"To be yourself in a world that is constantly trying to make you something else is the greatest accomplishment."
- Ralph Waldo Emerson

University of Alberta

Control Loop Performance Assessment with Closed-loop Subspace Identification

by

Nima Danesh Pour

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Master of Science

in

Process Control

Department of Chemical and Materials Engineering

Fall 2009
Edmonton, Alberta

Examining Committee

Prof. Biao Huang, Chemical and Materials Engineering

Prof. Sirish L. Shah, Chemical and Materials Engineering

Dr. Jong Min Lee, Chemical and Materials Engineering

Prof. Tongwen Chen, Electrical and Computer Engineering

To my *LOVE* and my parents

Abstract

This thesis is concerned with subspace identification and its applications for controller performance assessment and process modeling from closed-loop data.

A joint input-output closed-loop subspace identification method is developed which provides consistent estimation of the subspace matrices and the noise covariance matrix required for the LQG benchmark curve estimation.

Subspace LQG benchmark is also used for performance assessment of the cascade supervisory-regulatory control systems. Three possible scenarios for LQG control design and performance improvement are discussed for this structure. A closed-loop subspace identification method is also provided for estimation of the subspace matrices necessary for performance assessment.

A method of direct step model estimation from closed-loop data is provided using subspace identification. The variance calculation required for this purpose can be performed using the proposed method. The variances are used for weighted averaging on the estimated Markov parameters to attenuate the noise influence on the final step response estimation.

Acknowledgements

First of all, I express my gratitude to my Lord, the most High.

It has been a privilege to have both Prof. Biao Huang and Prof. Sirish L. Shah as supervisors. It is the encouragement, inspiration and ingenious direction of the former and the continuous guidance, constructive criticism and friendly support of the later that made this thesis a reality in a short time. Specifically, I have to appreciate all the help a I got in the the weekly meetings with Prof. Huang which was necessary for each and every step of this research. He took a lot of time to review and improve my work over and over.

I had the pleasure of working with Dr. Dave Shook as his assistant and I am also very thankful to him for all his help on performing my experiments in the CPC lab. I would also like to thank my colleagues in the process control group. Specially, I am thankful to my friends Anuj Narang, Dr. Mehrdad Sahebsara and Sadegh Amiri. Many thanks to my friend Dr. Iman Izadi for his help with writing of this thesis.

Financial support from Natural Sciences and Engineering Research Council of Canada (NSERC) and Syncrude Canada Ltd. is greatly acknowledged.

I am grateful to my parents. They always encouraged and supported me to seek out my own goals and for that I will be eternally grateful to them. Last but not the least, many many appreciations are due to my lovely wife Somayeh for all her support, loving care and enduring patience during this time.

Contents

List of Tables

List of Figures

List of Symbols

$\{A, B, C, D\}$	Dynamic state space system matrices of the process
$\{A_{cl}, B_{cl}, C_{cl}, D_{cl}\}$	Dynamic state space system matrices of the closed-loop system
$\{A_c, B_c, C_c, D_c\}$	Dynamic state space system matrices of the controller
$\hat{y}_t$	Predicted value of system output(s) at sampling instant t
d	Process time delay for the univariate process
E_f	Future data Hankel matrix for e_t
E_p	Past data Hankel matrix for e_t
e_t	White noise (innovations) sequence
f_i	Impulse response coefficient at ith sample
$G_c(z^{-1})$	Transfer function representation of the controller
$G_l(z^{-1})$	Transfer function representation of the stochastic part of the system or disturbance model
$G_p(z^{-1})$	Transfer function representation of the deterministic part of the system or process model
h_i^j	element (i, j) in the Toeplitz subspace matrix containing Markov parameters
I	Identity matrix
J	Optimization objective function
j	Number of block-columns in the block-Hankel matrices
J_{mvc}	Minimum variance control objective function
K	Kalman filter gain matrix

l Dimension of system input(s)

L_e Subspace matrix containing the noise model Markov parameters

$$\begin{pmatrix} I & 0 & \cdots & 0 \\ CK & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}K & CA^{N-3}K & \cdots & I \end{pmatrix}$$

L_u Subspace matrix containing the process Markov parameters

$$\begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}B & CA^{N-3}B & \cdots & D \end{pmatrix}$$

L_U Closed-loop subspace matrix between $W_p^{ur} \rightarrow U_f$

L_U^{cs} Closed-loop subspace matrix between $W_p^{ur} \rightarrow U_f$ in the cascade control structure

$L_{\bar{U}}^{cs}$ Closed-loop subspace matrix between $W_p^{\bar{u}r} \rightarrow \bar{U}_f$ in the cascade control structure

L_{UE},L_{ue}^{CL} Closed-loop subspace matrix between $E_f \rightarrow U_f$

L_{UE}^{cs} Closed-loop subspace matrix between $E_f \rightarrow U_f$ in the cascade control structure

$L_{\bar{U}E}^{cs}$ Closed-loop subspace matrix between $E_f \rightarrow \bar{U}_f$ in the cascade control structure

L_{UR},L_{ur}^{CL} Closed-loop subspace matrix between $R_f \rightarrow U_f$

L_{UR}^{cs} Closed-loop subspace matrix between $R_f \rightarrow U_f$ in the cascade control structure

$L_{\bar{U}R}^{cs}$ Closed-loop subspace matrix between $R_f \rightarrow \bar{U}_f$ in the cascade control structure

$L_{U\bar{U}}$ Closed-loop subspace matrix between $\bar{U}_f \rightarrow U_f$ in the cascade control structure

L_u^{CL} Closed-loop subspace matrix between $W_p^{CL} \rightarrow U_f$

L_w Subspace matrix corresponding to past inputs and outputs

L_Y Closed-loop subspace matrix between $W_p^{yr} \rightarrow Y_f$

L_Y^{cs}	Closed-loop subspace matrix between $W_p^{yr} \rightarrow Y_f$ in the cascade control structure
L_{YE},L_{ye}^{CL}	Closed-loop subspace matrix between $E_f \rightarrow Y_f$
L_{YE}^{cs}	Closed-loop subspace matrix between $E_f \rightarrow Y_f$ in the cascade control structure
L_{YR},L_{yr}^{CL}	Closed-loop subspace matrix between $R_f \rightarrow Y_f$
L_{YR}^{cs}	Closed-loop subspace matrix between $R_f \rightarrow Y_f$ in the cascade control structure
$L_{Y\bar{U}}$	Closed-loop subspace matrix between $\bar{U}_f \rightarrow Y_f$ in the cascade control structure
L_y^{CL}	Closed-loop subspace matrix between $W_p^{CL} \rightarrow Y_f$
L_y^c	Subspace matrix containing the process Markov parameters of the controller
L_y^{c1}	Subspace matrix containing the process Markov parameters of the controller C^1 in the cascade control structure
L_y^{c2}	Subspace matrix containing the process Markov parameters of the controller C^2 in the cascade control structure
L_y^{cc}	Defined for the cascade control structure as $L_y^{c2}L_y^{c1} + L_y^{c2}$
m	Dimension of system output(s)
N	Number of block-rows in the block-Hankel matrices
n	Dimension of system state(s)
n_c	Dimension of controller state(s)
P_v	Covariance matrix of the least squares residual
R	Covariance of innovation (white noise) sequence e_t
R_f	Future data Hankel matrix for r_t defined as $\begin{pmatrix} r_N & r_{N+1} & \cdots & r_{N+j-1} \\ r_{N+1} & r_{N+2} & \cdots & r_{N+j} \\ \cdots & \cdots & \cdots & \cdots \\ r_{2N-1} & r_{2N} & \cdots & r_{2N+j-2} \end{pmatrix}$

R_p Past data Hankel matrix for r_t defined as

$$\begin{pmatrix} r_0 & r_1 & \cdots & r_{j-1} \\ r_1 & r_2 & \cdots & r_j \\ \cdots & \cdots & \cdots & \cdots \\ r_{N-1} & r_N & \cdots & r_{N+j-2} \end{pmatrix}$$

r_t Setpoint (reference) value at sampling instant t

U, S, V Matrices of singular value decomposition

U_f Future data Hankel matrix for u_t

U_p Past data Hankel matrix for u_t

u_t System input(s) at sampling instant t

$\bar{u}_t$ High-level controller output in the cascade control structure at sampling instant t

W_p $\begin{pmatrix} Y_p \\ U_p \end{pmatrix}$

W_p^{CL} $\begin{pmatrix} Y_p \\ U_p \\ R_p \end{pmatrix}$

W_p^{ur} $\begin{pmatrix} U_p \\ R_p \end{pmatrix}$

$W_p^{u\bar{u}}$ $\begin{pmatrix} U_p \\ \bar{U}_p \end{pmatrix}$

$W_p^{\bar{u}r}$ $\begin{pmatrix} \bar{U}_p \\ R_p \end{pmatrix}$

W_p^{yr} $\begin{pmatrix} Y_p \\ R_p \end{pmatrix}$

$W_p^{y\bar{u}}$ $\begin{pmatrix} Y_p \\ \bar{U}_p \end{pmatrix}$

$W_p^{\xi u}$ $\begin{pmatrix} \Xi_p \\ U_p \end{pmatrix}$

X_f System future state matrix defined as

$$\begin{pmatrix} x_N & \cdots & x_{N+j-1} \end{pmatrix}$$

X_f^c Controller future state matrix defined as

$$\begin{pmatrix} x_N^c & \cdots & x_{N+j-1}^c \end{pmatrix}$$

X_f^{CL}	Future closed-loop state matrix defined as $\begin{pmatrix} X_f \\ X_f^c \end{pmatrix}$
X_f^{cc1}	Future state matrix defined in the cascade control structure as $\begin{pmatrix} X_f \\ X_f^{c1} \end{pmatrix}$
X_f^{cc2}	Future state matrix defined in the cascade control structure as $\begin{pmatrix} X_f \\ X_f^{c2} \end{pmatrix}$
X_f^{cs}	Future state matrix defined in the cascade control structure as $\begin{pmatrix} X_f \\ X_f^{c1} \\ X_f^{c2} \end{pmatrix}$
X_p	Past state matrix defined as $\begin{pmatrix} x_0 & \cdots & x_{j-1} \end{pmatrix}$
x_t	System states at sampling instant t
x_t^c	Controller states at sampling instant t
x_t^{c1}	High-level controller C^1 states at sampling instant t (defined in the cascade control structure)
x_t^{c2}	Low-level controller C^2 states at sampling instant t (defined in the cascade control structure)
Y_f	Future data Hankel matrix for y_t
Y_p	Past data Hankel matrix for y_t
y_t	System output(s) at sampling instant t
$\mathcal{F}$	Matrix defined as $[I + L_u L_y^c]^{-1}$
Δ_N^c	reversed extended controllability matrices of $\{A_c, B_c\}$ $= (A_c^{N-1}B_c \;\; A_c^{N-2}B_c \;\; \cdots \;\; A_cB_c \;\; B_c)$
Δ_N^d	reversed extended controllability matrices of $\{A, B\}$ $= (A^{N-1}B \;\; A^{N-2}B \;\; \cdots \;\; AB \;\; B)$
Δ_N^s	reversed extended controllability matrices of $\{A, K\}$ $= (A^{N-1}K \;\; A^{N-2}K \;\; \cdots \;\; AK \;\; K)$

η_y,η_u	Performance index
Γ_N	Extended observability matrix of the process up to order $N = \begin{pmatrix} C^T & (CA)^T & \cdots & (CA^{N-1})^T \end{pmatrix}^T$
Γ_N^{CL}	Extended observability matrix of the closed-loop system
Γ_N^c	Extended observability matrix of the controller
Γ_N^{c1}	Extended observability matrix of controller C^1 in the cascade control structure
Γ_N^{c2}	Extended observability matrix of controller C^2 in the cascade control structure
Γ_N^{cs}	Extended observability matrix defined for the cascade control structure as $\begin{pmatrix} \Gamma_N & L_u L_y^{c2} \Gamma_N^{c1} & L_u \Gamma_N^{c2} \end{pmatrix}$
λ	Weighting on the control effort
γ_i	Markov parameter used for calculating the cumulative effect of the noise e_t on the process output variance
ξ_t	System measured output(s) at sampling instant t
ψ_i	Markov parameter used for calculating the cumulative effect of the noise e_t on the process input variance
Υ	Pseudo-inverse of the subspace matrix containing past and future data $= \begin{pmatrix} W_p \\ U_f \end{pmatrix}$
†	Superscript; Moore-Penrose pseudo-inverse
^{T}	Superscript; Transpose transformation

List of Abbreviations

APC	Advanced Process Control
ARMAX	Auto Regressive Moving Average with eXogenous Input
ARX	Auto Regressive with eXogenous Input
CVA	Canonical Variate Analysis
FOPTD	First Order Plus Time Delay
IMC	Internal Model Control
LQG	Linear Quadratic Gaussian
MA	Moving Average
MIMO	Multi Input Multi Output
MOESP	MIMO Output Error State Space
MPC	Model Predictive Control
MVC	Minimum Variance Control
N4SID	Numerical Subspace State Space IDentification
OPC	OLE for Process Control
PID	Proportional Integral Derivative
RBS	Random Binary Sequence
SISO	Single Input Single output
SOPTD	Second Order Plus Time Delay

Chapter 1

Introduction

1.1 Objective of this thesis

In the past decades process control practice has progressed from conventional methods to more advanced methods in the most areas of control applications including control system design, process identification, fault detection and diagnosis and process integration. Research in the field of control theory has also been extended in many directions to provide new and advanced control algorithms for handling more complex processes with interactions and constraints, new open-loop and closed-loop identification required for process model development, control relevant identification methods, many model-based and signal-based fault detection and diagnosis approaches and a variety of controller performance assessment methods either model-based or data-driven.

In most of the control systems, many controllers initially perform well, but abrupt or gradual performance deterioration will occur as time goes by. It was reported by Ender (1993) [31] that as many as 60% of all industrial controllers have some kind of performance problem. This problem has received the attention of many researchers both from industry and academia. There has been considerable interest in developing methods for performance assessment of the control systems in the last 2 decades. Comprehensive review papers by Qin (1998, 2007) [80, 85] and Harris *et. al.* (1999) [38] provide a detailed review of the research on control loop performance assessment.

