

© Copyright by Shyam Panjwani 2016

All Rights Reserved

Data Driven Approaches to Model Building: Applications to Energy Industries

A Dissertation

Presented to

the Faculty of the Department of Chemical and Biomolecular Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

In Chemical Engineering

by

Shyam Panjwani

December 2016

Data Driven Approaches to Model Building: Applications to Energy Industries

Shyam Panjwani

Approved:

Chair of the Committee
Michael Nikolaou, Professor,
Chemical and Biomolecular Engineering

Committee Members:

Lars C. Grabow, Assistant Professor,
Chemical and Biomolecular Engineering

Jeremy Palmer, Assistant Professor,
Chemical and Biomolecular Engineering

Guan Qin, Associate Professor,
Petroleum Engineering

Mark L. Darby, CMiD Solutions

Suresh K. Khator, Associate Dean,
Cullen College of Engineering

Michael P. Harold, Professor and
Chair, Chemical and Biomolecular
Engineering

Acknowledgements

It would not have been possible to achieve this milestone without the help of all those who were the part of my journey from being an undergraduate student to a PhD candidate.

Foremost, I would like to express my gratitude for my advisor Dr. Michael Nikolaou for his continuous support, guidance and motivation. I was able to successfully handle difficult times in my PhD research only because of his constant help.

I would also like to thank other committee members Dr. Guan Qin, Dr. Lars Grabow, Dr. Jeremy Palmer and Dr. Mark Darby for serving in my committee. I would like to specially thank Dr. Tom Holley and Dr. Guan Qin for their valuable suggestions during the early phase of my PhD research.

I would like to thank industrial collaborators CSI Technology, Chesapeake Energy and Research Partnership to Secure Energy for America (RPSEA) for funding support. I have a special thanks for Jessica McDaniel for her valuable inputs.

I would like to thank my friends and lab-mates Kyle MacFarlane, Shobhit Mishra, and Sashank Kasiraju for helping me out during rough patches in the last four years of my life.

Lastly but most importantly, I have special regards to my family members for their love and support.

Data Driven Approaches to Model Building: Applications to Energy Industries

An Abstract

of a

Dissertation

Presented to

the Faculty of the Department of Chemical and Biomolecular Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

In Chemical Engineering

by

Shyam Panjwani

December 2016

Abstract

George Box's famous quote "All models are wrong, but some are useful" is now widely known. Mathematical models can be built based on a combination of first principles and available data. The focus of this work is on the application of data-driven modelling approaches in two specific instances of problems in upstream (oil & gas extraction) and downstream (refining & chemicals) industries, namely (a) cementing of wells drilled for production of oil and gas from unconventional resources, such as shales; and (b) design of robust control-relevant models for oil refineries and chemical plants.

Shale gas production from horizontal wells faces potential problems related to gas leakage from the cemented annulus of the well into the air and water reserves, with obvious environmental and productivity implications. Whether a well will leak or not depends on several factors, related to cement composition and preparation, the cementing process, well conditions, and others. A model would be useful in assessing ahead of time whether a cementing job will produce a non-leaking well or not. Such a model could be based on first principles, but would be extremely complicated. Alternatively, as done in this work, a model can be built using multivariate statistics and available data from several leaking and non-leaking wells, cemented under different enough scenarios. The model built has 35 input variables (in the broad categories of casing properties, cement and drilling mud properties, and operating conditions) and manages to correctly classify with confidence 81% of wells as leaking or non-leaking in cross-validation tests.

An advanced control system relies on a good control-relevant model that is not merely a good approximation of the actual process under control but also satisfies additional properties necessary for controller design. Control-relevant models are typically identified through industrial experiments whose design is considerably more involved than standard design for parameter estimation. The focus of this study is how to design control-relevant identification experiments when elements of the model are already known. A new theoretical framework is developed and its significant advantages over standard methods are illustrated through numerical simulations. Several possibilities for future development are suggested.

Table of Contents

Acknowledgements	v
Abstract	vii
Table of Contents	ix
List of Figures	xii
List of Tables	xvii
1 Introduction	1
1.1 Cementing shale gas wells.....	2
1.2 Designing experiments for control-relevant model identification.....	8
1.3 Organization of this Dissertation.....	12
2 Cementing Wells: Background Literature	14
2.1 Cement sheath quality	14
2.2 Cement sheath evaluation.....	19
3 Cementing Wells: A Modeling Study	22
3.1 Regression Models	22
3.2 Classification models.....	23
3.3 Modeling study	26
3.4 Results and Discussion	31
3.5 Summary.....	40
4 Ensuring Integral Controllability for Robust Multivariable Control ...	42

4.1	Introduction	42
4.2	Background.....	44
4.3	Main Results	49
4.4	Summary.....	66
5	Experiment Design for Control-Relevant Identification.....	68
5.1	Introduction	68
5.2	Background.....	71
5.3	Mathematical Problem Formulation and Main Results	74
5.4	Case Studies.....	88
5.5	Summary.....	104
6	Conclusion and Future Work	106
6.1	Cementing shale gas wells.....	106
6.2	Ensuring integral controllability	107
6.3	Experiment design for control-relevant identification of partially known stable multivariable systems	108
	References	110
	Appendix A	120
A.1	Classification using PLS-DA.....	120
A.2	Reliability of classification	122
A.3	Interpretation of PLS-DA weights	123

Appendix B	125
B.1 Proof of eqn. (4.10).....	125
Appendix C	127
List of Publications	138

List of Figures

Figure 1-1: Primary U.S energy production (Quadrillion BTU)	2
Figure 1-2: Use of natural gas in U.S in 2014	3
Figure 1-3: Geographical locations of shale basins in the US	3
Figure 1-4: Projections of dry natural gas production in the US	4
Figure 1-5: Potential leakage of hydrocarbons	5
Figure 1-6: Possibility for leakage of hydrocarbons into underground water or into the air. Gas migration can be prevented by proper cementing of the annular space between metal casing and the wall of the well.....	6
Figure 2-1: Standoff in the horizontal section of a well	15
Figure 2-2: Sustained casing pressure (SCP).....	21
Figure 3-1: Multistep PLSDA.....	25
Figure 3-2: VIP values of selected dimensionless variables for model-6.....	33
Figure 3-3: Latent variable weights for dimensionless inputs in the model-6.....	37
Figure 3-4: Latent variable x-scores for model-6	40
Figure 4-1: Visualization of the functions $\left \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1}) \right $ and $\left \frac{\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})}{2 + \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})} \right $ in a 3D plot (left) and contour plot (right). Note that violation of $\left \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1}) \right < 1$ for some $\mathbf{D} \in D$ does not imply violation of $\left \frac{\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})}{2 + \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})} \right < 1$ for some $\mathbf{D} \in D$	48
Figure 4-2: Input and output data in simulated identification experiment for \mathbf{G}_1 using PRBS inputs.....	53

Figure 4-3: Estimates and uncertainty regions resulting from least-squares estimation of the entries of \mathbf{G}_1 using the experimental data shown in Figure 4-2. All straight lines go through the origin..... 53

Figure 4-4: Graphical representation of IC satisfaction by $\hat{\mathbf{G}}_1$ and 630 randomly chosen $\mathbf{G}_1 \in U$ shown in Figure 4-3. For each vector, the point of origin corresponds to (g_{11}, g_{12}) and the head corresponds to (g_{21}, g_{22}) . Green and purple vectors correspond to pairs of $\hat{\mathbf{G}}_1$ and \mathbf{G}_1 for which eqn. (4.1) is or is not satisfied, respectively..... 54

Figure 4-5: Input and output data in simulated identification experiment for \mathbf{G}_1 using rotated PRBS inputs..... 55

Figure 4-6: Estimates and uncertainty regions resulting from least-squares estimation of the entries of \mathbf{G}_1 using the experimental data shown in Figure 4-5. All straight lines go through the origin..... 55

Figure 4-7: Percentage of IC violations varying with J for 2×2 ill-conditioned distillation column 56

Figure 4-8: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for a 2×2 ill-conditioned distillation column..... 57

Figure 4-9: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using rotated PRBS inputs for 2×2 ill-conditioned distillation column..... 58

Figure 4-10: Percentage of IC violation varying with J for 2×2 well-conditioned system 59

Figure 4-11: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for 2×2 well-conditioned distillation column	59
Figure 4-12: Percentage of IC violation varying with J for 3×3 distillation column....	60
Figure 4-13: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs 3×3 distillation column.	60
Figure 4-14: Percentage of IC violation varying with J for 4×4 distillation column with side stripper	61
Figure 4-15: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for 4×4 distillation column with side stripper.	62
Figure 4-16: Percentage of IC violation varying with J for 5×5 fluidized catalytic cracker (FCC)	63
Figure 4-17: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for 5×5 fluidized catalytic cracker (FCC).....	63
Figure 5-1. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D1	93
Figure 5-2. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D2.	94
Figure 5-3. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D3.	95
Figure 5-4. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D4.	96

Figure 5-5. Adaptive DOE for 5×5 FCC unit: Identification time required for IC satisfaction when inputs are produced by designs D1-D4. The vertical axis is J_c , eqn. (5.11), for designs D1 and D3, and J_u , eqn. (5.7), for designs D2 and D4.	99
Figure 5-6. Convergence of SSGM elements of 5×5 FCC unit for designs D1-D4. ...	100
Figure 5-7. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D1.	101
Figure 5-8. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D2.	102
Figure 5-9. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D3.	102
Figure 5-10. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D4.	102
Figure 5-11. Adaptive DOE for 2×2 two-stage absorber: Identification time required for IC satisfaction when inputs are produced by designs D1-D4. The vertical axis is J_c , eqn. (5.11), for designs D1 and D3, and J_u , eqn. (5.7), for designs D2 and D4.	104
Figure 5-12. Convergence of SSGM elements of 2×2 two-stage absorber for designs D1-D4.	104
Figure A-1: Posterior probabilities for two categories: category-1 and category-2	121
Figure A-2: Probability density functions for two categories: category -1 and category -2	122

Figure C-1 Satisfaction of the inequalities in eqn. (C.5), required by the Jury stability criterion for closed-loop stability, when $\hat{\mathbf{G}}_1$ (left) and $\hat{\mathbf{G}}_2$ (right) is used in the feedback controller, eqn. (C.1)..... 128

List of Tables

Table 3-1: Physical variables and respective ranges.....	28
Table 3-2: Affecting dimensionless variables for intermediate casing.....	29
Table 3-3: Intermediate casing PLS-DA model cross-validation results for different VIP cut-off values	32
Table 3-4: Selected dimensionless variables and their positive or negative effect on well leakage for model-6 in Table 3-3.....	37
Table 4-1: Parameter values for randomized simulations.....	51
Table 4-2: Parameter values for randomized simulations.....	51
Table 5-1: Case1: Parameters used in simulation for adaptive designs.....	92
Table 5-2. Characterization of inputs and outputs for designs D1-D4 for 5×5 FCC unit; Active constraints are in bold.	97
Table 5-3. Input and output correlations matrices for Designs D1-D4 for 5×5 FCC unit.	97
Table 5-4. Case 2: Parameters used in simulation for adaptive designs.....	101
Table 5-5. Results for designs D1-D4 for 2×2 two-stage absorber. Active constraints are in bold.	103

1 Introduction

Data-driven modeling has emerged as a new field of research in the last couple of decades. Owing to advancements in digital data storage and high speed computing, data-driven modeling is finding applications in a variety of industries. Exploration and production (E&P) and chemical process industries are also taking interest in data-driven modeling to address related problems. Seismic data analysis, reservoir surrogate modeling and injection-production optimization are major research areas in upstream E&P industries where data-driven modeling is being applied (Denney; Esmaili & Mohaghegh, 2016; Holdaway & Laing; Maucec et al., Mohaghegh, 2016; Stephen et al., Zhao et al., 2016). Fault diagnosis and control-relevant model identification are major application areas of data-driven modeling in oil refineries and chemical process industries (Kadlec et al., 2009; Rasmussen & Bay Jørgensen, 2005; Shu et al., 2016; Tabora, 2012; Xiong et al., 2014; Yamuna Rani & Patwardhan, 2007). The focus of this work is on the application of data-driven modelling approaches in two specific instances of problems in upstream (oil & gas extraction) and downstream (refining & chemicals) industries, namely (a) cementing of wells drilled for production of oil and gas from unconventional resources, such as shales; and (b) design of robust control-relevant models for oil refineries and chemical plants. In the subsequent sections, motivation behind this research is explained.

1.1 Cementing shale gas wells

As shown in Figure 1-1, oil and gas production from unconventional reservoirs has transformed the US energy landscape with profound economic and environmental implications (DOE, August 2016). Indeed, the US chemical industry now has the advantage of inexpensive natural gas as a raw material produced domestically; and the power industry is increasingly relying on natural gas, as an alternative to coal, for electricity generation with reduced greenhouse gas (GHG) and other emissions, higher flexibility, and lower cost (Figure 1-2) (Biello, 2015).

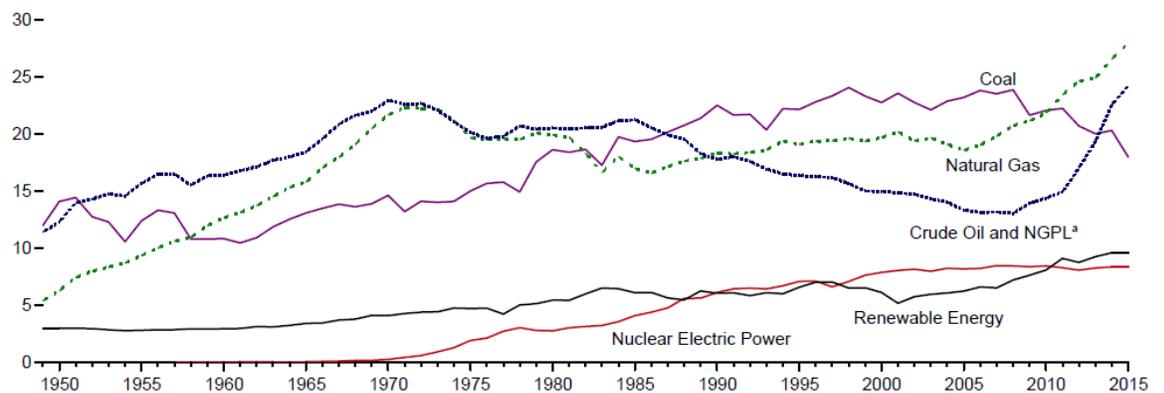
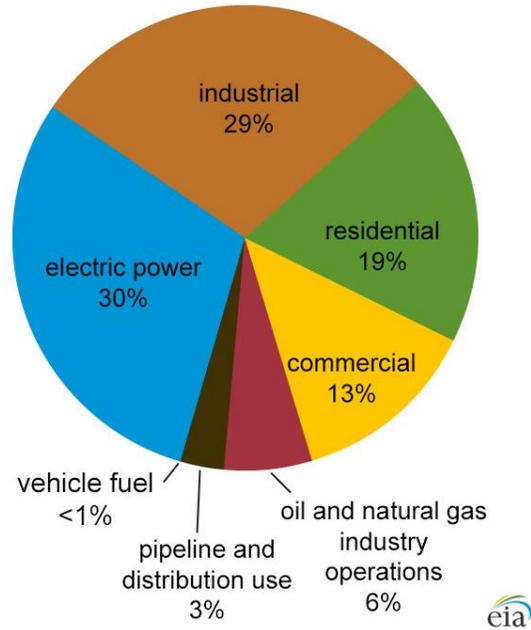


Figure 1-1: Primary U.S energy production (Quadrillion BTU)

Natural gas from shale formations known as shale gas, is playing a key role in natural gas energy sector. Figure 1-3 shows the presence of shale basins in lower 48 states of the United states (DOE, 2009). As shown in Figure 1-4, the share of natural gas from shale formations has increased from approximately 1% in early 2000 to more than 20 % in 2013 and is expected to increase to approximately 55 % by 2040 (DOE, April 2015).



Source: U.S. Energy Information Administration, *Natural Gas Monthly* (February 2015), preliminary data

Figure 1-2: Use of natural gas in U.S in 2014

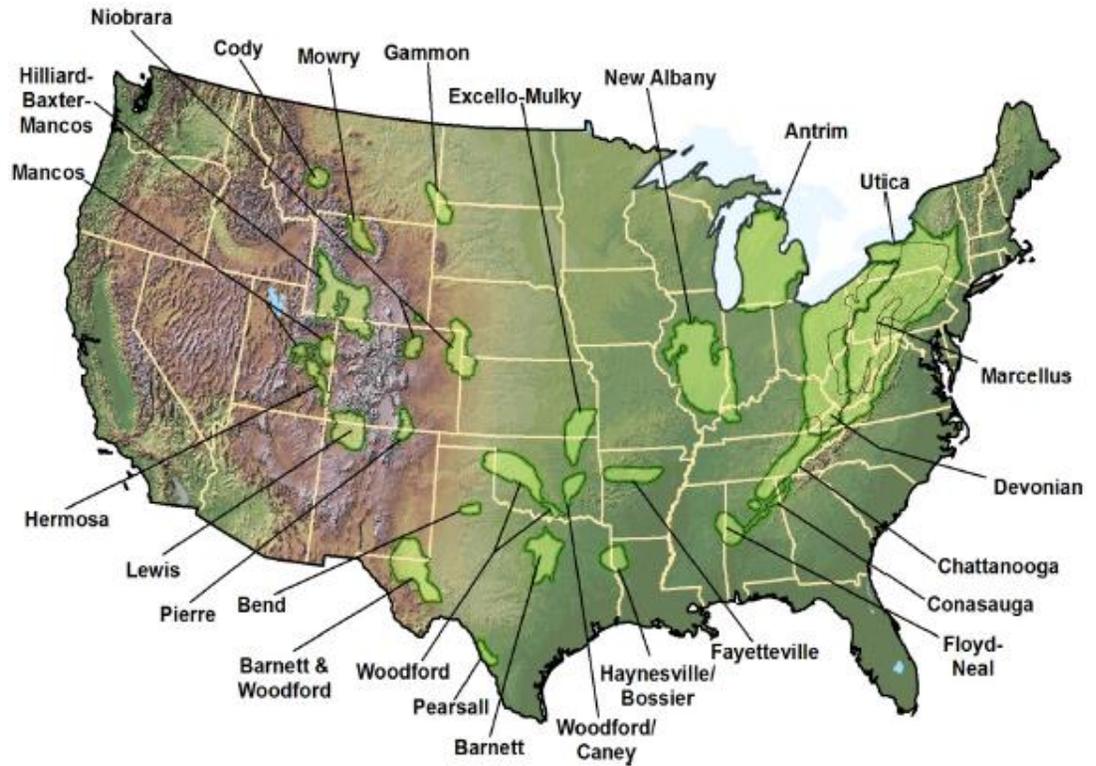
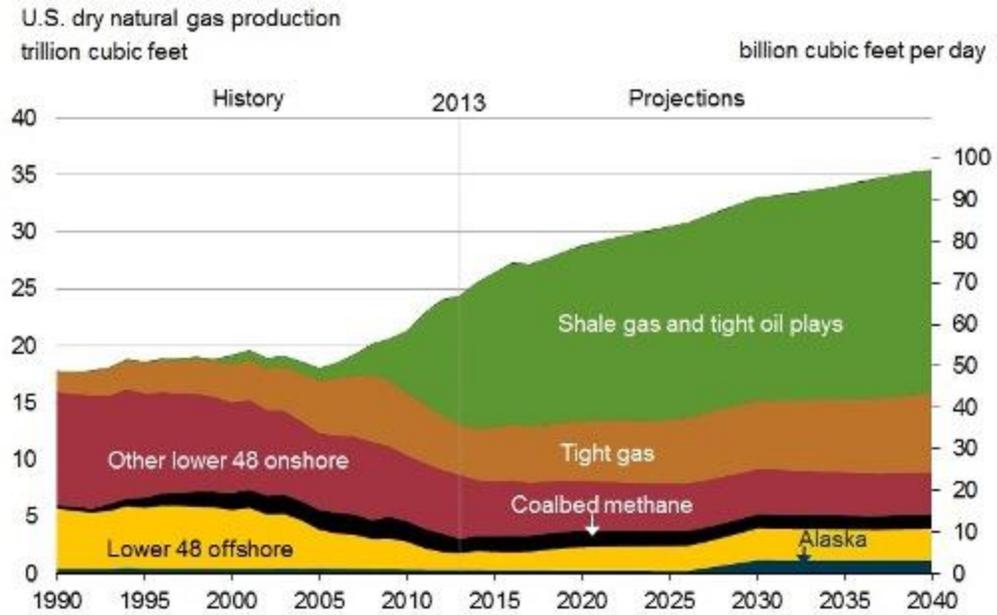


Figure 1-3: Geographical locations of shale basins in the US



Source: EIA, Annual Energy Outlook 2015 Reference case

Figure 1-4: Projections of dry natural gas production in the US

The extraction of large amounts of shale gas from almost impermeable rock formations became possible because of advancements in two key technologies: massive hydraulic fracturing and horizontal drilling (IEA, 2011). Massive hydraulic fracturing creates the flow paths needed for oil and gas to flow through an otherwise almost impermeable rock, within reasonable amount of time, to a production well and then to the surface. This task is assisted by horizontal drilling, which creates a large contact area between a very low-permeability rock formation and a production well (DOE, 2009). Although shale gas revolution has played a significant role in cutting down carbon-di-oxide emission but its production is not without environmental risks. An important consideration is potential gas leakage from various geological zones of a production well, either into the air or into underground water reserves. In fact, undesired gas emissions into the air – in addition to wasting a valuable resource – could potentially cancel any advantages on

greenhouse gas emissions reduction stemming from use of natural gas in power generation (Figure 1-5) (R. W. Howarth, D. Shindell, R. Santoro, A. Ingraffea, N. Phillips, and A. Townsend-Small., 2012), as methane has more than an order of magnitude higher global warming potential than CO₂ (IPCC, 2013). The acceptable threshold of natural gas leaks, above which total greenhouse gas emissions would actually increase, is vigorously debated (R. W. Howarth, Ingraffea, et al., 2011), on occasion among scientists within the same institution (Cathles et al., 2012; R. W. Howarth, Santoro, et al., 2011; R. W. Howarth et al., 2012), and has spawned high-profile scientific investigations (Caulton et al., 2014). Nevertheless, developing cost-effective solutions for unwanted natural gas leaks is an opportunity welcome by industry (Boling, 2015).

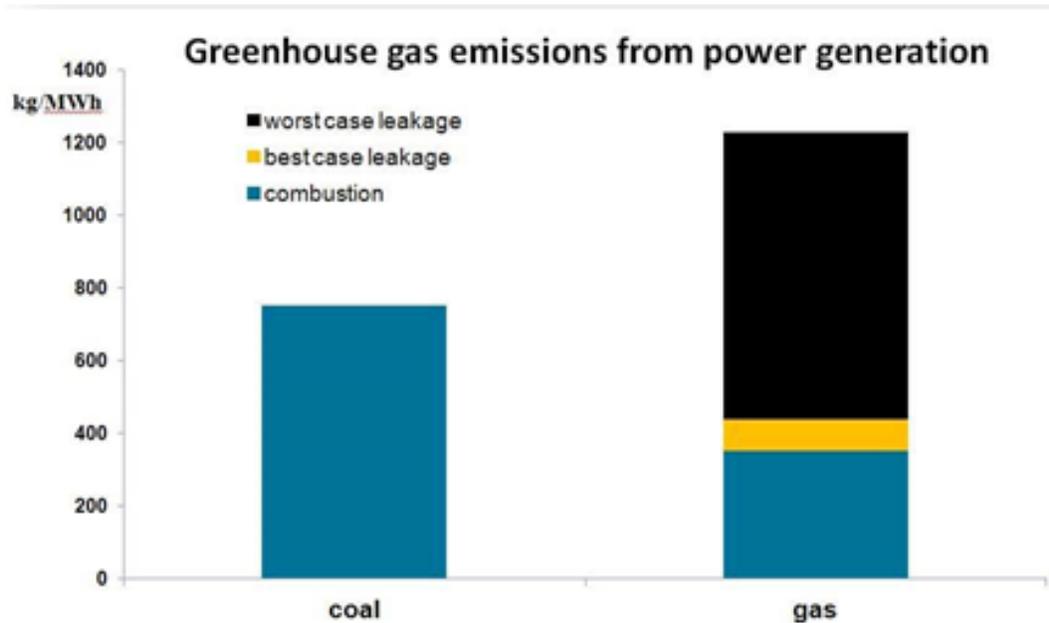


Figure 1-5: Potential leakage of hydrocarbons

Preventing the undesired flow of natural gas between various well zones and rock formations is known as zonal isolation, and is achieved by creating a tight seal through

cementing appropriate sections of a well (Figure 1-6). Providing adequate zonal isolation remains formidable for industry, because uniform placement of cement in the annulus between the metal casing and the wellbore formation is both technically challenging and costly. In current industrial practice, predicting zonal isolation mainly relies on laboratory investigations and/or predictions based on analytical models or finite-element analysis. Investigators have reported the effect of internal casing pressure and temperature on zonal isolation in experimental studies (Goodwin & Crook, 1992; Jackson & Murphey, 1993) and the effect of well events, such as completion and production, using finite-element methods (Tahmourpour & Griffith, 2004; Thiercelin et al., 1998). Correlation-based studies of the effect of formation properties on gas migration have also been presented (Wilkins & Free, 1989).

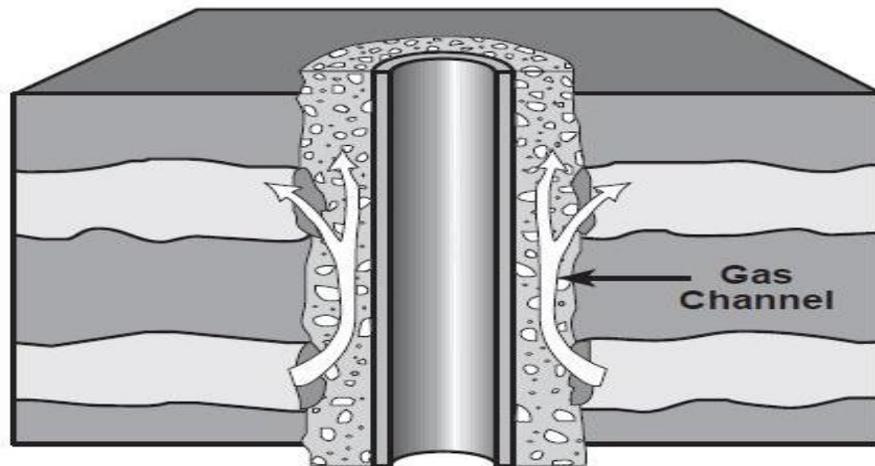


Figure 1-6: Possibility for leakage of hydrocarbons into underground water or into the air. Gas migration can be prevented by proper cementing of the annular space between metal casing and the wall of the well

Although such studies are quite useful for individual tasks related to well cementing and zonal isolation, no method has been presented to account for the combined effect of all

related factors on the quality of a cementing job. Yet such factors interact with one another, and their overall effect can be difficult to anticipate, either using fundamental equations (which become overly unwieldy) or by invoking experience (which, by itself, is insufficient). In particular, cementing operations and design for horizontal wells are generally more demanding than for vertical wells, and routinely require decisions on a large number of factors. Such decisions include cement-slurry design, spacer design, casing design, centralization, and evaluation of a cementing job after its completion, among others.

1.1.1 Problem statement

Assessing the effect of all such factors in a systematic way would be helpful for cementing engineers. In fact, a mathematical model that could make reasonable quantitative predictions would help the engineer perform what-if analysis and possibly optimize the cementing operation.

1.1.2 Proposed approach

Since the development of a mathematical model based on first principles would be extremely difficult in this case, because of the significant uncertainty and complexity in the fundamental equations governing the cementing process, a data-driven model could be built, provided that sufficient data is available.

A key focus of this work is the development of such a data-driven mathematical model. Input (predictor) variables to this model are factors that generally may have an effect on the quality of a cementing job. The output variable of this model is characterization of cementing quality of a well section, through classification of a section

as leaking or non-leaking, along with the relative confidence of such a prediction. The model is built using multivariate statistical analysis on data already available from a number of cemented wells; no first-principles equations are employed at all. The model can be used to guide the design of subsequent cementing jobs.

1.2 Designing experiments for control-relevant model identification

A good model is at the heart of good controller design. With uprisings applications of model predictive control in chemical industries, control-relevant model identification has become even more important. Identifying a right model structure requires right set of inputs. So experiment design has an important role to play in the model identification. Traditionally, experiments were designed to capture the effect of the input variable on the response variable by testing a large set of input variations. Generally, the objective of experiment design is to minimize errors in the estimated parameters so that the identified model is as close as possible to the actual process. Research in the field of experiment design with the only objective to minimize parameter covariance dates back to mid of 20th century (Federov, 1972; Kiefer, 1958). For multivariable systems, parameter covariance is not a scalar but a multidimensional matrix. Based on different scalar measures of the parameter covariance matrix, the following are vastly used experiment designs:

- D-optimal: minimization of determinant of the parameter covariance matrix
- L-optimal: minimization of weighted trace of the parameter covariance matrix (in case of unit weights, the design is called A-optimal)
- E-optimal: minimization of maximum eigenvalue of the parameter covariance matrix

Since finding covariance matrix may be difficult, inverse of the Fisher information matrix is often used as an approximation to the covariance matrix (Soderstrom & Stoica, 1989). The optimization problem is solved to find the optimal set of inputs. The optimal inputs obtained from the solution of the optimization problem may not be physically implementable. To avoid this situation, the objective function can be optimized subject to physical constraints on the inputs and the outputs (Gevers & Ljung, 1986).

The above mentioned and likewise designs only ensure the model closest to the actual process. However, control-relevant identification is slightly different as it requires certain control characteristics to be fulfilled as well (Gevers, 2005). One such control characteristic, is integral controllability (C. E. Garcia & M. Morari, 1985). Integral controllability (IC) is a desired property of multivariable models used in robust controller design. IC requires satisfaction of eigenvalue-based inequalities involving the real process and the inverse of the identified model. Since the IC satisfaction involves inversion of the identified model, identifying such models for ill-conditioned multivariable systems have been a challenge. A multivariable ill-conditioned system is characterized by a steady state gain matrix (SSGM) with very high condition number, which makes inversion of the SSGM of the identified model mathematically inaccurate. Although models for individual input-output pairs of a multivariable system can be identified via single-input testing. However researchers have shown that single-input testing does not produce good model particularly for ill-conditioned systems because of insufficient information in weak direction. This leaves no alternative other than MIMO testing and identification. MIMO testing and identification can be performed in open loop or closed loop manner.

Open loop testing is very common in industrial applications. Either one input at a time or multivariable random inputs can be used in open loop testing. A pseudo random binary sequence (PRBS) is persistently exciting input sequence and the most common type of input sequence used in open loop testing because of advantages associated with persistent excitation (Soderstrom & Stoica, 1989). A standard PRBS has binary states with zero mean and unit variance that can be approximated to a white noise signal. In a PRBS, minimum switching time, total sample time and upper/lower bounds can be varied to produce different input sequences. Open loop testing is useful except in those cases where unmeasured disturbances affect the process outputs. Closed loop testing and identification can be useful in those situations when keeping process outputs within desired range becomes a challenge. Under closed loop testing, a feedback controller is used to keep the process outputs within desired range. The challenge in closed loop testing is design of the feedback controller. The controller placed in closed loop testing is generally based on some preliminary model. The response of the controller depends on the goodness of the preliminary model. Since the purpose of the experiment design is to obtain a good model, dependency of the feedback controller design on the model quality makes this closed-loop approach iterative. If one can make the closed loop stable with the controller based on some good preliminary model, closed loop testing may lead to good model identification.

For identification of IC compliant model for ill-conditioned systems, (Koung & MacGregor, 1994) proposed a *rotated input design* for 2×2 systems based on principles of D-optimal design and provided extension of the *rotated input design* for $n \times n$ systems. The *rotated input design* uses singular value decomposition (SVD) of the gain matrix. Rotated inputs are set of correlated inputs that are obtained from rotation of PRBS inputs

of appropriate amplitudes. Since this design depends on the SVD of the unknown true gain matrix, desired inputs cannot be designed in one step. An iterative approach is required to design the inputs using this rotated input design. For identification of an IC compliant model, a number of other researchers have also established the usefulness of this rotated input design over a standard PRBS design (M L Darby & M. Nikolaou, 2014; Kulkarni, 2012; P. Misra & Nikolaou, 2003; S. Misra & Nikolaou, 2015; Panjwani & Nikolaou, 2016) . The rotated input design provides only sub-optimal when both input and output constraints are present (M L Darby & M. Nikolaou, 2014).

For real systems with input and output constraints present, designing experiments for the identification of an IC compliant model is a challenging task because of cumbersome eigenvalue based inequality check ($\text{Re}[\lambda(\mathbf{G}\hat{\mathbf{G}}^{-1})] > 0$ where \mathbf{G} is uncertain steady state gain matrix and $\hat{\mathbf{G}}$ is the identified model). To address this issue, Darby and Nikolaou (2009) developed a general mathematical framework, that relies on a much simpler inequality as its starting point. This simpler inequality is only sufficient but not necessary condition for IC satisfaction, which means it could possibly be a conservative check for IC satisfaction. Apart from this limitation, the DOE proposed by Darby and Nikolaou (2009) does not take into account any prior knowledge available about the system. It is not very uncommon to have partial knowledge available about the system. This prior knowledge could come from fundamental mass-energy balances or industrial experience. It is expected that inclusion of this partial knowledge in DOE would certainly improve the quality of the model.

1.2.1 Problem statement

In this work, the below two problems are addressed:

- 1) How conservative is the simpler inequality proposed by Darby and Nikolaou (2009)?
- 2) Is there a way to utilize partial knowledge in DOE, which can lead to better IC-compliance model?

1.2.2 Proposed approach

- 1) To address first problem, first sufficient condition was analyzed analytically and then numerical simulations were performed on a number of case-studies to check the conservatism of sufficient condition.
- 2) Since partial knowledge could be in any form, which makes this problem an open-ended problem. To begin with, partial knowledge in terms of linear equalities were considered in DOE. An optimization framework was setup to design experiment and identify an IC compliant model.

1.3 Organization of this Dissertation

The dissertation is organized as follows:

- In Chapter 2-3, firstly background literature on shale gas cementing is presented. Secondly, the data-driven modeling approach to address the well leakage problem is proposed. Finally, a case study is presented to demonstrate usefulness of the data-driven modeling approach.

- In Chapter 4-5, firstly conservatism of sufficient IC condition is checked through analytical study and numerical simulations on a number of case-studies. Secondly DOEs are developed incorporating partial knowledge available about the system and then a numerical optimization framework is presented that enables DOE for efficient identification of IC-compliant models for partially known systems.
- In Chapter 6, conclusions are drawn and future recommendations are made.

2 Cementing Wells: Background Literature

In this chapter, background and literature related to well cementing will be presented. Cement sheath quality evaluation and factors affecting cement sheath quality will be discussed in detail.

2.1 Cement sheath quality

Well cement sheath quality is affected by a number of factors. The significant factors are described as follows:

2.1.1 *Casing design.*

This includes the type of casings used for surface, intermediate, and production zones, internal diameter of casing, casing weight per unit length, number and type of centralizers, and casing/hole diameters ratio. The main purpose of using centralizers is to reduce eccentricity of casing in the borehole. Depending on the eccentricity of casing, a large number of centralizers may be required, but this creates a problem of extra drag while running casing, which is why spacing between centralizers needs to be calculated for minimizing the drag (Austin, 1988). Casing weight, hole-diameter, and hole-deviation play an important role in determining the minimum standoff (an eccentricity measure of casing pipe placement inside the borehole) required for uniform cement flow in the upper and lower parts of the horizontal section of the annulus (Figure 2-1 from Wilson and Sabins 1988). Minimum standoff required is 60% for uniform flow of cement inside the annulus (Wilson & Sabins, 1988). In general, the size of the annulus lies in a recommended range (0.75-1.5 inch) (Wilson & Sabins, 1988). Casing movement, i.e. rotation or reciprocation,

also affects the displacement efficiency as it breaks the agglomeration of gelled drilling mud (Wenande, 1987). Casing-connection is also an important part of casing-design, as they help in resisting leaks due to excessive internal or external fluid pressure and provide structural rigidity. There are a few different casing connections such as round thread and buttress thread. Casing connections may sometimes have a larger outer diameter than the body of the casing. Typically that only happens for the smaller casing sizes.

