of robust control.* Prentice hall Upper Saddle River, NJ, 1998, vol. 104.

[16] G. G. Kulikov and H. A. Thompson, *Dynamic Modelling of Gas Turbines: Identification, Simulation, Condition Monitoring and Optimal Control.* Springer, 2004.

[17] V. Arkov, C. Evans, P. Fleming, D. Hill, J. Norton, I. Pratt, D. Rees, and K. Rodrguez-Vzquez, "System identification strategies applied to aircraft gas turbine engines," *Annual Reviews in Control*, vol. 24, pp. 67–81, 2000.

[18] X. Dai, T. Breikin, and H. Wang, "An algorithm for identification of reduced-order dynamic models of gas turbines," in *First International Conference on Innovative Computing, Information and Control*, 2006, pp. 134–137.

[19] N. S. Nise, *Control Systems Engineering.* John Wiley & Sons, 2007.

[20] E.-W. Bai, "Identification of linear systems with hard input nonlinearities of known structure," in *Block-oriented Nonlinear System Identification*, ser. Lecture Notes in Control and Information Sciences, F. Giri and E.-W. Bai, Eds. Springer London, Jan. 2010, no. 404.

[21] Y. Han and R. A. de Callafon, "Closed-loop identification of Hammerstein systems using iterative instrumental variables," in *Proc. 18th IFAC World Congress*, 2011, pp. 13 930–13 935.

[22] D. Saez, A. Cipriano, and A. Ordys, *Optimization of Industrial Processes at Supervisory Level - Application to Control of Thermal*, ser. Advances in Industrial Control. Springer London, 2002.

[23] H. I. Saravanamuttoo, G. F. C. Rogers, and H. Cohen, *Gas turbine theory*, 6th ed. Harlow [u.a.]: Prentice Hall, 2009.

[24] F. A. Williams, *Combustion Theory: The Fundamental Theory of Chemically Reacting Flow Systems.* Westview Press, 1985.

[25] A. Lefebvre, *Gas Turbine Combustion.* Washington: Taylor & Francis, Jan. 1983.

[26] K. Narendra and P. Gallman, "An iterative method for the identification of nonlinear systems using a hammerstein model," *IEEE Transactions on Automatic Control*, vol. 11, no. 3, pp. 546–550, 1966.

[27] Y. Liu and E.-W. Bai, "Iterative identification of Hammerstein systems," *Automatica*, vol. 43, no. 2, pp. 346 – 354, 2007.

[28] E.-W. Bai, "An optimal two stage identification algorithm for Hammerstein-wiener nonlinear systems," in *American Control Conference, 1998. Proceedings of the 1998*, vol. 5, 1998, pp. 2756 – 2760.

[29] S. Billings and S. Fakhouri, "Identification of a class of nonlinear systems using correlation analysis," *Proceedings of the Institution of Electrical Engineers*, vol. 125, no. 7, pp. 691–697, July 1978.

[30] V. Laurain, M. Gilson, and H. Garnier, "Refined instrumental variable methods for identifying hammerstein models operating in closed loop," in *Proc. of the 48th IEEE Conference on Decision and Control*, Shanghai, China, 2009, pp. 3614 –3619.

[31] J. W. Wingerden and M. Verhaegen, "Closed loop identification of MIMO hammerstein models using LS-SVM," in *Proc. 16th IFAC World Congress*, W. Eric, Ed., July 2009, pp. 1650–1655.

[32] T. Van Pelt and D. Bernstein, "Nonlinear system identification using Hammerstein and nonlinear feedback models with piecewise linear static maps - part 1: Theory," in *Proc. of the American Control Conference*, vol. 1, Chicago, Illinois, 2000, pp. 225–229.

[33] W. C. Gibson, "The method of moments," in *The Method of Moments in Electromagnetics.* Chapman and Hall/CRC, 2007, pp. 33–62.

[34] R. Lippmann, "A critical overview of neural network pattern classifiers," in *IEEE Neural Networks for Signal Processing Workshop*, 1991, pp. 266–275.

[35] Y. Liu and B. D. O. Anderson, "Singular perturbation approximation of balanced systems," in *, Proceedings of the 28th IEEE Conference on Decision and Control, 1989*, 1989, pp. 1355–1360 vol.2.