Model-based performance assessment methods are on the basis of designing an optimal controller (mostly linear) for the process which is used as a benchmark for evaluation of the current controller performance. Such an optimal control design needs some information about the process dynamics. The first and most used benchmark is the minimum variance control (MVC) [4] benchmark presented by Harris (1989) [36]. LQG control has also been used as a benchmark for performance assessment which provides a more comprehensive benchmark by the cost of

requiring more information about the process. The benchmark in this method is the 'trade-off' curve [14] which presents the limit of optimal performance for linear control. Introduction to these methods are provided in the next chapter. Data-driven approaches do not require the process knowledge, but they can not provide the absolute limit of the optimal performance.

The required process knowledge for the purpose of controller performance assessment rises up the need for process modeling. Different benchmarks require different levels of the process knowledge for which many estimation and identification approaches have been proposed [45, 85]. The requirement of working with closed-loop data in real applications, motivated the researchers to focus mostly on the closed-loop identification methods for this purpose.

Classic open-loop and closed-loop identification theory has been well developed and presented in a great detail in the celebrated textbooks by Ljung (1999) [69] and Söderström and Stoica (1989) [87]. The most common classic identification methods are the prediction error method [69] and instrument variable method [87]. Subspace identification methods provide an alternative approach for system identification relying on numerically robust tools such as QR-factorization and singular value decomposition. Various methods of subspace identification have been developed in the past two decades [57, 64, 91, 96]. Subspace identification methods have also been applied in other areas of control applications such as performance assessment [43], predictive control [43, 29] and fault-tolerant control design [30].

The focus of this thesis is on the subspace identification and its application for process modeling and controller performance assessment from closed-loop data. The LQG trade-off curve is used as the benchmark for controller performance assessment in this work. Using the subspace framework removes the requirement of an explicit model for trade-off curve estimation. More details are provided in the next sections.

1.2 Outline of the thesis

In Chapter 2, an overview on the subspace identification and controller performance assessment is provided. The subspace identification approach proposed in the thesis is based on Regression Analysis approach [57]. This method is reviewed in this chapter and a numerical example is provided. The concept of LQG benchmark for performance assessment is reviewed in this chapter as well.

In Chapter 3, calculation of the LQG benchmark from closed-loop data is studied. Closed-loop subspace identification can be used to obtain the trade-off curve without requiring an explicit model of the process or disturbance. Also, the noise covariance plays an important roll in the calculation of the LQG trade-off curve and should be estimated via closed-loop identification. A previously proposed method

of closed-loop identification by Kadali and Huang (2002) [54] is revisited and then an alternative direct formulation of the joint input-output identification method is provided for which the consistency of noise variance estimation can be obtained. Monte-Carlo simulations are provided to investigate the consistency of the proposed method and some other relevant methods of closed-loop subspace identification. A subspace method of estimating the noise model and the noise covariance from closed-loop routine operating data (with no external excitations) is also presented in this chapter which is useful for performance assessment of model-based controllers in which a process model is normally available. A simulation example on a multivariate system is provided. Both methods are implemented in an experimental study on a pilot-scale 2×2 Continuous Stirred Tank Heater (CSTH) process.

Chapter 4 presents an extension of the subspace-based LQG benchmark to the case of a more general control structure consisting of supervisory advanced control and low-level regulatory control. This supervisory-regulatory control structure which comes from cascade implementation of advanced controllers [66] is very common in advanced process control applications. In this chapter, a review of the available options for implementation of advanced controllers is provided based on [66], and then performance assessment of the cascade control structure is studied using the LQG benchmark. For this purpose, three different scenarios are considered and LQG controller design is provided for each case. Following the subspace framework of Chapter 3, a control design is provided is free from the explicit process model requirement. Then, it is shown that LQG trade-off curves can be obtained for each case using certain subspace matrices. The curves can be used for performance assessment of supervisory and regulatory controllers and helps the engineer to decide which controller(s) to be re-tuned/re-designed for performance enhancement. The required subspace matrices as well as the noise covariance matrix can be estimated using a proposed closed-loop subspace identification approach for the cascade structure based on the idea of joint input-output identification. MATLAB and HYSYS simulations are used to demonstrate utility of the proposed method.

In Chapter 5, direct estimation of the process step response model from closed-loop data using subspace identification is studied. Necessary information concerning impulse response coefficients is embedded in subspace matrices. Therefore, the step response coefficients can be directly obtained from this matrix by integration without the need of extracting state space models as the conventional subspace identification does. Since the estimated subspace matrix contains more than one set of impulse response coefficients, a question raises about how to best synthesize them to obtain an optimal estimate of the impulse response coefficients. For this purpose, a reformulation of the subspace identification problem is required for which the variance of all elements in the related subspace matrix can be evaluated. The

calculated variances are then used to perform a weighted averaging on the impulse response coefficients in order to attenuate the noise influence on the step response estimation. Two Monte-Carlo simulations are provided followed by two pilot-scale experiments to verify the proposed method.

1.3 Contributions of the thesis

The contributions of this thesis are summarized as follows:

- An alternative formulation of the joint input-output subspace identification is proposed and used for performance assessment based on the LQG benchmark. Consistency analysis for the proposed closed-loop identification method is provided.
- The subspace LQG control is designed in a supervisory-regulatory control structure for three possible cases and the LQG trade-off curve for each case is developed using certain subspace matrices without the requirement of explicit models.
- A closed-loop subspace identification method based on the joint input-output identification approach is provided to estimate the required subspace matrices under the cascade control.
- Direct estimation of the process step response model from closed-loop data using a sequential closed-loop subspace identification method is provided.
- Two Graphic User Interfaces (GUI) are generated for closed-loop subspace identification and controller performance assessment based on the LQG benchmark.
- Experimental evaluation of the proposed methods on two pilot-scale processes.

Chapter 2

A Review of Subspace Identification and Performance Assessment

In this chapter brief introductions to the two major subjects of the thesis, subspace system identification and controller performance assessment, are provided. Our focus is on the approaches that are used in the following chapters which are Regression Analysis approach for subspace identification and LQG benchmark for performance assessment. However, to provide better understanding of advantages of the LQG benchmark, a brief review on the minimum variance control (MVC) benchmark is given as well. Since no simple numerical example on the procedure of subspace identification has been provided in the textbooks of system identification, we will describe regression analysis method procedure by use of a numerical example in this chapter. Also, an insightful study on the mechanism of the open-loop subspace identification is performed for the noise-free case. The reviews are mostly taken from [43, 57, 45].

2.1 A brief review on open-loop subspace identification

Subspace identification methods provide an alternative data-driven approach to classical system identification methods like the prediction error method [69] and instrument variable method [87]. Subspace methods use efficient computational algorithms such as QR-factorization and singular value decomposition, which makes them intrinsically robust from a numerical point of view. Various methods of subspace identification have been developed in the past two decades such as regression analysis approach, N4SID (Numerical SubSpace State Space IDentification), MOESP (MIMO Output Error State sPace) and CVA (Canonical Variate Analysis) [57, 64, 77, 91, 92, 90, 96]. Subspace identification methods are intrinsically suitable

for multivariate systems identification compared to the classic methods. This chapter gives an overview of the regression analysis method which is used in the next chapters. Although the principal goal of the method is to identify the state space system matrices {A, B, C, D, K}, certain subspace matrices are first calculated as an intermediate step. In this work, these intermediate matrices will be used for performance assessment and step model estimation,so the explicit model of the process is not of our interest. This issue will be elaborated more in the next section.

In the following, we provide a brief introduction to the subspace notation and the regression analysis approach. A simple numerical example is also provided as well as a study on the mechanism of the calculation in this method for noise-free case.

2.1.1 Subspace notation and preliminary

Consider the following state space representation for a linear system with l inputs and m outputs

$$\begin{cases} x_{t+1} &= Ax_t + Bu_t + Ke_t \\ y_t &= Cx_t + Du_t + e_t \end{cases} \tag{2.1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^l$, $y_t \in \mathbb{R}^m$ and $e_t \in \mathbb{R}^m$ is white noise. From Equation (2.1), for $t = N$

$$x_{N+1} = Ax_N + Bu_N + Ke_N \tag{2.2}$$

and for $t = N + 1$

$$\begin{aligned} x_{N+2} &= Ax_{N+1} + Bu_{N+1} + Ke_{N+1} \\ &= A^2 x_N + (AB \quad A) \begin{pmatrix} u_N \\ u_{N+1} \end{pmatrix} + (AK \quad B) \begin{pmatrix} e_N \\ e_{N+1} \end{pmatrix} \end{aligned}$$

This procedure can be continued to $t = 2N - 2$ which gives

$$\begin{aligned} x_{2N-1} = A^{N-1}x_N + (A^{N-2}B \quad A^{N-3}B \quad \cdots B) \begin{pmatrix} u_N \\ u_{N+1} \\ \cdots \\ u_{2N-2} \end{pmatrix} \\ + (A^{N-2}B \quad A^{N-3}B \quad \cdots B) \begin{pmatrix} e_N \\ e_{N+1} \\ \cdots \\ e_{2N-2} \end{pmatrix} \end{aligned} \tag{2.3}$$

Following the same procedure for the output equation in (3.1), similar results to (2.3) can be derived

$$y_{2N-1} = CA^{N_1}x_N + (CA^{N-2}B \quad CA^{N-3}B \quad \cdots D)\begin{pmatrix} u_N \\ u_{N+1} \\ \cdots \\ u_{2N-2} \end{pmatrix} + (CA^{N-2}K \quad CA^{N-3}K \quad \cdots I)\begin{pmatrix} e_N \\ e_{N+1} \\ \cdots \\ e_{2N-2} \end{pmatrix} \tag{2.4}$$

Assembling the results for $t = N, N+1, ..., 2N-1$, yields a matrix equation as follows:

$$\begin{pmatrix} y_N \\ y_{N+1} \\ y_{N+2} \\ \cdots \\ y_{2N-1} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \cdots \\ CA^{N-1} \end{pmatrix} x_N + \begin{pmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}B & CA^{N-3}B & CA^{N-4}B & \cdots & D \end{pmatrix}\begin{pmatrix} u_N \\ u_{N+1} \\ u_{N+2} \\ \cdots \\ u_{2N-1} \end{pmatrix} + \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ CK & I & 0 & \cdots & 0 \\ CAK & CK & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}K & CA^{N-3}K & CA^{N-4}K & \cdots & I \end{pmatrix}\begin{pmatrix} e_N \\ e_{N+1} \\ e_{N+2} \\ \cdots \\ e_{2N-1} \end{pmatrix} \tag{2.5}$$

Adding time index by 1, Equation (2.5) changes to

$$\begin{pmatrix} y_{N+1} \\ y_{N+2} \\ y_{N+3} \\ \cdots \\ y_{2N} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \cdots \\ CA^{N-1} \end{pmatrix} x_N + \begin{pmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}B & CA^{N-3}B & CA^{N-4}B & \cdots & D \end{pmatrix}\begin{pmatrix} u_{N+1} \\ u_{N+2} \\ u_{N+3} \\ \cdots \\ u_{2N} \end{pmatrix} + \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ CK & I & 0 & \cdots & 0 \\ CAK & CK & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}K & CA^{N-3}K & CA^{N-4}K & \cdots & I \end{pmatrix}\begin{pmatrix} e_{N+1} \\ e_{N+2} \\ e_{N+3} \\ \cdots \\ e_{2N} \end{pmatrix} \tag{2.6}$$

Continuing this procedure for time indexes 2, 3 until $j-1$ and assembling obtained equations gives

$$\begin{pmatrix} y_N & y_{N+1} & \cdots & y_{N+j-1} \\ y_{N+1} & y_{N+2} & \cdots & y_{N+j} \\ y_{N+2} & y_{N+3} & \cdots & y_{N+j+1} \\ \cdots & \cdots & \cdots & \cdots \\ y_{2N-1} & y_{2N} & \cdots & y_{2N+j-2} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \cdots \\ CA^{N-1} \end{pmatrix} \begin{pmatrix} x_N & x_{N+1} & \cdots & x_{N+j-1} \end{pmatrix} +$$
$$+ \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}B & CA^{N-3}B & \cdots & D \end{pmatrix} \begin{pmatrix} u_N & u_{N+1} & \cdots & u_{N+j-1} \\ u_{N+1} & u_{N+2} & \cdots & u_{N+j} \\ u_{N+2} & u_{N+3} & \cdots & u_{N+j+1} \\ \cdots & \cdots & \cdots & \cdots \\ u_{2N-1} & u_{2N} & \cdots & u_{2N+j-2} \end{pmatrix}$$
$$+ \begin{pmatrix} I & 0 & \cdots & 0 \\ CK & I & \cdots & 0 \\ CAK & CK & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}K & CA^{N-3}K & \cdots & I \end{pmatrix} \begin{pmatrix} e_N & e_{N+1} & \cdots & e_{N+j-1} \\ e_{N+1} & e_{N+2} & \cdots & e_{N+j} \\ e_{N+2} & e_{N+3} & \cdots & e_{N+j+1} \\ \cdots & \cdots & \cdots & \cdots \\ e_{2N-1} & e_{2N} & \cdots & e_{2N+j-2} \end{pmatrix} \tag{2.7}$$

Equation (2.7) provides the subspace equation for the outputs. Performing a similar procedure on the state equation gives the following subspace matrix equation for system states:

$$\begin{aligned} (x_N \quad x_{N+1} \quad \cdots \quad x_{N+j-1}) &= A^N \begin{pmatrix} x_0 & x_1 & \cdots & x_{j-1} \end{pmatrix} \\ &+ \begin{pmatrix} A^{N-1}B & A^{N-2}B & \cdots & B \end{pmatrix} \begin{pmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \cdots & \cdots & \cdots & \cdots \\ u_{N-1} & u_N & \cdots & u_{N+j-2} \end{pmatrix} \\ &+ \begin{pmatrix} A^{N-1}K & A^{N-2}K & \cdots & K \end{pmatrix} \begin{pmatrix} e_0 & e_1 & \cdots & e_{j-1} \\ e_1 & e_2 & \cdots & e_j \\ \cdots & \cdots & \cdots & \cdots \\ e_{N-1} & e_N & \cdots & e_{N+j-2} \end{pmatrix} \end{aligned} \tag{2.8}$$

Short-hand version of the above derivations are the following subspace equations:

$$Y_f = \Gamma_N X_f + L_u U_f + L_e E_f \tag{2.9}$$
$$Y_p = \Gamma_N X_p + L_u U_p + L_e E_p \tag{2.10}$$
$$X_f = A^N X_p + \Delta^d U_p + \Delta^s E_p \tag{2.11}$$