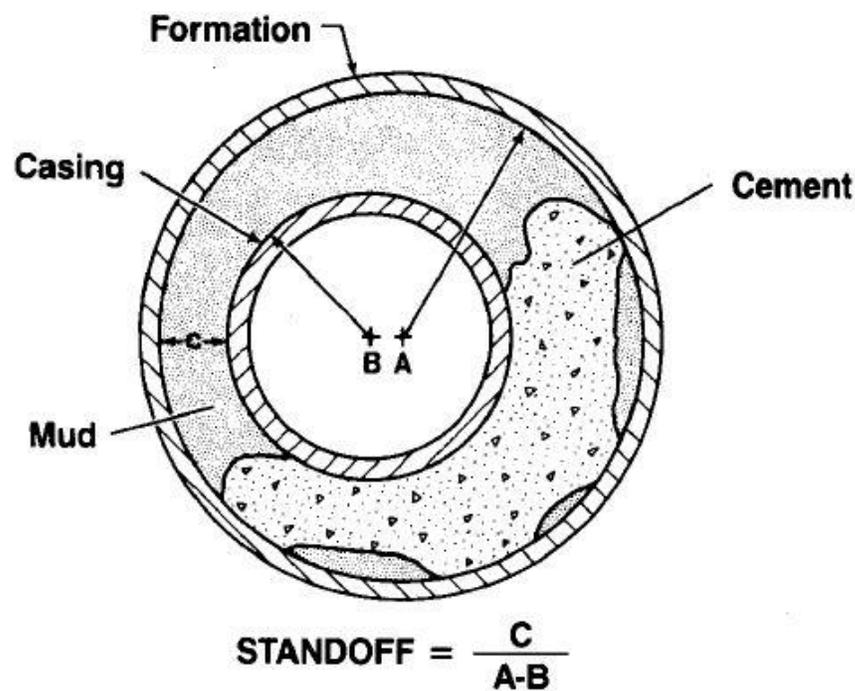


Figure 2-1: Standoff in the horizontal section of a well

2.1.2 Rheological properties.

Such properties include gel strength, plastic viscosity, and yield point. A number of researchers have suggested use the of drilling muds with ratio of yield point to plastic viscosity greater than one (Zurdo et al., 1986). These properties affect the cement

pumpability through the casing pipe and annulus. They also affect the drilling mud displacement by cement during cementing as they help to prevent settling of solids inside the horizontal annulus. It should be stressed that the key to successful cementing jobs is the properties of the slurries themselves, regardless of how such properties were achieved.

2.1.3 Cement properties and cementing procedure.

The properties of the cement slurry used in the lead and tail systems of a cemented section are generally different. In either case, important variables that ultimately affect cementing quality are slurry type, slurry density, slurry rheological properties, slurry free water content, fluid loss rate, thickening time, compressive strength, Young modulus and cement pumping rate. High fluid loss rate means more water is forced out of slurry to permeable formations which causes increase in cement slurry viscosity. High viscosity of cement slurry makes movement of the cement slurry more difficult. The Young modulus of the cement should be more than the Young modulus of the rock to achieve better cementing job quality (Thiercelin et al., 1998). The type and volume of a spacer, placed between the cement slurry front and the drilling mud displaced by the cement, play an important role in cementing, as they avert potential problems related to intermixing of drilling mud and cement. The presence of wiper plugs also helps in fluid displacement as it allows mechanical separation of fluids while fluids are inside the casing. Generally, two types of wiper plugs are used. The bottom plug is inserted before pumping the cement slurry. The diaphragm in the bottom plug ruptures after it reaches to the landing collar and allows the cement slurry to pass through it. The top pug is inserted at the end of cementing job.

2.1.4 Cement additives.

Accelerator, retarder, extender, fluid-loss, anti-foam, gas migration, dispersant, antigel, expansion additives and some special additives are added into the cement slurry in required proportions to achieve desired slurry properties. Each additive has its own role in affecting the cement sheath quality and there are multiple and complex interactions among these additives. The description of some of the additives are given below:

- i. Accelerator is added to accelerate the setting of the cement. It provides sufficient compressive strength, so that drilling operations can be continued without any problems. As the cement setting process is fast in high temperature conditions, this additive is generally used for near-surface cementing, where the temperature is not very high.
- ii. Retarder does exactly the opposite job of the accelerator additive. It is added in the cement slurry to increase the setting time for the purpose of proper placement. The retarder is needed in deep cementing operations.
- iii. Extender is added to decrease the density of a cement slurry. Reducing the density reduces the hydrostatic pressure of the cemented column, which helps in successful cementing of the weak zones of the well.
- iv. Fluid loss additive is added to control the loss of fluid from the cement slurry to the rock formation. The fluid loss from the cement slurry adversely affect the permeability of the reservoir formations and the cement setting process.
- v. Anti-foam additive is added to avoid the foam formation during the cement slurry preparation and mixing process. Excess foam may cause poor quality control of the cement slurry.

- vi. Gas migration additive is added to prevent the gas migration through and around the cement sheath.
- vii. Dispersant is added to reduce the cement slurry viscosity to enable easy flow of the cement slurry. It also helps in displacing drilling mud more efficiently.
- viii. Antigel additive is added to break conglomeration in the cement slurry, which enables the smooth flow of the cement slurry.

Several authors have studied the effect of special cement additives on cementing quality (Bach & Vijn, 2002; Elmarsafawi et al., 2006; Pollard, 1994).

2.1.5 Mixing water quality.

The quality of the mixing water also has an effect on cementing quality (Rowan & Stone). Contamination in the water affect thickening time, compressive strength, rheology, fluid loss and free water content in the water. The quality of the mixing water is primarily measured in terms of chloride content, sulfate content, pH, alkalinity and hardness of the water. High amount of chlorides in the mixing water may cause acceleration of the cement setting time. The major source of chloride contamination are water trucks that have previously carried brine water. High amount of sulfates in the mixing water may cause microstructural defects in the cement because of chemical changes in the cement. Hardness of the mixing water comes from magnesium ions, calcium ions and other dissolved carbonates.

2.1.6 Other variables.

Water temperature, bottom hole circulation temperature (BHCT), bottom hole static temperature (BHST), and hydrostatic pressure inside the well affect the reaction

kinetics of the cement slurry setting, rheology, and fluid loss, among others (Chow et al., 1988). Formation properties (e.g., porosity and permeability) as well as hole conditions (e.g., washouts) have an effect.

2.2 Cement sheath evaluation

The various techniques have been used for evaluation of the quality of a cement sheath. The following two techniques were used in this work.

2.2.1 Bond index.

Tools such as segmented cement bond tool (SBT) (Tyndall, 1990), cement bond log (CBL) (Fitzgerald et al., 1985), and the ultrasonic imager (USI) (Butsch, 1995) can be used for quantitative analysis of cement bond quality in terms of the bond index (BI), a continuous variable between zero and one, calculated using log data. The CBL tool functions on sonic log principals of refraction. Sound signals transmitted by the transmitter refracts along the casing-mud interface and returns back to receiver. The amplitude of the signal (in millivolts) or attenuation (in decibel/ft) is recorded. The bond index is calculated as the ratio of attenuation at any point and maximum attenuation. The maximum attenuation corresponds to the lowest amplitude in the CBL. BI values close to zero represent poor cementing quality whereas a BI value at one indicates perfect cementing. Although conventional CBL log display provides a good graphical way to evaluate cement sheath quality, CBL results can be analyzed using acoustic waveform display known as variable density logs (VDL). VDL helps in analyzing the effect of fast formation, decentralization and other problems. These waveforms have white-gray-black color band.

Zero amplitude, negative amplitude and positive amplitude are shown by grey, white and black colors respectively in CBL-VDL. Lower the amplitude better the quality of cementing.

The usage of CBL is limited to horizontal wells – not necessarily deep wells. The tool is typically run on a wireline and relies heavily on gravity to evaluate the entire well. Vertical sections are easier to log. Typically, the horizontal section requires a tractor-propelled CBL, which may be cost-prohibitive. The CBL is also not used often on less critical strings, because of the cost of the tool and associated rig time to run the tool. The operator may not elect to use a CBL on shallower strings but rather depend on other methods to determine whether there has been adequate isolation. Moreover, interpretation of CBL data in itself is not an easy task and it might produce erroneous results in the case of micro-annulus formation (Boyd et al., 2006). While CBL analyze the full 360deg, micro-annulus detection is difficult due to the way the tool functions. A micro annulus may only be indicated by performing runs with and without pressure on the casing to determine if there is a small gap. Other than micro-annulus formation, parameters like heavy mud weight in the wellbore, eccentricity of the CBL tool in the bore-hole, fast-formations, downhole pressure temperature, and directional holes affects the CBL functioning (Ashena et al.).

The USI tools are more useful as the tools runs under pressure to determine micro-annulus formation (Crain, 2000). The USI tools measure acoustic impedance of the material immediately behind the casing to evaluate casing-to-cement-bond quality. Color-coded cement maps provide a visual way to identify channels in the cement.

2.2.2 Sustained casing pressure.

A coarser alternative to the BI for assessment of cement sheath quality can be indirectly provided in terms of the sustained casing pressure (SCP). This is a measure of well annulus pressure that rebuilds after bleed-down (Figure 2-2). SCP is assumed to be caused by factors other than well temperature fluctuations or artificially imposed pressure. The major cause of such pressure buildup is poor primary cementing job (Rocha-Valadez et al., 2014). In ideal conditions, a cemented annulus should exhibit zero SCP. In practice, however, a small value of SCP is usually chosen as a threshold limit, above which gas leakage is considered to be appreciable.

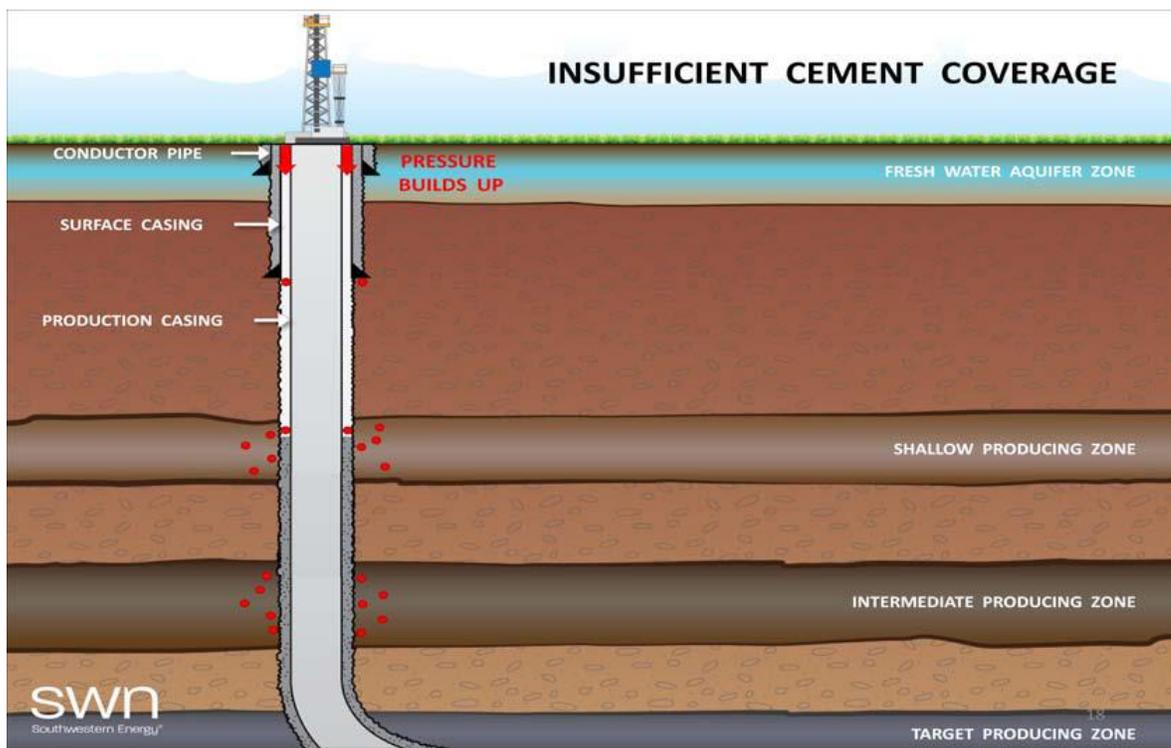


Figure 2-2: Sustained casing pressure (SCP)

3 Cementing Wells: A Modeling Study

As already mentioned in section 1.1.2 of chapter 1, the modeling exercise aims to generate a model that can provide guidelines on the effect of several (typically tens of) decision variables on the quality of a cementing job. In this chapter, model building will be attempted for either BI or SCP as the model output. In the first case a regression model can be built, since BI is a continuous variable (taking values in the interval $[0,1]$), whereas in the second case a classification model can be built, since SCP values below or above a threshold correspond to leaking or non-leaking cementing jobs (namely values 0 or 1). For either case a cross-validation method, such as 10-fold cross-validation, can be used, to avoid overfitting or underfitting the data (Seber & Lee, 2003).

Both regression and classification are broad subjects with vast literature. I summarize next the approaches that were considered in this work, as I deemed them to be most suitable for modeling the kind of data at hand.

3.1 Regression Models

Of the large array of standard methods available to build regression models, partial least squares (PLS) is a method that has found wide applicability in recent years (Svante Wold et al., 2001) to build models with collinear input (predictor) and/or output data, particularly from scant data. The data may correspond to multiple collinear variables, and may even be incomplete. PLS relies on constructing mutually uncorrelated latent input variables that are best correlated with mutually uncorrelated latent output variables. The latent variables (whether input or output) are constructed through linear combinations of

the original input and output variables. The number of latent variables is typically much smaller than the number of the original variables. This makes it easier to build a reliable model that connects all input with all output variables. To determine a reasonable (ideally optimal) number of latent variables, minimization of the cross-validation error (fitting error plus validation error) is usually employed. Both linear and nonlinear versions of the method are available (Holcomb & Morari, 1992; Malthouse et al., 1997; Qin & McAvoy, 1992; Taavitsainen & Korhonen, 1992; S. Wold, 1992). However, nonlinear models, because of their wide spectrum of possible structures, require substantially more data than linear models, and do not extrapolate reliably beyond the data used to build such a model.

3.2 Classification models

Some of the commonly used techniques for building classification models are linear discriminant analysis (LDA), support vector machines (SVM), discriminant analysis using principal component analysis (PCA) (Nguyen & Rocke, 2004), k-nearest neighbors (k-NN), and tree based classification (CART) (Wehrens, 2011), and partial least squares discriminant analysis (PLS-DA) (M. Barker & Rayens, 2003). Additionally, methods based on fuzzy logic and artificial neural networks (Haykin, 1999) provide higher flexibility, but require large amounts of data to produce models with useful predictive ability.

LDA is not suitable when the number of input variables is more than the numbers of data points or when the input variables are collinear. Methods based on dimensionality reduction, such as PLS or PCA have been found to be suitable for that kind of situation.

Partial least squares (PLS) or PCA are commonly used methods for the dimensionality reduction.

Partial least squares discriminant analysis (PLS-DA) builds a classification model through a combination of PLS regression (as mentioned above) and linear discriminant analysis. The PLS part of the analysis builds a linear model between continuous inputs and continuous outputs that are highly correlated with one another, whereas the DA part of the analysis maps the continuous outputs to discrete outputs, to complete the classification. In principle, standard linear regression could also be used in place of PLS if input and output variables were known not to be highly correlated.

PCA also reduces the dimensionality of predictor variable space by calculating principal components (PCs) but this method doesn't consider observable variables while calculating PCs. A number of studies prefer the use of PLS-DA over classification based on PCA (M. a. W. R. Barker, 2003; Nguyen & Rocke, 2004).

3.2.1 VIP variable selection and multi-step PLS-DA

To distinguish between input variables that have an appreciable effect on an observed output and those whose effect on the output is insignificant, one can calculate the VIP (variable importance in projection) value of each input variable after a model is estimated (Mehmood et al., 2011), retain in the list of inputs only those with $VIP > VIP_{\min}$, estimate the model again, and continue this cycle until convergence (Figure 3-1). While a typical value for VIP_{\min} is 1 (S. Wold, 1994) several values of VIP_{\min} can be examined and corresponding models built, so that the model with the lowest cross-validation error can be finally selected.

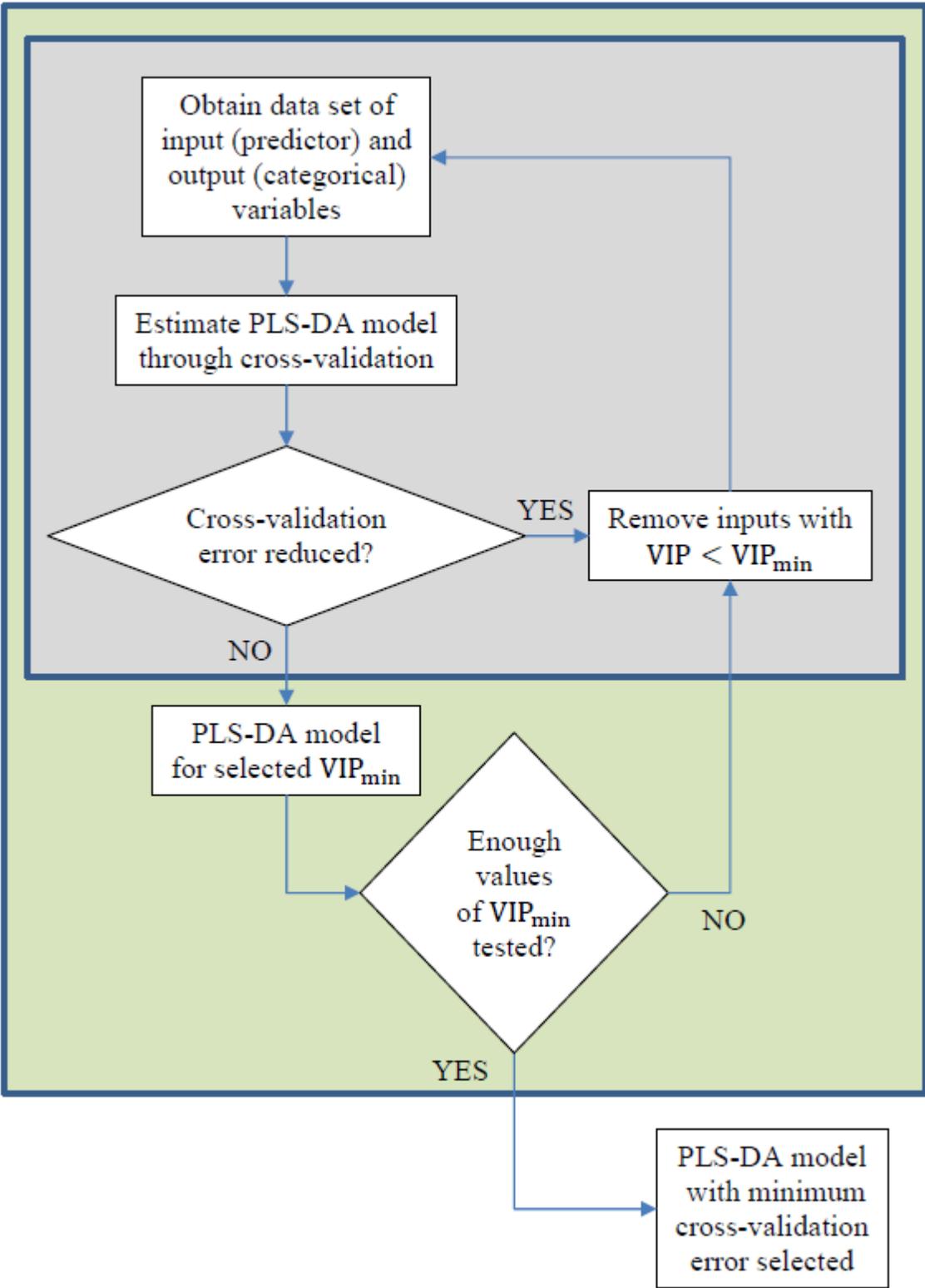


Figure 3-1: Multistep PLS-DA

3.2.2 Non-dimensionalization

To make the model as widely applicable as possible, and to potentially reduce the actual number of input variables in the model (thereby facilitating model building from scant data), the original input variables are clustered into dimensionless groups, using the standard pi theorem (Gibbins, 2011). The selection of such dimensionless groups is not unique, and has the flexibility to generate dimensionless groups that are intuitively simple to interpret and possibly correspond to choices already in practical use. After dimensionless groups are constructed for input variables, building corresponding regression or classification models can be attempted.

3.3 Modeling study

3.3.1 Available data

As already mentioned above, the proposed modeling approach relies on data alone to build a model that can predict the quality of a cementing job based on decisions made for a large number of factors. The data used for this study consisted of field data from the intermediate casing in the vertical section of horizontal shale-gas wells as well as data collected through laboratory experiments on the cement.

All 45 wells were in the same geological environment (Marcellus Shale). These wells were drilled by the same operator and cemented by two different service companies. Wiper plugs (one-plug system in most of the wells) were used in all the wells. Since all 45 wells were from the Marcellus shale region and drilled by the same operator, physical variables such as casing diameter, hole diameter, BHST and BHCT did not have a lot of variation in their values. Table 3-1 shows 35 physical variables related to cementing

quality and their respective ranges for all 45 wells. CBL data was available for 21 wells, which was used to calculate the average bond index (BI) value for these wells. Categorization of all 45 wells as leaking or non-leaking was also available, based on an SCP value threshold of 2 psi. Out of 45 wells, 19 showed SCP whereas 26 did not.

3.3.2 Dimensionless inputs

Starting with the 35 variables (Table 3-1) affecting cementing quality, 31 dimensionless groups were identified as shown in Table 3-2. Note that some of the original variables remained intact, while others moved to dimensionless groups. The dimensionless groups were created in such a way as to be close to well established dimensionless number already known in fluid flow and related disciplines. For example, dimensionless variable 2 resembles the skin friction coefficient (dimensionless number used in boundary-layer flow analysis of viscous flow); Variable 3 represents the ratio of viscous forces to inertial forces (equal to the inverse of the Reynolds number); Variables 4, 5, and 28 represent the ratio of shear forces to buoyancy forces at three different kinds of shear forces, namely pertinent to gel strength at two different times and at a critical point.

3.3.3 Regression analysis

Attempts were made to build a cross-validated linear regression model with predictive ability for the BI values recorded. All attempts proved unsuccessful. Therefore, the idea to build a linear regression model that could predict BI from available inputs was not pursued any further.

Table 3-1: Physical variables and respective ranges

No.	Physical Variable (unit)	Range
1	Displacement volume (barrel)	[58, 145]
2	Yield stress (lbf/ft ²)	[34, 94]
3	Plastic viscosity (centipoise)	[44, 175]
4	Gel strength[10 min] (lbf/100ft ²)	[12, 126]
5	Gel strength[10 sec] (lbf/100ft ²)	[8, 52]
6	BHCT (°F)	[70, 85]
7	BHST (°F)	[70, 85]
8	Spacer volume (barrel)	35
9	Average displacement rate (barrel/minute)	[3.7, 6]
10	Water to mix (barrel)	[48, 185]
11	Casing depth (ft)	[575, 2400]
12	Annulus size (inch)	[2, 3.625]
13	Cement additive: Accelerator (%)	[0, 0.2]
14	Cement additive: Fluid loss (%)	[0, 0.0085]
15	Cement additive: Anti foam (%)	[0, 0.005]
16	Cement additive: Gas migration (%)	[0, 0.002]
17	Cement additive: Anti gel (%)	[0, 0.0075]
18	Cement additive: Suspension (%)	[0, 0.005]
19	Cement additive: Extender (%)	[0, 0.09]
20	Cement additive: Dispersant (%)	[0, 0.012]
21	Cement additive: Retarder (%)	[0, 0.003]
22	Cement additive: Defoamer (%)	[0, 0.004]
23	Hardness of water (ppm)	[0, 280]
24	Sulfates in water (ppm)	[50, 200]
25	Chlorides in water (ppm)	[200, 1000]
26	Water pH	[6, 7]
27	Number of centralizers	[8, 23]
28	Critical Gel strength (lbf/100ft ²)	[257, 296]
29	Fluid loss rate (barrel/minute)	[5.4*10 ⁻⁵ , 6.2*10 ⁻⁵]
30	Displacement rate (standard deviation) (barrel/minute)	[0 2,.6]
31	Mud weight (lb/gal)	[8.3, 9.8]
32	Cement sacks	[388, 1140]
33	Cement density (lb/gal)	[15.4, 15.6]
34	Water temperature (°F)	[48, 81]
35	Cement pump rate (barrel/minute)	[3.1, 6]

Table 3-2: Affecting dimensionless variables for intermediate casing

No.	Dimensionless variable
1	Displacement volume/ Cement sacks
2	Yield stress \times Annulus size ⁴ /(Average displacement rate ² \times Cement density)
3	Plastic viscosity \times Annulus size/(Average displacement rate \times Cement density)
4	Gel strength[10 min]/(Cement density \times Annulus size)
5	Gel strength[10 sec]/(Cement density \times Annulus size)
6	BHCT/Water temperature
7	BHST/Water temperature
8	Spacer volume /Cement sacks
9	Average displacement rate/Cement pump rate
10	Water to mix /Cement sacks
11	Casing depth/Annulus size
12	Annulus size/Casing internal diameter
13	Cement additive: Accelerator
14	Cement additive: Fluid loss
15	Cement additive: Anti foam
16	Cement additive: Gas migration
17	Cement additive: Anti gel
18	Cement additive: Suspension
19	Cement additive: Extender
20	Cement additive: Dispersant
21	Cement additive: Retarder
22	Cement additive: Defoamer
23	Hardness of water
24	Sulfates in water
25	Chlorides in water
26	Water pH
27	Number of centralizers
28	Critical Gel strength/(Cement density \times Annulus size)
29	Fluid loss rate/Cement pump rate
30	Displacement rate (standard deviation)/ Average displacement rate
31	Mud weight/ Cement density

3.3.4 Linear discriminant analysis

To build a classification model, LDA performs inversion of the input variable matrix. Due to collinearity among the dimensionless input variables obtained from the data-set, LDA did not produce numerically accurate results.

3.3.5 Partial least squares discriminant analysis

Classification using PLS-DA was used with the 31 dimensionless variables shown in Table 3-2 as inputs and the designation Leak/No-Leak as output variable for each of the 45 wells in the database. Ten-fold cross-validation was used for model validation. In 10-fold cross-validation, the whole data-set is randomly partitioned into 10 equal-sized sub datasets. Out of 10 subsets, 9 subsets are used for training the model and a remaining subset is used for the model validation. This cross-validation process is repeated 10 times (folds), so that each subset can be used exactly once as the validation set. The PLS-Toolbox (Eigenvector, 2013) was used to build the PLS-DA model. All programming was done in MATLAB[®] (MATLAB, R2013b). Since the PLS-toolbox provides only the best estimate of the probability that the output belongs to either of the two categories, without a confidence assessment, the reliability of classification probability (as Leak or No-Leak) was calculated by considering the 68% confidence interval around observable variables (Pérez et al., 2009). Since these probabilities were calculated using Bayes formula (Duda et al., 2000) and cutoff was fixed at point where probability of belonging to either category was 50%, it is highly probable that some of the points lie very close to cutoff point although correctly classified. Classification results for such points should be used with caution. To make predictions more reliable, not-classified category was introduced. All those points

where probabilities lie in the range (50%-55%) were put into the not classified category. Details are provided in Appendix A.

3.4 Results and Discussion

3.4.1 Model accuracy and confidence in predictions

Table 3-3 shows cross-validated classification results for 14 models built using multistep PLS-DA (Figure 3-1). Each model uses a different VIP_{min} value (column 2 of Table 3-3) to select the number of significant inputs. Column 6 in Table 3-3 shows the percentage of all 45 cases for which each model *correctly* predicted the category of the output (leak/no-leak) with probability greater than 55% in cross-validation tests. Column 7 in Table 3-3 shows the percentage of all 45 cases for which each model *incorrectly* predicted the category of the output (leak/no-leak) with probability greater than 55% in cross-validation tests. Column 8 in Table 3-3 shows the percentage of all 45 cases for which each model *correctly or incorrectly* predicted the category of the output (leak/no-leak) with fairly low probability, namely below 55%, in cross-validation tests. Of the 14 models built, model 6 performed best, in that it has the highest percentage (75%) of correctly classified samples out of all samples (with probability above 55%), and the lowest percentage of non-classified samples (7%). Overall, model 6 performs best by predicting correctly $81\% \left(= \frac{75}{75+18} \times 100 \right)$ of all cases classified with probability above 55%.

For completeness, column 4 of Table 3-3 indicates that of the 31 original dimensionless inputs, a smaller number turned out to be significant and to contribute to each model's predictive ability. The corresponding VIP_{min} values used to determine the number of significant variables for each model are shown in column 2 of Table 3-3. Finally, column

5 of Table 3-3 shows the small number of latent variables used in the PLS part each model, determined by minimizing the cross-validation error.

Table 3-3: Intermediate casing PLS-DA model cross-validation results for different VIP cut-off values

Model No.	VIP _{min}	No. of iteration (inner loop, Figure 3-1)	No. of chosen inputs	No. of latent variable in PLS-DA	% correctly classified	% incorrectly classified	% non-classified
1	All inputs included		31	4	65	22	13
2	0.1	1	30	2	49	13	38
3	0.2	1	30	2	47	18	35
4	0.3	1	26	3	71	18	11
5	0.4	1	25	3	71	18	11
6	0.5	1	23	4	75	18	7
7		2	21	3	69	20	11
8	0.6	1	22	4	69	18	13
9		2	20	1	47	11	42
10	0.7	1	20	3	65	13	22
11		2	14	1	58	11	31
12	0.8	1	16	1	49	18	33
13	0.9	1	14	1	49	13	38
14	1.0	1	11	1	54	13	33

3.4.2 Significant variables for cementing quality

Model 6 used 23 dimensionless inputs which were selected using VIP_{min} equal to 0.5. The ranking of selected dimensionless variables is shown in Figure 3-2. This ranking can be useful in understanding the relative significance of selected variables for cementing quality. Note that this ranking is entirely based on data and is not influenced by any

preconceived notions of what variables may be significant or not. To our knowledge, no such ranking of the relative importance of these variables has appeared in literature.

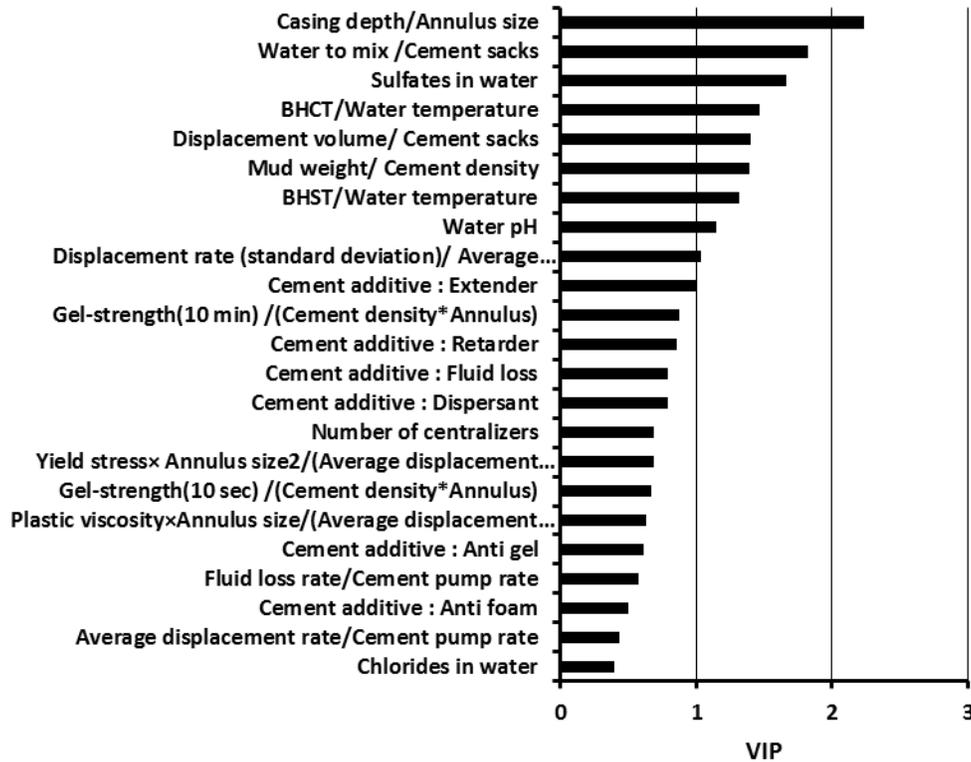


Figure 3-2: VIP values of selected dimensionless variables for model-6

3.4.3 Positive and negative effect of variables on cementing quality

Apart from the relative significance of input variables, the PLS-DA model developed here provides an indication in what direction a variable must be moved to decrease the chance of leak. While this can be determined numerically, there is also an insightful graphical tool that can provide quick visual understanding of how these physical variables are related to leakage of the wells. This tool makes use of the PLS part of the full PLS-DA model, as described in more detail in Appendix A. In summary, each latent input

variable t_j in the PLS model is a linear combination of the 23 input variables x_i , i.e.

$$t_j = \sum_{i=1}^{23} w_{ji}x_i ; \text{ and the output of the PLS model is } \hat{y} = q_1t_1 + q_2t_2 + \dots$$

Assuming that two latent input variables (t_1, t_2) capture enough variance of the output \hat{y} , one can check

whether the vectors $[w_{1,i} \quad w_{2,i}]$ and $[q_1 \quad q_2]$ form an acute angle or not in a 2D plot, as

shown in Figure 3-3. In case of acute angle, correlation is positive; otherwise correlation

is negative. For example, (Average displacement rate)/ (Cement pump rate) forms an acute

angle with the leak weight vector, corresponding to positive correlation, whereas Anti-

foam forms an obtuse angle and is negatively correlated with leaks. Similarly, Retarder,

forming an approximately right angle with Leak has little effect on leaks. In that figure,

the values of $[w_{1,i} \quad w_{2,i}]$ for $i = 1, \dots, 23$, as well as $[q_1 \quad q_2] = [0.28 \quad 0.17]$ are shown.

Variables that appear to have a fairly negative effect towards leaking are gel-strength (10

sec)/ (Cement density×Annulus size), gel-strength (10 min)/ (Cement density×Annulus

size) etc., whereas variables such as BHCT/Water temperature, BHST/Water temperature

etc., appear to have a positive effect towards leaking. The complete results of this analysis

are summarized in Table 3-4.

Using the correlations presented in Table 3-4, cementing variables should be altered

to ensure no leakage from the cemented annulus. However some of the variables like

BHCT, BHST, casing internal diameter and well depth, cannot be changed for a cementing

job. As it was mentioned earlier that geometry of the annulus (internal diameter, hole-

diameter depth etc.) affects the behavior of cement flow (laminar or turbulent) inside the

annulus and temperatures BHCT, BHST control setting of cement inside annulus, inclusion

of these variables makes our data-driven analysis more realistic.

Important observations from Table 3-4 and Figure 3-3 are as follows:

- For all dimensionless variables, except for variables 21, 27 and 30, the approximately computed sign of the corresponding variable's effect on leakage (positive or negative), as shown in Figure 3-3, agrees with the sign computed using the full PLS model (Appendix A). The sign discrepancy for these variables (positive/negative) is due to the fact that for these variables, latent variable weights associated with latent variables 3 and 4 are not small and hence the effect of latent variable weights associated with latent variables 3 and 4 should not be ignored to find the correlation with leakage.
- Variables 6 and 7 are inversely proportional to temperature of the mixing water and both variables are positively correlated with well leakage. Field observations also suggest that cold water leads to poor mixing (Watters, 2014).
- Cement slurries that develop high gel strength in short time span are useful to restrict fluid migration (Childs, 1984). Variables 4 and 5, which are directly proportional to gel-strength and negatively correlated with well leakage, also suggests that high gel-strength development in short time span is desirable for good cementing job.
- Negative correlation of variable 27 indicates that a high number of centralizers is required for better cementing job. However a very high number of centralizers may cause excess drag when running the casing resulting in extra washout or some negative effects on the wellbore. This may also affect the ability of the rig to rotate or reciprocate the pipe while

cementing, too many centralizers may prevent this pipe movement from occurring. Centralizers are necessary for uniform cement coverage by ensuring a centralized casing, so an optimized number of centralizers would be beneficial.