[36] P. Gill and W. Murray, "Algorithms for the solution of the nonlinear least-squares problem," *SIAM Journal on Numerical Analysis*, vol. 5, p. 977?992, 1978.

[37] P. Cézac, J.-P. Serin, J.-M. Reneaume, J. Mercadier, and G. Mouton, "Elemental sulphur deposition in natural gas transmission and distribution networks," *The Journal of Supercritical Fluids*, vol. 44, no. 2, pp. 115–122, 2008.

[38] C. Zhang and R. R. Bitmead, "Subspace system identification for training-based MIMO channel estimation," *Automatica*, vol. 41, no. 9, pp. 1623–1632, 2005.

[39] I. Markovsky, "Approximate system identification with missing data," in *2013 IEEE 52nd Annual Conference on Decision and Control (CDC)*, Dec. 2013, pp. 156–161.

[40] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. M. De Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325–329, Apr. 2005.

[41] P. Overschee and B. Moore, *Subspace Identification for Linear Systems Theory - Implementation - Applications.* Boston, MA: Springer US, 1996.

[42] T. Kailath, *Linear Systems*, 1st ed. Prentice Hall, 1980.

[43] E. D. Sontag, "On the length of inputs necessary in order to identify a deterministic linear system," *IEEE Transactions on Automatic Control*, vol. 25, no. 1, pp. 120–121, 1980.

[44] M. Viberg, B. Wahlberg, and B. Ottersten, "Analysis of state space system identification methods based on instrumental variables and subspace fitting," *Automatica*, vol. 33, no. 9, pp. 1603–1616, 1997.

[45] M. Verhaegen, "Identification of the deterministic part of MIMO state space models given in innovations form from input-output data," *Automatica*, vol. 30, no. 1, pp. 61–74, 1994.

[46] P. Van Overschee and B. De Moor, "N4sid: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.

[47] W. Larimore, "Canonical variate analysis in identification, filtering, and adaptive control," in *Proceedings of the 29th IEEE Conference on Decision and Control, 1990*, 1990.

[48] M. Viberg, "Subspace-based methods for the identification of linear time-invariant systems," *Automatica*, vol. 31, no. 12, pp. 1835–1851, 1995.

[49] T. McKelvey, H. Akcay, and L. Ljung, "Subspace-based multivariable system identification from frequency response data," *IEEE Transactions on Automatic Control*, vol. 41, no. 7, pp. 960–979, 1996.

[50] I. Markovsky, J. Willems, S. Van Huffel, and B. De Moor, *Exact and Approximate Modeling of Linear Systems*, ser. Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics, Jan. 2006.

[51] P. Carrette, G. Bastin, Y. Genin, and M. Gevers, "Discarding data may help in system identification," *Trans. Sig. Proc.*, vol. 44, no. 9, pp. 2300–2310, 1996.

[52] D. N. Miller, R. A. De Callafon, and M. J. Brenner, "Covariance-based realization algorithm for the identification of aeroelastic dynamics," *Journal of Guidance, Control, and Dynamics*, vol. 35, no. 4, pp. 1169–1177, 2012.

[53] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems," *Automatic Control, IEEE Transactions on*, vol. 34, no. 8, pp. 831–847, 1989.

[54] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*, ser. Feher/Prentice Hall Digital and. Prentice Hall, 1996.

[55] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*, ser. Multivariable Feedback Control: Analysis and Design. Wiley, 2005.

[56] B. Anderson and J. Moore, *Optimal Control: Linear Quadratic Methods*, ser. Dover Books on Engineering. Dover Publications, 2007.

[57] ——, *Optimal Filtering*, ser. Dover Books on Electrical Engineering. Dover Publications, 2012.

[58] H. Shu and T. Chen, "State-space approach to discrete-time $\mathcal{H}_2$-optimal control with a causality constraint," *Systems & Control Letters*, vol. 26, no. 1, pp. 69 – 77, 1995.

[59] A. Saberi, P. Sannuti, and B. Chen, $\mathcal{H}_2$ *Optimal Control*, ser. Prentice-Hall international series in systems and control engineering. Prentice Hall, 1995.

[60] P. Apkarian, P. Gahinet, and G. Becker, “Self-scheduled $\mathcal{H}_\infty$ control of linear parameter-varying systems: a design example,” *Automatica*, vol. 31, no. 9, pp. 1251–1261, Sept. 1995.