In the above equations, subscript 'p' stands for 'past' and 'f' stands for 'future'. Comparing Equation (2.9) with (2.7) gives the definition of subspace matrices Γ_N,

L_u and L_e as

$$\Gamma_N = \begin{pmatrix} C \\ CA \\ CA^2 \\ \cdots \\ CA^{N-1} \end{pmatrix} \tag{2.12}$$

$$L_u = \begin{pmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}B & CA^{N-3}B & CA^{N-4}B & \cdots & D \end{pmatrix} \tag{2.13}$$

$$L_e = \begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ CK & I & 0 & \cdots & 0 \\ CAK & CK & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}K & CA^{N-3}K & CA^{N-4}K & \cdots & I \end{pmatrix} \tag{2.14}$$

$\Gamma_N \in \mathbb{R}^{mN \times n}$ is the extended observability matrix and $L_u \in \mathbb{R}^{mN \times lN}$ and $L_e \in \mathbb{R}^{mN \times mN}$ contain the process and disturbance Markov parameters, respectively. In some subspace literatures H_N^d is used instead of L_u and H_N^s instead of L_e. The definition of data Hanlke matrices U_p, $U_f \in \mathbb{R}^{lN \times j}$, Y_p, $Y_f \in \mathbb{R}^{mN \times j}$ and E_p, $E_f \in \mathbb{R}^{mN \times j}$ can also be retrieved by comparing Equation (2.9) with (2.7) and Equation(2.11) with (2.8). As an example, U_p and U_f are defined as follows:

$$U_p = \begin{pmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \cdots & \cdots & \cdots & \cdots \\ u_{N-1} & u_N & \cdots & u_{N+j-2} \end{pmatrix}$$

$$U_f = \begin{pmatrix} u_N & u_{N+1} & \cdots & u_{N+j-1} \\ u_{N+1} & u_{N+2} & \cdots & u_{N+j} \\ \cdots & \cdots & \cdots & \cdots \\ u_{2N-1} & u_{2N} & \cdots & u_{2N+j-2} \end{pmatrix}$$

Typically, j should be much larger than $max(mN, lN)$ to reduce sensitivity to noise [90]. In fact, j plays the same role as the number of observations in regression analysis and N is related to the order of system. There are some common assumptions in all subspace identification methods; pair (A, C) is required to be observable and pair $(A, BKR^{1/2})$ should be controllable where R is the covariance matrix of white noise entering the process.

2.1.2 Regression analysis approach

The following derivation is taken from [57]. Consider the state space model (3.1) again. By shifting the time subscripts, we get

$$x_t = Ax_{t-1} + Bu_{t-1} + Ke_{t-1}$$
$$y_{t-1} = Cx_{t-1} + Du_{t-1} + e_{t-1}$$

Combining the two above equations yields

$$x_t = (A - KC)x_{t-1} + Ky_{t-1} + (B - KC)u_{t-1}$$

Doing the same combination recursively results in

$$\begin{aligned} x_t = (A - KC)^N x_{t-N} + (A - KC)^{N-1}(Ky_{t-N} + (B - KD)u_{t-N}) + \\ + \cdots + (A - KC)(Ky_{t-2} + (B - KD)u_{t-2}) + \\ + Ky_{t-1} + (B - KD)u_{t-1} \end{aligned} \tag{2.15}$$

The time index in Equation (2.15) can be replaced by $t = N, N+1, ..., N+j-1$. Collecting all of the resulting equations in a matrix format gives a subspace equation as follows:

$$X_f = \Phi_y Y_p + \Phi_u U_p + \Phi_x X_p \tag{2.16}$$

where

$$\Phi_x = (A - KC)^N$$
$$\Phi_y = \mathcal{C}(A - KC, K)$$
$$\Phi_u = \mathcal{C}(A - KC, B - KD)$$

and $\mathcal{C}$ operator is defined by

$$\mathcal{C}(A, B) \triangleq (A^{N-1}B \;\; A^{N-2}B \;\; \cdots \;\; AB \;\; B) \tag{2.17}$$

Φ_x represents error dynamics of the Kalman filter, so as $N \to \infty$, it results in $\Phi_x \to 0$, because of the stability of Kalman filter. As a result, for large N

$$X_f = \Phi_y Y_p + \Phi_u U_p \tag{2.18}$$

Substituting Equation (2.18) in (2.9) gives

$$Y_f = L_w W_p + L_u U_f + L_e E_f \tag{2.19}$$

where W_p is defined as:

$$W_p = \begin{pmatrix} Y_p \\ U_p \end{pmatrix} \tag{2.20}$$

and L_w can be represented by

$$L_w = \Gamma_N(\Phi_y \;\; \Phi_y) \tag{2.21}$$

Based on Equation (2.19), subspace matrices can be estimated by least squares or other regression methods

$$\begin{aligned}\left(\begin{array}{cc} \hat{L}_w & \hat{L}_u \end{array} \right) &= Y_f \left(\begin{array}{c} W_p \\ U_f \end{array} \right)^{\dagger} \\ &= Y_f \left(\begin{array}{cc} W_p^T & U_f^T \end{array} \right) \left(\left(\begin{array}{c} W_p \\ U_f \end{array} \right) \left(\begin{array}{cc} W_p^T & U_f^T \end{array} \right) \right)^{-1} \end{aligned} \tag{2.22}$$

This pseudo-inverse operation can be done in a more numerically robust way using QR-factorization [91, 90]. The residual of this least squares estimation is given by

$$V_f = Y_f - \hat{Y}_f \tag{2.23}$$

$$= L_e E_f \tag{2.24}$$

and its covariance can be estimated as

$$\hat{P}_v = \frac{1}{j} V_f V_f^T \tag{2.25}$$

Furthermore, the innovation e_f can be estimated from the first row of V_f as follows:

$$\hat{e}_f = V_f(1:m,:) \tag{2.26}$$

To retrieve the system matrices, the following SVD needs to be performed:

$$W_1 \hat{L}_w W_2 = (U_1 \;\; U_2) \left(\begin{array}{cc} S_1 & 0 \\ 0 & S_2 \end{array} \right) \left(\begin{array}{c} V_1 \\ V_2 \end{array} \right) \tag{2.27}$$

where W_1 and W_2 are proper weighting matrices. In Knudsen (2001) [57], the choices for them are given as

$$W_1 = I \tag{2.28}$$

$$W_2 = (W_p W_p^T)^{1/2} \tag{2.29}$$

Theoretically, for a finite dimension of state space models, there exists $S_2 = 0$ and the dimension of S_1 determines the dimension of the state space matrix A. From this decomposition, Γ_N can be obtained by

$$\hat{\Gamma}_N = U_1 S_1^{1/2} \tag{2.30}$$

System matrices now can be retrieved as follows [57]:

$$\hat{\Gamma}_N^1 = \hat{\Gamma}_N(1:(N-1)m,:) \tag{2.31}$$
$$\hat{\Gamma}_N^2 = \hat{\Gamma}_N(m+1:Nm,:) \tag{2.32}$$
$$\hat{C} = \hat{\Gamma}_N(1:m,1:n) \tag{2.33}$$
$$\hat{A} = (\hat{\Gamma}_N^1)^\dagger \hat{\Gamma}_N^2 \tag{2.34}$$
$$\hat{B} = (\Gamma_N^1)^\dagger \hat{L}_u(m+1:Nm,1:l) \tag{2.35}$$
$$\hat{D} = \hat{L}_u(1:m,1:l) \tag{2.36}$$
$$\hat{K} = (\hat{\Gamma}_N^1)^\dagger \hat{P}_v(m+1:Nm,1:m)\hat{R}^{-1} \tag{2.37}$$

There are some other approaches in subspace identification such as methods based on projection (e.g., N4SID) or statistical approaches (e.g., CVA) and MOESP, which are not used in this thesis. Details of these methods can be found in subspace identification literature (see [43], [90], [91], [96] and references therein).

Many methods have also been developed for closed-loop system identification based on subspace approach. Modified N4SID, ARX prediction approach, innovation estimation method, orthogonal projection approach and joint input-output identification are some of them and details can be found in the literature [42, 56, 64, 70, 84, 88, 94, 95, 98]. A comprehensive review is provided in [43]. We will review a method based on the joint input-output approach in the next chapter, and the consequent application for performance assessment will be studied.

In the following subsection, a simple numerical example of implementing the regression analysis method is given for better understanding of the method.

2.1.3 A numerical example

Consider the following state space representation:

$$x(t+1) = -0.1x(t) + u(t) + 0.2e(t) \tag{2.38}$$
$$y(t) = 0.5x(t) + e(t) \tag{2.39}$$

The noise variance is set to be 0.01. A small set of input-output data is generated for this process as provided in Table 2.1.

Table 2.1: Identification data for numerical example

u(t)	1	-1	1	-1	-1	1
y(t)	0.0021	0.5002	-0.5479	0.5562	-0.5536	-0.4435

Choosing $N = 2$ and $j = 3$, the data Hankel matrices can be built as follows:

$$U_p = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \tag{2.40}$$

$$U_f = \begin{pmatrix} 1 & -1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \tag{2.41}$$

$$Y_p = \begin{pmatrix} 0.0021 & 0.5002 & -0.5479 \\ 0.5002 & -0.5479 & 0.5562 \end{pmatrix} \tag{2.42}$$

$$Y_f = \begin{pmatrix} -0.5479 & 0.5562 & -0.5536 \\ 0.5562 & -0.5536 & -0.4435 \end{pmatrix} \tag{2.43}$$

Then, W_p is given by

$$W_p = \begin{pmatrix} 0.0021 & 0.5002 & -0.5479 \\ 0.5002 & -0.5479 & 0.5562 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \tag{2.44}$$

Using Equation (2.22), one can estimate L_w and L_u as follows:

$$\begin{pmatrix} \hat{L}_w & \hat{L}_u \end{pmatrix} = \begin{pmatrix} -0.5479 & 0.5562 & -0.5536 \\ 0.5562 & -0.5536 & -0.4435 \end{pmatrix} \begin{pmatrix} 0.0021 & 0.5002 & -0.5479 \\ 0.5002 & -0.5479 & 0.5562 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{pmatrix}^{\dagger}$$

$$= \begin{pmatrix} -0.5479 & 0.5562 & -0.5536 \\ 0.5562 & -0.5536 & -0.4435 \end{pmatrix} \begin{pmatrix} 0.0021 & 0.5002 & 1 & -1 & 1 & -1 \\ 0.5002 & -0.5479 & -1 & 1 & -1 & -1 \\ -0.5479 & 0.5562 & 1 & -1 & -1 & 1 \end{pmatrix} \times$$

$$\left(\begin{pmatrix} 0.0021 & 0.5002 & -0.5479 \\ 0.5002 & -0.5479 & 0.5562 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0.0021 & 0.5002 & 1 & -1 & 1 & -1 \\ 0.5002 & -0.5479 & -1 & 1 & -1 & -1 \\ -0.5479 & 0.5562 & 1 & -1 & -1 & 1 \end{pmatrix} \right)^{-1}$$

$$= \begin{pmatrix} 0.0604 & -0.1235 & -0.2345 & 0.2345 & -0.0032 & 0.0140 \\ 0.0049 & 0.0116 & 0.0258 & -0.0258 & 0.4985 & -0.0003 \end{pmatrix} \tag{2.45}$$

which gives

$$\hat{L}_w = \begin{pmatrix} 0.0604 & -0.1235 & -0.2345 & 0.2345 \\ 0.0049 & 0.0116 & 0.0258 & -0.0258 \end{pmatrix}$$

$$\hat{L}_u = \begin{pmatrix} -0.0032 & 0.0140 \\ 0.4985 & -0.0003 \end{pmatrix} \tag{2.46}$$

Equation (2.23) can now be used to estimate V_f as

$$V_f = \begin{pmatrix} 0.0142 & -0.0172 & 0.0126 \\ -0.1524 & 0.0186 & 0.2112 \end{pmatrix} \tag{2.47}$$

which gives

$$\begin{aligned} \hat{P}_v &= \begin{pmatrix} 0.0002 & 0.00006 \\ 0.00006 & 0.0226 \end{pmatrix} \\ e_f &= V_f(1,:) = \begin{pmatrix} 0.0142 & -0.0172 & 0.0126 \end{pmatrix} \\ \hat{R} &= cov(e_f) = 0.0003 \end{aligned} \tag{2.48}$$

Performing SVD on L_w yields

$$\hat{L}_w = \begin{pmatrix} 0.9948 & -0.1022 \\ -0.1022 & 0.9948 \end{pmatrix} \begin{pmatrix} 0.3609 & 0 \\ 0 & 0.0114 \end{pmatrix} \begin{pmatrix} -0.1651 & 0.9727 \\ 0.3437 & -0.0982 \end{pmatrix}$$

Using Equation (2.30), Γ_N can be estimated by

$$\hat{\Gamma}_N = \begin{pmatrix} 0.9948 \\ -0.1022 \end{pmatrix} (0.3609)^{1/2} = \begin{pmatrix} 0.5976 \\ -0.0614 \end{pmatrix} \tag{2.49}$$

Now, the system parameters {a,b,c,d,k} can be estimated using equations (2.33) to (2.37) as follows:

$$\hat{c} = \hat{\Gamma}_N(1,1) = 0.5976 \tag{2.50}$$

$$\hat{a} = (\Gamma_N^u)^\dagger \Gamma_N^y = (\hat{c})^{-1} \hat{\Gamma}_N(2,1) = (0.5976)^{-1}(-0.0614) = -0.1228 \tag{2.51}$$

$$\hat{b} = (\Gamma_N^u)^\dagger \hat{L}_u(2,1) = (\hat{c})^{-1} \hat{L}_u(2,1) = (0.5976)^{-1}(0.4995) = 0.9969 \tag{2.52}$$

$$\hat{d} = \hat{L}_u(1,1) = -0.0032 \tag{2.53}$$

$$\hat{k} = (\hat{\Gamma}_N^u)^\dagger \hat{P}(2,1)(\hat{R})^{-1} = (0.5976)^{-1}(0.00006)(0.0003)^{-1} = 0.3347 \tag{2.54}$$

$$\hat{R} = 0.0003 \tag{2.55}$$

Equations (2.55) and (2.54) show poor estimations of the noise dynamics and noise variance which can be as a result of using small number of input-output data points.