- Negative correlation of variable 30 suggests that high variation in the displacement rate is desirable for good cementing job. In fact, this correlation is in agreement with the field practice where the rate is slowed prior to bumping the plug to avoid over pressuring when the plug seals in the floats. The displacement rate is also slowed down to maintain the uniform density down hole. However the rate may be slowed because of mixing problems or equipment problems as well.
- Positive correlation of variable 10 suggests that very high volume of mixing water is not good for a cementing job.
- Variable 19 is negatively correlated with leakage (i.e. high extender implies low leakage). The model's prediction is also in agreement with the prescriptive suggestion to use extenders for high water/cement ratio.
- Positive correlation of variable 29 suggests that high fluid loss rate is not desirable for good cementing job. To reduce fluid loss rate, fluid loss cement additive is added and our model is also suggesting addition of fluid loss additive for better cementing
- Mud weight should be low (not very low) for better mud displacement. Positive correlation of variable 31 with well leakage, also suggests use of low mud weight for better cementing job.

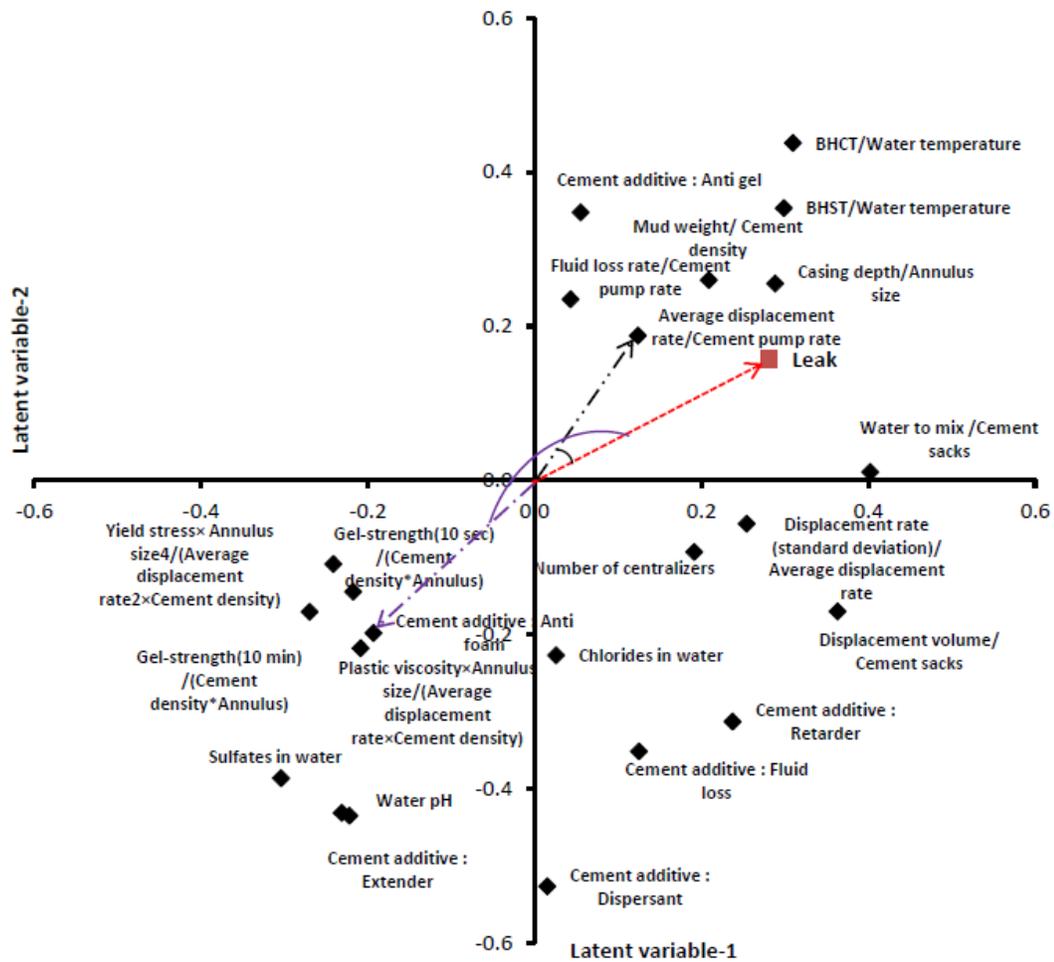


Figure 3-3: Latent variable weights for dimensionless inputs in the model-6

Table 3-4: Selected dimensionless variables and their positive or negative effect on well leakage for model-6 in Table 3-3

No. (As in Table 3-2)	Dimensionless Variable	Correlation with well leakage (using 2 latent variables, Figure 3-3)	Correlation with well leakage (using all 4 latent variables – Appendix A)
1	Displacement volume/ Cement sacks	Positive	Positive
2	Yield stress x Annulus size ⁴ / (Average displacement rate ² x Cement density)	Negative	Negative

Table 3-4 continued

3	Plastic viscosity×Annulus size/(Average displacement rate×Cement density)	Negative	Negative
4	Gel strength(10 min)/(Cement density×Annulus size)	Negative	Negative
5	Gel strength(10 sec)/(Cement density×Annulus size)	Negative	Negative
6	BHCT/Water temperature	Positive	Positive
7	BHST/Water temperature	Positive	Positive
9	Average displacement rate/Cement pump rate	Positive	Positive
10	Water to mix /Cement sacks	Positive	Positive
11	Casing depth/Annulus size	Positive	Positive
14	Cement additive: Fluid loss	Negative	Negative
15	Cement additive: Anti foam	Negative	Negative
17	Cement additive: Anti gel	Positive	Positive
19	Cement additive: Extender	Negative	Negative
20	Cement additive: Dispersant	Negative	Negative
21	Cement additive: Retarder	Positive	Negative
24	Sulfates in water	Negative	Negative
25	Chlorides in water	Negative	Negative
26	Water pH	Negative	Negative
27	Number of centralizers	Positive	Negative
29	Fluid loss rate/ Cement pump rate	Positive	Positive
30	Displacement rate (standard deviation)/ Average displacement rate	Positive	Negative
31	Mud weight/ Cement density	Positive	Positive

3.4.4 Visualizing and understanding differences between wells

The PLS-DA model can also provide intuitively appealing visual analysis of similarities and differences between wells. The idea is that the first two latent variables $t_1 = \sum_{i=1}^{23} w_{1i}x_i$ and $t_2 = \sum_{i=1}^{23} w_{2i}x_i$ are responsible for most of the variance seen in the output, as captured by the PLS-DA model. Now, each value of $[t_1 \ t_2]$ (also called x-scores) represents a point associated with a well in a 2D plot. Therefore, points close to each other refer to wells with fairly similar behavior. Using the preceding ideas, patterns, such as clusters or outliers, can be identified in a 2D plot with axes t_1 and t_2 . Such a plot is shown in Figure 3-4, where the following observations can be made.

- Wells 2, 10, 11, and 24 have very close scores for latent variables 1 and 2 (Figure 3-4). This hints toward the similarity among these wells. These wells can be distinguished from the rest of the wells by the value of dimensionless variable 10 (water to mix/cement sacks), which was found to be one of the most important variables according to VIP ranking (Figure 3-2). These wells have higher values for variable 10 (≈ 0.21) compared to the rest of the wells (≈ 0.12).
- Wells 5 and 18 lie on two extreme positions along latent variable-1 and belong to the no-leak and leak categories, respectively. One of the reasons for their extreme positions in the x-score plot is very high values of variables 4 and 5 for well 5 and very low values of the same variables for well 18. The same conclusion can be drawn just by observing the positions of well 5 in x-score plot (Figure 3-4) and positions of variables 4 and 5 in the weights plot (Figure 3-3). Variables 4 and 5 both are negatively correlated with well leakage.

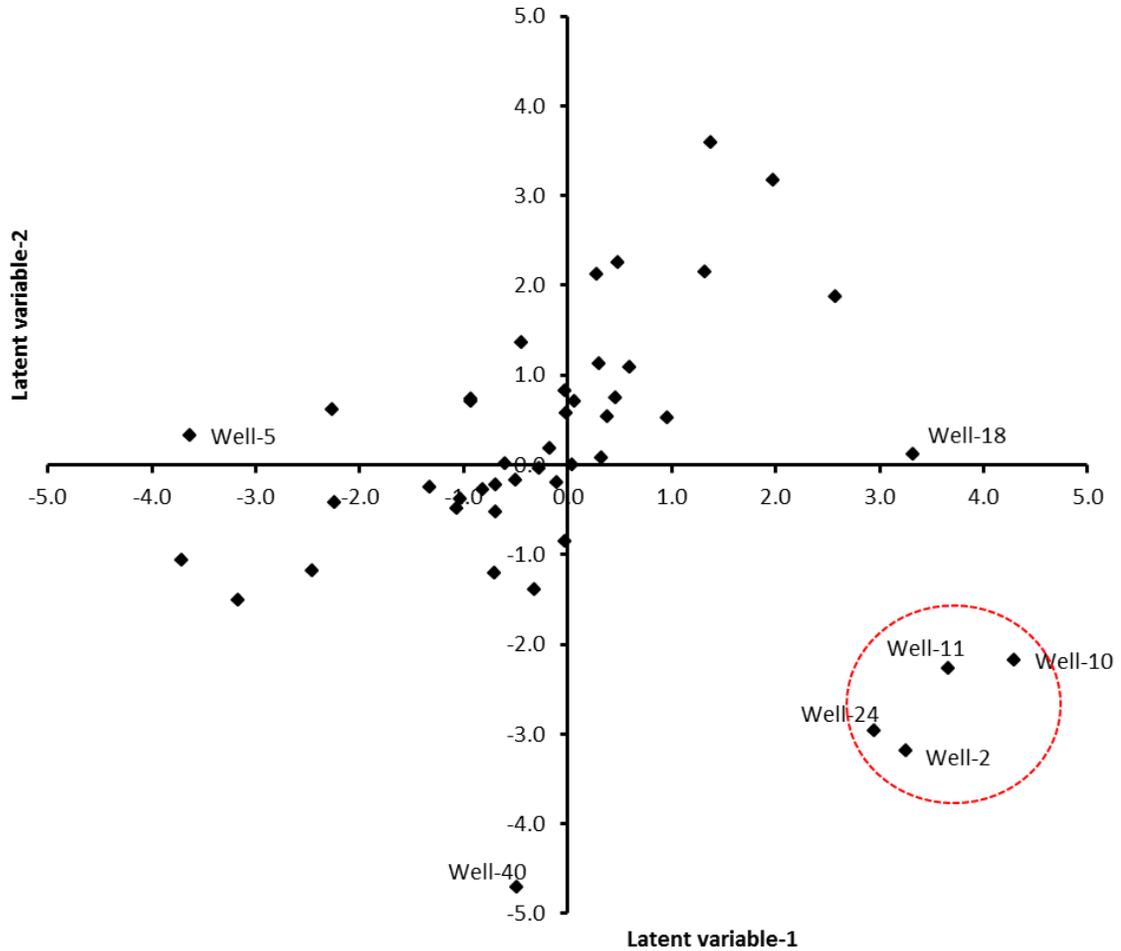


Figure 3-4: Latent variable x-scores for model-6

3.5 Summary

A data-driven classification model was developed using PLS-DA analysis. The model can make useful predictions on whether a set of values for a large number of cement and casing factors will result in a leaking or non-leaking well. The model classified correctly 81 % of the classified wells in a cross-validation study with reasonable confidence. In addition, the model provides a confidence with which each well prediction is classified as leaking or non-leaking.

The model is able to identify (in terms of the VIP statistic) the relative importance of cement and casing variables used as inputs to the model. Furthermore, the model also identifies whether there is positive or negative correlation between each input variable and leakage. Finally, a 2D score plot provides an intuitively appealing graphical way to cluster or distinguish wells based on their leaking behavior.

The model can be directly used in what-if analysis for future cementing jobs and achieve better cement integrity or avoid undesired gas leakage.

4 Ensuring Integral Controllability for Robust Multivariable Control

In this chapter, question raised in section 1.2.1 will be addressed. Namely, attempts will be made to check the conservatism of the integral controllability sufficient condition developed by Darby and Nikolaou (2009). The work presented in this chapter will clear some doubts before we attempt experiment design for control relevant model identification subject to integral controllability.

4.1 Introduction

A good model is at the heart of good controller design. Development of a control-relevant model for a multivariable process, particularly for an ill-conditioned one, is especially challenging. A variety of creative approaches have appeared towards design of informative experiments that facilitate control-relevant model development (M L Darby & M. Nikolaou, 2014, and references therein). A fundamental fact when developing a control-relevant model is that mere proximity between the model and the controlled process (e.g. in the sense of small errors in model parameter estimates) is not adequate. Conditions must also be satisfied that are relevant to the controller design method in which the model will be used. C E Garcia and M Morari (1985) showed that for robust multivariable controller design with the internal model control (IMC) method, a process with real steady-state gain matrix (SSGM) \mathbf{G} and model SSGM $\hat{\mathbf{G}}$ must be such that

$$\operatorname{Re}\left[\lambda(\mathbf{G}\hat{\mathbf{G}}^{-1})\right] > 0 \quad (4.1)$$

for all eigenvalues λ of $\mathbf{G}\hat{\mathbf{G}}^{-1}$. Satisfaction of eqn. (4.1) ensures integral controllability (IC), as defined precisely in the Background section.

Since the real process \mathbf{G} is never known with certainty, IC-compliance requires satisfaction of eqn. (4.1) for all \mathbf{G} in some uncertainty set U . Therefore, for a model to be used in multivariable IMC design, the model $\hat{\mathbf{G}}$ must satisfy eqn. (4.1) for all \mathbf{G} in U . However, it is not trivial to check whether eqn. (4.1) is satisfied for all \mathbf{G} in a typical set U , such as an ellipsoidal uncertainty set resulting from estimation of $\hat{\mathbf{G}}$ through linear least squares. To circumvent this difficulty, Darby and Nikolaou (2009) proposed an alternative condition to the inequality in eqn. (4.1), whose satisfaction is trivial to check, as summarized in the Background section. Based on this condition, Darby and Nikolaou (2009) developed a general mathematical framework for design of experiments that generate data for efficient identification of IC-compliant models under several circumstances. However, the condition proposed by Darby and Nikolaou (2009) is only sufficient and not necessary. This immediately raises the following question, Q1:

How conservative is the sufficient condition of Darby and Nikolaou (2009), namely how likely is it that a model $\hat{\mathbf{G}}$ violating that condition also violates eqn. (4.1)?

An even more fundamental question that can also be raised is the following question Q2:

If a model $\hat{\mathbf{G}}$ and real process \mathbf{G} with corresponding uncertainty set U violate eqn. (4.1), is there an alternative $\hat{\mathbf{G}}$ that can satisfy eqn. (4.1)?

The rationale behind the preceding question Q2 is that perhaps a model other than $\hat{\mathbf{G}}$, possibly a little farther from the real \mathbf{G} than $\hat{\mathbf{G}}$ is, but with the right structure, might satisfy eqn. (4.1) for all $\mathbf{G} \in U$, thus making the design of IC-compliant identification experiments potentially unnecessary.

The main purpose of this work is to address the preceding two questions Q1 and Q2. In summary, numerical simulations concerning question Q1 indicate that the corresponding mathematical framework appears not to be overly conservative for small systems, the conservatism increasing for larger systems. A counter-example for the 2×2 case indicates that the answer to question Q2 is negative, and design of experiments for efficient identification of IC-compliant models is indeed desirable.

In the rest of the chapter, firstly a brief background on integral controllability is provided, and then the above questions Q1 and Q2 are addressed.

4.2 Background

We consider an $n \times n$ stable, linear, time-invariant system with steady-state input-output behavior

$$\mathbf{y} = \mathbf{G}\mathbf{m}, \quad (4.2)$$

where $\mathbf{y}, \mathbf{m} \in \mathbb{R}^n$, and $\mathbf{G} \in \mathbb{R}^{n \times n}$. Both inputs and outputs are in deviation form from steady state; i.e., the steady state is at $\mathbf{y} = \mathbf{m} = \mathbf{0}$.

4.2.1 Integral controllability

The precise formulation of the result on integral controllability (C E Garcia & M Morari, 1985, Theorem 2) is as follows: Assume that internal model control (IMC) with

estimate $\hat{\mathbf{G}}$ of \mathbf{G} and a diagonal (decoupling) filter matrix $\mathbf{F}(z) = \text{diag} \left\{ (1-\alpha)/(1-\alpha z^{-1}) \right\}$ is used to control the system in eqn.(4.2). Then, there exists an $\alpha^* \in [0,1)$ such that the closed loop is stable for all $\alpha \in [\alpha^*,1)$, *if and only if* all eigenvalues $\lambda(\mathbf{G}\hat{\mathbf{G}}^{-1})$ of $\mathbf{G}\hat{\mathbf{G}}^{-1}$ are in the right half of the complex plane, i.e. they satisfy the IC condition, eqn. (4.1)

This result establishes the achievable robustness of decoupling multivariable controllers with integral action. Since \mathbf{G} is not known, eqn. (4.1) must be satisfied for all \mathbf{G} in an uncertainty set U .

4.2.2 Sufficient condition for integral controllability

Assuming that $\hat{\mathbf{G}}$ is the outcome of least-squares identification over t time steps, an uncertainty set U for \mathbf{G} can be defined as the standard ellipsoidal uncertainty set, resulting from least-squares estimation of \mathbf{G} , i.e.,

$$U = \left\{ \mathbf{G} = \begin{bmatrix} \mathbf{g}_1^T \\ \vdots \\ \mathbf{g}_n^T \end{bmatrix} \in \mathbb{R}^{n \times n} \left| \left(\mathbf{g}_i^T - \hat{\mathbf{g}}_i^T \right) \mathbf{M}^T \mathbf{M} \left(\mathbf{g}_i - \hat{\mathbf{g}}_i \right) \leq c^2, \quad i = 1, \dots, n \right. \right\}, \quad (4.3)$$

where the information matrix $\mathbf{M}^T \mathbf{M}$ results from the input matrix $\mathbf{M} \in \mathbb{R}^{t \times n}$ and $c^2 = s^2 n F_{1-\gamma}(n, t-n) \approx \sigma_{\text{noise}}^2 \chi_{1-\gamma}^2(n)$ for confidence level γ .

Assessing whether eqn. (4.1) holds for all \mathbf{G} in the uncertainty set U defined in eqn. (4.3) is challenging. To remedy that, Darby and Nikolaou (2009), Corollary 1 proved that eqn. (4.1) is satisfied for all \mathbf{G} in the set U of eqn. (4.3) if

$$J \triangleq \sum_{k=1}^n c \frac{\|\hat{\mathbf{u}}_k\|_1}{\hat{\sigma}_k} \sqrt{\hat{\mathbf{v}}_k^T (\mathbf{M}^T \mathbf{M})^{-1} \hat{\mathbf{v}}_k} < 1, \quad (4.4)$$

where

$$\hat{\mathbf{G}} = \hat{\mathbf{U}} \hat{\Sigma} \hat{\mathbf{V}}^T = \sum_{k=1}^n \hat{\sigma}_k \hat{\mathbf{u}}_k \hat{\mathbf{v}}_k^T \quad (4.5)$$

is the singular-value decomposition of the identified model $\hat{\mathbf{G}}$.

The above eqn. (4.4) is only sufficient, and potentially conservative, as already discussed. The main source of conservatism can be found in the proof of Theorem 1 in Darby and Nikolaou (2009), namely through the weakening of inequalities in the following steps:

- Satisfaction of the inequality

$$\operatorname{Re} \left[\lambda \left((\hat{\mathbf{G}} + \mathbf{D}) \hat{\mathbf{G}}^{-1} \right) \right] > 0 \Leftrightarrow 1 + \operatorname{Re} \left[\lambda \left(\mathbf{D} \hat{\mathbf{G}}^{-1} \right) \right] > 0 \text{ for all } \mathbf{G} \in U \text{ and} \quad (4.6)$$

$$\text{if } \left| \lambda \left(\mathbf{D} \hat{\mathbf{G}}^{-1} \right) \right| < 1 \text{ for all } \mathbf{D} \triangleq \mathbf{G} - \hat{\mathbf{G}} \in D, \quad (4.7)$$

where

$$D = \left\{ \mathbf{D} = \begin{bmatrix} \mathbf{d}_1^T \\ \vdots \\ \mathbf{d}_n^T \end{bmatrix} \in \mathbb{R}^{n \times n} \left| \mathbf{d}_i^T \mathbf{M}^T \mathbf{M} \mathbf{d}_i \leq c^2, i = 1, \dots, n \right. \right\}. \quad (4.8)$$

- The series of inequalities

$$\left\| \sum_{k=1}^n \frac{1}{\hat{\sigma}_k} \mathbf{D} \hat{\mathbf{v}}_k \hat{\mathbf{u}}_k^T \right\|_i \leq \sum_{k=1}^n \left\| \frac{1}{\hat{\sigma}_k} \mathbf{D} \hat{\mathbf{v}}_k \hat{\mathbf{u}}_k^T \right\|_i \leq \sum_{k=1}^n \frac{1}{\hat{\sigma}_k} \|\mathbf{D} \hat{\mathbf{v}}_k\|_i \|\hat{\mathbf{u}}_k^T\|_i. \quad (4.9)$$

Conservatism may arise in any of the weakened inequalities shown above. While a quantitative assessment of the conservatism introduced by each individual step is complicated, the following analysis, corroborated by the numerical simulations in section

4.3 suggests that eqn. (4.9) introduces more conservatism than eqn. (4.7). To begin with, eqn. (4.6) is equivalent to

$$\left| \frac{\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})}{2 + \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})} \right| < 1 \text{ for all } \mathbf{D} \in D \quad (4.10)$$

as can be shown from direct application of a Moebius transformation (Appendix B.). Eqn. (4.10) is trivially satisfied if eqn. (4.7) is true,

because

$$2|\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})| < 2 \Leftrightarrow |\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})| < 2 - |\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})| \leq |2 + \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})|. \quad (4.11)$$

However, the converse is not true, as shown in Figure 4-1. In Figure 4-1 (left) it is clear that if eqn. (4.7) is violated for some $\mathbf{D} \in D$ with $\text{Re}[\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})] < -1$, then eqn. (4.10) is also violated for the same $\mathbf{D} \in D$. Now, if there is some $\mathbf{D} \in D$ with $\text{Re}[\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})] > 1$, then, by eqn.(4.8), there also exists $-\mathbf{D} \in D$, for which $\text{Re}[\lambda(-\mathbf{D}\hat{\mathbf{G}}^{-1})] = -\text{Re}[\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})] < -1$, and consequently eqn. (4.10) is also violated for $-\mathbf{D} \in D$. Therefore, conservatism is introduced for $-1 < \text{Re}[\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})] < 1$ when violation of eqn. (4.7) does not necessarily imply violation of eqn. (4.10), as shown in Figure 4-1 (right).

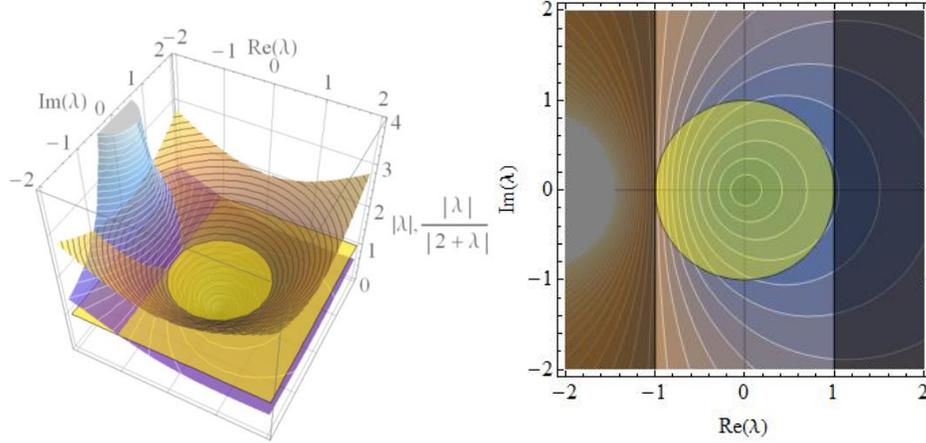


Figure 4-1: Visualization of the functions $\left| \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1}) \right|$ and $\left| \frac{\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})}{2 + \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})} \right|$ in a 3D plot (left) and contour plot (right). Note that violation of $\left| \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1}) \right| < 1$ for some $\mathbf{D} \in D$ does not imply violation of $\left| \frac{\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})}{2 + \lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})} \right| < 1$ for some $\mathbf{D} \in D$

Whether replacement of eqn. (4.6) by the weakened eqn. (4.7) introduces conservatism can also be tested numerically, namely for a large number of realizations of $\mathbf{D} \in D$, with D an uncertainty ellipsoid resulting from least-squares estimation of $\hat{\mathbf{G}}$, the eigenvalues of $\mathbf{D}\hat{\mathbf{G}}^{-1}$ can be plotted on the complex plane, and then one can check whether violation of eqn. (4.7) for some $\mathbf{D} \in D$ implies violation of eqn. (4.6) for some $\mathbf{D} \in D$. This test is carried further in the case studies of the next section.

4.3 Main Results

In this section we address the basic questions Q1 and Q2 raised earlier.

4.3.1 *How conservative is the sufficient condition for integral controllability?*

As discussed above, satisfaction of eqn. (4.4) guarantees that eqn.(4.1) is satisfied for all \mathbf{G} in the set U (eqn.(4.3)). However, if eqn. (4.4) is violated, eqn. (4.1) may or may not be violated. For a rigorous assessment of whether eqns. (4.1) and (4.4) are valid concurrently, one would have to check whether

$$\min_{\mathbf{G} \in U} \left\{ \text{Re} \left[\lambda(\mathbf{G}\hat{\mathbf{G}}^{-1}) \right] \right\} > 0 \quad (4.12)$$

coincides with satisfaction of eqn. (4.4). Unfortunately, finding the minimum in eqn. (4.12) is not trivial, because the function to be minimized is not convex (Appendix B.).

To circumvent this difficulty, we assess satisfaction of the inequality in eqn. (4.12) through randomized simulation, namely by checking whether eqn. (4.4) is satisfied for a large number of matrices \mathbf{G} , each randomly chosen from the set U in eqn. (4.3) As eqn. (4.3) suggests, the set U depends on the matrix $\mathbf{M}^T\mathbf{M}$, namely on the inputs \mathbf{m} chosen for the experiment. Therefore, the kind of inputs used in the identification experiment has a strong effect on the uncertainty ellipsoid corresponding to eqn. (4.3). Regarding such inputs, it is well known that

- (a) pseudo-random binary sequences (PRBS) (Ljung, 1999; Soderstrom & Stoica, 1989) are D-optimal (i.e. optimal when model-process proximity is the main objective of the identification experiment) whereas

(b) rotated PRBS inputs (Darby & Nikolaou, 2009; Koung & MacGregor, 1991) are IC-optimal (i.e., optimal when IC of the model-process pair is the main objective of the identification experiment). Such rotated inputs are defined as

$$\mathbf{m} = \hat{\mathbf{V}}\boldsymbol{\xi}, \quad (4.13)$$

where $\hat{\mathbf{V}}$ comes from the singular-value decomposition of $\hat{\mathbf{G}}$, eqn.(4.5), and the input vector $\boldsymbol{\xi}$ is chosen as a PRBS with individual entries of $\boldsymbol{\xi}$ proportioned as

$$\frac{\text{var}(\xi_k)}{\text{var}(\xi_j)} = \left(\frac{\hat{\sigma}_j^2}{\hat{\sigma}_k^2} \right)^x, \quad (4.14)$$

where $x = 1$ or $x = 1/3$ when the total output or input variance, respectively, must remain bounded during the experiment (Darby & Nikolaou, 2009, Theorems 4 and 5).

It is clear that the matrix $\mathbf{M}^T\mathbf{M}$ in eqn. (4.3) is approximately diagonal for PRBS inputs (above case (a)) and approximately equal to $\mathbf{V}^T \text{diag}\left(\frac{d}{\hat{\sigma}_1^{2x}}, \dots, \frac{d}{\hat{\sigma}_n^{2x}}\right)\mathbf{V}$ for rotated PRBS inputs (above case (b)).

Specifics of the randomized simulation procedure are as follows:

1. For noise variance σ_e^2 , perform an identification experiment by exciting the process with PRBS or rotated PRBS inputs of amplitude u_{\max} for t time steps and calculating the estimate $\hat{\mathbf{G}}$ of \mathbf{G} and the set U in which \mathbf{G} belongs, eqn. (4.3).
2. Randomly select n_s samples of the plant \mathbf{G} in the set U calculated in step 1.
3. Calculate the percentage of n_s samples that violate eqn. (4.1) for the model $\hat{\mathbf{G}}$ and all $\mathbf{G} \in U$ selected in step 2.

4. Repeat steps 2 and 3 n_R times.
5. Calculate the mean and standard deviation of IC violation percentages.
6. Increase the value of σ_e^2 and go to step 1, until finished.

We apply the above procedure to the following five cases. In all five cases, simulation parameter values are as indicated in Table 4-1.

Table 4-1: Parameter values for randomized simulations

Simulation parameter	Value
Number of samples in each repetition, n_s	50
Number of repetitions, n_R	10^4
Time steps in PRBS excitation experiment, t	50

Table 4-2: Parameter values for randomized simulations

Case-study	Simulation parameter	Value
Section 4.3.1.1	u_{\max}	1
	σ_e	7, 9, 9.5, 10, 11, 13, 15, 17, 19, 21
	c ($\gamma = 0.05$)	17.1, 22, 23.2, 24.5, 26.9, 31.8, 36.7, 41.6, 46.5, 51.4
Section 4.3.1.2	u_{\max}	0.01
	σ_e	1.5, 1.6, 1.65, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3
	c ($\gamma = 0.05$)	3.6, 3.9, 4.0, 4.1, 4.4, 4.6, 4.8, 5.1, 5.3, 5.6
Section 4.3.1.3	u_{\max}	1
	σ_e	0.18, 0.2, 0.22, 0.24, 0.25, 0.26, 0.28, 0.3, 0.32, 0.34, 0.36
	c ($\gamma = 0.05$)	0.5, 0.56, 0.62, 0.67, 0.69, 0.72, 0.78, 0.83, 0.9, 0.95, 1.0
Section 4.3.1.4	u_{\max}	1
	σ_e	0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.66, 0.67, 0.68, 0.69, 0.7, 0.75, 0.8, 0.85
	c ($\gamma = 0.05$)	1.2, 1.4, 1.5, 1.7, 1.8, 2.0, 2.03, 2.06, 2.09, 2.12, 2.15, 2.3, 2.5, 2.6

Table 4-2 continued

Section 4.3.1.5	u_{\max}	1
	σ_e	0.03, 0.033, 0.036, 0.04, 0.045, 0.055, 0.057, 0.058, 0.059, 0.06, 0.063, 0.066, 0.07
	c ($\gamma = 0.05$)	0.1, 0.11, 0.12, 0.13, 0.15, 0.18, 0.19, 0.192, 0.196, 0.2, 0.21, 0.22, 0.23

4.3.1.1 2×2 high-purity distillation column

A high-purity distillation column model studied extensively in literature has a following steady-state gain matrix (Sigurd Skogestad & Morari, 1988)

$$\mathbf{G}_1 = \begin{bmatrix} 87.8 & -86.4 \\ 108.2 & -109.6 \end{bmatrix}. \quad (4.15)$$

Before the procedure outlined above is implemented to test whether violation of eqn. (4.4) results in violation of eqn. (4.1), it is worth developing some insight into satisfaction of IC for different uncertainty sets U , resulting from identification experiments using different inputs \mathbf{m} , namely PRBS and rotated PRBS.

Figure 4-2 shows a simulated identification experiment with PRBS inputs, and resulting highly collinear outputs, owing to the high condition number of \mathbf{G}_1 ($\kappa = 143$). Using data from this experiment, Figure 4-3 shows the resulting estimates and uncertainty regions from least-squares estimation of the entries of \mathbf{G}_1 . Visual inspection of Figure 4-3 reveals that IC is not satisfied by all $\mathbf{G}_1 \in U$ and $\hat{\mathbf{G}}_1$, based on a simple criterion (Koung & MacGregor, 1991) according to which IC is equivalent to the existence of a straight line through the origin that separates the uncertainty regions for \mathbf{g}_1 and \mathbf{g}_2 in different half-planes. Indeed, it is visually obvious that no such line can be found in Figure 4-3.

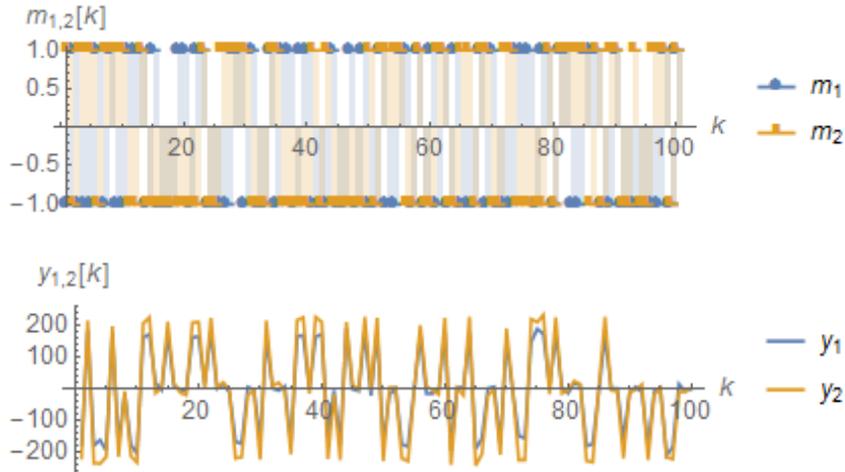


Figure 4-2: Input and output data in simulated identification experiment for G_1 using PRBS inputs.

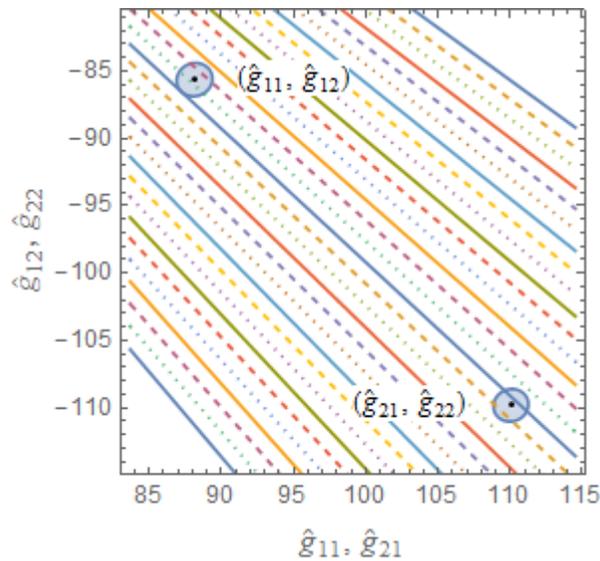


Figure 4-3: Estimates and uncertainty regions resulting from least-squares estimation of the entries of G_1 using the experimental data shown in Figure 4-2. All straight lines go through the origin.

Interestingly enough, while eqn. (4.1) is not satisfied by \hat{G}_1 and all $G_1 \in U$, it is violated by only a very small fraction of $G_1 \in U$, as shown in Figure 4-4. For each vector in Figure 4-4, the point of origin corresponds to the entries (g_{11}, g_{12}) and the head

corresponds to the entries (g_{21}, g_{22}) of $\mathbf{G}_1 \in U$. Green and purple vectors correspond to pairs of $\hat{\mathbf{G}}_1$ and \mathbf{G}_1 for which eqn. (4.1) is or is not satisfied, respectively. Note that only 10 out of 630 instances of \mathbf{G}_1 violate eqn. (4.1), and that all 10 are fairly collinear and close to the boundaries of their respective uncertainty regions.

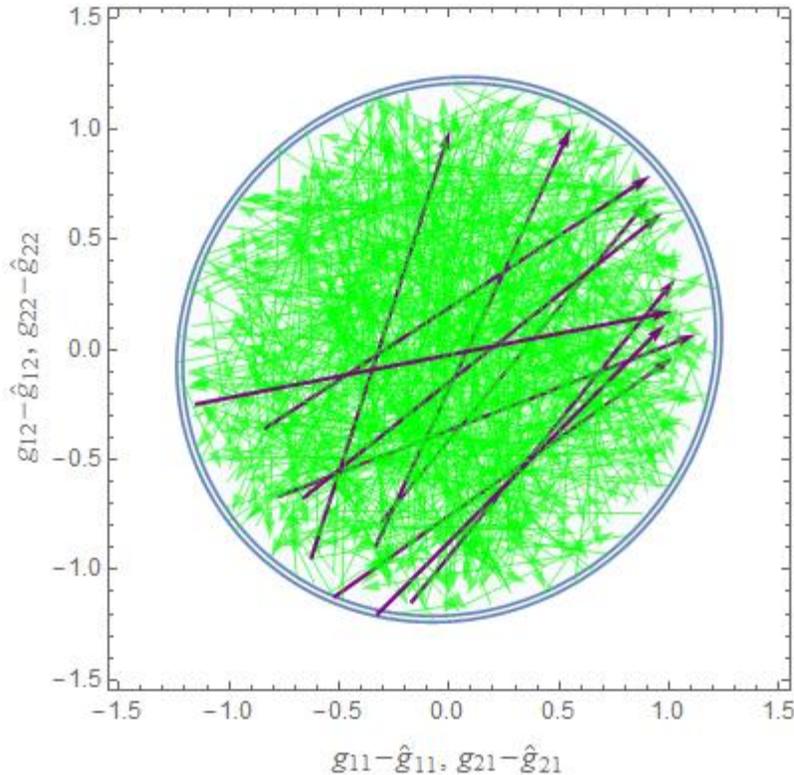


Figure 4-4: Graphical representation of IC satisfaction by $\hat{\mathbf{G}}_1$ and 630 randomly chosen $\mathbf{G}_1 \in U$ shown in Figure 4-3. For each vector, the point of origin corresponds to (g_{11}, g_{12}) and the head corresponds to (g_{21}, g_{22}) . Green and purple vectors correspond to pairs of $\hat{\mathbf{G}}_1$ and \mathbf{G}_1 for which eqn. (4.1) is or is not satisfied, respectively.