2.1.4 An insight to the noise-free identification

In this section, the mechanism of regression analysis method in the noise-free case is studied using simulation to provide more insight to the subspace identification approach. This simple yet insightful study is worthy because normally it is nontrivial to understand the calculation procedure of the subspace identification since it involves projection and inversion of some large matrices. Our result shows that in the absence of the noise, Markov parameters in the subspace matrix L_u are naturally

estimated by differentiation and averaging procedure. We have not found this issue being studied in the subspace literature.

Recall Equation (2.22) from Section 2.1.2. We define

$$\Upsilon \triangleq \begin{pmatrix} W_p \\ U_f \end{pmatrix}^{\dagger} \tag{2.56}$$

$$U_f^{\ddagger} \triangleq \Upsilon(:, 2lN+1 : 3lN) \tag{2.57}$$

Equation (2.22) shows that $\hat{L}_u$ is calculated by

$$\hat{L}_u = Y_f \times U_f^{\ddagger} \tag{2.58}$$

To avoid dealing with large matrices, we shall discuss the SISO case only. Our simulations show that $U_f^{\ddagger}$ has a special structure which helps to understand the mechanism of calculation of the Markov parameters in L_u. Assume that the excitation signal, u_t, is a 'well-designed' RBS signal as shown in Figure 2.1. We choose $N = 3$ and $j = 10$. Then, $U_f^{\ddagger} \in \mathbb{R}^{10\times 3}$ has the following structure:

$$U_f^{\ddagger} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \alpha \\ 0 & \alpha & -\alpha \\ \alpha & -\alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & -\alpha \\ 0 & -\alpha & \alpha \\ -\alpha & \alpha & 0 \\ \alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.59}$$

Therefore, $\hat{L}_u$ is calculated by

$$\begin{aligned} \hat{L}_u &= Y_f \times U_f^{\ddagger} \\ &= \begin{pmatrix} y_3 & y_4 & y_5 & y_6 & y_7 & y_8 & y_9 & y_{10} & y_{11} & y_{12} \\ y_4 & y_5 & y_6 & y_7 & y_8 & y_9 & y_{10} & y_{11} & y_{12} & y_{13} \\ y_5 & y_6 & y_7 & y_8 & y_9 & y_{10} & y_{11} & y_{12} & y_{13} & y_{14} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \alpha \\ 0 & \alpha & -\alpha \\ \alpha & -\alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & -\alpha \\ 0 & -\alpha & \alpha \\ -\alpha & \alpha & 0 \\ \alpha & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} \alpha[(y_7 - y_6) + (y_{10} - y_{11})] & \hat{L}_u(1,2) & \hat{L}_u(1,3) \\ \alpha[(y_8 - y_7) + (y_{11} - y_{12})] & \alpha[(y_7 - y_6) + (y_{10} - y_{11})] & \hat{L}_u(2,3) \\ \alpha[(y_9 - y_8) + (y_{12} - y_{13})] & \alpha[(y_8 - y_7) + (y_{11} - y_{12})] & \alpha[(y_7 - y_6) + (y_{10} - y_{11})] \end{pmatrix} \end{aligned} \tag{2.60}$$

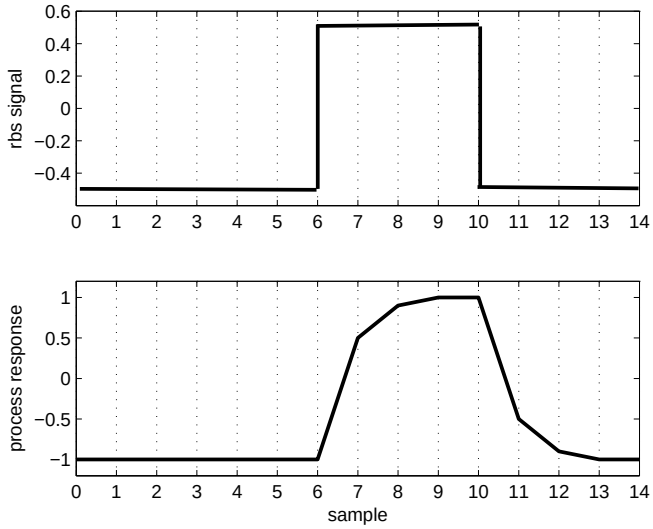


Figure 2.1: RBS test signal and the process response

It shows that the lower triangular part of L_u should have a Toeplitz structure. Now, three natural questions are: (1) Where do the non-zero elements appear in $U_f^\ddagger$? (2) What is the value of α and does the lower triangular part of $\hat{L_u}$ contain Markov parameters? (3) Is the upper triangular part of $\hat{L}_u$ zero?

1. Position of the non-zero elements in $U_f^\ddagger$ is determined by the time of switchings in the RBS signal. If at sample $t = i$ a positive step excitation occurs in u_t, the $(i - N + 1)th$ element in the first column of $U_f^\ddagger$ is α and the next element, $(i - N + 2)th$, is $-\alpha$. Next columns follow the structure of (2.59). If $-\alpha$ appears first, it is as a result of a negative step switching in the test signal. Looking at the test signal shown in Figure 2.1 explains the structure of $U_f^\ddagger$ in (2.59). It shows that a positive step change occurs at $t = 6$ and a negative step at $t = 10$. The first one results in non-zero elements at $U_f^\ddagger(4, 1)$ and $U_f^\ddagger(5, 1)$ and the second one makes $U_f^\ddagger(8, 1)$ and $U_f^\ddagger(9, 1)$ non-zero in the first column (Note that any step changes before time $t = N$ does not affect $U_f^\ddagger$).

2. Equation (2.60) shows that, for this example, two terms have contribution in each of the impulse response coefficients. For instance, $(y_7 - y_6)$ and $-(y_{11} - y_{10})$ for the first Markov parameter. Figure 2.1 clarifies that each of these two terms is indeed the first impulse response coefficient obtained by differentiating the first and second step response coefficients. The multiplier α appears to take the average of the two estimated impulse response coefficients and therefore, its value is given by

$$\alpha = \frac{1}{\#switchings} \tag{2.61}$$

where the RBS signal has magnitude of 1. If not, α will be

$$\alpha = \frac{1}{(\#switchings)(magnitude\,of\,RBS)} \tag{2.62}$$

3. The upper part of $\hat{L}_u$ contains $\hat{L}_u(1,2)$, $\hat{L}_u(1,3)$ and $\hat{L}_u(2,3)$ which are given by the following relations:

$$\hat{L}_u(1,2) = (y_6 - y_5) + (y_9 - y_{10}) = 0 + 0 \tag{2.63}$$

$$\hat{L}_u(1,3) = (y_5 - y_4) + (y_8 - y_9) = 0 + 0 \tag{2.64}$$

$$\hat{L}_u(2,3) = (y_6 - y_5) + (y_9 - y_{10}) = 0 + 0 \tag{2.65}$$

which show that the upper part elements are indeed zero.

Remark. It should be noted that this result is only valid when the identification test signal is step-type and noise-free. Furthermore, the Nyquist frequency of the test signal should be designed in a way that allows the process to show its low-frequency dynamics. In the absence of these conditions, matrix $U_f^{\ddagger}$ no longer has the mentioned structure.

Remark. Even when all the above conditions are satisfied, no clear conclusion can be proposed for the mechanism of the calculation of $\hat{L}_w$.

In the next section of this chapter, a brief review on some important concepts in controller performance assessment is presented. The focus will be on the methods based on MVC and LQG benchmarks.

2.2 A brief review on performance assessment

The design of automated and effective strategies for control performance assessment and monitoring has become a necessity in many industries [13, 31, 39, 61, 76]. Controller performance assessment has been one of the most active areas of research in the field of process control during the past two decades. Considerable academic as well as commercial interests have been devoted to the monitoring of both univariate and multivariate control systems [47, 55, 59, 75, 80, 85, 89, 101]. The first important step was taken by Harris (1989) [36] who proposed the minimum variance control benchmark for performance assessment. This benchmark has been widely used as a reference bound on achievable performance since his work. Using this MVC performance bound as a performance benchmark provides the absolute lower bound on the output variations and it can be evaluated without complete knowledge of the process model. Therefore, methodologies for the assessment of a MVC benchmark

have been reported in a variety of control applications including single-loop feedback control (Harris, (1989) [36]; Lynch and Dumont, (1996) [73]), feedforward/feedback control (Desborough and Harris, (1993) [28]), cascade control (Ko and Edgar, (2000) [59]), and multivariable feedback control (Harris *et. al.* (1996) [37]; Huang *et. al.* (1997,1999) [47, 45]). In the next subsection, we review this approach briefly.

2.2.1 Performance assessment using MVC benchmark

For a system with time delay d, a part of output variance is independent of feedback and is achievable by minimum variance control. This part which can be estimated from routine operating data, is called feedback control invariant part [36]. To separate this feedback invariant term, closed-loop output y should be represented as a moving average model such as

$$y_t = \underbrace{f_0 e_t + f_1 e_{t-1} \ldots + f_{d-1} e_{t-(d-1)}}_{y_{mv}} + f_d e_{t-d} + f_{d+1} e_{t-(d+1)} + \ldots \tag{2.66}$$

where y_{mv} is the feedback invariant portion of y_t and e_t is white noise. y_{mv} is a measure of theoretically achievable lower bound of output variance.

Note that this variance may or may not be achievable in practice, however as a benchmark, MVC provides some useful information about how well current controller works in comparison to minimum variance controller. The following is a short review on SISO feedback control performance assessment using MVC benchmark.

Consider a SISO system under feedback control where G_p and G_l are the process and disturbance transfer functions, respectively. One can represent the process transfer function as $G_p = z^{-d}\tilde{G}_p$, where d is the time-delay and $\tilde{G}_p$ is the delay-free model. Defining G_l as the disturbance transfer function and G_c as the controller transfer function, y_t can be represented by [45]

$$y_t = \frac{G_l}{1 + G_p G_c}\ e_t = \frac{G_l}{1 + z^{-d}\tilde{G}_p G_c}\ e_t \tag{2.67}$$

Using Diaphontine identity, G_l can be written as

$$G_l = \underbrace{f_0 + f_1 z^{-1} + \cdots + f_{d-1} z^{-d+1}}_{F_1} + F_2 z^{-d} \tag{2.68}$$

where f_i's are impulse response coefficients of G_l and F_2 is the remaining rational

proper transfer function. Now, Equation (2.67) can be written as

$$\begin{aligned} y_t &= \frac{F_1 + z^{-d}F_2}{1 + z^{-d}\tilde{G}_p G_c}\, e_t \\ &= [F_1 + \frac{F_2 - F_1\tilde{G}_p G_c}{1 + z^{-d}\tilde{G}_p G_c} z^{-d}]\, e_t \\ &= F_1 e_t + L e_{t-d} \end{aligned} \tag{2.69}$$

where L is appropriately defined. Since $F_1 e_t$ in independent of the white noise occurring before time $t - d + 1$, one can write

$$Var(y_t) = Var(F_1 e_t) + Var(L r_{t-d}) \tag{2.70}$$

Therefore

$$Var(y_t) > Var(F_1 e_t) \tag{2.71}$$

and the equality holds when $L = 0$ which gives

$$F_2 - F_1\tilde{G}_p G_c = 0$$

and results in the following MVC law:

$$G_c = \frac{F_2}{\tilde{G}_p F_1} \tag{2.72}$$

Therefore, if a stable process output is modeled by an infinite moving average (MA) model, the first d terms is an estimate of the MVC term Fe_t. It should be noted that the above results does not hold for the case on non-minimum phase systems because it results in unstable pole-zero cancelation. The modified control derivation is provided in [45].

For implementing the MVC benchmark in performance assessment applications, one needs to have some *a priori* knowledge about the process time-delay. This can be problematic in the case of MIMO systems. In this case, a matrix of delays, named '*interactor matrix*', should be known. To circumvent this problem, some methods for estimating the interactor matrix or even avoiding the use of it have been proposed. Huang *et al.* (1997) [46] developed a filtering and correlation (FCOR) analysis algorithm. Harris *et al.* (1996) [37] developed another method based on spectral factorization. The algorithm by Ko and Edgar (2001) [59] avoided the direct use of interactor matrix. In more recent works, McNabb and Qin (2003) [76] proposed an alternative approach for MVC benchmark based on state space model. In a data-driven framework, Kadali and Huang (2003) [55] developed an algorithm using subspace approach.

2.2.2 LQG benchmark for performance assessment

The MVC benchmark may not be directly applicable for assessing the performance of those control systems whose objective is not just minimizing process output variance but also keeping the input variability within some specified range. This may be considered for reducing the upset to the process, conserve energy and and lessen the equipment damage. The LQG benchmark is a more appropriate benchmark for this type of control systems. However, obtaining this benchmark in a traditional way needs a complete model of the process which is a demanding requirement. To reduce the dependency of the method to the process model, some efforts have been done in using subspace approaches to find the LQG benchmark from closed-loop data. This idea has been explored in [53, 43]. The final results of this method are reviewed at the end of this section. First, we provide a review on the concept of performance assessment using LQG benchmark.

Performance assessment using the LQG benchmark is to determine how far the output variance is from the best achievable variance for a given input variance. In other words, one should solve the following problem:

$$Given\ E[u_t^2] < \alpha,\ \ what\ is\ \ min\ E[y_t^2]?$$

Solution of this problem is given by a curve named 'trade-off' curve [14] which can be obtained by solving the LQG problem

$$J(\lambda) = E[y_t^2] + \lambda E[u_t^2] \tag{2.73}$$

by varying values of λ. Detailed discussion on the equivalence of the above statement and Equation (2.73 is provided in [14] based on the concept of Pareto optimization. In this way, various solutions for $E[u_t^2]$ and $E[y_t^2]$ can be calculated and a curve with the optimal $E[y_t^2]$ as the abscissa and $E[u_t^2]$ as the ordinate is formed by minimizing Equation (2.73). This curve provides the limit of performance for linear controllers in terms of variance [14]. A typical LQG curve is shown in Figure 2.2.