Figure 4-5 shows a simulated identification experiment with highly correlated rotated PRBS inputs, and resulting highly uncorrelated outputs, owing to the design based on eqns. (4.7) and (4.8). Using data from this experiment, Figure 4-6 shows the resulting estimates and uncertainty regions from least-squares estimation of the entries of \mathbf{G}_1 .

Visual inspection of Figure 4-6 reveals that IC is satisfied by all $\mathbf{G}_1 \in U$ and $\hat{\mathbf{G}}_1$, based on the simple criterion of Koung and MacGregor (1991) referred to above. Note the sharp difference between the shape of the uncertainty ellipsoids in Figure 4-3 and Figure 4-6.

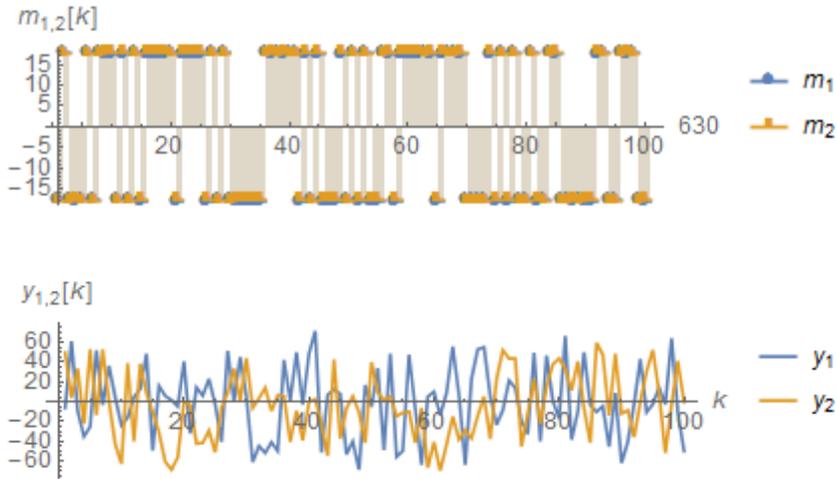


Figure 4-5: Input and output data in simulated identification experiment for \mathbf{G}_1 using rotated PRBS inputs.

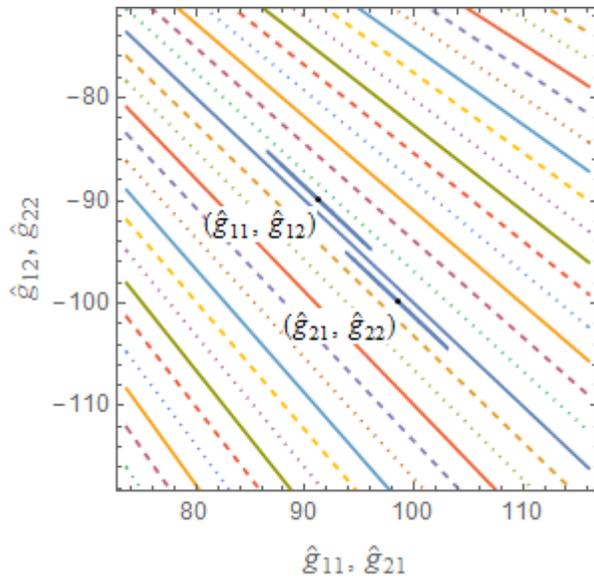


Figure 4-6: Estimates and uncertainty regions resulting from least-squares estimation of the entries of \mathbf{G}_1 using the experimental data shown in Figure 4-5. All straight lines go through the origin.

Turning now to the testing procedure outlined above, Figure 4-7 shows that even a relatively small increase in the value of J above 1 ($J = 1.05$) results in IC violation, which suggests that eqn. (4.4) is not overly conservative.

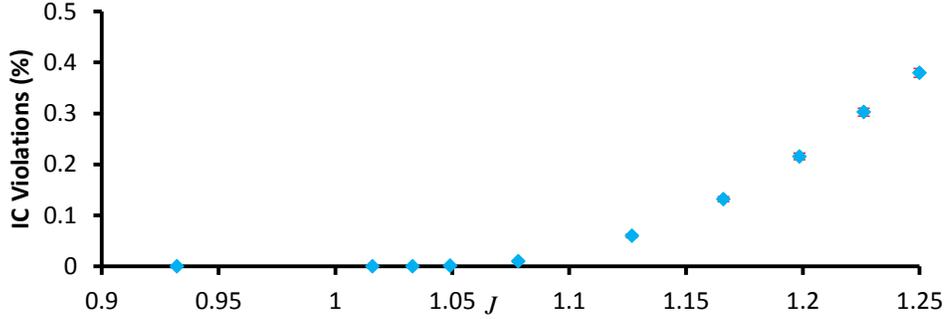


Figure 4-7: Percentage of IC violations varying with J for 2×2 ill-conditioned distillation column

To determine where the (small) conservatism of eqn. (4.4) is coming from, namely whether from eqn. (4.7) or from eqn. (4.9), we check the conservatism of eqn. (4.7) as a weakened form of eqn. (4.6). We do this by plotting the loci of $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ on the complex plane for a large number of $\mathbf{D} \in D$, and check whether the existence of a $\mathbf{D} \in D$ with $|\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})| > 1$ implies the existence of $\mathbf{D} \in D$ with $\text{Re}[\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})] < -1$. The results of 5,000 simulated identification experiments with PRBS and rotated PRBS inputs are shown in Figure 4-8 and Figure 4-9, respectively. As can be seen by inspection of the eigenvalue loci patterns in either Figure, the existence of a $\mathbf{D} \in D$ with $|\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})| > 1$ implies the existence of a $\mathbf{D} \in D$ with $\text{Re}[\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})] < -1$. The preceding observation can be justified as follows: With scaling of \mathbf{D} (depending on noise level and number of experimental data points), the eigenvalues of $\mathbf{D}\hat{\mathbf{G}}^{-1}$ are scaled accordingly. Therefore, if any eigenvalue

$\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ exits the unit disk as a result of scaling of \mathbf{D} , it will be a real eigenvalue, as the smallest of the co-centric circles enveloping all real eigenvalues shown in either Figure 4-8 or Figure 4-9 also encircles all complex eigenvalues, but not vice versa.

Therefore, replacement of eqn. (4.6) by eqn. (4.7) does not seem to introduce any conservatism at all in this case. Whatever conservatism appears in Figure 4-7 appears to be due to eqn. (4.9).

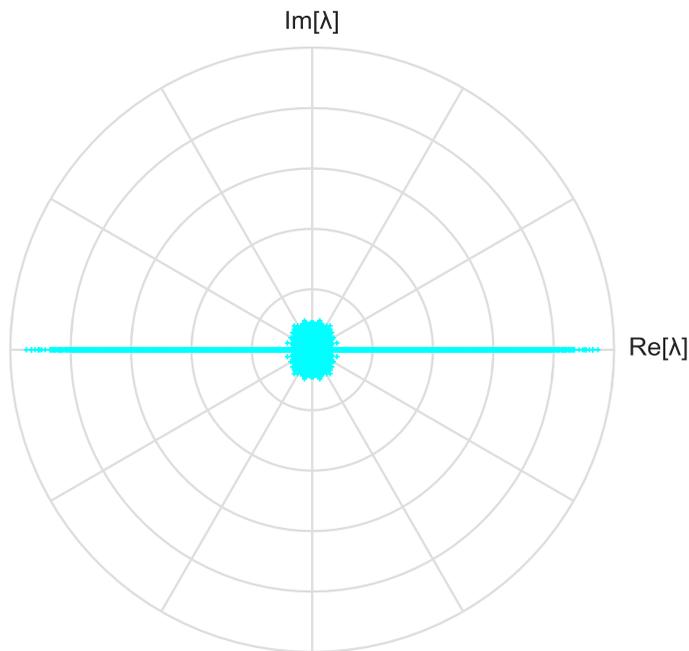


Figure 4-8: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for a 2×2 ill-conditioned distillation column.

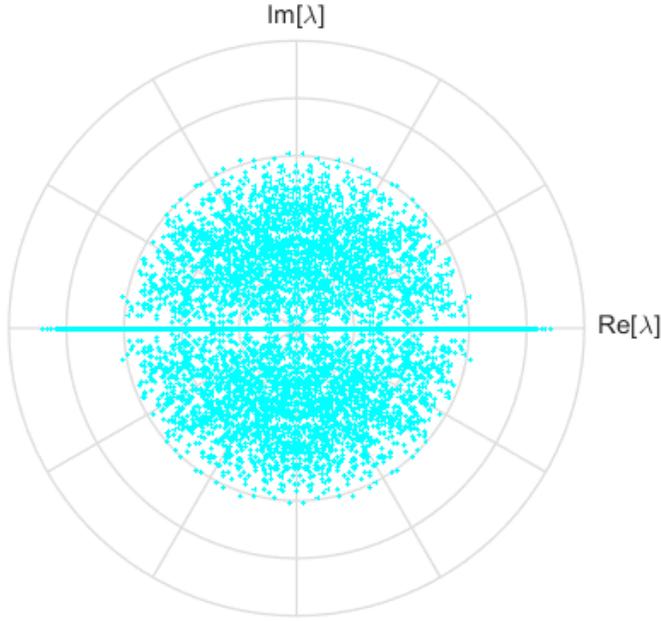


Figure 4-9: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using rotated PRBS inputs for 2×2 ill-conditioned distillation column

4.3.1.2 2×2 low-purity distillation column

A well-conditioned 2×2 distillation column (Wood & Berry, 1973) is

$$\mathbf{G}_2(0) = \begin{bmatrix} 341.3 & -378 \\ 176 & -388 \end{bmatrix}. \quad (4.16)$$

Simulation results for this column are shown in Figure 4-10. The first violation of IC appears around $J = 1.15$, suggesting again that eqn. (4.4) is not overly conservative. As shown in Figure 4-11, it can be concluded here as well that whatever conservatism appears in Figure 4-10 it is due to eqn. (4.9)

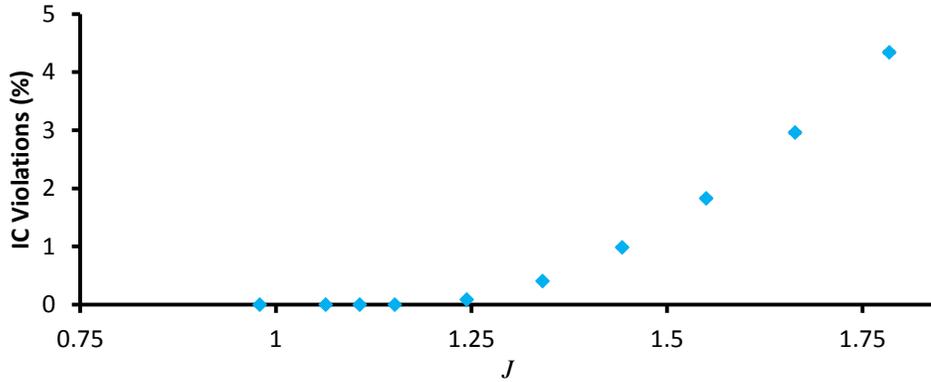


Figure 4-10: Percentage of IC violation varying with J for 2×2 well-conditioned system

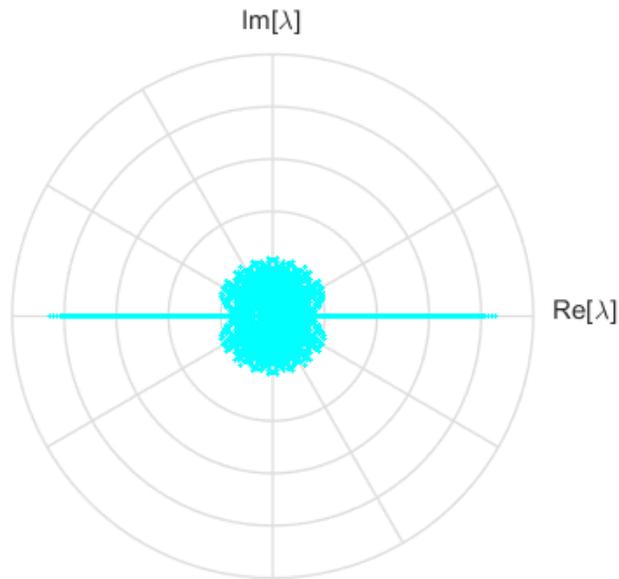


Figure 4-11: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for 2×2 well-conditioned distillation column

4.3.1.3 3×3 distillation column

A 3×3 distillation column with

$$\mathbf{G}_3 = \begin{bmatrix} -0.64 & -0.21 & 1.82 \\ -0.6 & 1.19 & -0.34 \\ 0.55 & -1.12 & 1.14 \end{bmatrix} \quad (4.17)$$

was studied by Hovd and Skogestad (1994). Simulation results for this column are shown in Figure 4-12. The first violation of IC appears around $J = 1.29$, suggesting once more that eqn. (4.4) is not overly conservative, although it is evidently more conservative than for the previous two 2×2 cases. Figure 4-13 once again establishes that the conservatism appearing in Figure 4-12 is due to eqn. (4.9).

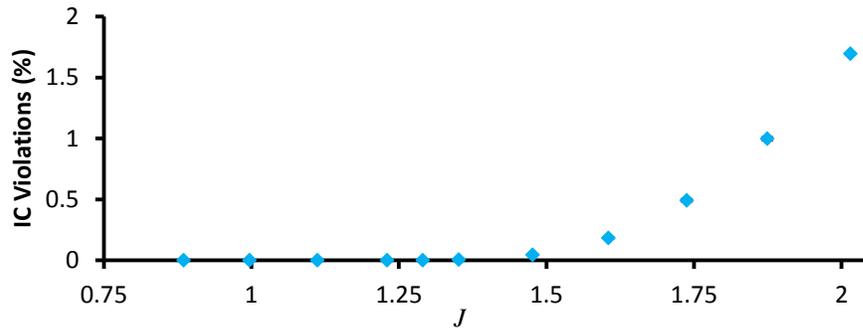


Figure 4-12: Percentage of IC violation varying with J for 3×3 distillation column

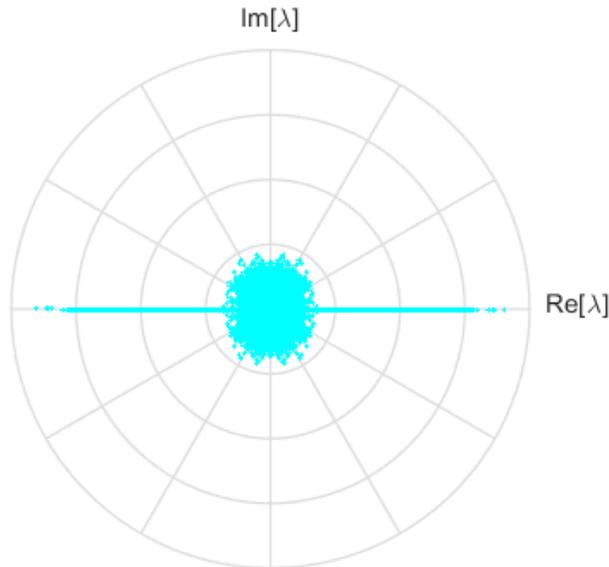


Figure 4-13: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs 3×3 distillation column.

4.3.1.4 4×4 distillation column with side stripper

A 4×4 complex distillation column with side stream stripper is described as

$$\mathbf{G}_4 = \begin{bmatrix} 4.09 & -6.36 & -0.25 & -0.49 \\ -4.17 & 6.93 & -0.05 & 1.53 \\ 1.73 & 5.11 & 4.61 & -5.49 \\ -11.2 & 14 & 0.1 & 4.49 \end{bmatrix} \quad (4.18)$$

(Sigurd Skogestad et al., 1990). Simulation results for this column are shown in Figure 4-14. The first violation of IC appears around $J = 1.56$, suggesting that the conservatism of eqn. (4.4) is higher for this 4×4 system in comparison to the three previous ones. Similar to the three previous cases, Figure 4-15 establishes here that the conservatism appears to be due to eqn. (4.9) as well.

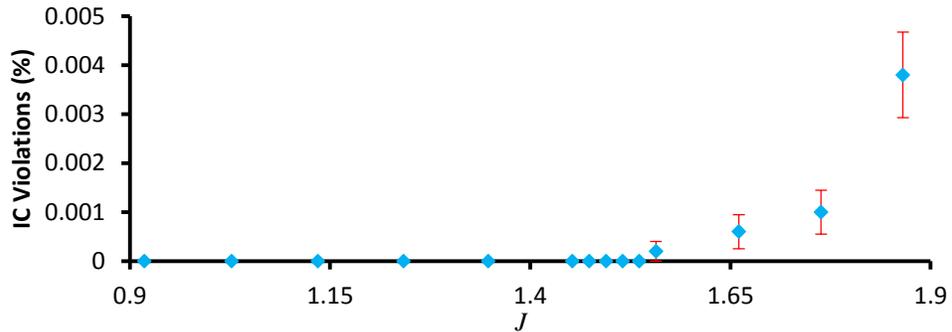


Figure 4-14: Percentage of IC violation varying with J for 4×4 distillation column with side stripper

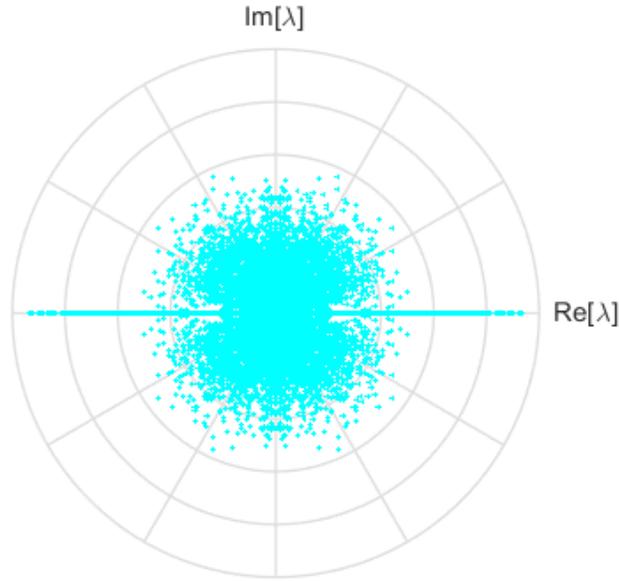


Figure 4-15: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for 4×4 distillation column with side stripper.

4.3.1.5 5×5 fluidized catalytic cracker

A 5×5 fluidized catalytic cracker (FCC) is described as

$$\mathbf{G}_5 = \begin{bmatrix} 0.3866 & 0 & 0 & 0 & 0 \\ 0 & -0.6935 & 0 & 0 & -0.5805 \\ 0.1192 & 1.5461 & 0.5224 & 0 & -0.3667 \\ 0 & -0.1313 & -0.1298 & 0.1058 & -0.2057 \\ 0.0631 & -0.2462 & 0 & 0 & -0.4435 \end{bmatrix} \quad (4.19)$$

(M. L. Darby, 2008). Simulation results for this FCC are shown in Figure 4-16. The first violation of IC appears around $J = 1.69$, suggesting that the conservatism of eqn. (4.4) is higher for this 5×5 system in comparison to the four previous ones. Similar to the four previous cases, Figure 4-17 establishes again that the conservatism appears to be due to eqn. (4.9).

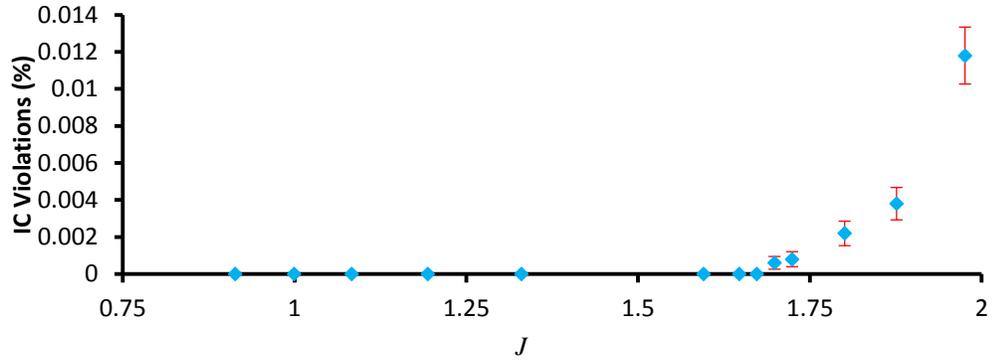


Figure 4-16: Percentage of IC violation varying with J for 5×5 fluidized catalytic cracker (FCC)

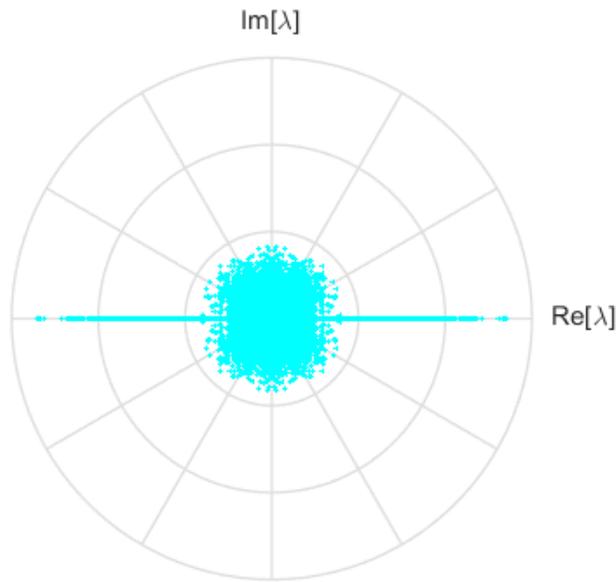


Figure 4-17: Loci of eigenvalues $\lambda(\mathbf{D}\hat{\mathbf{G}}^{-1})$ for least squares estimate $\hat{\mathbf{G}}$ and $\mathbf{D} \in D$ using PRBS inputs for 5×5 fluidized catalytic cracker (FCC).

4.3.2 *Can an alternative model help ensure integral controllability?*

As discussed above, if the standard least-squares estimate $\hat{\mathbf{G}}$ does not satisfy eqn. (4.1) for all $\mathbf{G} \in U$ in eqn. (4.3), it is natural to ask whether a model other than $\hat{\mathbf{G}}$ could be used for that purpose. The rationale behind the preceding question is that perhaps a

model other than $\hat{\mathbf{G}}$, possibly a little farther from the real \mathbf{G} than $\hat{\mathbf{G}}$ is but with the right structure, might satisfy eqn. (4.1) for all $\mathbf{G} \in U$ in eqn. (4.3). We show below that this is generally impossible, by developing a rigorous counter-example for $\mathbf{G} \in \mathbb{R}^{2 \times 2}$.

It is a standard elementary fact that the eigenvalues of a 2×2 matrix are in the right half of the complex plane if and only if both the trace and the determinant of that matrix are positive. To apply this fact to the matrix $\mathbf{G}\hat{\mathbf{G}}^{-1}$ in eqn. (4.1), we follow an approach similar to the one followed by Koung and MacGregor (1991).

Parametrize $\hat{\mathbf{G}} \in \mathbb{R}^2$ as

$$\hat{\mathbf{G}} = \begin{bmatrix} \hat{l}_1 \cos \hat{\theta}_1 & \hat{l}_1 \sin \hat{\theta}_1 \\ \hat{l}_2 \cos \hat{\theta}_2 & \hat{l}_2 \sin \hat{\theta}_2 \end{bmatrix}, \quad (4.20)$$

where $\hat{l}_i > 0$ and $\hat{\theta}_i$, are the length and angle of each nominal row vector $\hat{\mathbf{g}}_i^T$, $i = 1, 2$. Then the real steady state gain matrix $\mathbf{G} \in U$ can be expressed as

$$\mathbf{G} = \begin{bmatrix} n_1 \hat{l}_1 \cos(\hat{\theta}_1 + \alpha_1) & n_1 \hat{l}_1 \sin(\hat{\theta}_1 + \alpha_1) \\ n_2 \hat{l}_2 \cos(\hat{\theta}_2 + \alpha_2) & n_2 \hat{l}_2 \sin(\hat{\theta}_2 + \alpha_2) \end{bmatrix} \triangleq \begin{bmatrix} l_1 \cos \theta_1 & l_1 \sin \theta_1 \\ l_2 \cos \theta_2 & l_2 \sin \theta_2 \end{bmatrix}, \quad (4.21)$$

where $n_i > 0$ and α_i are the multiplicative error in length and additive error in the angle of each row vector \mathbf{g}_i^T , $i = 1, 2$, respectively. Therefore, $\mathbf{G}\hat{\mathbf{G}}^{-1}$ can be expressed as (eqn.

13 of Koung and MacGregor (1991))

$$\mathbf{G}\hat{\mathbf{G}}^{-1} = \frac{1}{\hat{l}_1 \hat{l}_2 \sin(\hat{\theta}_2 - \hat{\theta}_1)} \begin{bmatrix} n_1 \hat{l}_1 \hat{l}_2 \sin(\hat{\theta}_2 - \theta_1) & n_1 \hat{l}_1^2 \sin(\alpha_1) \\ -n_2 \hat{l}_2^2 \sin(\alpha_2) & -n_2 \hat{l}_1 \hat{l}_2 \sin(\theta_2 - \hat{\theta}_1) \end{bmatrix}. \quad (4.22)$$

It follows that

$$\det(\mathbf{G}\hat{\mathbf{G}}^{-1}) = n_1 n_2 \frac{\sin(\theta_2 - \theta_1)}{\sin(\hat{\theta}_2 - \hat{\theta}_1)} \quad (4.23)$$

and

$$\text{tr}(\mathbf{G}\hat{\mathbf{G}}^{-1}) = n_1 \frac{\sin(\hat{\theta}_2 - \theta_1)}{\sin(\hat{\theta}_2 - \hat{\theta}_1)} + n_2 \frac{\sin(\theta_2 - \hat{\theta}_1)}{\sin(\hat{\theta}_2 - \hat{\theta}_1)}. \quad (4.24)$$

If $\det(\mathbf{G}\hat{\mathbf{G}}^{-1}) > 0$ for all $\{\theta_1, \theta_2\}$ corresponding to $\mathbf{G} \in U$, then it immediately follows that $\text{tr}(\mathbf{G}\hat{\mathbf{G}}^{-1}) > 0$ for all $\{\theta_1, \theta_2\}$ corresponding to $\mathbf{G} \in U$.

In case of violation of eqn. (4.1) for some $\mathbf{G} \in U$ given $\hat{\mathbf{G}}$, then IC is not satisfied. In that case, a new model $\hat{\mathbf{X}}$ would be desirable that can make both the determinant and trace of $\mathbf{G}\hat{\mathbf{X}}^{-1}$ positive for all $\mathbf{G} \in U$, while being as close to $\hat{\mathbf{G}}$ as possible. Parametrizing the new model $\hat{\mathbf{X}}$ as

$$\hat{\mathbf{X}} = \begin{bmatrix} \hat{q}_1 \cos \hat{\phi}_1 & \hat{q}_1 \sin \hat{\phi}_1 \\ \hat{q}_2 \cos \hat{\phi}_2 & \hat{q}_2 \sin \hat{\phi}_2 \end{bmatrix} \hat{=} \begin{bmatrix} \hat{\mathbf{x}}_1^T \\ \hat{\mathbf{x}}_2^T \end{bmatrix} \quad (4.25)$$

in analogy to eqn. (4.22), $\mathbf{G}\hat{\mathbf{X}}^{-1}$ can be written as

$$\mathbf{G}\hat{\mathbf{X}}^{-1} = \frac{1}{\hat{q}_1 \hat{q}_2 \sin(\hat{\phi}_2 - \hat{\phi}_1)} \begin{bmatrix} n_1 \hat{l}_1 \hat{q}_2 \sin(\hat{\phi}_2 - \theta_1) & -n_1 \hat{l}_1 \hat{q}_1 \sin(\hat{\phi}_1 - \theta_1) \\ n_2 \hat{l}_2 \hat{q}_2 \sin(\hat{\phi}_2 - \theta_2) & -n_2 \hat{l}_2 \hat{q}_1 \sin(\hat{\phi}_1 - \theta_2) \end{bmatrix} \quad (4.26)$$

It follows that

$$\det(\mathbf{G}\hat{\mathbf{X}}^{-1}) = \frac{n_1 n_2 \hat{l}_1 \hat{l}_2 \sin(\theta_2 - \theta_1)}{\hat{q}_1 \hat{q}_2 \sin(\hat{\phi}_2 - \hat{\phi}_1)} \quad (4.27)$$

and

$$\text{tr}(\mathbf{G}\hat{\mathbf{X}}^{-1}) = \frac{n_1 \hat{l}_1 \hat{q}_2 \sin(\hat{\phi}_2 - \theta_1)}{\hat{q}_1 \hat{q}_2 \sin(\hat{\phi}_2 - \hat{\phi}_1)} + \frac{n_2 \hat{l}_2 \hat{q}_1 \sin(\theta_2 - \hat{\phi}_1)}{\hat{q}_1 \hat{q}_2 \sin(\hat{\phi}_2 - \hat{\phi}_1)}. \quad (4.28)$$

in analogy to eqns. (4.23) and (4.24), respectively.

For IC compliance of the model $\hat{\mathbf{X}}$ with $\mathbf{G} \in U$, both the determinant and trace of the matrix $\mathbf{G}\hat{\mathbf{X}}^{-1}$, eqns. (4.27) and (4.28), respectively, must be positive for $\hat{\mathbf{X}}$ and for all $\mathbf{G} \in U$.

Now, because $n_i > 0$, $\hat{l}_i > 0$, $\hat{q}_i > 0$, $i = 1, 2$, eqns. (4.27) and (4.23) suggest that the sign of $\det(\mathbf{G}\hat{\mathbf{X}}^{-1})$ varies for $\mathbf{G} \in U$ similarly to the sign of the term $\sin(\theta_2 - \theta_1) / \sin(\hat{\phi}_2 - \hat{\phi}_1)$ for angles $\{\theta_2, \theta_1\}$ corresponding to $\mathbf{G} \in U$. Therefore, if $\det(\mathbf{G}\hat{\mathbf{G}}^{-1})$ does not keep the same sign for all $\mathbf{G} \in U$, neither does $\det(\mathbf{G}\hat{\mathbf{X}}^{-1})$, and, consequently, IC is not satisfied.

Furthermore, if it were $\det(\mathbf{G}\hat{\mathbf{G}}^{-1}) > 0$ for all $\mathbf{G} \in U$, then $\text{tr}(\mathbf{G}\hat{\mathbf{G}}^{-1}) > 0$ for all $\mathbf{G} \in U$, as stated above, and no $\hat{\mathbf{X}}$ would be needed, since the model $\hat{\mathbf{G}}$ with $\mathbf{G} \in U$ would be IC-compliant.

In conclusion, if the model $\hat{\mathbf{G}}$ with $\mathbf{G} \in U$ is *not* IC-compliant, then no other model $\hat{\mathbf{X}}$ can be found that is IC-compliant for $\mathbf{G} \in U$. To build a model $\hat{\mathbf{G}}$ that is IC-compliant for $\mathbf{G} \in U$, *it is necessary to rely on identification which employs suitable data*. Hence the need for better design of experiments for identification of IC-compliant models.

4.4 Summary

In developing a mathematical framework for design of experiments that facilitate identification of IC-compliant models, Darby and Nikolaou (2009) relied on a sufficient condition (eqn. (4.4)). In this paper we present a numerical assessment of the conservatism introduced by that sufficient condition, along with some analysis, to pinpoint the sources

of whatever conservatism may be present. In the numerical simulations presented, conservatism turned out not to be overly high, but increasing, nevertheless, as the size of the identified system increases. The corresponding steps that introduced conservatism in the sufficient condition for IC were illustrated. Interestingly enough, the patterns shown in all five case studies (Figure 4-8, Figure 4-9, Figure 4-11, Figure 4-13, Figure 4-15, Figure 4-17) show striking similarities for fairly different systems, suggesting that a fundamental pattern may underlie all such cases. This warrants further investigation in the future.

Finally, it was demonstrated rigorously, by counter-example, that avoiding IC-compliant design of experiments through replacement of a standard least-squares model by another model that might satisfy IC is generally infeasible. Therefore, it appears that research efforts to tame the eigenvalue inequalities (eqn. (4.4)) whose *"main weakness is that they consist of a coupling between the plant model and the true plant, which is highly cumbersome"* (Featherstone & Braatz, 1998) are well warranted.

5 Experiment Design for Control-Relevant Identification

After successfully addressing issues raised in chapter 4, attempts will be made to address second question raised in section 1.2.1 of chapter 1. In this chapter, new theoretical framework will be developed to utilize partial knowledge in designing experiments for control relevant model identification.

5.1 Introduction

A good mathematical model is useful for the design of a good automatic controller, whether the model is used explicitly or implicitly. However, mere proximity between the dynamic behavior of a mathematical model and that of the controlled system does not necessarily imply that the model can be used effectively with a certain controller design method. For control of multivariable systems, a simple classic example is a case in point: Consider a (stable) system with steady-state gain matrix (SSGM) (S. Skogestad & Morari, 1987)

$$\mathbf{G} = \begin{bmatrix} 0.878 & -0.864 \\ 1.082 & -1.096 \end{bmatrix} \quad (5.1)$$

with two alternative models:

$$\hat{\mathbf{G}}_1 = \begin{bmatrix} 0.870 & -0.880 \\ 1.092 & -1.096 \end{bmatrix} \quad (5.2)$$

and

$$\hat{\mathbf{G}}_2 = \begin{bmatrix} 1.054 & -0.691 \\ 1.298 & -0.877 \end{bmatrix}, \quad (5.3)$$

where the respective relative errors for $\hat{\mathbf{G}}_1$ and $\hat{\mathbf{G}}_2$ are approximately 1% and 20% compared to \mathbf{G} . It turns out (see Appendix C for details) that \mathbf{G} can be controlled robustly (i.e. retains closed-loop stability for a wide range of tunings) by a fully decoupling controller that employs the far more inaccurate model $\hat{\mathbf{G}}_2$, rather than $\hat{\mathbf{G}}_1$. In fact, closed-loop stability is problematic for any tuning, when the far more accurate model $\hat{\mathbf{G}}_1$ is used. The underlying reason for this seeming paradox is that $\hat{\mathbf{G}}_2$, along with \mathbf{G} , satisfies the *integral controllability* (IC) condition (Carlos E. Garcia & Manfred Morari, 1985)

$$-\text{Re} \left[\lambda_i \left(\mathbf{G} \hat{\mathbf{G}}^{-1} \right) \right] < 0 \text{ for all } i, \quad (5.4)$$

(where $\lambda_i \left(\mathbf{G} \hat{\mathbf{G}}^{-1} \right)$ is an eigenvalue of $\mathbf{G} \hat{\mathbf{G}}^{-1}$) whereas $\hat{\mathbf{G}}_1$ does not. Consequently, despite its much larger approximation error, $\hat{\mathbf{G}}_2$ is preferable to $\hat{\mathbf{G}}_1$ for the design of the intended multivariable controller.

The preceding realization suggests that generating data for identification of models satisfying the IC condition through deliberately designed experiments requires an approach that departs from standard design of experiments (DOE) and takes IC explicitly into account. A fairly general mathematical framework for such design was developed by Darby and Nikolaou (2009), inspired by pioneering ideas of Koung and MacGregor (1991, 1993, 1994). While these investigations have provided valuable insight (as well as simple recipes, in some occasions, involving *rotated inputs with appropriately proportioned amplitudes*) for DOE that enables efficient identification of IC-compliant models, the focus

of the investigations was entirely on identification of complete models, without assuming any partial prior knowledge. Yet, it is not uncommon to encounter situations where part of a multivariable model to be identified is already known (Abonyi et al., 2000; Johansen, 1996; Kothare et al., 2004; Timmons et al., 1997). For instance, structural knowledge about a model may dictate that a number of elements in a transfer matrix are identically zero, or that certain entries should be trivially equal to each other, or that they add up to zero, and so on. One could reasonably anticipate that incorporation of such partial knowledge into DOE for identification of IC-compliant models would have distinct benefits.

In this work, it is shown that this intuition is indeed correct. That is, explicitly incorporating partial knowledge in DOE for identification of IC-compliant models produces data from which such models can be identified a lot faster than from data generated through DOE that does not take partial knowledge into account. The proposed approach expands substantially on the mathematical framework presented by Darby and Nikolaou (2009). Specifically, we present theoretical results and a numerical optimization framework that enables DOE for efficient identification of IC-compliant models for systems that are partially known at the outset. Through numerical simulations on industrial and literature models, it is demonstrated that the proposed approach results in significant efficiency improvements over standard approaches.

In the rest of chapter, firstly IC and DOE are briefly introduced and then main results are presented. Further, numerical simulations exemplify the main results.

5.2 Background

5.2.1 Integral controllability

The precise formulation of the result on integral controllability proved by García and Morari (1985, Theorem 2) is as follows: Assume that internal model control (IMC) with a diagonal (decoupling) filter matrix $\mathbf{F}(z) = \text{diag} \left\{ \frac{1-\alpha}{1-\alpha z^{-1}} \right\}$ is used to control an $n \times n$ stable, linear, time-invariant system with steady-state input-output behavior $\mathbf{y} = \mathbf{G}\mathbf{m}$, $\mathbf{y}, \mathbf{m} \in \mathbb{R}^n$, $\mathbf{G} \in \mathbb{R}^{n \times n}$. Then, there exists an $\alpha^* \in [0, 1)$ such that the closed loop remains stable for all $\alpha \in [\alpha^*, 1)$ if and only if the matrix $\mathbf{G}\hat{\mathbf{G}}^{-1}$ (where \mathbf{G} and $\hat{\mathbf{G}}$ are the actual and estimated steady-state gain matrices, respectively) has eigenvalues $\lambda(\mathbf{G}\hat{\mathbf{G}}^{-1})$ in the right half of the complex plane, i.e. satisfies the IC condition, eqn. (5.4).

This result establishes the achievable robustness of decoupling multivariable controllers with integral action.

5.2.2 Uncertainty description

Since \mathbf{G} is not known exactly, eqn. (5.4) must be satisfied for all \mathbf{G} in an uncertainty set U . Assuming that $\hat{\mathbf{G}}$ is the outcome of standard least-squares identification using data over T_f time steps, the uncertainty set U for \mathbf{G} can be defined as the standard ellipsoidal uncertainty set

$$U = \left\{ \mathbf{G} = \begin{bmatrix} \mathbf{g}_1^T \\ \vdots \\ \mathbf{g}_n^T \end{bmatrix} \in \mathbb{R}^{n \times n} : (\mathbf{g}_i^T - \hat{\mathbf{g}}_i^T) \mathbf{M}^T \mathbf{M} (\mathbf{g}_i - \hat{\mathbf{g}}_i) \leq c^2, i = 1, \dots, n \right\}, \quad (5.5)$$

where the information matrix $\mathbf{M}^T\mathbf{M}$ results from the input matrix $\mathbf{M} \in \mathbb{R}^{T_f \times n}$ and

$$c^2 = s^2 n F_{1-\gamma}(n, T_f - n) \approx \sigma_{\text{noise}}^2 \chi_{1-\gamma}^2(n) \quad (5.6)$$

for confidence level γ , and for the F-distribution with $(n, T_f - n)$ degrees of freedom or, approximately for $T_f \gg n$ for the chi-square distribution with n degrees of freedom.

5.2.3 Ensuring integral controllability

Assessing whether eqn. holds for all \mathbf{G} in the set U defined in eqn. (5.5) is not trivial. Worse yet, the IC condition involves eigenvalue inequalities whose "main weakness is that they consist of a coupling between the plant model and the true plant, which is highly cumbersome" for design of IC-compliant identification experiments, in that the plant inputs to be selected by DOE do not appear in that inequality (Featherstone & Braatz, 1998). To remedy that problem, Darby and Nikolaou (Darby & Nikolaou, 2009) proved that the IC inequality in eqn. (5.7) is satisfied for all \mathbf{G} in the set U defined in eqn. (5.5) if

$$J_u \triangleq \sum_{k=1}^n c \frac{\|\hat{\mathbf{u}}_k\|_1}{\hat{\sigma}_k} \sqrt{\hat{\mathbf{v}}_k^T (\mathbf{M}^T \mathbf{M})^{-1} \hat{\mathbf{v}}_k} < 1, \quad (5.7)$$

where

$$\hat{\mathbf{G}} = \sum_{k=1}^n \hat{\sigma}_k \hat{\mathbf{u}}_k \hat{\mathbf{v}}_k^T \quad (5.8)$$

is the singular-value decomposition (SVD) of the steady-state gain matrix of the identified model.

The advantage of eqn. (5.7), compared to eqn. (5.4), is that it does not include the uncertain matrix \mathbf{G} directly, whereas the manipulated inputs appear directly and explicitly

in the matrix $\mathbf{M}^T\mathbf{M}$. Therefore, eqn. (5.7) can be used (as a constraint or objective) in DOE much more easily than eqn. (5.4), as was demonstrated by the theoretical and numerical results developed in Darby and Nikolaou (Darby & Nikolaou, 2009). In fact, in some cases it leads to analytical results that are in full agreement with recipes that have appeared in literature, and in other cases it leads to novel such recipes of similar simplicity (Darby & Nikolaou, 2009). Such recipes rely on the idea of *rotated inputs*, with *appropriately proportioned amplitudes* (Koung & Macgregor, 1993).

Note also that, while the inequality in eqn. (5.7) is only sufficient for IC, numerical tests in the previous chapter suggest that it is not overly conservative (Panjwani & Nikolaou, 2016).

5.2.4 Design of experiments for identification of models satisfying integral controllability

To design an identification experiment numerically based on eqn. (5.7), a standard approximation of the information matrix $\mathbf{M}^T\mathbf{M}$ can first be considered in terms of the input covariance matrix, \mathbf{C}_m , and duration of the experiment, T_f , as $\mathbf{M}^T\mathbf{M} \approx (T_f - 1)\mathbf{C}_m$; then, \mathbf{C}_m can be parametrized in terms of a triangular matrix \mathbf{Q} through a Cholesky factorization $\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T$; finally, eqn. (5.7) can be used either as a constraint, or the left-hand side of eqn. (5.7),

$$\beta = \sum_{k=1}^n \frac{\|\hat{\mathbf{u}}_k\|_1}{\hat{\sigma}_k} \sqrt{\hat{\mathbf{v}}_k^T (\mathbf{Q}\mathbf{Q}^T)^{-1} \hat{\mathbf{v}}_k}, \quad (5.9)$$

can be used as an objective to be minimized in an optimization problem that produces \mathbf{Q} and from that the corresponding input covariance matrix \mathbf{C}_m .

5.2.5 Adaptive design

The inequality in eqn. (5.7) characterizes the optimal inputs, dictated by DOE, in terms of the SVD of the estimate $\hat{\mathbf{G}}$, eqn. (5.8), which results from the identification experiment. However, because the identification experiment has not yet been conducted, and consequently the estimate $\hat{\mathbf{G}}$ is not yet known at the time of the design, optimal input *design* – rather than mere *characterization* – requires some sort of an adaptive approach. This was already proposed by M. L Darby (2008), was later modified by Kulkarni (2012), and was found to work quite satisfactorily on a number of simulations (S. Misra & Nikolaou, 2015). The general idea of the adaptive approach is as follows.

1. Develop a preliminary model from input-output data using standard PRBS inputs for limited time.
2. Based on the available model, perform DOE that complies with IC (e.g. by minimizing β , subject to input and/or output constraints).
3. Implement the inputs determined in the above step for limited time, and collect input-output data, to update the model.
4. If the updated model does not satisfy eqn. (5.7) (therefore IC), go to step 2. Else, stop.

Details can be found in Appendix C.

5.3 Mathematical Problem Formulation and Main Results

In this section we first formulate mathematically the problem at hand, and then explain how this formulation leads to a numerical solution.

We consider that the steady-state gain matrix \mathbf{G} of the system to be identified is partially known, in terms of linear equality constraints. A typical kind of such constraints arises when matrix elements are known a priori to be equal to zero, because pertinent inputs are known to have no effect on related outputs. Another kind arises from fundamental balance equations. Examples of each kind are discussed in the Case Studies section.

From a mathematical viewpoint, linear equality constraints capturing partial process knowledge involve elements either from individual rows of \mathbf{G} or from multiple rows of \mathbf{G} . It turns out that each of these two classes of constraints needs to be handled differently mathematically.

The linear constraints mentioned above are certainly just one form of partial knowledge available. Other forms of partial knowledge include linear inequality constraints, nonlinear constraints, constraints for the full dynamic model, and so on. While all such possibilities are worth examining – both from a theoretical and from a practical viewpoint – the possibility examined here, namely linear equality constraints on the steady-state gain matrix, is not uncommon and is of practical interest. At the same time, it is challenging enough to warrant investigation on its own right. Therefore, we quickly summarize linear-equality constrained least-squares next, and proceed to develop our main results afterwards.

5.3.1 Least-squares identification for partially known systems

Consider an $n \times n$ multivariable system around a nominal steady state, and steady-state deviations of inputs and outputs $\mathbf{m}, \mathbf{y} \in \mathbb{R}^n$, respectively. Assume that the steady-state input-output behavior of the system is described by the equation

$$\mathbf{y}(t+1) = \mathbf{G}\mathbf{m}(t) + \mathbf{e}(t), \quad (5.10)$$

where $\mathbf{G} \triangleq [\mathbf{g}_1 \dots \mathbf{g}_n]^T \in \mathbb{R}^{n \times n}$; t is discrete time; and $\mathbf{e}(t)$ is Gaussian white noise with zero mean and covariance matrix $\sigma_e^2 \mathbf{I}$. Note that the sampling period has been assumed to be long enough, to eliminate transient dynamics in eqn. (10). This assumption is reasonable for most plants, which have a fairly smooth low-pass type of frequency response without pronounced resonance frequencies, and makes the problem manageable enough to allow for a realistic solution to be obtained with reasonable effort (Koung & MacGregor, 1991, 1993, 1994). The more general case of a full dynamic model will be examined in the future, along the lines of Darby and Nikolaou (2014).

5.3.1.1 Identification with linear equality constraints for each individual row of \mathbf{G} .

Assume that there is partial knowledge about \mathbf{G} in terms of equality constraints for each row \mathbf{g}_i^T of \mathbf{G} , namely

$$\mathbf{H}_i \mathbf{g}_i = \mathbf{h}_i, \quad i = 1, \dots, n, \quad (5.11)$$

where $\mathbf{H}_i \in \mathbb{R}^{n_i \times n}$, $\mathbf{h}_i \in \mathbb{R}^{n_i}$ with n_i equal to the number of equality constraints for that row. Then, the identification problem can be viewed as a collection of n multi-input-single output (MISO) problems. The least-squares estimate $\hat{\mathbf{g}}_i$ of \mathbf{g}_i for the model of eqn. (5.10) subject to the equality constraints in eqn. (5.11) is (Seber & Lee, 2003)

$$\begin{aligned} \hat{\mathbf{g}}_i &= (\mathbf{M}^T \mathbf{M})^{-1} \left(\mathbf{M}^T \mathbf{y}_i + \mathbf{H}_i^T \left[\mathbf{H}_i (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{H}_i^T \right]^{-1} \left[\mathbf{h}_i - \mathbf{H}_i (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{y}_i \right] \right) \\ &= \hat{\mathbf{g}}_{i, \text{LS}} + (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{H}_i^T \left[\mathbf{H}_i (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{H}_i^T \right]^{-1} (\mathbf{h}_i - \mathbf{H}_i \hat{\mathbf{g}}_{i, \text{LS}}) \end{aligned} \quad (5.12)$$

and the ellipsoidal uncertainty region of \mathbf{g}_i for confidence level γ is

$$(\mathbf{g}_i - \hat{\mathbf{g}}_i)^T \mathbf{A}_i (\mathbf{g}_i - \hat{\mathbf{g}}_i) \leq r_A^2, \quad i=1, \dots, n, \quad (5.13)$$

where

$$\mathbf{M} = [\mathbf{m}(0) \quad \dots \quad \mathbf{m}(T-1)]^T, \quad (5.14)$$

$$\mathbf{y}_i = [y_i(1) \quad \dots \quad y_i(T)]^T, \quad i=1, \dots, n, \text{ and} \quad (5.15)$$

$$\hat{\mathbf{g}}_{i, \text{LS}} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{y}_i, \quad i=1, \dots, n \quad (5.16)$$

is the unconstrained least-squares estimate of \mathbf{g}_i ;

$$r_A^2 \approx \sigma_e^2 \chi_{1-\gamma}^2(n), \quad (5.17)$$

$$\mathbf{A}_i = \mathbf{R}_{1,i} \mathbf{D}_{1,i}^{-1} \mathbf{R}_{1,i}^T, \quad i=1, \dots, n \quad (5.18)$$

and $\mathbf{D}_{1,i}$ is the diagonal matrix of non-zero singular values and $\mathbf{R}_{1,i}$ is the matrix of corresponding singular vectors in the singular value decomposition of the matrix

$$(\mathbf{M}_i^T \mathbf{M}_i)^{-1} \triangleq (\mathbf{M}^T \mathbf{M})^{-1} \left(\mathbf{I}_n - \mathbf{H}_i^T [\mathbf{H}_i (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{H}_i^T]^{-1} \mathbf{H}_i (\mathbf{M}^T \mathbf{M})^{-1} \right) = [\mathbf{R}_{1,i} \quad \mathbf{R}_{2,i}] \begin{bmatrix} \mathbf{D}_{1,i} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{1,i}^T \\ \mathbf{R}_{2,i}^T \end{bmatrix}. \quad (5.19)$$

In terms of the input covariance matrix \mathbf{C}_m and total experimentation time T_f , eqn.

(5.19) can be rewritten as

$$(\mathbf{M}_i^T \mathbf{M}_i)^{-1} \approx \frac{1}{T_f - 1} \mathbf{C}_m^{-1} \left(\mathbf{I}_n - \mathbf{H}_i^T [\mathbf{H}_i \mathbf{C}_m^{-1} \mathbf{H}_i^T]^{-1} \mathbf{H}_i \mathbf{C}_m^{-1} \right) \triangleq \frac{1}{T_f - 1} \mathbf{C}_{m,i}^{-1}, \quad i=1, \dots, n. \quad (5.20)$$

5.3.1.2 Identification with linear equality constraints relating multiple rows of \mathbf{G}

Since there are constraints relating multiple rows of the matrix \mathbf{G} , a different formulation of the constrained estimation problem is needed, to include the constraints in

the least-squares problem. To accomplish this, one can vectorize \mathbf{G} and proceed with its identification as a full multi-input-multi-output (MIMO) model, as follows:

The input-output system described by eqn. (5.10) can be rewritten as

$$\mathbf{y}(t+1) = \mathbf{W}(t)^T \text{vec}(\mathbf{G}^T) + \mathbf{e}(t), \quad (5.21)$$

where

$$\text{vec}(\mathbf{G}^T) \triangleq [\mathbf{g}_1^T \quad \dots \quad \mathbf{g}_n^T]^T \in \mathbb{R}^{n^2} \text{ and} \quad (5.22)$$

$$\mathbf{W}(t)^T = \begin{bmatrix} \mathbf{m}(t)^T & \mathbf{0} & \vdots \\ \mathbf{0} & \ddots & \mathbf{0} \\ \vdots & \mathbf{0} & \mathbf{m}(t)^T \end{bmatrix} \triangleq \mathbf{I}_n \otimes \mathbf{m}(t)^T \in \mathbb{R}^{n \times n^2} \quad (5.23)$$

with \otimes denoting the standard Kronecker product.

Partial knowledge about \mathbf{G} in terms of equality constraints relating multiple rows of the matrix \mathbf{G} can be written as

$$\mathbf{K} \text{vec}(\mathbf{G}^T) = \mathbf{k}, \quad (5.24)$$

where $\mathbf{K} \in \mathbb{R}^{p \times n^2}$, $\mathbf{k} \in \mathbb{R}^p$, with p equal to the total number of linear equality constraints.

The least-squares estimate $\text{vec}(\hat{\mathbf{G}}^T) \triangleq [\hat{\mathbf{g}}_1^T \quad \dots \quad \hat{\mathbf{g}}_n^T]^T$ of $\text{vec}(\mathbf{G}^T)$ for the model of eqn. (5.21) subject to the equality constraints in eqn. (5.24) is (Seber & Lee, 2003)

$$\begin{aligned} \text{vec}(\hat{\mathbf{G}}^T) &= (\mathbf{X}^T \mathbf{X})^{-1} \left(\mathbf{X}^T \mathbf{Y} + \mathbf{K}^T \left[\mathbf{K} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{K}^T \right]^{-1} \left[\mathbf{k} - \mathbf{K} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \right] \right) \\ &= \text{vec}(\hat{\mathbf{G}}_{\text{LS}}^T) + (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{K}^T \left[\mathbf{K} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{K}^T \right]^{-1} \left[\mathbf{k} - \mathbf{K} \text{vec}(\hat{\mathbf{G}}_{\text{LS}}^T) \right] \end{aligned} \quad (5.25)$$

and the ellipsoidal uncertainty region of $\text{vec}(\mathbf{G}^T)$ for confidence level γ is

$$\left(\text{vec}(\mathbf{G}^T) - \text{vec}(\hat{\mathbf{G}}^T) \right)^T \mathbf{B} \left(\text{vec}(\mathbf{G}^T) - \text{vec}(\hat{\mathbf{G}}^T) \right) \leq r_{\mathbf{B}}^2, \quad (5.26)$$

where

$$\mathbf{X} = \begin{bmatrix} \mathbf{W}(0)^T \\ \vdots \\ \mathbf{W}(T_f - 1)^T \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n \otimes \mathbf{m}(0)^T \\ \vdots \\ \mathbf{I}_n \otimes \mathbf{m}(T_f - 1)^T \end{bmatrix} \in \mathbb{R}^{n \cdot T_f \times n^2}, \quad (5.27)$$

$$\tilde{\mathbf{y}} = [\mathbf{y}^T(1) \ \cdots \ \mathbf{y}^T(T_f)]^T \in \mathbb{R}^{n \cdot T_f}, \quad (5.28)$$

$$\mathbf{X}^T \mathbf{X} = \mathbf{I}_n \otimes (\mathbf{M}^T \mathbf{M}) \in \mathbb{R}^{n^2 \times n^2}, \text{ and} \quad (5.29)$$

$$\text{vec}(\hat{\mathbf{G}}_{\text{LS}}^T) = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \tilde{\mathbf{y}} \quad (5.30)$$

is the unconstrained least-squares estimate of $\text{vec}(\mathbf{G}^T)$,

$$r_{\mathbf{B}}^2 \approx \sigma_e^2 \chi_{1-\gamma}^2(n^2), \quad (5.31)$$

$$\mathbf{B} = \tilde{\mathbf{R}}_1 \tilde{\mathbf{D}}_1^{-1} \tilde{\mathbf{R}}_1^T \quad (5.32)$$

and $\tilde{\mathbf{D}}_1$, $\tilde{\mathbf{R}}_1$ are the diagonal matrix of non-zero singular values and matrix of corresponding singular vectors, respectively, obtained from singular value decomposition of the symmetric matrix.

$$\mathbf{C} \triangleq (\mathbf{X}^T \mathbf{X})^{-1} \left(\mathbf{I}_{n^2} - \mathbf{K}^T [\mathbf{K} (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{K}^T]^{-1} \mathbf{K} (\mathbf{X}^T \mathbf{X})^{-1} \right) = [\tilde{\mathbf{R}}_1 \ \tilde{\mathbf{R}}_2] \begin{bmatrix} \tilde{\mathbf{D}}_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{R}}_1^T \\ \tilde{\mathbf{R}}_2^T \end{bmatrix} \quad (5.33)$$

with \mathbf{I}_k the identity matrix of dimensions $k \times k$.

Eqn. (5.33) can be rewritten in terms of the input covariance matrix \mathbf{C}_m and total experimentation time T_f as

$$\mathbf{C} \approx \mathbf{I}_n \otimes \frac{1}{T_f - 1} \mathbf{C}_m^{-1} \left(\mathbf{I}_{n^2} - \mathbf{K}^T [\mathbf{K} (\mathbf{I}_n \otimes \mathbf{C}_m^{-1}) \mathbf{K}^T]^{-1} \mathbf{K} (\mathbf{I}_n \otimes \mathbf{C}_m^{-1}) \right). \quad (5.34)$$

5.3.2 Ensuring IC in identification of partially known models

As already mentioned above, the purpose of DOE is to determine inputs that will generate data enabling the identification of an IC-compliant model, i.e. a model $\hat{\mathbf{G}}$, determined by eqns. (5.12) or (5.25), which satisfies eqn. (5.4) along with the real system \mathbf{G} constrained in eqns. (5.13) or (5.26) respectively.

As has been argued already (Darby & Nikolaou, 2009; Featherstone & Braatz, 1998), directly relying on eqn. (5.4) results in an overly complicated DOE task. To circumvent that difficulty, in the following theorems we develop inequalities that guarantee satisfaction of eqn. (5.4) yet are much simpler in that they

- (a) directly involve the inputs to be determined by DOE, and
- (b) do not contain the uncertain matrix \mathbf{G} .

Consequently, these inequalities are much easier to use in DOE for identification of IC-compliant models, as will be explained afterwards.

Theorem 1 – Sufficient condition for IC of model identified as multiple, partially known MISO models

A model $\hat{\mathbf{G}}$, partially known through eqn. (5.11) and identified according to eqn. (5.12), satisfies IC for all potential \mathbf{G} in the set

$$D = \left\{ \mathbf{G} = [\mathbf{g}_1 \dots \mathbf{g}_n]^T \in \mathbb{R}^{n \times n} : (\mathbf{g}_k - \hat{\mathbf{g}}_k)^T \mathbf{A}_k (\mathbf{g}_k - \hat{\mathbf{g}}_k) \leq r_A^2, 1 \leq k \leq n \right\} \quad (5.35)$$

suggested by eqn. (5.13), where \mathbf{A}_k and r_A are as in eqns.(5.18) - (5.20)-, if

$$J_c \triangleq \sum_{k=1}^n r_A \frac{\|\hat{\mathbf{u}}_k\|_1}{\hat{\sigma}_k} \sqrt{\hat{\mathbf{v}}_k^T \mathbf{A}_k^{-1} \hat{\mathbf{v}}_k} < 1. \quad (5.36)$$

Proof: Placing the matrix \mathbf{A}_k in place of $\mathbf{M}^T\mathbf{M}$ in Theorem 1 in Darby and Nikolaou (Darby & Nikolaou, 2009), the proof of this Theorem follows the exact same pattern, and is omitted for brevity.

Note the similarity between the sets D in eqn. (5.35) and U in eqn. (5.5). Note also that, in contrast to eqn. (5.4), the above eqn. (5.36) in Theorem 1 directly involves plant inputs through the matrix \mathbf{A}_k , which, as eqns. (5.18) and (5.19) indicate, is a function of both the data matrix, \mathbf{M} , and the constraint matrix, \mathbf{H}_k . In addition, the uncertain matrix \mathbf{G} , which is present in eqn. (5.4), has been eliminated in eqn. (5.36).

The above Theorem 1 is applicable to those cases where individual rows of the plant steady-state gain matrix \mathbf{G} are not related to each other through constraining equalities. However, as already argued, situations with constraints involving multiple rows of \mathbf{G} are not uncommon. The subsequent Theorem 2, following the intermediate Lemma 1 in Appendix C, is developed for such situations.

Theorem 2 – Sufficient condition for IC of model identified as a single, partially known MIMO model

A model $\hat{\mathbf{G}}$, partially known through eqn. (5.24) and identified according to eqn. (5.25), satisfies IC for all potential \mathbf{G} in the set

$$\tilde{D} = \left\{ \mathbf{G} = [\mathbf{g}_1 \dots \mathbf{g}_n]^T \in \mathbb{R}^{n \times n} : \left(\text{vec}(\mathbf{G}^T) - \text{vec}(\hat{\mathbf{G}}^T) \right)^T \mathbf{B} \text{vec}(\text{vec}(\mathbf{G}^T) - \text{vec}(\hat{\mathbf{G}}^T)) \leq r_{\mathbf{B}}^2 \right\} \quad (5.37)$$

suggested by eqn. (5.26), where \mathbf{B} and $r_{\mathbf{B}}$ given in eqns. (5.32) - (5.34), if

$$\mu_{\max}(\mathbf{B}^{-1}\Phi) < \frac{1}{nr_{\mathbf{B}}^2}, \quad (5.38)$$

where $\Phi = \mathbf{I}_n \otimes (\hat{\mathbf{G}}^T \hat{\mathbf{G}})^{-1}$, and $\mu_{\max}(\mathbf{B}^{-1}\Phi)$ is the largest eigenvalue of the matrix $\mathbf{B}^{-1}\Phi$.

Proof: See Appendix C

Theorem 1 and Theorem 2 immediately suggest how DOE for identification of an IC-compliant model can be formulated as an optimization problem. The optimization resulting from Theorem 1 is not far from the formalism in Darby and Nikolaou (Darby & Nikolaou, 2009). By contrast, Theorem 2 departs significantly from that formalism and is more challenging. Therefore, in the following we will first discuss briefly the DOE implications of Theorem 1, and will concentrate more on discussing the DOE implications of Theorem 2.

5.3.3 Design of experiments for IC-compliant model to be identified as multiple, partially known MISO models

Theorem 1 suggests that DOE for identification of an IC-compliant model as in eqn. (5.10), with partial model knowledge captured in constraints as in eqn. (5.11), can be carried out using eqn. (5.36), as follows. Find

$$\mathbf{C}_m^{\text{opt}} = \arg \left(\min_{\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T} \sum_{k=1}^n r_A \frac{\|\hat{\mathbf{u}}_k\|_1}{\hat{\sigma}_k} \sqrt{\hat{\mathbf{v}}_k^T \mathbf{A}_k^{-1} \hat{\mathbf{v}}_k} \right) \quad (5.39)$$

subject to variance constraints on individual outputs and inputs, such as

$$\text{var}(\mathbf{m}_i) = [\mathbf{C}_m]_{kk} \leq M_k^2, \quad k = 1, \dots, n \quad \text{and} \quad (5.40)$$

$$\text{var}(\mathbf{y}_i) = [\hat{\mathbf{G}}\mathbf{C}_m\hat{\mathbf{G}}^T]_{kk} \leq Y_k^2, \quad k = 1, \dots, n, \quad (5.41)$$

where the input covariance matrix is parametrized through a Cholesky factorization as $\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T$. Other kinds of constraints, in place of eqns. (5.40) and (5.41) may also be considered, as discussed in (Darby & Nikolaou, 2009).

It should be noted, that eqn. (5.36) rigorously *characterizes* (rather than *prescribes*) the optimal inputs under the respective set of identification constraints, since it involves the elements $\hat{\mathbf{u}}_k, \hat{\mathbf{v}}_k, \hat{\sigma}_k$ of the SVD of the identified model $\hat{\mathbf{G}}$. As already mentioned above, an adaptive DOE can address this issue, as follows.

1. Develop a preliminary model from input-output data using standard PRBS inputs for limited time.
2. Based on the available model, perform DOE that complies with IC by minimizing J_c , eqn. (5.36), subject to input and/or output constraints.
3. Implement the inputs determined in the above step for limited time, and collect input-output data, to update the model.
4. If the updated model does not satisfy eqn. (5.36) (therefore IC), go to step 2. Else, stop.

5.3.4 Design of experiments for IC-compliant model to be identified as a single, partially known MIMO model

Theorem 2 suggests that DOE for identification of an IC-compliant model as in eqn.(5.24), with partial model knowledge captured in constraints as in eqn. (5.21), can be carried out using eqn. (5.38), as follows: Find

$$\mathbf{C}_m^{\text{opt}} = \arg \left(\min_{\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T} \mu_{\max}(\mathbf{B}^{-1}\mathbf{\Phi}) \right) \quad (5.42)$$

subject to variance constraints on outputs and inputs, as in eqn. (5.40) or (5.41).

While the task in eqn. (5.42) is challenging, we explain below that minimization of an upper bound on $\mu_{\max}(\mathbf{B}^{-1}\mathbf{\Phi})$ is manageable in realistic situations. In fact the analysis that follows establishes that, in the absence of partial knowledge about the model, DOE

based on rotated inputs emerges naturally as the result of minimizing that upper bound on $\mu_{\max}(\mathbf{B}^{-1}\Phi)$, subject to various constraints on inputs and outputs. The corresponding results are summarized in Theorem 3 and Theorem 4, after the related context is set first.

5.3.5 Design of experiments for IC-compliant model to be identified as a single MIMO model.

In standard linear regression, absent any partial knowledge on \mathbf{G} , the matrix \mathbf{B} in eqn. (5.37) is

$$\mathbf{B} = \mathbf{I} \otimes \mathbf{M}^T \mathbf{M}, \quad (5.43)$$

where $\mathbf{M}^T \mathbf{M}$ is the information matrix containing the inputs that must be designed, and can be diagonalized as

$$\mathbf{M}^T \mathbf{M} = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T. \quad (5.44)$$

Given this, it can be shown (Appendix C) that the DOE problem in eqn. (5.42) can be cast as the following optimization problem: Find

$$\min_{\mathbf{\Lambda}} \left(\min_{\mathbf{Q}=\hat{\mathbf{V}}^T \mathbf{P}} \text{tr}(\mathbf{Q} \mathbf{\Lambda}^{-2} \mathbf{Q}^T \hat{\mathbf{\Sigma}}^{-4}) \right) \quad (5.45)$$

with respect to the *orthonormal* matrix $\mathbf{Q} = \hat{\mathbf{V}}^T \mathbf{P}$ and the *diagonal* matrix $\mathbf{\Lambda}$, subject to pertinent input and output constraints. More specifically, such constraints can be

$$(T_f - 1) \text{var}(m_k) \approx [\mathbf{M}^T \mathbf{M}]_{kk} = [\mathbf{P} \mathbf{\Lambda} \mathbf{P}^T]_{kk} \leq (T_f - 1) M_k^2, \quad k = 1, \dots, n, \quad (5.46)$$

$$(T_f - 1) \text{var}(y_k) \approx [\mathbf{Y}^T \mathbf{Y}]_{kk} = [\hat{\mathbf{T}} \hat{\mathbf{V}}^T \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T \hat{\mathbf{V}} \hat{\mathbf{T}}^T]_{kk} \leq (T_f - 1) Y_k^2, \quad k = 1, \dots, n, \quad (5.47)$$

$$(T_f - 1) \sum_{k=1}^n \text{var}(m_k) \approx \text{tr}[\mathbf{M}^T \mathbf{M}] = \sum_{k=1}^n \lambda_k^2 \leq (T_f - 1) M^2, \quad (5.48)$$

$$(T_f - 1) \sum_{k=1}^n \text{var}(y_k) \approx \text{tr}[\mathbf{Y}^T \mathbf{Y}] = \text{tr}[\hat{\mathbf{\Sigma}}^2 \hat{\mathbf{V}}^T \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T \hat{\mathbf{V}}] \leq (T_f - 1) Y^2, \quad \text{and} \quad (5.49)$$

$$\chi \underbrace{\text{tr}[\hat{\Sigma}^2 \hat{\mathbf{V}}^T \mathbf{P} \mathbf{A} \mathbf{P}^T \hat{\mathbf{V}}]}_{\approx (T_f - 1) \text{var}(\mathbf{y})} + (1 - \chi) \underbrace{\text{tr}[\mathbf{P} \mathbf{A} \mathbf{P}^T]}_{\approx (T_f - 1) \text{var}(\mathbf{m})} \leq w^2, \quad (5.50)$$

where $\hat{\mathbf{T}} \triangleq \hat{\mathbf{U}} \text{diag}(\hat{\sigma}_1, \dots, \hat{\sigma}_n)$ and $\chi, 1 - \chi$ are relative weights of the output and input variances, respectively, with $0 \leq \chi \leq 1$.

5.3.5.1 Analytical solution to eqn. (5.45) in the absence of partial model knowledge:

Rotated inputs. To ensure that eqn. (5.45) poses a meaningful numerical optimization problem, it would be useful to examine the outcome of this optimization when it accepts an analytical solution. As shown below (Theorem 3 and Theorem 4) that in the case of constraints on the total input or output variance, as in eqn. (5.50), an explicit solution to the minimization of $\text{tr}(\mathbf{Q} \mathbf{\Lambda}^{-2} \mathbf{Q}^T \hat{\Sigma}^{-4})$ can be found, in terms of rotated inputs, as discussed in the Introduction section. Specifically, the minimization problem in eqn. (5.45) can proceed in two steps:

- 1) by finding the optimal (orthonormal) \mathbf{P} , given $\mathbf{\Lambda}$, in the first step, and
- 2) by finding the optimal $\mathbf{\Lambda}$, in the second step.

The first step is carried out in Theorem 3 and the second step in Theorem 4 below.

Theorem 3 – Emergence of rotated inputs as optimal choice for identification of IC-compliant models

Given $\hat{\mathbf{V}} = [\hat{\mathbf{v}}_1 \ \cdots \ \hat{\mathbf{v}}_n] \in \mathbb{R}^{n \times n}$ orthonormal, $\mathbf{\Lambda} = \text{diag}(\lambda_1^2, \dots, \lambda_n^2)$ with diagonal entries indexed in increasing order as

$$0 < \lambda_{\ell_1} < \cdots < \lambda_{\ell_n}, \quad (5.51)$$

and $\hat{\Sigma} = \text{diag}(\hat{\sigma}_1, \dots, \hat{\sigma}_n)$ with diagonal entries indexed in decreasing order as

$$\hat{\sigma}_1 > \cdots > \hat{\sigma}_n > 0, \quad (5.52)$$

the minimum of $\text{tr}(\mathbf{Q}\mathbf{\Lambda}^{-2}\mathbf{Q}^T\hat{\mathbf{\Sigma}}^{-4})$ with respect to the orthonormal matrix $\mathbf{P} = \hat{\mathbf{V}}\mathbf{Q} \in \mathbb{R}^{n \times n}$ is

$$\min_{\mathbf{P}=\hat{\mathbf{V}}\mathbf{Q}} \text{tr}(\mathbf{Q}\mathbf{\Lambda}^{-2}\mathbf{Q}^T\hat{\mathbf{\Sigma}}^{-4}) = \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^4 \lambda_{\ell_k}^4} \quad (5.53)$$

and is obtained for

$$\mathbf{P}^{\text{opt}} = \hat{\mathbf{V}}\mathbf{\Pi} = [\hat{\mathbf{v}}_{\ell_1} \quad \dots \quad \hat{\mathbf{v}}_{\ell_n}] \in \mathbb{R}^{n \times n}, \quad (5.54)$$

where $\mathbf{\Pi}$ is a permutation matrix such that

$$[\ell_1, \dots, \ell_n] = [1, \dots, n]\mathbf{\Pi}. \quad (5.55)$$

Proof: See Appendix C

Corollary 1 – Rotated inputs as optimal choice for identification of IC-compliant models

Under the conditions of Theorem 3, optimal inputs (in the sense of $\min_{\mathbf{P}=\hat{\mathbf{V}}\mathbf{Q}} \text{tr}(\mathbf{Q}\mathbf{\Lambda}^{-2}\mathbf{Q}^T\hat{\mathbf{\Sigma}}^{-4})$) for the process to be identified have the form

$$\mathbf{m} = \hat{\mathbf{V}}\boldsymbol{\xi}, \quad (5.56)$$

where $\hat{\mathbf{V}}$ is a rotation (orthonormal) matrix in the SVD of eqn. (5.8), and

$$\text{var}(\boldsymbol{\xi}_k) = \lambda_{\ell_k}^2 \quad (5.57)$$

Proof: Eqns. (5.44) and (5.54) imply

$$\underbrace{\hat{\mathbf{V}}^T \mathbf{M}^T}_{\boldsymbol{\Xi}^T} \mathbf{M} \hat{\mathbf{V}} = \boldsymbol{\Xi}^T \boldsymbol{\Xi} = \mathbf{\Pi} \mathbf{\Lambda} \mathbf{\Pi}^T = \text{diag}(\lambda_{\ell_1}^2, \dots, \lambda_{\ell_n}^2). \quad (5.58)$$

with the rotated inputs $\boldsymbol{\xi} = \hat{\mathbf{V}}^T \mathbf{m}$, rather than the original inputs \mathbf{m} , being

uncorrelated, with covariance matrix $\mathbf{C}_{\boldsymbol{\xi}} \approx \frac{1}{T_f - 1} \boldsymbol{\Xi}^T \boldsymbol{\Xi} = \frac{1}{T_f - 1} \text{diag}(\lambda_{\ell_1}^2, \dots, \lambda_{\ell_n}^2)$.