Having the curve and current input and output variances, one can determine certain performance indices to compare the performance of the current controller to the optimal LQG controller [43]. As an example, assume that the actual input and output variances are represented by V_u and V_y, respectively. The optimal output variance corresponding to V_u is shown by V_y^o and the optimal input variance corresponding to V_y is V_u^o (see Figure 2.2). Then the following performance indices for output and input variances can be defined:

$$\eta_y = \frac{(V_y^o)}{(V_y)} \qquad , \qquad \eta_u = \frac{(V_u^o)}{(V_u)}$$

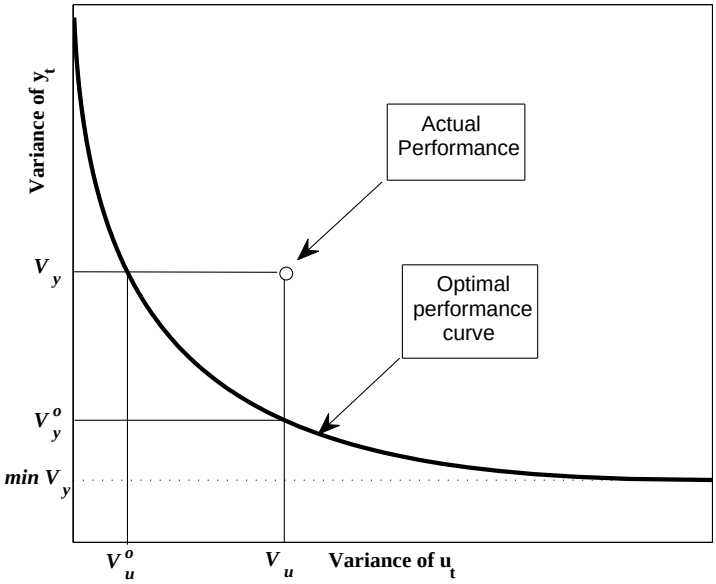


Figure 2.2: A sample LQG trade-off curve

The achievable MVC benchmark can also be found from this curve, as shown in the figure. Based on Equation (2.73), this is the case when $\lambda \to 0$.

A state-space or input-output model of the process can be used to obtain the LQG curve [45]. For instance, consider the following input-output model in ARMAX form:

$$Ay_t = Bu_t + Ce_t$$

where e_t is white noise with unity variance. If a regulatory LQG control law is given by $u_t = -\frac{E}{F}y_t$, then

$$y_t = \frac{CF}{AF + BE}\ e_t = G_y e_t \tag{2.74}$$

and

$$u_t = \frac{CE}{AF + BE}\ e_t = G_u e_t \tag{2.75}$$

where E and F are functions of λ. The stability of the above control design is shown in [62]. Using Parseval's Theorem, the output and input variances can be expressed as

$$Var(y_t) = \| G_y \|_2^2$$
$$Var(u_t) = \| G_u \|_2^2$$

Therefore, $Var(y_t)$ and $Var(u_t)$ are the H_2 norms of the closed-loop transfer functions from noise to the output and input, respectively. The LQG solution cane also be found approximately using MPC solvers [45].

Huang and Kadali (2002) [53] presented a method to approximate solution of the LQG problem using subspace matrices L_u and L_e. This approach does not require the explicit model of the process. The final results of the method are summarized below:

Define

$$\begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{N-1} \end{pmatrix} = -(L_u^T L_u + \lambda I)^{-1} L_u^T L_{e,1}$$

and

$$\begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{N-1} \end{pmatrix} = [I - L_u(L_u^T L_u + \lambda I)^{-1} L_u^T] L_{e,1}$$

where $L_{e,1}$ is the first block column of L_e. It has been shown that the optimal control action can be obtained by

$$u_t^{opt} = \sum_{i=0}^{N-1} \psi_i e_{t-i}$$

which results in an output as follows:

$$y_t^{opt} = \sum_{i=0}^{N-1} \gamma_i e_{t-i}$$

Finally, input and output variances under LQG control can be obtained using the following relations:

$$Var[u_t] = \sum_{i=0}^{N-1} \psi_i \ Var[e_t] \ \psi_i^T, \tag{2.76}$$

$$Var[y_t] = \sum_{i=0}^{N-1} \gamma_i \ Var[e_t] \ \gamma_i^T \tag{2.77}$$

It is shown that above LQG design is equivalent to the classic LQG control design for the case of $N \to \infty$ [32], so the stability results [14] can be also extended for the subspace design. It should also be noted that the classic LQG design requires a form of parametric model, but the above subspace expression relies on non-parametric expressions of the process and disturbance. As s result, using the subspace framework reduces the bias error of the parametric modeling, but it has higher variance.

Calculation of the LQG benchmark using this method requires the knowledge of L_u, L_e and the noise variance. For control performance monitoring, it is desired to identify the required subspace matrices and the noise covariance from closed-loop data which brings the requirement of a proper closed-loop subspace identification method. This issue is discussed in details in the next chapter.

2.3 Concluding Remarks

In this chapter, an introduction to the subspace identification and controller performance assessment was provided. Focusing on the open-loop subspace identification, we reviewed the method of regression analysis approach which will be used in the following chapters for the purpose of closed-loop identification. A simple numerical example was also provided to show the procedure of subspace identification. A closer look into the mechanism of the open-loop subspace identification for the noise-free case was also presented. To the best our knowledge, this analysis has not been provided in the subspace literature. This study showed that in the absence of noise, the Markov parameters of the process (in L_u) are naturally calculated by averaging over a series of estimated impulse response coefficients coming from step response differentiation.

A review of controller performance assessment based on the MVC benchmark and LQG benchmark was provided to show the advantages and drawbacks of each method. Since the focus of the next chapters is on the LQG benchmark, a more detailed review on this approach was presented.

Chapter 3

Closed-loop Subspace Identification for Performance Assessment Using the LQG benchmark

3.1 Introduction

High performance control systems require healthy controllers. However, survey shows that sixty percent of industrial controllers have some performance problems [86]. Even if the controllers work properly at the time of initial commissioning, many of them encounter performance deterioration after some time of service. It is therefore necessary to design automated and effective strategies for control performance assessment and monitoring.

Minimum variance control (MVC) benchmark originally proposed by Harris (1989) [36] provides the most fundamental step for modern algorithms to measure the performance. For multi-input multi-output (MIMO) systems, the extended method needs the knowledge of an interactor matrix [27, 37, 38, 46]. Although the minimum variance control provides valuable information about a lower bound on the process variance, it is often not a practically implemented controller, because of its aggressive control actions. Furthermore, the minimum variance control benchmark only focuses on the output without considering limitations on the inputs.

The LQG benchmark approach considers variances of both input and output and provides a 'trade-off' curve which represents a feasible range of performance for linear controllers [14, 40, 45]. In other words, this curve provides the lowest achievable variance of the controlled variable corresponding to different values of the variance of the manipulated variable. Performance of the existing controllers can be evaluated by comparing current input and output variances against the lowest achievable bounds represented by the trade-off curve. This method is based on a

model of the process which has to be obtained through process identification.

Subspace identification methods provide an alternative data-driven approach to classical system identification methods such as the prediction error method [69] and instrument variable method [87]. Subspace methods offer several advantages over classical transfer function based methods. For example the parametrization issue related to MIMO system is avoided, and they use efficient computational algorithms such as QR-factorization and singular value decomposition, which makes them intrinsically robust from a numerical point of view. Various methods of subspace identification have been developed over the past two decades such as regression analysis approach, N4SID (Numerical SubSpace State Space IDentification), MOESP (MIMO Output Error State sPace) and CVA (Canonical Variate Analysis) [57, 64, 77, 91, 92, 90, 96]. Most of these methods have been extended to closed-loop identification (see [42, 56, 64, 70, 84, 88, 94, 95, 98]). Consistency analysis for different open-loop and closed-loop subspace identification methods has been proposed in the subspace literature. Several papers have presented studies on the consistency analysis of open-loop methods [5, 8, 9, 11, 51, 57, 71, 74]. Consistency analysis results for the closed-loop problem are also presented by many researchers [18, 20, 22, 42, 68, 98].

Kadali and Huang [53] presented a method for calculating the LQG trade-off curve in a subspace framework without using an explicit model. A relevant closed-loop subspace identification method is also developed in [54]. The subspace approach to estimating the LQG benchmark needs certain intermediate matrices in subspace identification and the covariance matrix of the noise. Consistency of the estimation of the noise covariance matrix has remained an outstanding issue in the previous work. We study this issue in this chapter.

In a typical model-based control application, the plant model is usually known through process identification, but the disturbance model is typically assumed to take a certain fixed form such as an integrated white noise. Thus the actual disturbance model is often not available. Even if a disturbance model is identified, it is unlikely that this model could be useful later owing to the more likely time-varying nature of the disturbance dynamics. Therefore, disturbance model should ideally be identified from routine operating data for the sake of evaluating current control performance. Here, we develop a procedure to estimate the noise model subspace matrix without explicitly identifying parametric models. This issue also arises when the consistency requirements, as discussed in this chapter, may dictate the need for a large set of experiment data for consistent estimation of the noise covariance.

The consistency of the LQG benchmark hinges on consistency of estimation of the noise covariance matrix. Through several Monte Carlo simulations, we found that several closed-loop subspace identification methods that are suitable for the

estimation of the LQG benchmark do not provide consistent estimation of the noise covariance as they may have intended. Motivated by this finding, we first develop an alternative closed-loop subspace identification method to achieve consistency of estimation of the noise covariance, and then go on to utilize these results for computing the LQG trade-off curve for SISO and MIMO examples.

This chapter is organized as follows: In the next section, a review of the existing method of joint input-output subspace identification is provided and a lack of consistency is pointed out. Section 3 provides a modified joint input-output formulation and the consistency analysis for the noise covariance followed by a simulation example. In section 4, we demonstrate how the noise subspace matrix and noise covariance can be estimated from the routine operating data. Section 5 provides implementation results in a pilot-scale experiment. Concluding remarks are provided in the Section 6.

3.2 Subspace approach to the LQG benchmark estimation

To implement LQG benchmark for performance assessment in subspace framework, certain subspace matrices are required. In this section we provide a brief review of the existing joint input-output identification method presented in [54].

3.2.1 Joint input-output closed-loop identification

In the method presented in Section 2.2.2, calculation of the LQG benchmark requires the knowledge of L_u and L_e. For control performance monitoring, it is desired to identify the required subspace matrices and the noise covariance from closed-loop data. This subsection provides a review of the identification method proposed by Kadali and Huang [54] to estimate these matrices from closed-loop data.

Consider the following state space representation for a linear system with l inputs and m outputs:

$$\begin{cases} x_{t+1} &= Ax_t + Bu_t + Ke_t \\ y_t &= Cx_t + Du_t + e_t \end{cases} \tag{3.1}$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^l$, $y_t \in \mathbb{R}^m$ and $e_t \in \mathbb{R}^m$ is white noise. It was shown in Chapter 2 that system (3.1) can be represented by the following subspace equation:

$$Y_f = L_w W_p + L_u U_f + L_e E_f \tag{3.2}$$

where the definitions of Y_f, U_f, E_f and the subspace matrices, L_w, L_u and L_e, are given in Chapter 2.

A linear controller used in a regular feedback loop can be represented by

$$\begin{cases} x_{t+1}^c &= A_c x_t^c + B_c(r_t - y_t) \\ u_t &= C_c x_t^c + D_c(r_t - y_t) \end{cases} \tag{3.3}$$

where $x_t \in \mathbb{R}^{n_c}$, $u_t \in \mathbb{R}^l$, $y_t \in \mathbb{R}^m$ and $r_t \in \mathbb{R}^m$ is the setpoint. Similar to system (3.1), subspace equations for the controller can be obtained as follows:

$$X_f^c = A_c^N X_p^c + \Delta_N^c(R_p - Y_p)$$
$$U_f = \Gamma_N^c X_f^c + L_y^c(R_f - Y_f)$$

where $\Gamma_N^c \in \mathbb{R}^{lN \times n_c}$ is the controller extended observability matrix, $L_y^c \in \mathbb{R}^{lN \times mN}$ contains the Markov parameters of the controller and $\Delta_N^c \in \mathbb{R}^{n_c \times mN}$ is the reversed extended controllability matrix of $\{A_c, B_c\}$. U_p and $U_f \in \mathbb{R}^{lN \times j}$ and Y_p, Y_f, R_p, $R_f \in \mathbb{R}^{mN \times j}$ are data Hankel matrices for past and future inputs, outputs and setpoints, respectively.

Applying regression analysis [57] on these equations, one can get the following input-output relation by eliminating the states:

$$U_f = L_w^c W_p^c + L_y^c R_f - L_y^c Y_f \tag{3.4}$$

where $L_w^c \in R^{lN \times (l+2m)N}$ and W_p^c is defined as

$$W_p^c = \begin{pmatrix} Y_p \\ U_p \\ R_p \end{pmatrix} \tag{3.5}$$

Substituting Equation (3.4) in Equation (3.2) with some rearrangements yields

$$Y_f = L_y^{CL} W_p^{CL} + L_{yr}^{CL} R_f + L_{ye}^{CL} E_f \tag{3.6}$$

where

$$\begin{aligned} L_y^{CL} W_p^{CL} &= (I + L_u L_y^c)^{-1}(L_w W_p + L_u L_w^c W_p^c) \\ L_{yr}^{CL} &= (I + L_u L_y^c)^{-1} L_u L_y^c \\ L_{ye}^{CL} &= (I + L_u L_y^c)^{-1} L_e \end{aligned}$$

and $W_p^{CL} = W_p^c$. Similarly, substituting Equation (3.2) in Equation (3.4) yields

$$U_f = L_u^{CL} W_p^{CL} + L_{ur}^{CL} R_f + L_{ue}^{CL} E_f \tag{3.7}$$

where

$$\begin{aligned} L_u^{CL} W_p^{CL} &= (I + L_y^c L_u)^{-1}(L_w^c W_p^c + L_y^c L_w W_p) \\ L_{ur}^{CL} &= (I + L_y^c L_u)^{-1} L_y^c \\ L_{ue}^{CL} &= -(I + L_y^c L_u)^{-1} L_y^c L_e \end{aligned}$$

With the above results, a joint input-output identification can be formulated as follows:

$$\begin{pmatrix} Y_f \\ U_f \end{pmatrix} = \begin{pmatrix} L_y^{CL} \\ L_u^{CL} \end{pmatrix} W_p^{CL} + \begin{pmatrix} L_{yr}^{CL} \\ L_{ur}^{CL} \end{pmatrix} R_f + \begin{pmatrix} L_{ye}^{CL} \\ L_{ue}^{CL} \end{pmatrix} E_f \tag{3.8}$$