Corollary 2 – Sufficient condition for IC in process identification without partial knowledge

IC is guaranteed for identification of a process as in eqn. (5.10) if

$$\sum_{k=1}^n \frac{1}{\hat{\sigma}_k^4 \lambda_{\ell_k}^4} < \frac{1}{nr_{\mathbf{B}}^2} \quad (5.59)$$

Proof: Immediate, from combination of eqns. (5.53) in Theorem 3, (5.38) in Theorem 2, (5.42), and. (5.45)

Theorem 4 – Emergence of optimal scaling of rotated inputs for identification of IC-compliant models

$$\min_{\lambda_{\ell_k}} \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^4 \lambda_{\ell_k}^4} \quad (5.60)$$

subject to the constraint in eqn. (5.50) and the ordering of λ_{ℓ_k} in eqn. (5.51), is obtained at

$$\lambda_{\ell_k} = \frac{w}{(\hat{\sigma}_k^2 b_k)^{1/3}} \frac{1}{\left[\sum_{i=1}^n \left(\frac{b_i}{\hat{\sigma}_i} \right)^{4/3} \right]^{1/2}}, \quad (5.61)$$

where $b_k = (\chi \hat{\sigma}_k^2 + 1 - \chi)^{1/2}$.

Proof: See Appendix C

Corollary 3 – Optimal scaling of rotated inputs for identification of IC-compliant models

Under the conditions of Theorem 4,

$$r_{jk}(\chi) \triangleq \frac{\sqrt{\text{var}(\xi_j)}}{\sqrt{\text{var}(\xi_k)}} = \frac{\lambda_{\ell_j}}{\lambda_{\ell_k}} = \frac{(\chi \hat{\sigma}_k^6 + (1 - \chi) \hat{\sigma}_k^4)^{1/6}}{(\chi \hat{\sigma}_j^6 + (1 - \chi) \hat{\sigma}_j^4)^{1/6}}. \quad (5.62)$$

Proof: Immediate from eqns. (5.61) and (5.57).

Note that for $\chi = 1$, i.e. constraint only on total output variance through eqn. (5.50), the optimal ratio $r_{jk}(\chi = 1)$ of rotated-input amplitudes expressed in eqn. (5.62) recovers the ratio $\hat{\sigma}_k / \hat{\sigma}_j$ suggested by Koung and MacGregor (1994). However, for $\chi = 0$, i.e. constraint only on total input variance through eqn. (5.50), the optimal ratio $r_{jk}(\chi = 0)$ of rotated-input amplitudes expressed in eqn. (5.62) turns out to be $(\hat{\sigma}_k / \hat{\sigma}_j)^{2/3}$. This result is related to the design $r_{jk}(\chi = 0) = (\hat{\sigma}_k / \hat{\sigma}_j)^{1/3}$ suggested by Darby and Nikolaou (Darby & Nikolaou, 2009). In fact, both designs, i.e. $r_{jk}(\chi = 0) = (\hat{\sigma}_k / \hat{\sigma}_j)^{2/3}$ and $r_{jk}(\chi = 0) = (\hat{\sigma}_k / \hat{\sigma}_j)^{1/3}$, provide rigorous justification to the heuristic suggestion of using a ratio $r_{jk} = \varepsilon \hat{\sigma}_k / \hat{\sigma}_j$ with $0 < \varepsilon < 1$ at the beginning of an identification experiment (to account for uncertainty) proposed by Bruwer and MacGregor (Bruwer & MacGregor, 2006).

5.4 Case Studies

The IC-optimal designs proposed in the previous section will be illustrated next with numerical simulations on two multivariable systems, namely a 5×5 industrial fluid catalytic cracker (FCC) reactor/regenerator unit and a 2×2 multi-stage absorber unit, presented below.

The reason for selecting the FCC unit is that its steady-state gain matrix has a number of entries known to be identically zero. This kind of partial system knowledge

conforms with eqn. (5.11). Therefore, the corresponding theory, epitomized by eqn. (5.36) will be used.

The reason for selecting the two-stage absorber unit is that entries of its steady-state gain matrix are related through linear constraints that conform with eqn. (5.24). Therefore, the corresponding theory, epitomized by eqn. (5.42), will be used.

In addition to IC-optimal designs, D-optimal designs (Mehra, 1974) will be illustrated, for comparison. The reason for this comparison is that D-optimal designs maximize the determinant of the input covariance matrix, $\mathbf{C}_m \approx \frac{1}{T_f - 1} \mathbf{M}^T \mathbf{M}$. By contrast, IC-optimal designs involve the quantities indicated in eqns. (5.36) and (5.42), which are different functions of \mathbf{C}_m .

Four different designs will be illustrated, summarized as follows:

- a. Design D1 is the IC-optimal DOE approach based on minimization of J_c , eqn. (5.39) (or minimization of $\mu_{\max}(\mathbf{B}^{-1}\Phi)$, eqn. (5.42)), taking into account partial knowledge, eqn. (5.11) (or eqn. (5.24)), and input-output variance constraints, e.g. eqn.(5.40), (5.41) (or any of (5.46) -(5.50)).
- b. Design D2 is similar to D1, but without use of partial knowledge; i.e. D2 minimizes J_u , eqn.(5.7), subject to input-output variance constraints, as above.
- c. Design D3 is D-optimal DOE (Mehra, 1974) with partial knowledge; i.e. D3 minimizes the covariance of the parameter estimator, $\text{diag}\left(\left(\mathbf{M}_1^T \mathbf{M}_1\right)^{-1}, \dots, \left(\mathbf{M}_n^T \mathbf{M}_n\right)^{-1}\right)$ as in eqns. (5.18) - (5.20), for least squares identification with partial knowledge as in eqn. (5.11):

$$\min_{\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T} \left(\sum_{i=1}^n \log(\det(\mathbf{D}_{1,i})) \right) \quad (5.63)$$

subject to input-output variance bounds, as above; or minimization of the covariance of the parameter estimator, \mathbf{C} , as in eqns. (5.32) -(5.34) for least squares identification with partial knowledge as in eqn.:

$$\min_{\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T} \left(\log(\det(\tilde{\mathbf{D}}_1)) \right) \quad (5.64)$$

subject to input-output variance bounds, as above.

d. Design D4 is similar to D3 but without use of partial knowledge; i.e. D4 minimizes

the covariance of the parameter estimator, $\mathbf{C}_m \approx \frac{1}{T_f - 1} \mathbf{M}^T \mathbf{M}$:

$$\min_{\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T} \left(-\log(\det(\mathbf{C}_m)) \right) \quad (5.65)$$

subject to input-output variance constraints, as above.

For all four designs, identification is performed through constrained least-squares, namely minimization of the sum of the squared errors subject to the equality constraints, eqn. (5.11) (or eqn. (5.24)), emanating from partial knowledge.

Given that the IC-optimal designs D1 and D2 through eqns. (5.39) or (5.42), respectively, require knowledge of $\hat{\mathbf{G}}$, the plant estimate *after* the identification, two different kinds of designs D1 and D2 are performed in the simulations that follow:

- a. DOE assuming perfect knowledge of \mathbf{G} , and replacing $\hat{\mathbf{G}}$ by \mathbf{G} in eqns. . (5.39) or (5.42) at the time of DOE. For numerical simulations resulting from such DOE, satisfaction of IC is *indirectly* checked through eqns. (5.36) or (5.38), since checking satisfaction of eqn. (5.4) is not trivial. While this design is obviously unrealistic from a practical viewpoint, it provides a clear *characterization* of the

optimal inputs and illustrates what can be anticipated in a best-case, if unrealistic, scenario.

- b. Adaptive DOE, as outlined in Adaptive design of the Background section, using the latest estimate of $\hat{\mathbf{G}}$ to design optimal inputs for the next time-segment of identification, at the end of which $\hat{\mathbf{G}}$ will be updated and DOE repeated. After an initial period of identification with PRBS inputs, eqns. (5.39) or (5.42) are adaptively used at the time of DOE, with $\hat{\mathbf{G}}$ being the latest estimate. Satisfaction of IC is again checked through eqns. (5.36) or (5.38). While this design confounds the IC-optimal design with the adaptation method used, it provides a clear illustration of how the proposed approach would work in practice. It is also clear that for very long identification experiments, this design will eventually converge towards producing inputs similar to those produced by design (a). The effect of adaptation on convergence warrants its own investigation, and is examined elsewhere (S. Misra & Nikolaou, 2015).

For numerical solution of the optimization problems in D-optimal designs D3 and D4, we use the semidefinite programming solver SeDuMi (Sturm, 1999) of YALMIP-MATLAB[®]. For the IC-optimal designs D1 and D2, we use the nonlinear optimization function `fmincon` available in MATLAB[®], with initial guesses obtained from corresponding D-optimal designs. To avoid local minima, the Multistart algorithm available in MATLAB[®] is also employed with `fmincon`. The computational time required for `fmincon` with Multistart was found to increase by a number approximately equal to the number of initial guesses considered by Multistart. However, no improvement in the solution was observed, suggesting small probability of convergence to local minima.

5.4.1 Fluid catalytic cracking reactor-regenerator

A 5×5 fluid catalytic cracking (FCC) reactor-regenerator unit identified from plant testing (Harmse, 2007) is used here to compare performances of designs D1-D4, with emphasis on eqn. (5.39). The state-space model of the system can be found in M. L Darby (2008). The SSGM \mathbf{G}_1 of the system is

$$\mathbf{G}_1 = \begin{bmatrix} 0.3866 & 0 & 0 & 0 & 0 \\ 0 & -0.6935 & 0 & 0 & -0.5805 \\ 0.1192 & 1.5461 & 0.5224 & 0 & -0.3667 \\ 0 & -0.1313 & -0.1298 & 0.1058 & -0.2057 \\ 0.0631 & -0.2462 & 0 & 0 & -0.4435 \end{bmatrix} \quad (5.66)$$

with identically zero elements evident. This partial knowledge can be included in DOE in the form of linear constraints on the parameters, as shown in eqn. (5.11). All four experiment designs D1-D4 use the following variance constraints on individual inputs (m_i) and outputs (y_i).

$$\begin{aligned} \text{var}(y_1) &\leq 0.35 & \text{var}(m_1) &\leq 1.5 \\ \text{var}(y_2) &\leq 0.35 & \text{var}(m_2) &\leq 1.5 \\ \text{var}(y_3) &\leq 0.65 & \text{var}(m_3) &\leq 3.0 \\ \text{var}(y_4) &\leq 0.35 & \text{var}(m_4) &\leq 1.5 \\ \text{var}(y_5) &\leq 0.35 & \text{var}(m_5) &\leq 1.5 \end{aligned} \quad (5.67)$$

Parameters used in the simulation for all adaptive designs are given in Table 1.

Table 5-1: Case1: Parameters used in simulation for adaptive designs

Time steps of initial identification with PRBS inputs	55
Standard deviation of output noise	0.25
Length of time segment at the end of which $\hat{\mathbf{G}}$ is updated and DOE repeated	1
Total number of identification steps	1000

The optimal inputs and outputs for 5×5 FCC unit as characterized by designs D1-D4 are shown in Figure 5-1 - Figure 5-4 respectively.

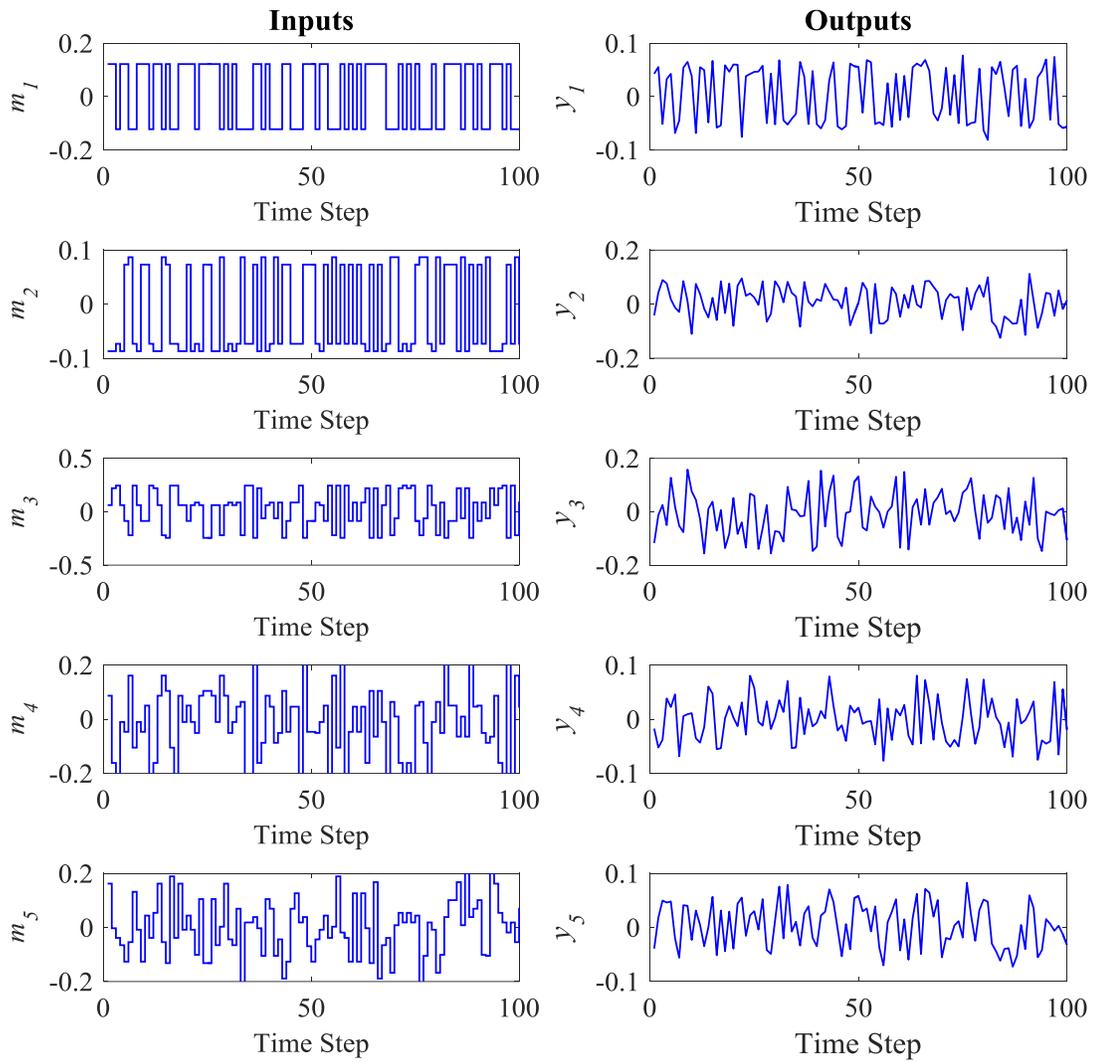


Figure 5-1. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D1

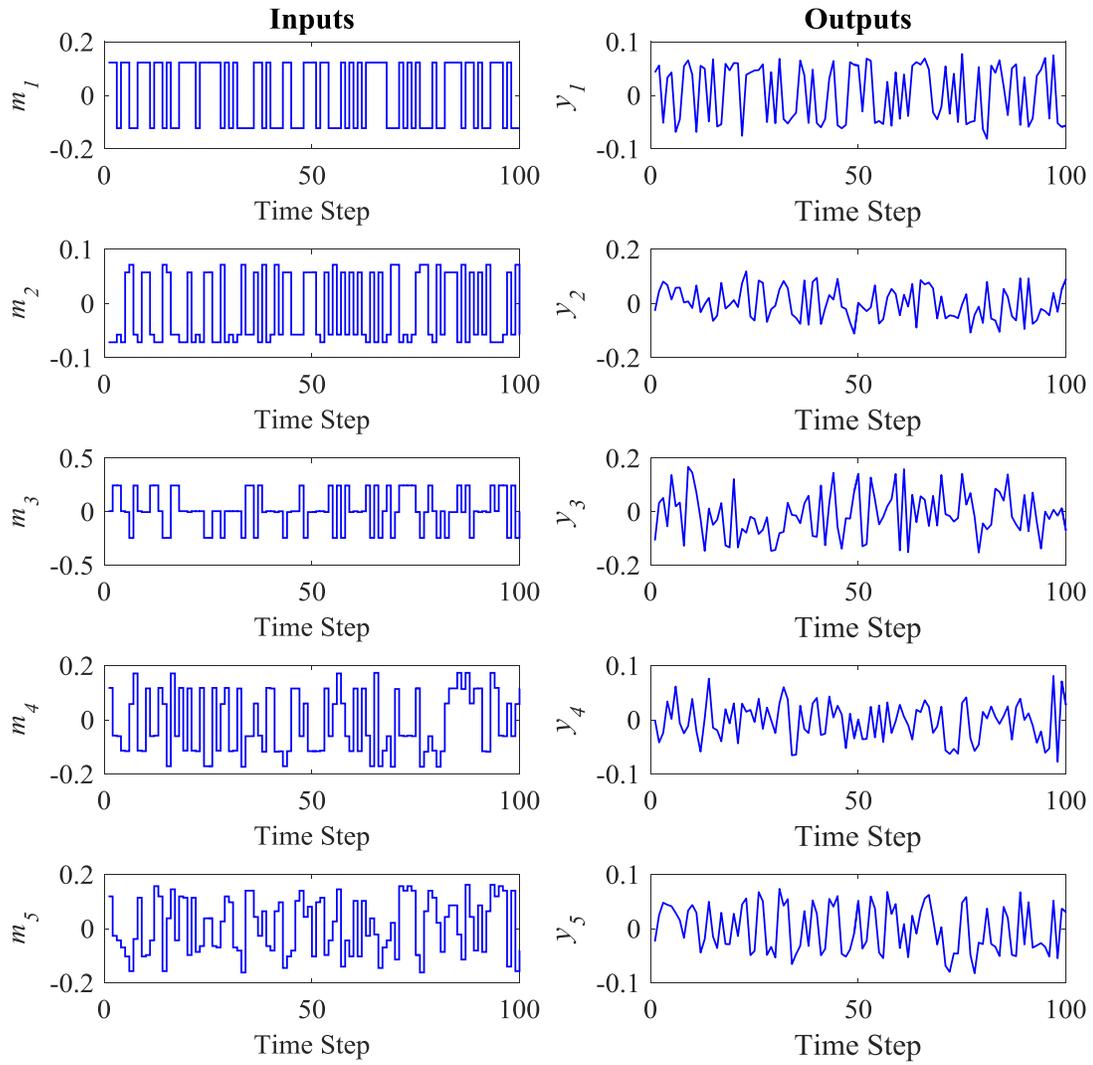


Figure 5-2. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D2.

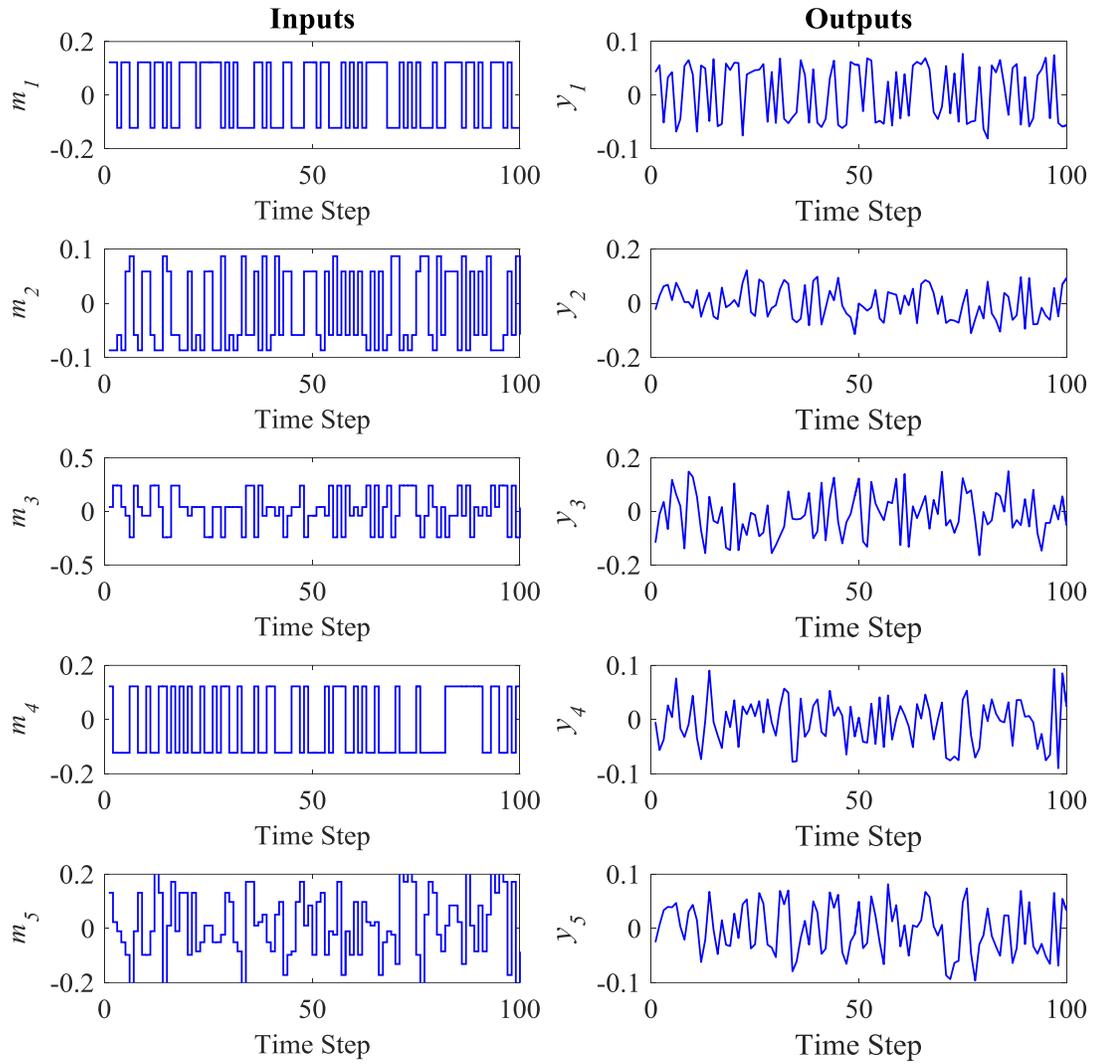


Figure 5-3. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D3.

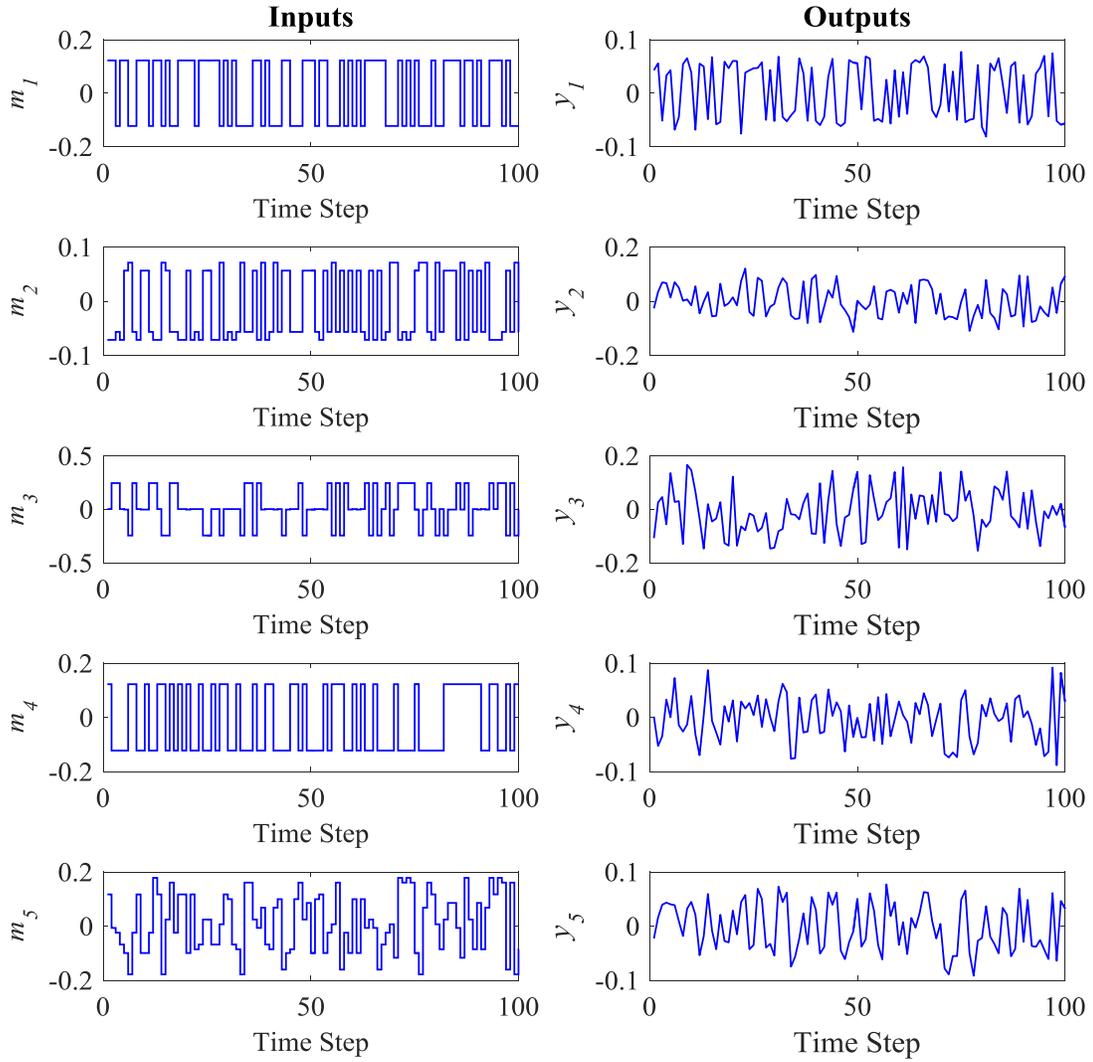


Figure 5-4. Optimal inputs and outputs for 5×5 FCC unit as characterized by design D4.

Table 5-2. Characterization of inputs and outputs for designs D1-D4 for 5×5 FCC unit; Active constraints are in bold.

Design	$\det(\mathbf{C}_m)$	$\det(\mathbf{A}_i), i = 1, \dots, 5$	$\text{var}(m_i), i = 1, \dots, 5$	$\text{var}(y_i), i = 1, \dots, 5$
D1	0.02	1.49, 0.60, 0.49, 0.01, 0.88	1.50 , 0.65, 3.00 , 1.16, 1.34	0.22, 0.35 , 0.65 , 0.17, 0.19
D2	1.01	1.5, 0.37, 0.75, 0.69, 0.55	1.50 , 0.42, 3.00 , 1.50 , 1.04	0.22, 0.35 , 0.65 , 0.09, 0.18
D3	0.99	1.5, 0.53, 0.66, 0.75, 0.77	1.50 , 0.55, 3.00 , 1.50 , 1.36	0.23, 0.35 , 0.65 , 0.14, 0.20
D4	1.16	1.5, 0.39, 0.77, 0.80, 0.58	1.50 , 0.41, 3.00 , 1.50 , 1.17	0.22, 0.35 , 0.65 , 0.13, 0.19

Table 5-3. Input and output correlations matrices for Designs D1-D4 for 5×5 FCC unit.

Design	\mathbf{R}_m	\mathbf{R}_y
D1	$\begin{bmatrix} 1 & -0.08 & -0.07 & 0.24 & 0.15 \\ & 1 & -0.87 & 0.49 & -0.55 \\ & & 1 & -0.84 & 0.59 \\ & & & 1 & -0.34 \\ & & & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.09 & -0.11 & 0.04 & 0.03 \\ & 1 & -0.01 & 0.33 & 0.86 \\ & & 1 & 0.49 & 0.41 \\ & & & 1 & 0.70 \\ & & & & 1 \end{bmatrix}$
D2	$\begin{bmatrix} 1 & -0.11 & 0.002 & 0.009 & 0.082 \\ & 1 & -0.71 & -0.23 & -0.38 \\ & & 1 & 0.33 & 0.41 \\ & & & 1 & 0.16 \\ & & & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0.001 & 0.01 & -0.02 & 0.14 \\ & 1 & 0.002 & 0.50 & 0.91 \\ & & 1 & 0.12 & 0.26 \\ & & & 1 & 0.71 \\ & & & & 1 \end{bmatrix}$

Table 5-3 continued

D3	$\begin{bmatrix} 1 & -0.19 & 0.003 & -3 \times 10^{-4} & 0.14 \\ & 1 & -0.80 & 5 \times 10^{-8} & -0.53 \\ & & 1 & 6 \times 10^{-9} & 0.60 \\ & & & 1 & -3 \times 10^{-9} \\ & & & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 4 \times 10^{-9} & -0.16 & -0.04 & 0.08 \\ & 1 & -1 \times 10^{-8} & 0.50 & 0.89 \\ & & 1 & 0.37 & 0.33 \\ & & & 1 & 0.77 \\ & & & & 1 \end{bmatrix}$
D4	$\begin{bmatrix} 1 & -0.11 & -1 \times 10^{-8} & 1 \times 10^{-9} & 0.08 \\ & 1 & -0.71 & -9 \times 10^{-10} & -0.43 \\ & & 1 & -5 \times 10^{-10} & 0.50 \\ & & & 1 & 1 \times 10^{-9} \\ & & & & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & -7 \times 10^{-9} & 2 \times 10^{-8} & -0.02 & 0.13 \\ & 1 & 7 \times 10^{-9} & 0.51 & 0.91 \\ & & 1 & 0.13 & 0.26 \\ & & & 1 & 0.73 \\ & & & & 1 \end{bmatrix}$

Important observations from Figure 5-1 - Figure 5-4, Table 5-2 and Table 5-3 are as follows:

- For designs D2 and D4, \mathbf{A}_k , $k = 1, \dots, 5$, are information matrices, according to eqn.(5.35). Larger value of $\det(\mathbf{A}_k)$ imply more accurate parameter estimation. Because the design D4 is D-optimal, $\det(\mathbf{A}_k)$ for that design is larger than $\det(\mathbf{A}_k)$ for design D2, which is IC-optimal. This is clearly in agreement with the fact that IC-optimal designs sacrifice some accuracy in parameter estimation to achieve IC satisfaction.
- All designs result in input or output pairs that may be from highly correlated to fairly uncorrelated. This is due to the fact that in the presence of input and/or output constraints, optimal experiments are obtained through constrained numerical optimization. These observations suggest that the frequently mentioned rule of thumb “opt for correlated inputs and uncorrelated outputs when identifying ill-conditioned systems” *is not universally applicable*. Similar findings were also stated by Darby and Nikolaou (M L Darby & M. Nikolaou, 2014).

- Figure 5-5 shows comparison of identification time required for IC satisfaction ($J < 1$) for designs D1-D4. Designs D1 and D3, which use partial knowledge about the system, require fewer time steps than designs D2 and D4. This clearly establishes the usefulness of using partial knowledge in designing the experiments. Design D1, being an IC-optimal design, performs slightly better than D-optimal design D3.
- Figure 5-6 shows the convergence of elements of identified SSGM using given input/output constraints. Evidently, designs D1-D4 adaptively produce models which are close to the actual system \mathbf{G}_1 . However, IC-optimal designs D1 and D2 deliberately compromise the accuracy of some of the parameters, in exchange of IC satisfaction.

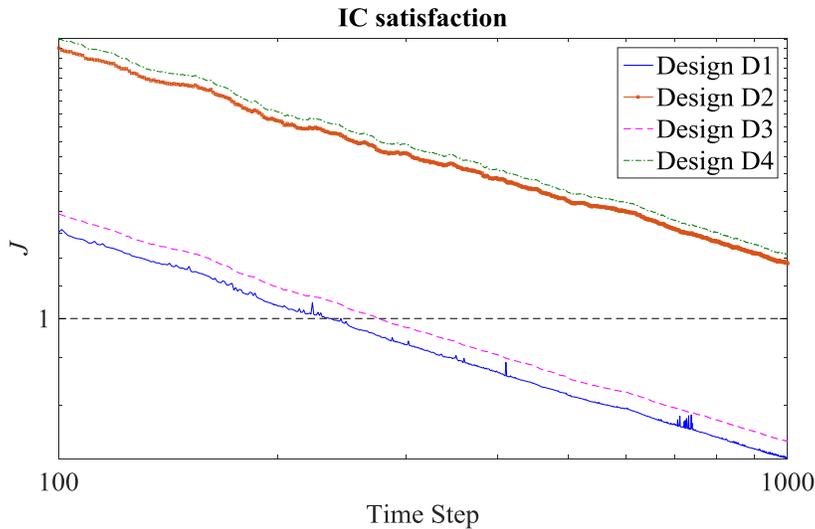


Figure 5-5. Adaptive DOE for 5×5 FCC unit: Identification time required for IC satisfaction when inputs are produced by designs D1-D4. The vertical axis is J_c , eqn. (5.11), for designs D1 and D3, and J_u , eqn. (5.7), for designs D2 and D4.

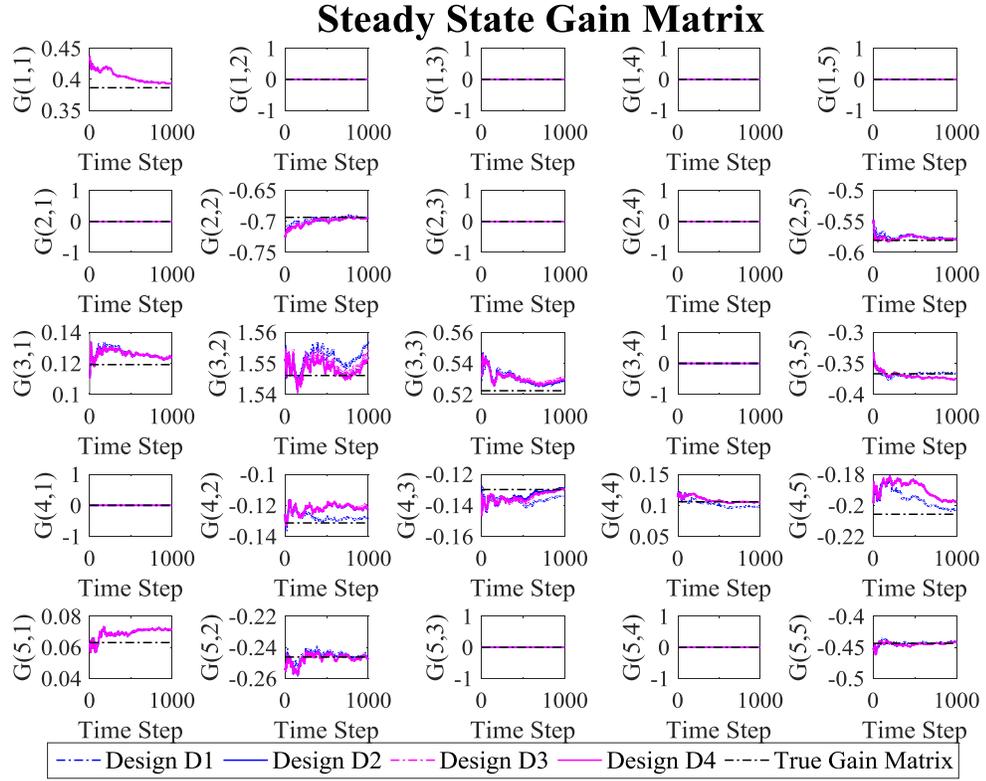


Figure 5-6. Convergence of SSGM elements of 5×5 FCC unit for designs D1-D4.

5.4.2 Case 2: Two-Stage Absorber Unit

A 2×2 two-stage absorber (Amundson, 1966) is used here to compare performances of designs D1-D4, with emphasis on eqn. (5.42). Input variables are solute concentrations in the liquid and vapor streams entering the absorber, and output variables are concentrations of solute in the liquid at each of the two stages. The SSGM \mathbf{G}_2 for this system is

$$\mathbf{G}_2 = \begin{bmatrix} 0.2632 & 0.1053 \\ 0.1579 & 0.2632 \end{bmatrix}. \quad (5.68)$$

Partial knowledge, in terms of the linear equality between the first and second row of the matrix \mathbf{G}_2 is given as

$$\begin{aligned} \mathbf{G}_2(1,1) &= \mathbf{G}_2(2,2) \\ \mathbf{G}_2(1,2) + \mathbf{G}_2(2,1) &= \mathbf{G}_2(2,2) \end{aligned} \quad (5.69)$$

Input and output variance constraints are

$$\text{var}(y_i) \leq 0.1, \quad \text{var}(m_i) \leq 0.5, \quad i = 1, 2. \quad (5.70)$$

Parameters used in the simulation for all adaptive designs are given in Table 5-4.