Since R_f is the Hankel matrix of the external perturbations uncorrelated with W_p^{CL} and E_f, closed-loop subspace matrices L_y^{CL}, L_u^{CL}, L_{yr}^{CL} and L_{ur}^{CL} can be estimated using least square estimation as

$$(\hat{L}_u^{CL} \quad \hat{L}_{ur}) = U_f \begin{pmatrix} W_p^{CL} \\ R_f \end{pmatrix}^{\dagger} \tag{3.9}$$

$$(\hat{L}_y^{CL} \quad \hat{L}_{yr}) = Y_f \begin{pmatrix} W_p^{CL} \\ R_f \end{pmatrix}^{\dagger} \tag{3.10}$$

$\hat{U}_f$ and $\hat{Y}_f$ are found by the orthogonal projection of the row space of U_f and Y_f. The first row of $\hat{Y}_f$ is the one step ahead estimation of the output. So the innovation sequence can be estimated by

$$\begin{aligned} \hat{e}_f &= (e_N \quad e_{N+1} \quad \cdots \quad e_{N+j-1})^T \\ &= Y_f(1:m,:) - \hat{Y}_f(1:m,:) \end{aligned} \tag{3.11}$$

where $(1:m,:)$ represents the first m rows and all columns of the matrix (following MATLAB notation). Using e_f, the block Hankel matrix for noise, E_f, can be built. Define

$$\Xi_f \triangleq U_f - \hat{U}_f = L_{ue}^{CL} E_f$$

and, for L_{ue}^{CL} we have

$$\hat{L}_{ue}^{CL} = \Xi_f / E_f = \Xi_f E_f^{\dagger} \tag{3.12}$$

Using matrix inversion lemma, it has been shown in [54] that $\hat{L}_u$ can be obtained from closed-loop matrices by the following equation:

$$\hat{L}_u = \hat{L}_{yr}^{CL} (\hat{L}_{ur}^{CL})^{-1} \tag{3.13}$$

Calculation of $\hat{L}_e$ is straightforward from the definition of L_{ur}^{CL} and L_{ue}^{CL}:

$$\hat{L}_e = -(\hat{L}_{ur}^{CL})^{-1} \hat{L}_{ue}^{CL} \tag{3.14}$$

Estimated subspace matrices $\hat{L}_u$ and $\hat{L}_e$ and the covariance matrix of the estimated noise sequence are used for input and output variance calculation under LQG control which provides the trade-off curve [53]. However, our simulation studies show a potential lack of consistency in the noise covariance estimation using this method which may result in a biased LQG trade-off curve.

There are several other methods of closed-loop subspace identification that may be used to estimate the LQG benchmark. However, through Monte-Carlo simulations, we found these methods do not provide consistent estimation of the covariance matrix too, which motives this work. Before introducing the formulation of our method, a comparison to the other methods of closed-loop subspace identification is necessary.

One of the earliest methods is the joint input-output identification by Verhaegen [95] which can not be used for estimation of the subspace matrices. Ljung and McKelvey (1996) [70] proposed another method of closed-loop subspace identification which requires a preliminary ARX modeling step. A similar requirement exists in some other methods such as Shi (2001) [82], Jansson (2003) [50] and Larimore (2004) [65] for removing the effect of undesired terms due to the feedback. The method driven by Van Overschee and De Moor [94] needs Markov parameters of the controller and it does not provide direct estimation of the subspace matrices L_u and L_e. The closed-loop method of Chou and Verhaegen (1997) [24] is extremely sensitive to noise [21] and the presented algorithm can not be used for estimating the noise model. Based on the idea of [50], Chiuso and Picci [21] presented another method by replacing the first step with an oblique projection step and provided the consistency analysis of their method in [20, 22]. However, this method as well as the method of Huang et. al (2005) [42] based on orthogonal projection do not provide direct estimation of the noise subspace matrix , L_e. It is shown in [42] that the method of Wang and Qin (2002) [97] based on principal component analysis may deliver a bias for closed-loop data. The 'innovation estimation' method by Qin and Ljung (2003) [83, 68] is based on performing N least squares estimations to identify a set of casual models [81] which provides consistent estimation of L_u and L_e directly. This method is shown to be sensitive to unstable open-loop systems [20]. Also, our simulations show that this method does not provide consistent estimation of the noise variance estimation. Wang and Qin (2006) proposed another method similar to [42], using parity space. This methods requires a first step of principal component analysis and SVD to estimate L_u followed by a least squares and QR factorization step to estimate L_e and the noise covariance matrix. However, the consistency of noise variance is not discussed in the paper.

The most pertinent closed-loop subspace identification methods that provide direct estimates of the subspace matrices and the covariance matrix are the joint input-output approach [54] and the innovation estimation approach [83, 68]. In the next section, we will conduct a Monte-Carlo simulation to evaluate the consistency of the covariance estimation of these methods. For more comprehensive study, we also simulate classical N4SID closed-loop identification algorithm [93, 94] and the CVA method for which the codes where available.

A comparative Monte-Carlo simulation

In this example, we show the results of Monte-Carlo simulation using four different methods of subspace closed-loop identification. As mentioned before, our focus is on the inconsistency of noise variance estimation. These methods include: the joint input-output identification by Kadali and Huang (2002) [54], the 'innovation estimation' method by Qin and Ljung (2003) [83], the closed-loop method by Van Overschee and De Moor [93, 94] and CVA method from MATLAB® toolbox.

The following system is taken from [90] with some minor modifications:

$$x_{t+1} = \begin{pmatrix} 0.6 & 0.6 & 0 \\ -0.6 & 0.6 & 0 \\ 0 & 0 & 0.7 \end{pmatrix} x_t + \begin{pmatrix} 1.616 \\ -0.348 \\ 2.631 \end{pmatrix} u_t - \begin{pmatrix} 1.147 \\ 1.520 \\ 3.199 \end{pmatrix} e_t$$
$$y_t = \begin{pmatrix} -0.437 & -0.504 & 0.093 \end{pmatrix} x_t - 0.775 u_t + e_t$$

A PI controller, $[0.1 + 0.05/s]$, is used to control the process. Variance of the input

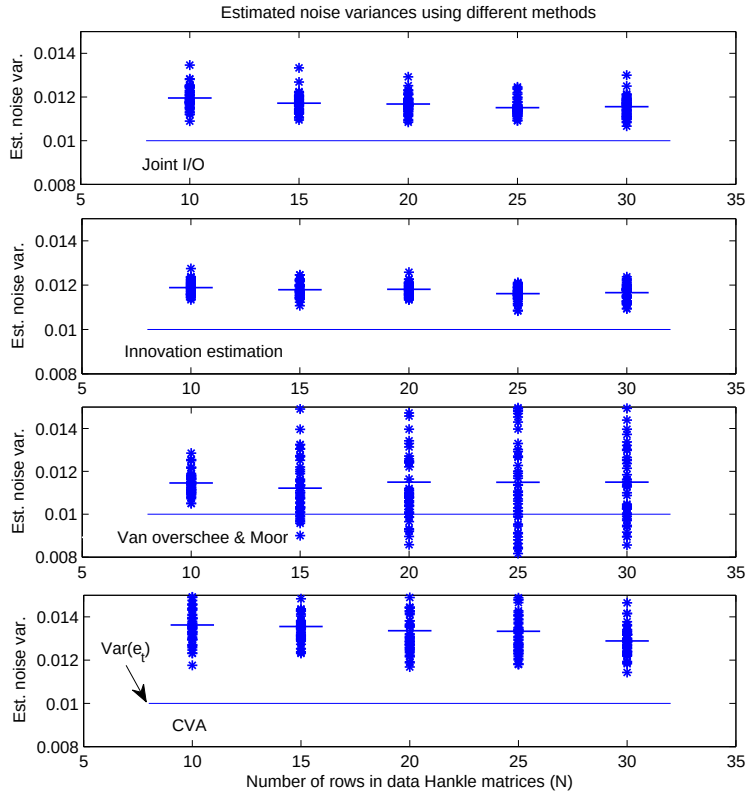


Figure 3.1: Results of Monte-Carlo simulation using four different closed-loop subspace identification methods

noise is 0.01. The test signal is designed by MATLAB® command '*idinput*' with Nyquist frequency of 0.12 and magnitude of 0.5. The variance-based signal to noise ratio is approximately 10.

Each simulation run generates 3000 data points. We generate 50 data sets, each time with the same reference input r_t but with a different noise sequence e_t. Since the number of rows in the data Hankle matrices, N, is an important factor in the consistency of subspace methods, we run the Monte-Carlo simulations on each method for 5 different values of N. The result of the noise variance estimations is shown in Figure 3.1. The average value of the estimated noise variances for each value of N and the true value of the noise variance are also indicated in the figure.

As the figure shows, none of these methods provide consistent estimation of the noise variance.

In the following section, we present a modified version of the method of [54] which divides the closed-loop identification problem into two open-loop identification problems for which the noise covariance consistency is guaranteed.

3.3 A direct formulation of joint input-output closed-loop identification

In this section, we present an alternative formulation of the joint input-output closed-loop identification followed by the consistency analysis and a simulation study. In [57] it has been established that consistency can be achieved for estimation of the subspace matrices and noise variance using regression analysis approach for open-loop system identification. We will use this result to address the consistency of noise covariance estimation in our formulation.

3.3.1 A direct formulation of joint input-output closed-loop subspace identification

Recall the basic subspace relation for the process in Equation (2.9), and similarly for the controller:

$$Y_f = \Gamma_N X_f + L_u U_f + L_e E_f \tag{3.15}$$

$$U_f = \Gamma_N^c X_f^c + L_y^c (R_f - Y_f) \tag{3.16}$$

Substituting Equation (3.16) in (3.15) gives

$$\begin{aligned} Y_f &= \Gamma_N X_f + L_u U_f + L_e E_f \\ &= \Gamma_N X_f + L_u[\Gamma_N^c X_f^c + L_y^c R_f - L_y^c Y_f] + L_e E_f \\ &= \Gamma_N X_f + L_u \Gamma_N^c X_f^c + L_u L_y^c R_f - L_u L_y^c Y_f + L_e E_f \\ &= (I + L_u L_y^c)^{-1}[\Gamma_N X_f + L_u \Gamma_N^c X_f^c + L_u L_y^c R_f + L_e E_f] \end{aligned} \tag{3.17}$$

For simplicity in the notation, define $\mathcal{F} \triangleq (I + L_u L_y^c)^{-1}$. Substituting Y_f obtained from (3.17) in Equation (3.16) yields

$$\begin{aligned} U_f =& \Gamma_N^c X_f^c + L_y^c R_f - L_y^c \mathcal{F} \Gamma_N X_f - L_y^c \mathcal{F} L_u \Gamma_N^c X_f^c - L_y^c \mathcal{F} L_u L_y^c R_f - L_y^c \mathcal{F} L_e E_f \\ =& - L_y^c \mathcal{F} \Gamma_N X_f + (\Gamma_N^c - L_y^c \mathcal{F} L_u \Gamma_N^c) X_f^c + (L_y^c - L_y^c \mathcal{F} L_u L_y^c) R_f - L_y^c \mathcal{F} L_e E_f \end{aligned} \tag{3.18}$$

Equations (3.17) and (3.18) can be represented as follows:

$$\begin{aligned} Y_f =& (\mathcal{F}\Gamma_N \quad \mathcal{F} L_u \Gamma_N^c) X_f^{CL} + (\mathcal{F} L_u L_y^c) R_f + \mathcal{F} L_e E_f \\ U_f =& (-L_y^c \mathcal{F} \Gamma_N \quad \Gamma_N^c - L_y^c \mathcal{F} L_u \Gamma_N^c) X_f^{CL} + \\ & (L_y^c - L_y^c \mathcal{F} L_u L_y^c) R_f - L_y^c \mathcal{F} L_e E_f \end{aligned}$$

or in a compact form

$$Y_f = \Gamma_N^Y X_f^{CL} + L_{YR} R_f + L_{YE} E_f \tag{3.19}$$

$$U_f = \Gamma_N^U X_f^{CL} + L_{UR} R_f + L_{UE} E_f \tag{3.20}$$

where

$$X_f^{CL} = [(X_f)^T \quad (X_f^c)^T]^T \tag{3.21}$$

$$\Gamma_N^Y = (\mathcal{F}\Gamma_N \quad \mathcal{F} L_u \Gamma_N^c)$$

$$\Gamma_N^U = (-L_y^c \mathcal{F} \Gamma_N \quad \Gamma_N^c - L_y^c \mathcal{F} L_u \Gamma_N^c)$$

$$\begin{aligned} L_{YR} &= \mathcal{F} L_u L_y^c \\ &= (I + L_u L_y^c)^{-1} L_u L_y^c \end{aligned} \tag{3.22}$$

$$\begin{aligned} L_{UR} &= L_y^c - L_y^c \mathcal{F} L_u L_y^c = [I - L_y^c (I + L_u L_y^c)^{-1} L_u] L_y^c \\ &= (I + L_y^c L_u)^{-1} L_y^c \end{aligned} \tag{3.23}$$

$$\begin{aligned} L_{YE} &= \mathcal{F} L_e \\ &= (I + L_u L_y^c)^{-1} L_e \end{aligned} \tag{3.24}$$

$$\begin{aligned} L_{UE} &= -L_y^c \mathcal{F} L_e = -L_y^c (I + L_u L_y^c)^{-1} L_e \\ &= -(I + L_y^c L_u)^{-1} L_y^c L_e \end{aligned} \tag{3.25}$$

Regression analysis [57] can now be performed on equations (3.19) and (3.20) after replacing future states with past inputs and outputs to obtain equations (3.26) and (3.27) as follows

$$Y_f = L_Y W_p^{yr} + L_{YR} R_f + L_{YE} E_f \tag{3.26}$$

$$U_f = L_U W_p^{ur} + L_{UR} R_f + L_{UE} E_f \tag{3.27}$$

where W_p^{yr} and W_p^{ur} are defined by

$$W_p^{yr} = \begin{pmatrix} Y_p \\ R_p \end{pmatrix} \quad \textit{and} \quad W_p^{ur} = \begin{pmatrix} U_p \\ R_p \end{pmatrix} \tag{3.28}$$

and $L_Y \in \mathbb{R}^{mN\times 2mN}$ and $L_U \in \mathbb{R}^{lN\times(m+l)N}$ have the same role as L_w for the open-loop system (3.1). Choosing R_f to be uncorrelated with E_f, W_p^{ur} and W_p^{yr}, each of these open-loop identification problems can be solved by least squares estimation.

Note that these two estimation problems are indeed open-loop identification problems. Similar to other subspace identification methods, some assumptions are required [57]: *i*) pair $\{A_{cl}, C_{cl}\}$ is observable, *ii*) pair $\{A_{cl}, [B_{cl}\ K_{cl}]\}$ is controllable (definitions are given in (3.39)), *iii*) the transfer function from e_t to y_t has all its zeros strictly inside unit circle and *iv*) r_t and e_t are jointly quasi-stationary and uncorrelated.

$\hat{U}_f$ and $\hat{Y}_f$ are found by the orthogonal projection of the row space of U_f and Y_f.

For estimating open-loop subspace matrices L_u and L_e, three of the the subspace matrices in equations (3.26) and (3.27) are required: L_{YR}, L_{UR} and L_{UE}. The first two are already available from least square estimation. The last one is estimated as follows [43]:

$$\begin{aligned}\Xi_f &\triangleq U_f - \hat{U}_f\\ \hat{L}_{UE} &= \Xi_f E_f^{\dagger}\end{aligned}$$

The required noise vector in the above equation (for constructing E_f) can be estimated based on Equation (4.75). But, a closer look of Equation (3.24) shows that term $Y_f(1:m,:) - \hat{Y}_f(1:m,:)$ in Equation (4.75) gives an estimation of $\mathcal{F}(1:m,:)e_f$ not e_f:

$$Y_f(1:m,:) - \hat{Y}_f(1:m,:) = \mathcal{F}(1:m,1:m)e_f$$

and e_f should be estimated as

$$\hat{e}_f = \mathcal{F}^{-1}(1:m,1:m)[Y_f(1:m,:) - \hat{Y}_f(1:m,:)] \tag{3.29}$$

It means that we need a correction term in the future noise estimation, comparing to Equation (4.75), before constructing E_f. This task can be done using closed-loop subspace matrices that we have already identified. Using the definition of L_{YR} in Equation. (3.22) we have

$$\begin{aligned}&L_{YR} = (I + L_u L_y^c)^{-1} L_u L_y^c\\ &\Rightarrow (I + L_u L_y^c) L_{YR} = L_u L_y^c\\ &\Rightarrow L_{YR} = L_u L_y^c (I - L_{YR})\end{aligned}$$

which gives the estimation of $\widehat{L_u L_y^c}$ by

$$\widehat{L_u L_y^c} = \hat{L}_{YR}(I - \hat{L}_{YR})^{-1} \tag{3.30}$$

This results in the following estimation for $\mathcal{F}^{-1}$:

$$\hat{\mathcal{F}}^{-1} = I + \hat{L}_{YR}(I - \hat{L}_{YR})^{-1}$$

Since the estimation of $\hat{L}_{YR}$ is consistent [57], $\hat{\mathcal{F}}^{-1}$ estimation is also consistent.

In summary, in this alternative method, two least squares estimations should be performed

$$(\hat{L}_U \quad \hat{L}_{UR}) = U_f \begin{pmatrix} W_p^{ur} \\ R_f \end{pmatrix}^{\dagger} \tag{3.31}$$

$$(\hat{L}_Y \quad \hat{L}_{YR}) = Y_f \begin{pmatrix} W_p^{yr} \\ R_f \end{pmatrix}^{\dagger} \tag{3.32}$$

Note that these two estimation problems are indeed open-loop identification problems. Then, innovation estimation is performed as follows:

$$\hat{e}_f = [I + \hat{L}_{YR}(I - \hat{L}_{YR})^{-1}]_{(1:\,m,\,:)}[Y_f(1:m,:) - \hat{Y}_f(1:m,1:m)] \tag{3.33}$$

which results in *consistent estimation* of noise variance to be shown shortly.

Based on the definitions of L_u and L_y^c, it is clear that if either the process or the controller has at least one sample delay (D or D_c is zero), the correction factor will be I.

After constructing E_f by the use of (3.33), L_{UE} can be estimated as:

$$\hat{L}_{UE} = (U_f - \hat{U}_f)E_f^{\dagger} \tag{3.34}$$

Finally, L_u and L_e, needed for calculating the LQG benchmark, are estimated [54]:

$$\hat{L}_u = \hat{L}_{YR}(\hat{L}_{UR})^{-1} \tag{3.35}$$

$$\hat{L}_e = -(\hat{L}_{UR})^{-1}\hat{L}_{UE} \tag{3.36}$$

Although the previous method [54] is also joint input and output closed-loop identification, the main difference between the new formulation and the method of [54] is in defining two different data matrices in (5.4), where in [54], $W_p^{yr} = W_p^{ur} = [Y_p^T \;\; U_p^T \;\; R_p^T]^T$. The direct extension from the conventional joint input-output closed-loop identification to that of subspace, as proposed in this chapter, has truly decomposed closed-loop identification problem into two open-loop identification problems, for which the consistency of noise covariance estimation, required for the calculation of the LQG trade-off curve, can be proven (to be shown shortly).

For better understanding of the proposed method, let's revisit classic joint input-output identification briefly.

Assume that process and disturbance models and controller are defined by $G_p(z^{-1})$, $G_l(z^{-1})$ and $G_c(z^{-1})$, respectively. In a closed-loop system the following relations can be obtained [69]:

$$\begin{aligned} y_t &= \frac{G_c(z^{-1})G_p(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1})}r_t + \frac{G_l(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1})}e_t \\ &\triangleq M(z^{-1})r_t + N(z^{-1})e_t \end{aligned}$$

$$\begin{aligned} u_t &= \frac{G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1})}r_t + \frac{-G_l(z^{-1})G_c(z^{-1})}{1+G_c(z^{-1})G_p(z^{-1})}e_t \\ &\triangleq P(z^{-1})r_t + Q(z^{-1})e_t \end{aligned} \tag{3.37}$$

or in a matrix format:

$$\begin{pmatrix} y_t \\ u_t \end{pmatrix} = \begin{pmatrix} M(z^{-1}) \\ P(z^{-1}) \end{pmatrix} r_t + \begin{pmatrix} N(z^{-1}) \\ Q(z^{-1}) \end{pmatrix} e_t \tag{3.38}$$

Clearly, $M(z^{-1})$,$N(z^{-1})$,$P(z^{-1})$ and $Q(z^{-1})$ can be estimated as an open-loop identification problem and then process and disturbance models can be constructed as follows:

$$\begin{aligned} \hat{G}_p(z^{-1}) &= \hat{M}(z^{-1})\hat{P}(z^{-1})^{-1} \\ \hat{G}_l(z^{-1}) &= -\hat{P}(z^{-1})^{-1}\hat{Q}(z^{-1}) \end{aligned}$$

Looking again at equations (3.26) and (3.27) in the modified approach shows that Equation (3.38) has been transformed into subspace framework where $M(z^{-1})$ has the same role as L_{YR} and $P(z^{-1})$ has the same role as L_{UR}. The same can be said for $Q(z^{-1})$ and L_{UE}. So we can estimate L_u and L_e in a similar way as that of $G_p(z^{-1})$ and $G_l(z^{-1})$

$$\begin{aligned} \hat{L}_u &= \hat{L}_{YR}\ (\hat{L}_{UR})^{-1} \\ \hat{L}_e &= -(\hat{L}_{UR})^{-1}\hat{L}_{UE} \end{aligned}$$

Using open-loop identification technique brings the advantage of consistency in the noise covariance estimation which is shown next.

3.3.2 Consistency of noise covariance estimation

For the proof of consistency in the estimation of noise covariance, we follow the framework of [57]. Consider the system with r_t as input and y_t as output:

$$\begin{cases} x_{t+1}^{cl} & = A_{cl}x_t^{cl} + B_{cl}r_t + K_{cl}e_t \\ y_t & = C_{cl}x_t^{cl} + D_{cl}r_t + e_t \end{cases} \tag{3.39}$$

where $x_t^{cl} \in \mathbb{R}^{n+n_c}$. Using the standard procedure of subspace approach [43, 57, 96], the state vector of this system, X_f^{CL}, defined in Equation (3.21), can be written in the subspace form as:

$$X_f^{CL} = \phi_y Y_p + \phi_r R_p + L_{x^{cl}} X_p^{CL} \tag{3.40}$$

where

$$\begin{aligned} \phi_y &= \mathcal{C}(A_{cl} - K_{cl}C_{cl}, K_{cl}) \\ \phi_r &= \mathcal{C}(A_{cl} - K_{cl}C_{cl}, B_{cl} - K_{cl}D_{cl}) \\ L_{x^{cl}} &= (A_{cl} - K_{cl}C_{cl})^N \end{aligned} \tag{3.41}$$

where $\mathcal{C}$ operator is defined by (2.17).

Substituting X_f^{CL} in Equation (3.19), for Y_f we have

$$\begin{aligned} Y_f &= L_Y W_p^{yr} + L_{YR} R_f + \Gamma_N^{CL} L_{x^{cl}} X_p^{CL} + L_{YE} E_f \\ &= \Theta Z + \Gamma_N^{CL} L_{x^{cl}} X_p^{CL} + L_{YE} E_f \end{aligned} \tag{3.42}$$

where $\Gamma_N^{CL} \in \mathbb{R}^{mN \times (n+n_c)}$ is the extended observability matrix and

$$\Theta = \begin{pmatrix} L_Y & L_{YR} \end{pmatrix} \quad and \quad Z = \begin{pmatrix} W_p^{yr} \\ R_f \end{pmatrix}$$

For convenience in the notation we use L_x instead of $L_{x^{cl}}$ and X_p instead of X_p^{CL} in the rest of this section. The term with L_x in (3.42) goes to zero for a large value of N and the following regression model appears

$$Y_f = \Theta Z + V \tag{3.43}$$

Θ can be estimated by $\hat{\Theta} = Y_f Z^\dagger$ and the residual of the estimation is $v = L_{YE} e_f$. Covariance of $L_{YE} e_f$ is estimated by

$$\hat{P}_v = \frac{1}{j} V V^T \tag{3.44}$$

Note that the noise covariance matrix estimation, $\hat{R}$, is calculated by $\hat{R} = \hat{P}_v(1 : m, 1 : m)$, so consistency of noise covariance estimation can be shown by means of $\hat{P}_v$.

From least squares estimation property, for $j \to \infty$, $Y_f \to \Theta Z$ in (3.42) which gives

$$L_{YE} e_f = -\Gamma_N^{CL} L_x x_p$$

where $x_p(k) = x(k)$ for $k = 1, \cdots, N$. The covariance of the residual, v, can be expressed in terms of L_x, x_p and Γ_N^{CL} by

$$\begin{aligned} P_v &= Cov(L_{YE} e_f) \\ &= E[(L_{YE} e_f)(L_{YE} e_f)^T] \\ &= \Gamma_N^{CL} L_x [Cov(x_p x_p^T)] L_x^T (\Gamma_N^{CL})^T \end{aligned} \tag{3.45}$$

Equation (3.45) will be used later in the proof of consistency.

Now consider the problem of estimation of x_p having information z using the least squares method where $z(k) = [y(k)^T \; r(k)^T]^T$ for $k = 1, \cdots, N$. This problem can be formulated as follows:

$$x_p = \Pi z + \tilde{x} \tag{3.46}$$

where $\tilde{x}$ is the estimation error. The following is the standard procedure of least squares estimation:

$$\begin{aligned} \hat{\Pi} =& X_p Z^T (ZZ^T)^{-1} \\ \Rightarrow \tilde{X} =& X_p - \hat{\Pi} Z = X_p - X_p Z^T (ZZ^T)^{-1} Z \end{aligned}$$

which gives

$$\begin{aligned} \tilde{X}\tilde{X}^T =& [X_p - X_p Z^T (ZZ^T)^{-1} Z][X_p^T - Z^T (ZZ^T)^{-1} Z X_p^T] \\ =& X_p X_p^T - X_p Z^T (ZZ^T)^{-1} Z X_p - X_p Z^T (ZZ^T)^{-1} Z X_p \\ &+ X_p Z^T (ZZ^T)^{-1} [ZZ^T (ZZ^T)^{-1}] Z X_p \\ =& X_p X_p^T - X_p Z^T (ZZ^T)^{-1} Z X_p \end{aligned}$$

Dividing both side by $1/j$

$$\begin{aligned} \frac{1}{j}\tilde{X}\tilde{X}^T =& \frac{1}{j} X_p X_p^T - (\frac{1}{j} X_p Z^T)(\frac{1}{j} ZZ^T)^{-1} (\frac{1}{j} Z X_p^T) \\ =& \frac{1}{j} \sum_{k=1}^{j} x_p(k) x_p(k)^T - (\frac{1}{j} \sum_{k=1}^{j} x_p(k) z(k)^T) \\ & (\frac{1}{j} \sum_{k=1}^{j} z(k) z(k)^T)^{-1} (\frac{1}{j} \sum_{k=1}^{j} z(k) x_p(k)^T) \end{aligned} \tag{3.47}$$

and for $j \to \infty$ it gives

$$Cov(\tilde{x}) \triangleq P_{\tilde{x}} = \bar{E}(x_p x_p^T) - \bar{E}(x_p z^T) \bar{E}(zz^T)^{-1} \bar{E}(z x_p^T) \tag{3.48}$$

where $\bar{E}$ is defined as

$$\bar{E}((\bullet)) = \lim_{j \to \infty} \frac{1}{j} \sum_{k=1}^{j} E((\bullet)_k)$$

Now we can state the following theorem for the consistency of noise covariance estimation for the closed-loop identification proposed in this chapter.

Theorem 1: Assuming under closed-loop condition, e_t and r_t are uncorrelated and r_t is persistently exciting of sufficient order, then estimation of noise variance $(\hat{P}_v)$ for $N \to \infty$ converges to P_v asymptotically.