Table 5-4. Case 2: Parameters used in simulation for adaptive designs

Time steps of initial identification with PRBS inputs	55
Standard deviation of output noise	0.5
Length of time segment at the end of which $\hat{\mathbf{G}}$ is updated and DOE repeated	1
Total number of identification steps	300

Optimal inputs and outputs for 2×2 two stage absorber unit as characterized by designs D1-D4 are shown in Figure 5-7 - Figure 5-10, respectively.

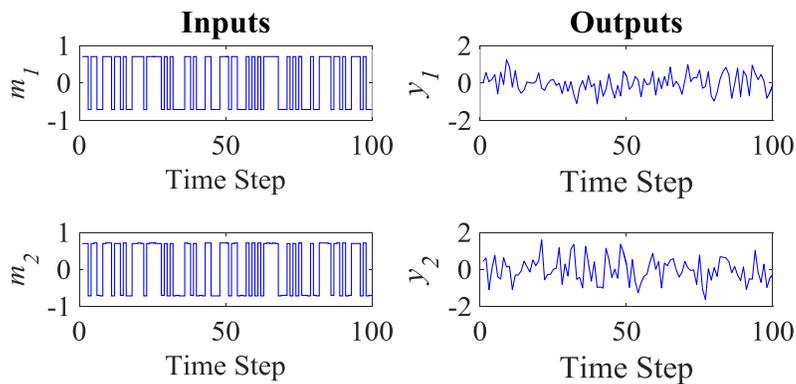


Figure 5-7. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D1.

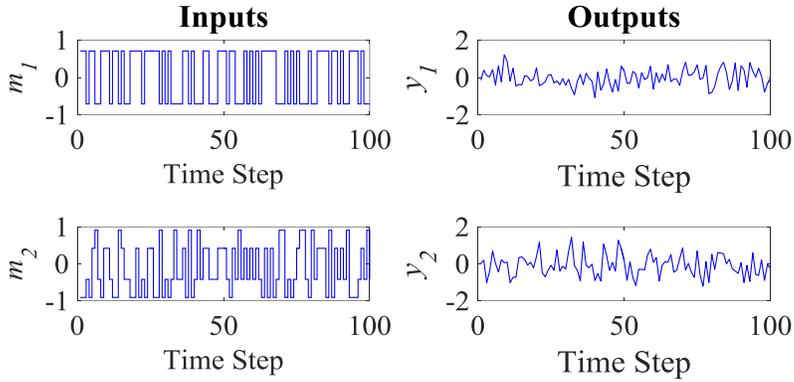


Figure 5-8. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D2.

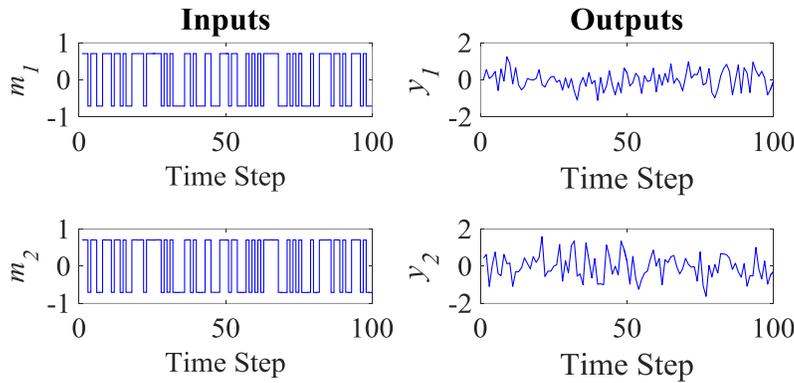


Figure 5-9. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D3.

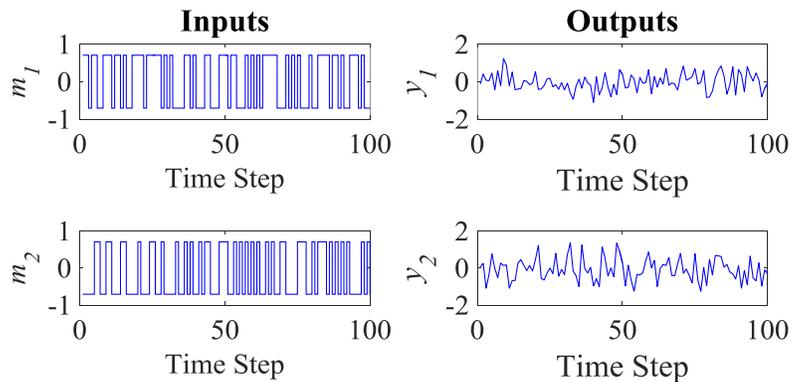


Figure 5-10. Optimal inputs and outputs for 2×2 two-stage absorber unit as characterized by design D4.

Table 5-5. Results for designs D1-D4 for 2×2 two-stage absorber. Active constraints are in bold.

Design	$\det(\mathbf{C}_m)$	$\det(\mathbf{B})$	$\text{var}(m_i)$	$\text{var}(y_i)$	ρ_{m_1, m_2}	ρ_{y_1, y_2}
D1	9×10^{-5}	0.45	0.5, 0.5	0.07, 0.09	0.99	1.00
D2	0.22	0.18	0.5, 0.5	0.07, 0.09	-0.35	1.00
D3	1×10^{-8}	0.45	0.5, 0.5	0.07, 0.09	1.00	1.00
D4	0.25	0.25	0.5, 0.5	0.04, 0.05	-4×10^{-9}	0.79

As in the first case study, it can be observed here that $\det(\mathbf{B})$ (a measure of parameter estimation accuracy) for design D4 is larger than $\det(\mathbf{B})$ for design D2 (Table 5-5). This again establishes the fact that IC-optimal designs sacrifice some accuracy in parameter estimation to achieve IC satisfaction.

Figure 5-11 compares the identification time required for IC satisfaction ($J < 1$) for designs D1-D4. Designs with partial knowledge (D1 and D3) lead to faster IC satisfaction than designs without partial knowledge (D2 and D4) as anticipated. The overlapping profiles for design D1 and D3 show close agreement between the IC-optimal design D1 and the D-optimal design D3.

Figure 5-12 shows the convergence of the SSGM elements for the identified models. Despite nearly the same closeness of the matrix elements to the actual process over the adaptations, some models satisfy IC and some do not. This again illustrates the claim made in the Introduction about closeness of the model to the real system and control relevance.

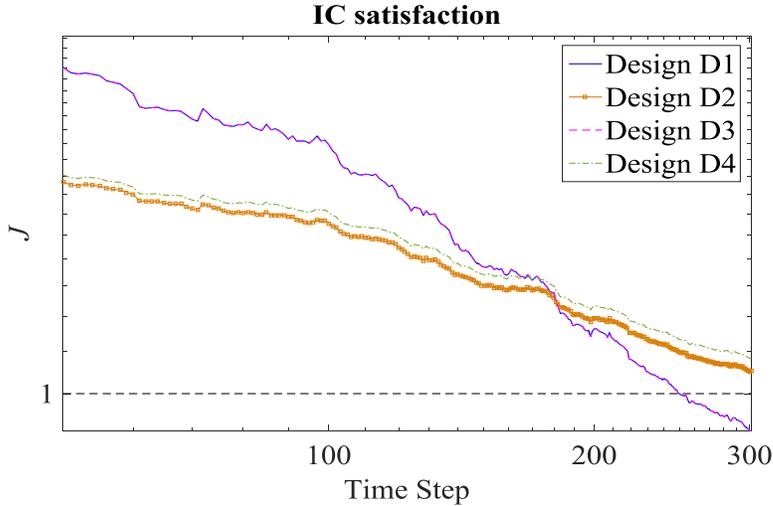


Figure 5-11. Adaptive DOE for 2×2 two-stage absorber: Identification time required for IC satisfaction when inputs are produced by designs D1-D4. The vertical axis is J_c , eqn. (5.11), for designs D1 and D3, and J_u , eqn. (5.7), for designs D2 and D4.

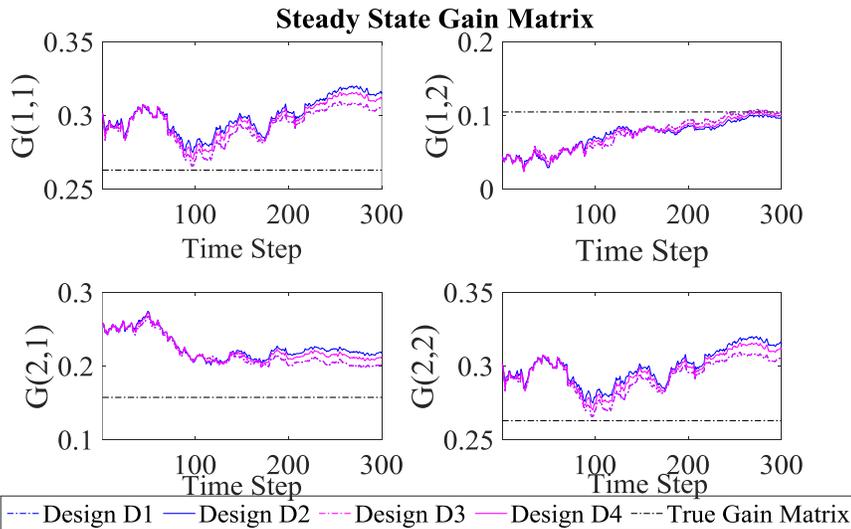


Figure 5-12. Convergence of SSGM elements of 2×2 two-stage absorber for designs D1-D4.

5.5 Summary

A general mathematical framework was presented to design experiments for efficient identification of partially known models that are required to satisfy the integral

controllability condition. The mathematical framework relies on guaranteeing integral controllability (eqn. (5.4)) by satisfaction of simpler inequalities (eqns. (5.36) or (5.38), depending on the nature of partial knowledge of the identified model), in which process inputs appear explicitly. This framework produces experiment designs either analytically, in simple cases, or through solution of an appropriately formulated constrained optimization problem, when input and/or output constraints are present.

The proposed framework was illustrated with numerical simulations on two multivariable systems, namely a 5×5 industrial fluid catalytic cracker (FCC) reactor/regenerator unit and a 2×2 multi-stage absorber unit. These simulations showed agreement of the observed results with results in literature under similar conditions. They also demonstrated novel results, under conditions not examined in literature before.

6 Conclusion and Future Work

The previous chapters elaborated the application of data-driven modeling in upstream and downstream chemical industrial purposes. Under upstream industry application, data-driven modeling was performed for shale gas well cementing. A new experiment design was proposed for control relevant model identification for partially known systems under downstream industry application. Statistical simulations were performed to ensure the suitability of the control-relevant model in terms of satisfaction of integral controllability.

In this chapter, summary and future extension of the work done are presented in the subsequent sections.

6.1 Cementing shale gas wells

Successful cementing of shale gas wells is an important step in stopping gas leaks during natural gas production. Since building a model based on fundamental knowledge only is not possible because of missing mathematical relationships between most of the affecting factors and cement bond quality. A data-driven classification model was developed using PLS-DA analysis. The model was validated using ten-fold cross-validation tests. The model is not only able to correctly classify 81 % of the classified wells but it also helps in identifying the relative importance of cement and casing variables used as inputs to the model. The model also provides correlations between input variables and well leakage. Interestingly, some of findings based on these correlations match with the common industrial practices for good cementing. The model also provides a graphical

way to easily distinguish leaking wells and possible reasons behind the leaking behavior of wells. In totality, the model provides a guideline for safe cementing operations.

The complete methodology presented here would certainly supplement the existing knowledge about well cementing. There is a scope to extend this work along the ideas presented below.

- Although the optimal model presented here produced reasonably good results from relatively scant data, it is anticipated that availability of more data points will improve the predictive ability of the model and will increase its utility.
- In fact, availability of data for additional variables, such as related to *lithology*, *mechanical properties of cement sheath*, or others would also be expected to improve the predictive ability of the model.
- Finally, the possibility of other statistical modeling techniques or a hybrid modeling approach that combines first principles and empirical data would be worth investigating.

6.2 Ensuring integral controllability

The framework developed by Darby and Nikolaou (2009) to design experiments for identification of IC-compliant models, relied on the sufficient condition instead of eigenvalue based inequality. Since the condition was only sufficient, the approach based on the sufficient condition could be potentially be a conservative. So in this study, the conservatism of the sufficient condition was assessed numerically for a number of cases. The numerical assessment showed that conservatism is not very high however

conservatism increases as the size of the identified system increases. Graphical analysis was also presented to pinpoint the source of conservatism in the sufficient condition.

An even more important question arose whether any other model can satisfy IC if least-squares identified model cannot. It was shown through counter example that it is generally not possible to find an IC compliant model without designing IC-compliant experiments.

Since the step causing conservatism in the derivation of the sufficient condition, has been pinpointed, one could think of avoiding that step. More work is needed here to derive less conservative condition or perhaps not conservative at all.

6.3 Experiment design for control-relevant identification of partially known stable multivariable systems

Here a general optimization framework was proposed to design experiments for identification of IC compliant models, which utilizes partial knowledge available about the system. The design of experiments developed in this study are based on simpler inequalities instead of cumbersome eigenvalue based condition. The process inputs appear explicitly in both these simpler inequalities that makes them suitable for optimization based framework. The framework produces analytical expressions for input design in simple case while the systems with complex input/output constraints can be handled using numerical optimization. The proposed framework was tested on two multivariable cases. Numerical results in both the cases established the usefulness of newly designed experiments in terms of less experimentation time required for identification of IC compliant model.

A number of items related to this work can be examined in the future, such as the following:

- D-optimal and IC-optimal designs produced inputs with different but not entirely dissimilar characteristics. What are the reasons underlying this outcome?
- In the present work, partial knowledge of the process to be identified is expressed in terms of equality constraints. Can these ideas be extended to inequality constraints, and would that be worthwhile from a practical viewpoint?
- The framework presented emphasizes steady-state behavior. Can it be extended to the design of experiments for identification of partially known dynamic systems along the lines of Darby and Nikolaou (M L Darby & M. Nikolaou, 2014)?

References

- Abonyi, J., Babuska, R., Verbruggen, H. B., & Szeifert, F. (2000). "Incorporating prior knowledge in fuzzy model identification." *International Journal of Systems Science*, 31, 657-667.
- Amundson, N. (1966). Mathematical Methods in Chemical Engineering: Prentice-Hall.
- Ashena, R., Thonhauser, G., & Dianati, D. "Assessment of Reliability of Cement Bond Evaluation with Some Interesting Field Case Studies". In: *Society of Petroleum Engineers*.
- Austin, C., Zimmerman, Chris, Sullaway, Bob, and Sabins, Fred. (1988). Halliburton Services: "Fundamentals of Horizontal Well Completion," *Drilling*. In (pp. 28-31).
- Bach, D., & Vijn, P. (2002). "Environmentally Acceptable Cement Fluid Loss Additive. In: *Society of Petroleum Engineers*.
- Barker, M., & Rayens, W. (2003). "Partial least squares for discrimination." *Journal of Chemometrics*, 17, 166-173.
- Biello, D. (2015). "Fact or Fiction?: Natural Gas Will Reduce Global Warming Pollution." *Scientific American*.
- Boling, M. (2015). "The EPA's Natural Gas Problem." In P. McKenna (Ed.), *NOVA*.
- Boyd, D. A., Al-Kubti, S. A. R., Khedr, O. H., Khan, N., Al-Nayadi, K. G., Degouy, D., Elkadi, A., & Al-Kindi, Z. L. (2006). "Reliability of Cement Bond Log Interpretations Compared to Physical Communication Tests Between Formations." In: *Society of Petroleum Engineers*.
- Bruwer, M. J., & MacGregor, J. F. (2006). "Robust multi-variable identification: Optimal experimental design with constraints." *Journal of Process Control*, 16, 581-600.

- Butsch, R. J. (1995). "Overcoming Interpretation Problems of Gas-Contaminated Cement Using Ultrasonic Cement Logs." In: *Society of Petroleum Engineers*.
- Cathles, L. M. I., Brown, L., Taam, M., & Hunter, A. (2012). A commentary on "The greenhouse-gas footprint of natural gas in shale formations" by R.W. Howarth, R. Santoro, and Anthony Ingraffea. *Climatic Change*.
- Caulton, D. R., Shepson, P. B., Santoro, R. L., Sparks, J. P., Howarth, R. W., Ingraffea, A. R., Cambaliza, M. O. L., Sweeney, C., Karion, A., Davis, K. J., Stirm, B. H., Montzka, S. A., & Miller, B. R. (2014). "Toward a better understanding and quantification of methane emissions from shale gas development." *Proceedings of the National Academy of Sciences of the United States of America*, 111, 6237-6242.
- Childs, J. D. T., Mary Jo. (1984). "Thixotropic cements for use in wells." In E. Patent (Ed.). US.
- Chow, T. W., McIntire, L. V., Kunze, K. R., & Cooke, C. E. (1988). "The Rheological Properties of Cement Slurries: Effects of Vibration, Hydration Conditions, and Additives." Crain. (2000). "CEMENT INTEGRITY LOGS - PART 3 -- USB / CBIL." In (pp. Petrophysical Handbook).
- Darby, M. L. (2008). Studies of online optimization methods for experimental test design and state estimation. PhD, University of Houston, Houston.
- Darby, M. L., & Nikolaou, M. (2009). "Multivariable system identification for integral controllability." *Automatica*, 45, 2194-2204.
- Darby, M. L., & Nikolaou, M. (2014). "Identification test design for multivariable model-based control: An industrial perspective." *Control Engineering Practice*, 22, 165-180.

de Jong, S. (1993). "SIMPLS: An alternative approach to partial least squares regression." *Chemometrics and Intelligent Laboratory Systems*, 18, 251-263.

Denney, D. "Reservoir Models Built With 4D-Seismic and Well Data."

DOE. (2009). Modern Shale Gas Development in the United States: A Primer In.

DOE. (April 2015). Annual Energy Outlook 2015 with projections to 2040. In.

DOE. (August 2016). U.S. Energy Information Administration / Monthly Energy Review. In.

Duda, R. O., Hart, P. E., & Stork, D. G. (2000) Pattern Recognition (pp. 20-84). New York, NY, USA: John Wiley & Sons, Inc.

Eigenvector. (2013). PLS Toolbox 7.3. In.

Elmarsafawi, Y. A., Al-Yami, A. S., Nasr-El-Din, H. A., Al-Jeffri, A. M., Misran, M., Hasan, A. A., & Jain, B. (2006). "A New Cementing Approach to Improve and Provide Long-Term Zonal Isolation." In: *Society of Petroleum Engineers*.

Esbensen, K. H. (2000). Multivariate Data Analysis - in practice (4th ed.). Oslo, Norway: CAMO AS.

Esmaili, S., & Mohaghegh, S. D. (2016). "Full field reservoir modeling of shale assets using advanced data-driven analytics." *Geoscience Frontiers*, 7, 11-20.

Featherstone, A. P., & Braatz, R. D. (1998). "Integrated robust identification and control of large-scale processes." *Industrial & Engineering Chemistry Research*, 37, 97-106.

Federov, V. V. (1972). Theory of Optimal Experiments, Academic Press, New York.

Fitzgerald, D. D., Mcghee, B. F., & Mcguire, J. A. (1985). "Guidelines for 90-Percent Accuracy in Zone-Isolation Decisions." *Journal of Petroleum Technology*, 37, 2013-2022.

- Garcia, C. E., & Morari, M. (1985). "Internal model control: 2. design procedure for multivariate systems." *Ind. Eng. Chem. Process Des. Dev.*, 24, 472-484.
- Gevers, M. (2005). "Identification for Control: From the Early Achievements to the Revival of Experiment Design*." *European Journal of Control*, 11, 335-352.
- Gevers, M., & Ljung, L. (1986). "Optimal experiment designs with respect to the intended model application." *Automatica*, 22, 543-554.
- Gibbins, J. C. (2011). Dimensional Analysis, Springer.
- Goodwin, K. J., & Crook, R. J. (1992). "Cement Sheath Stress Failure."
- Harmse, M. (2007). AptiTune multivariable PID tuning software. In.
- Haykin, S. (1999). Neural Networks: A Comprehensive Foundation, Prentice Hall.
- Holcomb, T. R., & Morari, M. (1992). "PLS/Neural Networks." *Comp. Chem. Engng*, 16, 393-411.
- Holdaway, K. R., & Laing, M. L. "Drilling and Completion Optimization in Unconventional Reservoirs with Data-Driven Models." In: *Society of Petroleum Engineers*.
- Hovd, M., & Skogestad, S. (1994). "Sequential design of decentralized controllers." *Automatica*, 30, 1601 – 1607.
- Howarth, R. W., D. Shindell, R. Santoro, A. Ingraffea, N. Phillips, and A. Townsend-Small. (2012). "Methane emissions from natural gas systems." Background paper prepared for the National Climate Assessment. In.
- Howarth, R. W., Ingraffea, A., & Engelder, T. (2011). "Natural gas: Should fracking stop?" *Nature*, 477, 271-275.

Howarth, R. W., Santoro, R., & Ingraffea, A. (2011). "Methane and the greenhouse-gas footprint of natural gas from shale formations." *Climatic Change*, 106, 679-690.

Howarth, R. W., Santoro, R., & Ingraffea, A. (2012). "Venting and leaking of methane from shale gas development: response to Cathles et al." *Climatic Change*, 113, 537-549.

IEA. (2011). "Are we entering a golden age of gas?", Paris: International Energy Agency. In.

IPCC. (2013). "Climate Change 2013: The Physical Science Basis." Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Ch.8, Table 8.7. In (pp. 711-714).

Jackson, P. B., & Murphey, C. E. (1993). "Effect of Casing Pressure on Gas Flow Through a Sheath of Set Cement." In: *Society of Petroleum Engineers*.

Johansen, T. A. (1996). "Identification of non-linear systems using empirical data and prior knowledge - An optimization approach." *Automatica*, 32, 337-356.

Kadlec, P., Gabrys, B., & Strandt, S. (2009). "Data-driven Soft Sensors in the process industry." *Computers & Chemical Engineering*, 33, 795-814.

Kiefer, J. (1958). "On the Nonrandomized Optimality and Randomized Nonoptimality of Symmetrical Designs." *The Annals of Mathematical Statistics*, 29, 675-699.

Kothare, S. L., Lu, Y., & Mandler, J. A. (2004). "Constrained System Identification for Incorporation of a priori Knowledge." In U. S. P. A. Publication (Ed.), (G06F 7/60; G06F 17/10; G06N 5/00; G06F 17/00; G06N 5/02 ed.). United States.

Koung, C. W., & Macgregor, J. F. (1993). "Design of Identification Experiments for Robust-Control - a Geometric Approach for Bivariate Processes." *Industrial & Engineering Chemistry Research*, 32, 1658-1666.

- Koung, C. W., & MacGregor, J. F. (1994). "Identification For Robust Multivariable Control - The Design Of Experiments." *Automatica*, 30, 1541-1554.
- Kulkarni, P. (2012). Design of experiments for identification of dynamic models satisfying the condition of integral controllability. Unpublished M.S. Thesis, University of Houston, Houston.
- Ljung, L. (1999). System Identification: Theory for the User, Prentice Hall.
- Malthouse, E. C., Tamhane, A. C., & Mah, R. S. H. (1997). "Nonlinear partial least squares." *Computers in Chemical Engineering*, 21., 875-890.
- MATLAB. (R2013b). MATLAB R2013b, The MathWorks, Inc., Natick, Massachusetts, United States. In.
- Maucec, M., Singh, A. P., Bhattacharya, S., Yarus, J. M., Fulton, D. D., & Orth, J. M. "Multivariate Analysis and Data Mining of Well-Stimulation Data by Use of Classification-and-Regression Tree with Enhanced Interpretation and Prediction Capabilities."
- Mehmood, T., Martens, H., Saebo, S., Warringer, J., & Snipen, L. (2011). "A Partial Least Squares based algorithm for parsimonious variable selection." *Algorithms Mol Biol*, 6, 27.
- Mehra, R. K. (1974). "Optimal Input Signals for Parameter Estimation in Dynamic Systems Survey and New Results." *Ieee Transactions on Automatic Control*, Ac19, 753-768.
- Misra, P., & Nikolaou, M. (2003). "Input design for model order determination in subspace identification." *Aiche Journal*, 49, 2124-2132.
- Misra, S., & Nikolaou, M. (2015). "Adaptive Design of Experiments to Identify Dynamic Models for Integral Controllability." In preparation.

- Mohaghegh, S. D. (2016). "Determining the main drivers in hydrocarbon production from shale using advanced data-driven analytics – A case study in Marcellus shale." *Journal of Unconventional Oil and Gas Resources*, 15, 146-157.
- Nguyen, D. V., & Rocke, D. M. (2004). "On partial least squares dimension reduction for microarray-based classification: a simulation study." *Computational Statistics & Data Analysis*, 46, 407-425.
- Panjwani, S., & Nikolaou, M. (2016). "Ensuring Integral Controllability for Robust Multivariable Control." *Computers & Chemical Engineering*, 92, 172-179.
- Pérez, N. F., Ferré, J., & Boqué, R. (2009). "Calculation of the reliability of classification in discriminant partial least-squares binary classification." *Chemometrics and Intelligent Laboratory Systems*, 95, 122-128.
- Pollard, R. (1994). "New Cement Additive Improves Slurry Properties and Saves Cost." In: *Society of Petroleum Engineers*.
- Qin, S. J., & McAvoy, T. J. (1992). "Nonlinear PLS Modeling Using Neural Networks." *Comp. Chem. Engng*, 16, 379-391.
- Rasmussen, J. K., & Bay Jørgensen, S. (2005). "EXPERIMENTAL INVESTIGATION OF DATADRIVEN MODELLING FOR CONTROL OF FED-BATCH CULTIVATION." *IFAC Proceedings Volumes*, 38, 19-26.
- Rocha-Valadez, T., Hasan, A. R., Mannan, S., & Kabir, C. S. (2014). "Assessing Wellbore Integrity in Sustained-Casing-Pressure Annulus."
- Rowan, T., & Stone, M. L. "Water Quality: Is It Important for Cement?" In: *Society of Petroleum Engineers*.
- Seber, G. A. F., & Lee, A. J. (2003). Linear Regression Analysis, Wiley.

Shu, Y., Ming, L., Cheng, F., Zhang, Z., & Zhao, J. (2016). "Abnormal situation management: Challenges and opportunities in the big data era." *Computers & Chemical Engineering*, 91, 104-113.

Skogestad, S., Lundstrom, P., & Jacobsen, E. W. (1990). "Selecting the best Distillation Control Configuration." *AIChE J.*, 36, 753-761.

Skogestad, S., & Morari, M. (1987). "Control Configuration Selection For Distillation-Columns." *Aiche Journal*, 33, 1620-1635.

Skogestad, S., & Morari, M. (1988). "Understanding the Dynamic Behavior of Distillation Columns." *Ind. Eng. Chem. Res.*, 27, 1848-1862.

Soderstrom, T., & Stoica, P. (1989). System Identification, Prentice Hall.

Stephen, K. D., Soldo, J., Macbeth, C., & Christie, M. A. "Multiple Model Seismic and Production History Matching: A Case Study."

Sturm, J. F. (1999). "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones." *Optimization Methods & Software*, 11-2, 625-653.

Taavitsainen, V., & Korhonen, P. (1992). "Nonlinear Data-Analysis With Latent-Variables." *Chemometrics and Intelligent Laboratory Systems*, 14, 185-194.

Tabora, J. E. (2012). "8th IFAC Symposium on Advanced Control of Chemical Processes Data Driven Modelling and Control of Batch Processes in the Pharmaceutical Industry." IFAC Proceedings Volumes, 45, 708-714.

Tahmourpour, F., & Griffith, J. E. (2004). "Use of Finite Element Analysis to Engineer the Cement Sheath for Production Operations." In: *Petroleum Society of Canada*.

Thiercelin, M. J., Dargaud, B., Baret, J. F., & Rodriguez, W. J. (1998). "Cement Design Based on Cement Mechanical Response."

- Timmons, W. D., Chizeck, H. J., Casas, F., Chankong, V., & Katona, P. G. (1997). "Parameter-constrained adaptive control." *Industrial & Engineering Chemistry Research*, 36, 4894-4905.
- Tyndall, J. H. (1990). "Segmented Bond Tool - A New Generation Cement Bond Logging Device." In: *Petroleum Society of Canada*.
- Watters, J. (2014). Lowering Drilling Cost, Improving Operational Safety, and Reducing Environmental Impact through Zonal Isolation Improvements for Horizontal Wells Drilled in the Marcellus Shale
- Wehrens, R. (2011). Chemometrics with R: Multivariate Data Analysis in the Natural Sciences and Life Sciences. New York: Springer.
- Wenande, B. (1987). "A Review of the First Horizontal Well-Denmark." In *Offshore Drilling Conference*. London.
- Wilkins, R. P., & Free, D. (1989). "A New Approach to the Prediction of Gas Flow After Cementing." In: *Society of Petroleum Engineers*.
- Wilson, M. A., & Sabins, F. L. (1988). "A Laboratory Investigation of Cementing Horizontal Wells."
- Wold, S. (1992). "Nonlinear partial least squares modelling II. spline inner relation." *Neural Computation*, 1, 425-464.
- Wold, S. (1994). "Chemometric Methods in Molecular Design." In P. K.-L. R. Mannhold, H. Timmerman (Ed.), Methods and Principles in Medicinal Chemistry (Vol. 2, pp. 195-218).
- Wold, S., Sjöström, M., & Eriksson, L. (2001). "PLS-regression: a basic tool of chemometrics." *Chemometrics and Intelligent Laboratory Systems*, 58, 109-130.

- Wood, R. K., & Berry, M. W. (1973). "Terminal composition control of a binary distillation column." *Chemical Engineering Science*, 28, 1707-1717.
- Xiong, R., Sun, F., Chen, Z., & He, H. (2014). "A data-driven multi-scale extended Kalman filtering based parameter and state estimation approach of lithium-ion polymer battery in electric vehicles." *Applied Energy*, 113, 463-476.
- Yamuna Rani, K., & Patwardhan, S. C. (2007). "Data-Driven Model Based Control of a Multi-Product Semi-Batch Polymerization Reactor." *Chemical Engineering Research and Design*, 85, 1397-1406.
- Zhao, H., Kang, Z., Zhang, X., Sun, H., Cao, L., & Reynolds, A. C. INSIM: "A Data-Driven Model for History Matching and Prediction for Waterflooding Monitoring and Management with a Field Application." In: *Society of Petroleum Engineers*.
- Zhao, H., Li, Y., Cui, S., Shang, G., Reynolds, A. C., Guo, Z., & Li, H. A. (2016). "History matching and production optimization of water flooding based on a data-driven interwell numerical simulation model." *Journal of Natural Gas Science and Engineering*, 31, 48-66.
- Zurdo, C., Georges, C., & Martin, M. (1986). "Mud and Cement for Horizontal Wells." In SPE 15464, *SPE Annual Technical Conference and Exhibition*.

Appendix A

A.1 Classification using PLS-DA

To calculate whether a vector \mathbf{x} (i.e. the 31 dimensionless inputs shown in Table 3-2) will result in a leaking well (i.e., a value of the output $y = 1$) or not ($y = 0$), the PLS-DA model uses Bayes' rule (Duda et al., 2000) to calculate whether the probability (Figure A-1).

$$P(\omega_1 | \hat{y}_s) = \frac{P(\hat{y}_s | \omega_1)P(\omega_1)}{P(\hat{y}_s | \omega_1)P(\omega_1) + P(\hat{y}_s | \omega_2)P(\omega_2)} \quad (\text{A.1})$$

is greater than or less than 0.5, respectively. The quantities that appear in eqn. (A.1) are calculated as follows: $P(\omega_1), P(\omega_2)$ are the probabilities of a leaking or non-leaking well; $P(\hat{y}_s | \omega_1), P(\hat{y}_s | \omega_2)$ are the conditional densities of the continuous variable

$$\hat{y}_s = \hat{\boldsymbol{\beta}}^T \mathbf{x} \quad (\text{A.2})$$

if a well is leaking or non-leaking, respectively; the vector $\hat{\boldsymbol{\beta}}$ is the outcome of PLS estimation; and

$$P(\hat{y} | \omega_c) = \frac{\sum_{i=1}^{n_c} g_i(\hat{y})}{n_c} \quad (\text{A.3})$$

(Figure A-2) with $c = 1, 2$ and n_c the number of samples in category c in the training set;

$$g_i(\hat{y}) = \frac{1}{\text{SEP}_i \sqrt{2\pi}} \exp\left(\frac{-1}{2} \left(\frac{\hat{y} - \hat{y}_{s,i}}{\text{SEP}_i}\right)^2\right) \dots\dots\dots (\text{A.4})$$

$$\text{SEP}_i = [(1 + h_i) \times \text{MSEC}_{bc}]^{1/2} \dots\dots\dots (\text{A.5})$$

is the standard error in prediction (SEP_i) for the i^{th} sample.

$$h_i = \mathbf{t}_i^T (\mathbf{T}^T \mathbf{T})^{-1} \mathbf{t}_i \dots\dots\dots (A.6)$$

is the leverage and \mathbf{t}_i is the score vector for the i^{th} sample on r latent vectors in the PLS model ;

$$MSEC_{bc} = \frac{\sum_1^n (\hat{y}_{s,i} - y_i - bias_c)^2}{n - r - 1} \dots\dots\dots (A.7)$$

is the bias corrected mean squared error of calibration ($MSEC_{bc}$) (Esbensen, 2000) where

$$bias_c = \frac{\sum_1^{n_c} (\hat{y}_{s,i} - y_i)}{n_c} \dots\dots\dots (A.8)$$

with y_i dummy values and $\hat{y}_{s,i} = \boldsymbol{\beta}^T \mathbf{x}_i$.

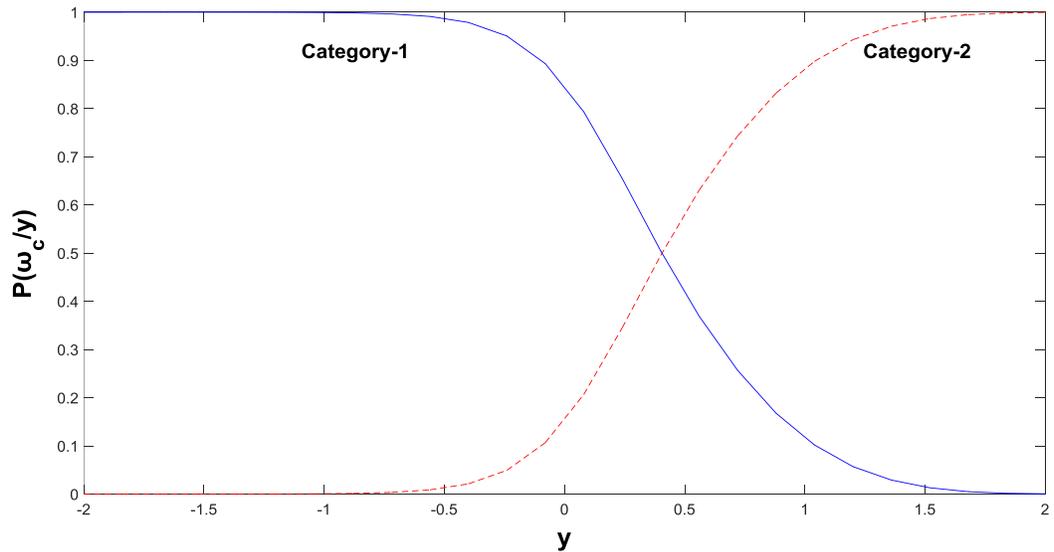


Figure A-1: Posterior probabilities for two categories: category-1 and category-2

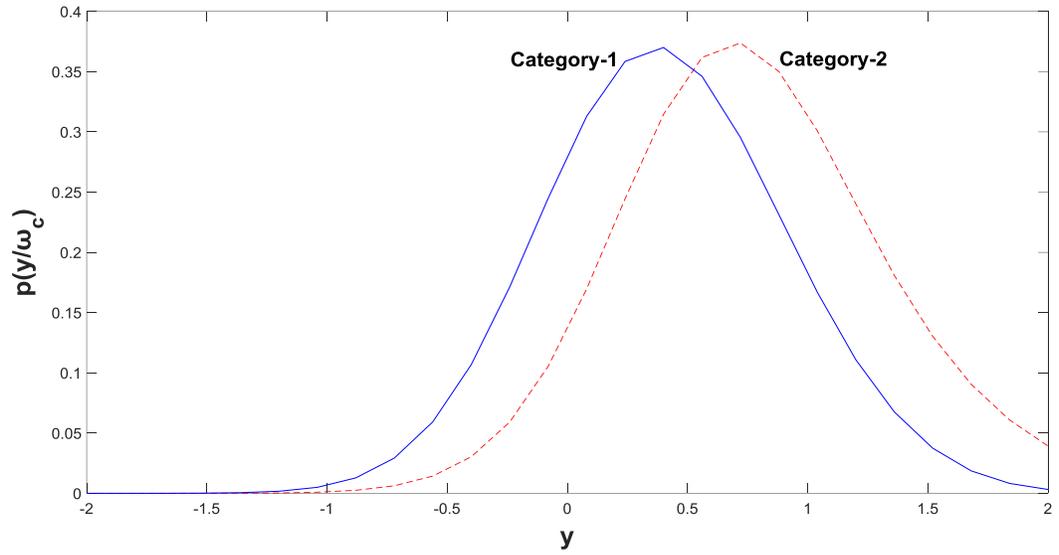


Figure A-2: Probability density functions for two categories: category -1 and category -2

In this study each of the priori probabilities $P(\omega_1)$ and $P(\omega_2)$ are assumed to be 0.5. The objective of this exercise is to minimize the number of incorrect classifications in 10-fold cross validation test. The number of latent variables, r , corresponds to minimum classification error.

A.2 Reliability of classification

Classification according to the method presented in previous section (method used by PLS-Toolbox (Eigenvector, 2013)) does not consider the uncertainty in prediction of \hat{y}_s . If we consider uncertainty, eqn. (A.1) cannot be used to decide the category of the sample. To accommodate this uncertainty Bayes decision rule was changed as

$$P[\omega_c | \hat{y}_{s,l} \leq \hat{y}_s \leq \hat{y}_{s,u}] = \frac{P[\hat{y}_{s,l} \leq \hat{y}_s \leq \hat{y}_{s,u} | \omega_c]}{P[\hat{y}_{s,l} \leq \hat{y}_s \leq \hat{y}_{s,u}]} P(\omega_c) = \frac{\int_{\hat{y}_{s,l}}^{\hat{y}_{s,u}} p(\hat{y}_s | \omega_c) d\hat{y}_s}{\int_{\hat{y}_{s,l}}^{\hat{y}_{s,u}} p(\hat{y}_s) d\hat{y}_s} \dots\dots\dots (A.9)$$

where $\hat{y}_{s,l} = \hat{y}_s - k\text{SEP}_s$ and $\hat{y}_{s,u} = \hat{y}_s + k\text{SEP}_s$ $P(\hat{y}_s) = P(\hat{y}_s | \omega_1)P(\omega_1) + P(\hat{y}_s | \omega_2)P(\omega_2)$; $k = 1, 2$ correspond to 68% and 95% confidence intervals respectively.

The numerator in eqn. (A.9) corresponds to the area under the curve of category- c and the denominator to the total area under both curves. A new sample is assigned to the category for which value of $P[\omega_c | \hat{y}_{s,l} \leq \hat{y}_s \leq \hat{y}_{s,u}]$ is higher.

Because the approach shown in eqn. (A.9) is not built into the PLS-Toolbox (Eigenvector, 2013), code for related calculations was written.

A.3 Interpretation of PLS-DA weights

In summary, the PLS part in the PLS-DA model develops latent inputs $t_1 = \mathbf{w}_1^T \mathbf{x}$, $t_2 = \mathbf{w}_2^T \mathbf{x}$, ..., and coefficients q_1, q_2, \dots such that the output is

$$\begin{aligned} \hat{y} &= q_1 t_1 + q_2 t_2 + \dots \\ &= q_1 \mathbf{w}_1^T \mathbf{x} + q_2 \mathbf{w}_2^T \mathbf{x} + \dots = (q_1 \mathbf{w}_1^T + q_2 \mathbf{w}_2^T + \dots) \mathbf{x} \\ &= \boldsymbol{\beta}^T \mathbf{x} \\ &= \beta_1 x_1 + \beta_2 x_2 + \dots, \end{aligned} \tag{A.10}$$

where

$$\boldsymbol{\beta} = [\mathbf{w}_1 \quad \mathbf{w}_2 \quad \dots] \mathbf{q} \tag{A.11}$$

It is clear from eqn. (A.11) that an input variable x_i for $i = 1, 2, \dots$ will affect \hat{y} positively or negatively based on the sign of the corresponding β_i . But eqn. (A.11) implies that

$$\beta_i = w_{1,i} q_1 + w_{2,i} q_2 + \dots + w_{r,i} q_r \tag{A.12}$$

Now, if a PLS model with two latent input variables ($r = 2$) captures enough variance of the output \hat{y} , then a simple 2D plot can be used to visualize when β_i is positive or negative by checking whether the vectors $[w_{1,i} \quad w_{2,i}]$ and $[q_1 \quad q_2]$ form an acute angle or not (as shown in Figure 3-3).

For the actual calculation of a PLS model there is a number of variants, such as the SIMPLS algorithm (de Jong, 1993).

Appendix B.

B.1 Proof of eqn. (4.10)

It is well known (e.g., through the Jury stability criterion) that the transformation $z = \frac{1-s}{1+s}$ maps the right-half of the complex plane ($\text{Re}[s] > 0$) to the inside of the unit disk ($|z| < 1$). The translation $s = w+1$ immediately yields that the transformation $z = \frac{-w}{2+w}$ maps the area to the right of the vertical axis ($-1, jx$) of the complex plane ($1 + \text{Re}[w] > 0$) to the inside of the unit disk ($|z| < 1$).

B.2 Proof of non-convexity of eqn. (4.12)

For $\mathbf{D} \hat{=} \mathbf{G} - \hat{\mathbf{G}}$ belonging to an uncertainty set D , eqn. (4.12) can be alternatively written as

$$\text{Re} \left[\lambda \left((\hat{\mathbf{G}} + \mathbf{D}) \hat{\mathbf{G}}^{-1} \right) \right] > 0. \quad (\text{B.1})$$

For a 2×2 system it can be easily shown that $\text{Re} \left[\lambda \left((\hat{\mathbf{G}} + \mathbf{D}) \hat{\mathbf{G}}^{-1} \right) \right]$ is a non-convex function of $\mathbf{D} \in D$, as follows: Given a model

$$\hat{\mathbf{G}} \hat{=} \begin{bmatrix} \hat{g}_{11} & \hat{g}_{12} \\ \hat{g}_{21} & \hat{g}_{22} \end{bmatrix} \quad (\text{B.2})$$

and process

$$\mathbf{G} \hat{=} \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} \hat{g}_{11} + d_{11} & \hat{g}_{12} + d_{12} \\ \hat{g}_{21} + d_{21} & \hat{g}_{22} + d_{22} \end{bmatrix}, \quad (\text{B.3})$$

eqn. (B.1) is true for all $\mathbf{D} \in D$ if and only if the following two inequalities are satisfied:

$$\text{tr} \left[\lambda \left((\hat{\mathbf{G}} + \mathbf{D}) \hat{\mathbf{G}}^{-1} \right) \right] = \frac{d_{11} \hat{g}_{22} - d_{12} \hat{g}_{21} + d_{22} \hat{g}_{11} - d_{21} \hat{g}_{12} + 2 \det(\hat{\mathbf{G}})}{\det(\hat{\mathbf{G}})} > 0 \quad (\text{B.4})$$

$$\det \left[\lambda \left((\hat{\mathbf{G}} + \mathbf{D}) \hat{\mathbf{G}}^{-1} \right) \right] = \frac{\overbrace{d_{11} d_{22} - d_{12} d_{21}} + d_{11} \hat{g}_{22} - d_{12} \hat{g}_{21} - d_{21} \hat{g}_{12} + d_{22} \hat{g}_{11} + \det(\hat{\mathbf{G}})}{\det(\hat{\mathbf{G}})} > 0 \quad (\text{B.5})$$

Eqn. (B.4) is linear in terms of the elements of the matrix \mathbf{D} , but eqn. (B.5) contains the bilinear term $d_{11} d_{22} - d_{12} d_{21}$, which is generally non-convex.

The preceding analysis can be easily generalized for any higher-order system, as the Routh-Hurwitz theorem for eqn. (B.1) will contain terms that are bilinear or multilinear in the entries of $\mathbf{D} \in D$, hence generally non-convex.

Appendix C

C.1 Controller design for motivating example in Introduction

Given a transfer matrix model $\hat{\mathbf{G}}f(z)$, where $f(z)$ is a stable transfer function with stable inverse, the internal-model control (IMC) method can be used to design a feedback controller $\mathbf{C}(z) = \mathbf{Q}(z) \left(\mathbf{I} - \hat{\mathbf{G}}f(z)\mathbf{Q}(z) \right)^{-1}$ by selecting $\mathbf{Q}(z) = \hat{\mathbf{G}}^{-1}\mathbf{F}(z)/f(z)$, which yields

$$\mathbf{C}(z) = \hat{\mathbf{G}}^{-1}\mathbf{F}(z) \left(\mathbf{I} - \mathbf{F}(z) \right)^{-1} / f(z) \quad (\text{C.1})$$

For a diagonal IMC filter $\mathbf{F}(z) = \frac{1-\alpha}{z-\alpha} \mathbf{I}$, the closed-loop sensitivity function

becomes

$$\left(\mathbf{I} + \mathbf{G}f(z)\mathbf{C}(z) \right)^{-1} = \left(\mathbf{I} + \mathbf{G}\hat{\mathbf{G}}^{-1} \frac{1-\alpha}{z-1} \right)^{-1} = (z-1)\hat{\mathbf{G}}^{-1} \left(\hat{\mathbf{G}}(z-1) + \mathbf{G}(1-\alpha) \right)^{-1} \quad (\text{C.2})$$

with closed-loop characteristic equation $a_2 + a_1z + z^2 = 0$, where the parameters a_1, a_2 are

$$a_2 = 0.0215 + 4.67\alpha - 3.69\alpha^2, \quad a_1 = 4.71 + 2.71\alpha \quad (\text{C.3})$$

for $\hat{\mathbf{G}}_1$, and

$$a_2 = -0.0417 + 1.04\alpha^2, \quad a_1 = 0.833 - 2.08\alpha \quad (\text{C.4})$$

for $\hat{\mathbf{G}}_2$.

For the roots of the characteristic equation to be inside the unit disk, the Jury stability criterion requires

$$\begin{aligned}
 a_2 - 1 &< 0 \\
 -a_1 - a_2 - 1 &< 0 \\
 a_1 - a_2 - 1 &< 0
 \end{aligned}
 \tag{C.5}$$

Figure C-1 indicates that these inequalities cannot be satisfied simultaneously for any value of α in the interval $[0,1)$ when $\hat{\mathbf{G}}_1$ is used in the controller structure of eqn. (C.1), but they can be easily satisfied for all α in the interval $[0,1)$ when $\hat{\mathbf{G}}_2$ is used in that controller structure.

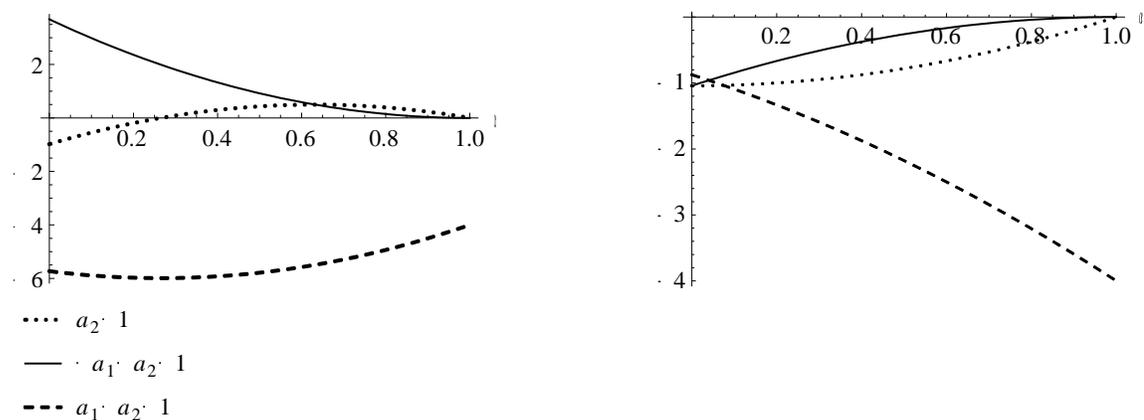


Figure C-1 Satisfaction of the inequalities in eqn. (C.5), required by the Jury stability criterion for closed-loop stability, when $\hat{\mathbf{G}}_1$ (left) and $\hat{\mathbf{G}}_2$ (right) is used in the feedback controller, eqn. (C.1).

C.2 Adaptive design to build an IC-compliant model

1. Develop a preliminary model $\hat{\mathbf{G}}$ from input-output data using standard PRBS inputs for limited time.
2. Calculate \mathbf{C}_m by solving below optimization problem:

$$\mathbf{C}_m^{opt} = \min_{\mathbf{C}_m} \beta \quad (\text{C.6})$$

subject to variance constraints on individual outputs and inputs

$$\begin{aligned} \text{var}(\mathbf{m}_i) &= [\mathbf{C}_m]_{ii} \leq M_i^2 \\ \text{var}(\mathbf{y}_i) &= [\hat{\mathbf{G}}\mathbf{C}_m\hat{\mathbf{G}}^T]_{ii} \leq Y_i^2 \end{aligned} \quad (\text{C.7})$$

3. Calculate \mathbf{Q}^{opt} by Cholesky decomposition of $\mathbf{C}_m^{opt} = \mathbf{Q}^{opt}(\mathbf{Q}^{opt})^T$ and design input $\mathbf{m} = \mathbf{Q}^{opt}\mathbf{z}$ where \mathbf{z} is a zero-mean PRBS with unit covariance.
4. Add new input \mathbf{m} in the set of input signals; perform identification and update model $\hat{\mathbf{G}}$.
5. Check for satisfaction of IC sufficient condition (eqn. (5.9)); stop if maximum number of iterations achieved, otherwise go to step-2 and repeat subsequent steps.

C.3 Proof of Lemma 1

The following Lemma 1 is proved first, to make it easier to follow the proof of the subsequent Theorem 2.

Lemma 1 – Sufficient condition for IC of model identified as a single, partially known MIMO model

For a model $\hat{\mathbf{G}} \in \mathbb{R}^{n \times n}$ and a real plant $\mathbf{G} \in \mathbb{R}^{n \times n}$ with uncertainty $\mathbf{D} = \mathbf{G} - \hat{\mathbf{G}}$, the integral controllability condition (eqn. (5.4)) is satisfied if

$$J_c \triangleq \sum_{k=1}^n \frac{1}{\hat{\sigma}_k} \left\| \hat{\mathbf{z}}_k^T \text{vec}(\mathbf{D}^T) \right\|_2 < 1, \quad (\text{C.8})$$

where

$$\hat{\mathbf{Z}}_k^T = \begin{bmatrix} \hat{\mathbf{v}}_k^T & \mathbf{0} & \ddots \\ \mathbf{0} & \ddots & \mathbf{0} \\ \ddots & \mathbf{0} & \hat{\mathbf{v}}_k^T \end{bmatrix} = \mathbf{I}_n \otimes \hat{\mathbf{v}}_k^T \in \mathbb{R}^{n \times n^2}, \quad k = 1, \dots, n \quad (\text{C.9})$$

$\hat{\mathbf{v}}_k$, $k = 1, \dots, n$ are the right singular vectors of the SVD of

$$\hat{\mathbf{G}} = \sum_{k=1}^n \sigma_k \hat{\mathbf{u}}_k \hat{\mathbf{v}}_k^T \quad (\text{C.10})$$

and

$$\text{vec}(\mathbf{D}^T) = \text{vec}(\mathbf{G}^T) - \text{vec}(\hat{\mathbf{G}}^T) \triangleq [\mathbf{d}_1^T \quad \dots \quad \mathbf{d}_n^T]^T \quad (\text{C.11})$$

Proof: Eqn. (5.4) is equivalent to

$$\text{Re} \left[\lambda \left((\hat{\mathbf{G}} + \mathbf{D}) \hat{\mathbf{G}}^{-1} \right) \right] > 0 \Leftrightarrow 1 + \text{Re} \left[\lambda \left(\mathbf{D} \hat{\mathbf{G}}^{-1} \right) \right] > 0 \quad (\text{C.12})$$

which is satisfied if

$$\left| \lambda \left(\mathbf{D} \hat{\mathbf{G}}^{-1} \right) \right| < 1 \quad (\text{C.13})$$

By the matrix spectral radius theorem we have

$$\left| \lambda \left(\mathbf{D} \hat{\mathbf{G}}^{-1} \right) \right| \leq \left\| \mathbf{D} \hat{\mathbf{G}}^{-1} \right\|_i \quad (\text{C.14})$$

for any induced matrix norm $\|\cdot\|_i$. Therefore, eqn. (C.14) is satisfied if

$$\left\| \mathbf{D} \hat{\mathbf{G}}^{-1} \right\|_i < 1 \quad (\text{C.15})$$

for some induced matrix norm $\|\cdot\|_i$.

Using the singular-value decomposition $\hat{\mathbf{G}} = \sum_{k=1}^n \hat{\sigma}_k \hat{\mathbf{u}}_k \hat{\mathbf{v}}_k^T$, eqn. (C.15) is equivalent

to

$$\left\| \mathbf{D} \left(\sum_{k=1}^n \hat{\sigma}_k \hat{\mathbf{u}}_k \hat{\mathbf{v}}_k^T \right)^{-1} \right\|_i < 1 \Leftrightarrow \left\| \sum_{k=1}^n \frac{\mathbf{D} \hat{\mathbf{v}}_k \hat{\mathbf{u}}_k^T}{\hat{\sigma}_k} \right\|_i < 1 \quad (\text{C.16})$$

which is satisfied if

$$\sum_{k=1}^n \frac{1}{\hat{\sigma}_k} \|\mathbf{D}\hat{\mathbf{v}}_k\hat{\mathbf{u}}_k^T\|_i < 1 \quad (\text{C.17})$$

for any induced matrix norm $\|\cdot\|_i$.

Using the vec operator, the term $\mathbf{D}\hat{\mathbf{v}}_k$ in the above inequality can be written as

$$\begin{aligned} \mathbf{D}\hat{\mathbf{v}}_k &= \begin{bmatrix} \mathbf{d}_1^T \\ \vdots \\ \mathbf{d}_n^T \end{bmatrix} \hat{\mathbf{v}}_k = \begin{bmatrix} \hat{\mathbf{v}}_k^T \mathbf{d}_1 \\ \vdots \\ \hat{\mathbf{v}}_k^T \mathbf{d}_n \end{bmatrix} \\ &= \underbrace{\begin{bmatrix} \hat{\mathbf{v}}_k^T & \mathbf{0} & \ddots \\ \mathbf{0} & \ddots & \mathbf{0} \\ \ddots & \mathbf{0} & \hat{\mathbf{v}}_k^T \end{bmatrix}}_{\hat{\mathbf{Z}}_k^T} \underbrace{\begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_n \end{bmatrix}}_{\text{vec}(\mathbf{D}^T)} = \hat{\mathbf{Z}}_k^T \text{vec}(\mathbf{D}^T) \end{aligned} \quad (\text{C.18})$$

Therefore

$$\|\mathbf{D}\hat{\mathbf{v}}_k\hat{\mathbf{u}}_k^T\|_{i_2} = \|\mathbf{D}\hat{\mathbf{v}}_k\|_{i_2} = \|\hat{\mathbf{Z}}_k^T \text{vec}(\mathbf{D}^T)\|_{i_2} = \|\hat{\mathbf{Z}}_k^T \text{vec}(\mathbf{D}^T)\|_2, \quad (\text{C.19})$$

where $\|\cdot\|_{i_2}$ is the induced 2-norm of a matrix.

Consequently, eqn. (C.19) implies that the inequality in eqn. (C.17) is satisfied if

$$\sum_{k=1}^n \frac{1}{\hat{\sigma}_k} \|\hat{\mathbf{Z}}_k^T \text{vec}(\mathbf{D}^T)\|_2 < 1. \quad (\text{C.20})$$

C.4 Proof of Theorem 2

By the Cauchy-Schwarz inequality, we have that

$$\left(\sum_{k=1}^n \sqrt{\mathbf{x}_k^T \mathbf{x}_k} \right)^2 \leq n \sum_{k=1}^n \mathbf{x}_k^T \mathbf{x}_k. \quad (\text{C.21})$$

Applied to the left-hand side of eqn. (C.20) with $\mathbf{x}_k \triangleq \frac{1}{\hat{\sigma}_k} \hat{\mathbf{Z}}_k^T \text{vec}(\mathbf{D}^T)$, eqn. (C.21)

implies that the inequality in eqn. (C.20) is satisfied if

$$\begin{aligned} n \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} \text{vec}(\mathbf{D}^T)^T \hat{\mathbf{Z}}_k \hat{\mathbf{Z}}_k^T \text{vec}(\mathbf{D}^T) < 1 &\Leftrightarrow \\ \text{vec}(\mathbf{D}^T)^T \underbrace{\left(\sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} \hat{\mathbf{Z}}_k \hat{\mathbf{Z}}_k^T \right)}_{\Phi} \text{vec}(\mathbf{D}^T) < \frac{1}{n} &\Leftrightarrow \\ \text{vec}(\mathbf{D}^T)^T \Phi \text{vec}(\mathbf{D}^T) < \frac{1}{n}. \end{aligned} \quad (\text{C.22})$$

The worst case of model uncertainty corresponds to the maximum in the left-hand side of the final inequality in eqn. (C.22), namely

$$\max_{\mathbf{D} \in \Delta} \underbrace{\text{vec}(\mathbf{D}^T)^T}_{\mathbf{x}^T} \Phi \underbrace{\text{vec}(\mathbf{D}^T)}_{\mathbf{x}} \quad (\text{C.23})$$

with

$$\Delta = \left\{ \mathbf{D} : \underbrace{\text{vec}(\mathbf{D}^T)^T}_{\mathbf{x}^T} \mathbf{B} \underbrace{\text{vec}(\mathbf{D}^T)}_{\mathbf{x}} \leq r_{\mathbf{B}}^2 \right\} \quad (\text{C.24})$$

according to eqn. (5.26).

Using the method of Lagrange multipliers (see below), the maximum of $\mathbf{x}^T \Phi \mathbf{x}$ in the above optimization problem of eqns. (C.23) and (C.24) can be easily shown (see below) to be equal to $\mu_{\max}(\mathbf{B}^{-1} \Phi) r_{\mathbf{B}}^2$ and to be attained for \mathbf{x}_{\max} being the eigenvector corresponding to the largest eigenvalue of the matrix $\mathbf{B}^{-1} \Phi$, $\mu_{\max}(\mathbf{B}^{-1} \Phi)$.

Finally, it is straightforward to show that $\Phi \hat{=} \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} \hat{\mathbf{z}}_k \hat{\mathbf{z}}_k^T = \mathbf{I}_n \otimes (\hat{\mathbf{G}}^T \hat{\mathbf{G}})^{-1}$,

because

$$(\hat{\mathbf{G}}^T \hat{\mathbf{G}})^{-1} = (\hat{\mathbf{V}} \hat{\mathbf{S}}^T \hat{\mathbf{U}} \hat{\mathbf{S}} \hat{\mathbf{V}}^T)^{-1} = \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} \hat{\mathbf{v}}_k \hat{\mathbf{v}}_k^T \quad (\text{C.25})$$

and the mixed-product property of the Kronecker product,

$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$, implies

$$\begin{aligned} \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} \hat{\mathbf{z}}_k \hat{\mathbf{z}}_k^T &= \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} (\mathbf{I}_n \otimes \hat{\mathbf{v}}_k) (\mathbf{I}_n \otimes \hat{\mathbf{v}}_k^T) \\ &= \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} (\mathbf{I}_n \mathbf{I}_n) \otimes (\hat{\mathbf{v}}_k \hat{\mathbf{v}}_k^T) \\ &= \mathbf{I}_n \otimes \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^2} (\hat{\mathbf{v}}_k \hat{\mathbf{v}}_k^T) \\ &= \mathbf{I}_n \otimes (\hat{\mathbf{G}}^T \hat{\mathbf{G}})^{-1}. \end{aligned} \quad (\text{C.26})$$

Lagrange multiplier method:

First, $\max_{\mathbf{x} \in \Delta} \mathbf{x}^T \Phi \mathbf{x}$ such that $\mathbf{x}^T \mathbf{B} \mathbf{x} \leq r_{\mathbf{B}}^2$ is attained for

$$\mathbf{x}^T \mathbf{B} \mathbf{x} = r_{\mathbf{B}}^2 \quad (\text{C.27})$$

because, if the optimum were attained for $(\mathbf{x}^{\text{opt}})^T \mathbf{B}(\mathbf{x}^{\text{opt}}) < r_{\mathbf{B}}^2$, then scaling up \mathbf{x}^{opt} would trivially yield a larger value for $\mathbf{x}^T \Phi \mathbf{x}$.

Then, the Lagrangian is $L = \mathbf{x}^T \Phi \mathbf{x} - \eta(\mathbf{x}^T \mathbf{B} \mathbf{x} - r_{\mathbf{B}}^2)$, which implies that at the optimum

$$\frac{dL}{d\mathbf{x}^T} = 2(\Phi - \eta \mathbf{B})\mathbf{x} = \mathbf{0} \Leftrightarrow \mathbf{B}^{-1} \Phi \mathbf{x} = \eta \mathbf{x} \quad (\text{C.28})$$

which, in turn, implies that (η, \mathbf{x}) is an eigenvalue-eigenvector pair for the matrix $\mathbf{B}^{-1} \Phi$.

Now, at the optimum, the objective function will be

$$\mathbf{x}^T \Phi \mathbf{x} = \mathbf{x}^T \mathbf{B} \mathbf{B}^{-1} \Phi \mathbf{x} = \eta \mathbf{x}^T \mathbf{B} \mathbf{x} = \eta r_{\mathbf{B}}^2 \quad (\text{C.29})$$

which implies that η must be the largest of all eigenvalues of $\mathbf{B}^{-1} \Phi$, $\mu_{\max}(\mathbf{B}^{-1} \Phi)$, \mathbf{x} its corresponding eigenvector, and the maximum of $\mathbf{x}^T \Phi \mathbf{x}$ is $\mathbf{x}_{\max}^T \Phi \mathbf{x}_{\max} = \mu_{\max}(\mathbf{B}^{-1} \Phi) r_{\mathbf{B}}^2$.

C.5 Proof of eqn. (5.45)

It is straightforward to show that the eigenvalues of $\mu(\mathbf{B}^{-1} \Phi)$ satisfy the equalities

$$\mu(\mathbf{B}^{-1} \Phi) = \mu\left(\left(\mathbf{M}^T \mathbf{M}\right)^{-1} \left(\hat{\mathbf{G}}^T \hat{\mathbf{G}}\right)^{-1}\right) = \mu\left(\mathbf{P} \Lambda^{-1} \mathbf{P}^T \hat{\mathbf{V}} \hat{\Sigma}^{-2} \hat{\mathbf{V}}^T\right) = \mu\left(\begin{array}{c} \hat{\mathbf{V}}^T \mathbf{P} \Lambda^{-1} \mathbf{P}^T \hat{\mathbf{V}} \hat{\Sigma}^{-2} \\ \mathbf{Q} \quad \mathbf{Q}^T \end{array}\right), \quad (\text{C.30})$$

the last equality owing to the similarity of the matrices $\mathbf{P} \Lambda^{-1} \mathbf{P}^T \hat{\mathbf{V}} \hat{\Sigma}^{-2} \hat{\mathbf{V}}^T$ and $\hat{\mathbf{V}}^T \mathbf{P} \Lambda^{-1} \mathbf{P}^T \hat{\mathbf{V}} \hat{\Sigma}^{-2} = \hat{\mathbf{V}}^T (\mathbf{P} \Lambda^{-1} \mathbf{P}^T \hat{\mathbf{V}} \hat{\Sigma}^{-2} \hat{\mathbf{V}}) \hat{\mathbf{V}}^T$.

Because $\mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \hat{\Sigma}^{-2}$ is positive definite, its largest eigenvalue $\mu_{\max}(\mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \hat{\Sigma}^{-2})$ is equal to its spectral radius $\rho(\mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \hat{\Sigma}^{-2})$. A standard upper bound for the spectral radius is the Frobenius norm of the corresponding matrix, which yields the following:

$$\begin{aligned} \mu_{\max}(\mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \hat{\Sigma}^{-2}) &= \rho(\mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \hat{\Sigma}^{-2}) \\ &\leq \left\| \mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \hat{\Sigma}^{-2} \right\|_F \\ &= \sqrt{\text{tr}\left(\hat{\Sigma}^{-2} \mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \mathbf{Q} \Lambda^{-1} \mathbf{Q}^T \hat{\Sigma}^{-2}\right)} \\ &= \sqrt{\text{tr}\left(\mathbf{Q} \Lambda^{-2} \mathbf{Q}^T \hat{\Sigma}^{-4}\right)} \end{aligned} \quad (\text{C.31})$$

C.6 Proof of Theorem 3

By the properties of matrix trace it follows that

$$\text{tr}(\mathbf{Q}\mathbf{\Lambda}^{-2}\mathbf{Q}^T\hat{\mathbf{\Sigma}}^{-4}) = \text{tr}\left(\sum_{k=1}^n \frac{\mathbf{q}_k\mathbf{q}_k^T}{\lambda_k^4}\hat{\mathbf{\Sigma}}^{-4}\right) = \sum_{j=1}^n \frac{\mathbf{q}_j^T\hat{\mathbf{\Sigma}}^{-4}\mathbf{q}_j}{\lambda_j^4} \quad (\text{C.32})$$

Defining the Lagrangian as $L = \sum_{j=1}^n \frac{\mathbf{q}_j^T\hat{\mathbf{\Sigma}}^{-4}\mathbf{q}_j}{\lambda_j^4} - \sum_{i=1}^n \sum_{k=1}^i \eta_{ik} (\mathbf{q}_i^T\mathbf{q}_k - \delta_{ik})$ and setting its

partial derivatives with respect to \mathbf{q}_ℓ equal to zero yields

$$\cdot \left(\frac{\partial L}{\partial \mathbf{q}_\ell}\right)^T = 2 \frac{\hat{\mathbf{\Sigma}}^{-4}\mathbf{q}_\ell}{\lambda_\ell^4} - \sum_{k=1}^{\ell} \eta_{\ell k} \mathbf{q}_k - \sum_{i=\ell}^n \eta_{i\ell} \mathbf{q}_i = \mathbf{0}. \quad (\text{C.33})$$

Premultiplying eqn. (C.33) by \mathbf{q}_ℓ^T for $\ell = 1, \dots, n$ gives

$$\mathbf{q}_\ell^T \underbrace{\left(\frac{\hat{\mathbf{\Sigma}}^{-4}}{\lambda_\ell^4} - \eta_{\ell\ell} \mathbf{I}\right)}_{\mathbf{B}^{(\ell)}} \mathbf{q}_\ell = 0 \Leftrightarrow \sum_{k=1}^n b_{kk}^{(\ell)} q_{\ell,k}^2 = 0, \quad \ell = 1, \dots, n \quad (\text{C.34})$$

With singular values in $\hat{\mathbf{\Sigma}}$ distinct (eqn. (5.52)), the diagonal matrix

$\mathbf{B}^{(\ell)} = \frac{\hat{\mathbf{\Sigma}}^{-4}}{\lambda_\ell^4} - \eta_{\ell\ell} \mathbf{I}$ can only have one entry equal to zero, for a corresponding choice of $\eta_{\ell\ell}$

. Consequently, \mathbf{q}_ℓ in eqn. (C.34) can only have one nonzero entry, corresponding to the

diagonal zero entry of $\mathbf{B}^{(\ell)}$; that is \mathbf{q}_ℓ must be a unit vector $\mathbf{e}_{\pi_\ell} \hat{=} [0, \dots, 0, \pi_\ell, 0, \dots, 0]^T$ with

$\pi_\ell \neq 0$. Substituting \mathbf{q}_ℓ in eqn. (C.32) yields $\sum_{k=1}^n \frac{1}{\lambda_{\ell_k}^4 \sigma_k^4}$, where index

$[\ell_1, \dots, \ell_n] = [1, \dots, n] \mathbf{\Pi}$. By the rearrangement inequality, it can be trivially shown that the

minimum value of $\sum_{k=1}^n \frac{1}{\lambda_{\ell_k}^4 \sigma_k^4}$ is obtained when σ_k and λ_{ℓ_k} are matched in reverse order.

Therefore,

$$\mathbf{P}^{\text{opt}} = \arg \left[\min_{\mathbf{P} \in \mathbb{R}^{n \times n}, \text{ orthonormal}} \text{tr}(\mathbf{A}^{-1} \mathbf{H}) \right] = \hat{\mathbf{V}} \mathbf{Q}^{\text{opt}} = \hat{\mathbf{V}} \mathbf{\Pi} \quad (\text{C.35})$$

C.7 Proof of Theorem 4

Substituting $\hat{\mathbf{V}}$ from eqn. (5.54) into eqn. (5.50) yields

$$\sum_{k=1}^n \underbrace{(\chi \hat{\sigma}_k^2 + 1 - \chi) \lambda_{\ell_k}^2}_{b_k^2} \leq w^2 \quad (\text{C.36})$$

Using the Lagrange multiplier method, minimization of the objective function (eqn. (5.60)) subject to constraints as in eqn. (C.36) and ordering of λ_{ℓ_k} in eqn. (5.51) proceeds as follows:

$$\text{Define the Lagrangian as } L = \sum_{k=1}^n \frac{1}{\hat{\sigma}_k^4 \lambda_{\ell_k}^4} + \sum_{i=0}^n \mu_i g_i,$$

where

$$\begin{aligned} g_0 &= \sum_{k=1}^n \underbrace{(\chi \hat{\sigma}_k^2 + 1 - \chi) \lambda_{\ell_k}^2}_{b_k^2} - w^2 \leq 0 \\ g_1 &= -\lambda_{\ell_1} < 0 \\ g_i &= \lambda_{\ell_{i-1}} - \lambda_{\ell_i} < 0, \quad 2 \leq i \leq n \end{aligned} \quad (\text{C.37})$$

Applying the Karush-Kuhn-Tucker (KKT) conditions

$$\frac{\partial L}{\partial \lambda_{\ell_k}} = -\frac{4}{\hat{\sigma}_k^4 \lambda_{\ell_k}^5} + \mu_0 (2 \lambda_{\ell_k} b_k^2) + (\mu_{k+1} - \mu_k) = 0, \quad (\text{C.38})$$

$$\mu_i g_i(\lambda_{\ell_k}) = 0, \quad 0 \leq i \leq n, \text{ and} \quad (\text{C.39})$$

$$\mu_i \geq 0, \quad 0 \leq i \leq n \quad (\text{C.40})$$

yields $\lambda_{\ell_k} = \left(\frac{2}{\mu_0 \hat{\sigma}_k^4 b_k^2} \right)^{1/6}$, $g_0 = 0$, and $\mu_i = 0$, for $1 \leq i \leq n$.

Solving $\lambda_{\ell_k} = \left(\frac{2}{\mu_0 \hat{\sigma}_k^4 b_k^2} \right)^{1/6}$ with $g_0 = 0$, yields the optimal solution

$$\lambda_{\ell_k} = \frac{w}{(\hat{\sigma}_k^2 b_k)^{1/3}} \frac{1}{\left[\sum_{i=1}^n \left(\frac{b_i}{\hat{\sigma}_i} \right)^{4/3} \right]^{1/2}}. \quad (\text{C.41})$$

List of Publications

1. **Panjwani, S.**, Nikolaou, M. (2016), “*Experiment design for control-relevant identification of partially known stable multivariable systems*”. *AIChE J.* doi:10.1002/aic.15212 (Chapter-5)
2. **Panjwani, S.**, Nikolaou, M., “*Ensuring Integral Controllability for Robust Multivariable Control*”, *Computers and Chemical Engineering* 92 (2016) 172–179 (Chapter-4)
3. **Panjwani, S.**, Nikolaou, M., “*Improvement of Zonal Isolation in Horizontal Shale Gas Wells: A Data-Driven Model-Based Approach*”, *SPE Drilling and Completion*, 2015 (submitted revised manuscript) (Chapter-2 and 3)