Proof.
Based on Equation (3.43), and using the definition of pseudo-inverse and Equation (3.42), $\hat{\Theta}$ can be expressed as

$$\begin{aligned}\hat{\Theta} &= Y_f Z^\dagger = Y_f Z^T (ZZ^T)^{-1} \\ &= (\Theta Z + \Gamma_N^{CL} L_x X_p + L_{YE} E_f) Z^T (ZZ^T)^{-1} \\ &= \Theta + \Gamma_N^{CL} L_x X_p Z^T (ZZ^T)^{-1} + L_{YE} E_f Z^T (ZZ^T)^{-1} \end{aligned} \tag{3.49}$$

The last term of Equation (3.49) goes to zero when $j \to \infty$, because $\bar{E}(e_f z^T) = 0$ and $\bar{E}(zz^T)$ exists, owing to the persistent excitation. So, $\bar{E}(e_f z^T)\bar{E}(zz^T)^{-1} = 0$ and therefore we can show that

$$\begin{aligned} E_f Z^T (ZZ^T)^{-1} &= \frac{1}{j} E_f Z^T (\frac{1}{j} ZZ^T)^{-1} \\ &= \frac{1}{j} \sum_{k=1}^{j} e_f(k) z(k)^T (\frac{1}{j} \sum_{k=1}^{j} z(k) z(k)^T)^{-1} \\ &= \bar{E}(e_f z^T)\bar{E}(zz^T)^{-1} = 0 \quad (for j \to \infty) \end{aligned} \tag{3.50}$$

Based on the results of (3.50), Equation (3.49) gives the $\hat{\Theta}$ as follows:

$$\hat{\Theta} = \Theta + \Gamma_N^{CL} L_x X_p Z^T (ZZ^T)^{-1}$$

which yields

$$\hat{Y}_f = \Theta Z + \Gamma_N^{CL} L_x X_p Z^T (ZZ^T)^{-1} Z$$

Therefore, residual V can be estimated by

$$\hat{V} = Y_f - \hat{Y}_f = -\Gamma_N^{CL} L_x X_p Z^T (ZZ^T)^{-1} Z$$

Based on Equation (3.44), for estimation of the covariance of v, $\hat{P}_v$, one needs to calculate $\hat{V}\hat{V}^T$ as

$$\begin{aligned} \hat{V}\hat{V}^T = \\ &= \Gamma_N^{CL} L_x X_p Z^T (ZZ^T)^{-1} [ZZ^T (ZZ^T)^{-1}] Z X_p^T L_x^T (\Gamma_N^{CL})^T \\ &= \Gamma_N^{CL} L_x [X_p Z^T (ZZ^T)^{-1} Z X_p^T] L_x^T (\Gamma_N^{CL})^T \\ &= \Gamma_N^{CL} L_x [X_p X_p^T + X_p Z^T (ZZ^T)^{-1} Z X_p^T - X_p X_p^T] L_x^T (\Gamma_N^{CL})^T \\ &= \Gamma_N^{CL} L_x X_p X_p^T L_x^T (\Gamma_N^{CL})^T - \Gamma_N^{CL} L_x [X_p X_p^T - X_p Z^T (ZZ^T)^{-1} Z X_p^T] L_x^T (\Gamma_N^{CL})^T \end{aligned}$$

Using derivations of P_v and $P_{\tilde{x}}$ in equations (3.45) and (3.48), $\hat{P}_v$ is given by

$$\hat{P}_v = \frac{1}{j}\hat{V}\hat{V}^T = P_v - \Gamma_N^{CL} L_x P_{\tilde{x}} L_x^T (\Gamma_N^{CL})^T \tag{3.51}$$

$L_x \to 0$ when N increases because of the stability of Kalman filter and $P_{\tilde{x}}$ also decreases with N, so convergence of $\hat{P}_v$ with respect to N is fast due to the three factors in $L_x P_{\tilde{x}} L_x^T$ decreasing with N.

Remark. Selection of row dimension in data Hankel matrices, N, has been addressed in some subspace identification papers [9, 74, 25, 78] for the open-loop identification case. In the literature, this parameter is often said to be selected 'large enough' to cover the dynamics of system, but not too large to cause 'over-parametrization' problem. The most cited method was presented in [78] based on the analogy between subspace identification and ARX identification. It examines AIC or BIC criteria on a set of ARX models of order N (obtained from the first row block of Equation (2.19)) for $n \leqslant N \leqslant log(j)^{\alpha}$, to find the best value of N. n is the order of system and $\alpha < \infty$. However, there is no detailed analysis on the selection of N for closed-loop identification.

Based on the above proof procedure, consistency depends on both N and the number of columns in data Hankel matrices, j. In the analysis of consistency, the value of j is considered to be infinite. The value of N is required to be sufficiently large such that L_x in Equation (3.41) approaches zero. It should be noted that $(A_{cl} - K_{cl}C_{cl})$ represents dynamics of the optimal observer which is stable. Thus, it is always possible to find a sufficiently large N for $\{L_x \to 0\}$ to hold. But, finding a proper value for N based on this analogy needs the optimal observer dynamics which is unknown before the identification is performed. One approach to solve the problem is to use an iterative method which starts from a small value of N (larger than the system order) and at each iteration identifies the closed-loop system matrices and evaluates L_x. By increasing N at each iteration, one can find a value of N for which L_x is ignorable.

To avoid this iteration approach, we use upper bound of N by using the closed-loop process dynamics. Dynamics of the optimal observer should normally be faster than the process dynamics to be observed. Thus, a value of N that assures $(A_{cl})^N$ to be close to zero will generally guarantee that L_x is close to zero. Note that the convergence of term $\{\Gamma_N^{CL} L_x P_{\tilde{x}} L_x^T (\Gamma_N^{CL})^T\}$ in Equation (3.51) is faster than L_x. This value of N can be obtained from a correlation analysis on the closed-loop data, and provides an upper bound on N. Note that the closed-loop system is stable. So, impulse response coefficients of the system must converge to zero. In the joint input-output identification approach we consider two closed-loop systems, one with y_t as the output and one with u_t as the output. Therefore, two correlation analysis

should be performed and the maximum of the two calculated values of N can be used as common N. We use this approach in the next example to elaborate the idea.

Remark. Although we show that the proposed method of this chapter is consistent, it should be noted that contrary to direct identification methods, joint input-output identification requires the assumption that the controller is linear with no constraints on the control signal. Fortunately, the potential problem of solving non-trivial model reduction problem in joint input-output identification [22] does not appear here, because the explicit model of the process is not required for LQG curve estimation.

3.3.3 Simulation

Consider the system and controller given in the previous example. We run a Monte-Carlo simulation using the proposed closed-loop identification method of this chapter. The same conditions as the previous example are applied here for the simulation.

Figure 3.2 shows the correlation analysis results on the closed-loop data. Based on this graph, a proper upper bound on N can be chosen as 25.

The results of Monte-Carlo simulation shown in Figure 3.3 indicate the consistent estimation of the noise variance for large enough value of N.

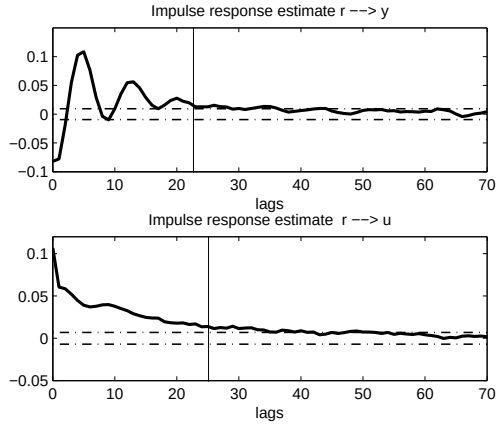


Figure 3.2: Results of correlation analysis on closed-loop data

Figure 3.4 shows the effect of value of N on the estimated LQG curve which agrees with 3.3. The average value of the estimated noise variance has been used for each N. Figure 3.5 presents the final results for this example which shows that the estimated LQG curves converge to the 'True curve' asymptotically using the proposed method of this chapter (for $N = 25$). The 'True Curve' comes from solving the LQG problem using the true model of the process.

In Section 5, this method will be implemented in a pilot-scale application.

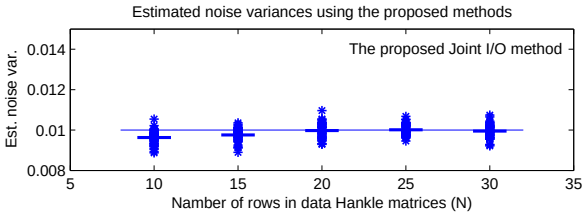


Figure 3.3: Results of Monte-Carlo simulation using the proposed method of joint I/O identification.

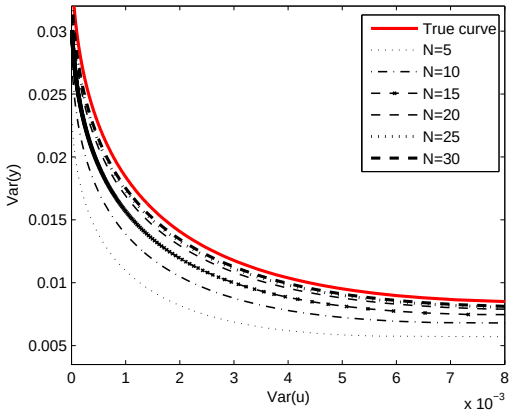


Figure 3.4: Trade-off curves for different values of N (j=4000)

3.4 Practical considerations

In many control applications, specifically when model predictive controllers are designed for a process, the disturbance model is typically 'selected' (as a tuning parameter) rather than the real disturbance model. Disturbance dynamics changes with time in many applications. As discussed before, both process and disturbance models are needed for calculating the LQG trade-off curve. On the other hand, if one intends to identify both process and noise models by performing identification test, a large set of experiment data might be required, because the value of N for having consistent estimation of the noise covariance can be large, specifically for MIMO processes.

In this section, we will present a simple yet practical method of using closed-loop routine operating data and available subspace matrices of plant dynamics to calculate the noise subspace matrix. Note that number of data samples is typically not a problem when using routine operating data. We provide the proof of consistency for the noise model and noise covariance estimation using this method. A simulation example is provided in this section followed by a pilot-scale experiment in the next section.

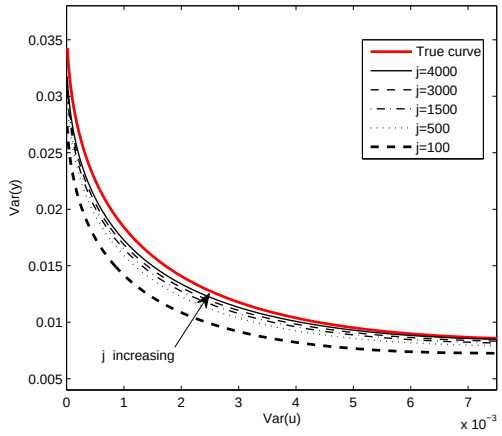


Figure 3.5: trade-off curves for increasing values of j (N=25)

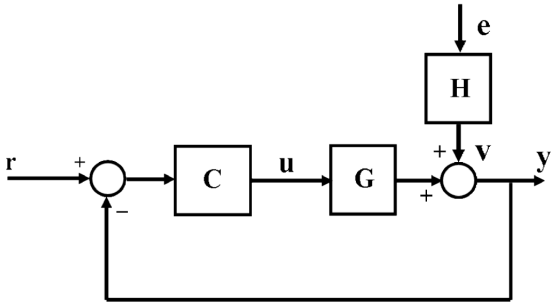


Figure 3.6: A typical closed-loop process

Consider a linear dynamic model of the process available in a format such as transfer function, finite impulse response (FIR), finite step response (FSR) or state space. The subspace matrix of the process, L_u, can be obtained from any of these formats. Some of them can be transformed directly to L_u, such as FIR and state space models. Subspace matrices for open-loop process, L_u and L_w, can be identified from open-loop or closed-loop experiments as well [43, 57].

Now, consider the typical closed-loop system shown in Figure 3.6 in which the contribution of noise in the output is defined by 'v'. Having a set of input-output data from routine operation of the closed-loop system, one can predict the process output based on Equation (3.2) as

$$\hat{y}_f = L_w w_p + L_u u_f$$

Having the output prediction, v_f can be found from

$$v_f = y_f - \hat{y}_f \tag{3.52}$$

The first half of this data set can be used to form V_p and the second half for

constructing V_f, so that both have the half number of columns comparing to that of E_f. Therefore, in the rest of computations, E_f will be replaced by $E'_f \triangleq E_f(:, j/2 : j)$. Note that this modification only means using less number of data points and does not affect the size of L_e in the estimation. The details of estimation and consistency analysis are provided in Section 4.1. The term with X_p^h in Equation (3.59) goes to zero with $N \to \infty$, and the following regression model appears:

$$V_f = L_v V_p + L_e E'_f$$

Then, least square estimation can be used to obtain $\hat{L}_v$ as

$$\hat{L}_v = V_f V_p^{\dagger} \tag{3.53}$$

The residual of estimation is $s = L_e e'_f$ and its covariance can be estimated by

$$\hat{P}_s = \frac{1}{j} S S^T \tag{3.54}$$

Note that the noise covariance matrix estimation, $\hat{R}$, is calculated by $\hat{R} = \hat{P}_s(1 : m, 1 : m)$.

$\hat{V}_f$ is found by the orthogonal projection of the row space of V_f and Y_f. One can define

$$\begin{aligned} \Omega &\triangleq V_f - \hat{V}_f \\ &= V_f - \hat{L}_v V_p = L_e E'_f \end{aligned} \tag{3.55}$$

So, the innovation term, e'_f, can be estimated as

$$\hat{e}'_f = \Omega(1 : m, :)$$

E'_f can be constructed from $\hat{e}'_f$ and $\hat{L}_e$ is given by

$$\hat{L}_e = \Omega(E'_f)^{\dagger}$$

3.4.1 Consistency Analysis

The disturbance model $H(z)$ in Figure 3.6 can be represented in the state space form as:

$$\begin{cases} x_{t+1}^h &= A^h x^h(t) + K^h e(t) \\ v_t &= C^h x^h(t) + e(t) \end{cases} \tag{3.56}$$

where $x_t^h \in R^{n_h}$, $v_t \in R^m$ and $e_t \in R^m$ is white noise. Similar to system (3.1), the basic subspace equation for this system is as follows:

$$V_f = \Gamma_N^h X_f^h + L_e E_f \tag{3.57}$$

Using regression analysis, future states can be represented by past outputs and states in subspace format as

$$X_f^h = \phi_v V_p + L_x^h X_p^h \tag{3.58}$$

where

$$\phi_v = \mathcal{C}(A^h - K^h C^h, K^h)$$
$$L_x^h = (A^h - K^h C^h)^N$$

Substituting (3.58) in (3.57) yields

$$\begin{aligned} V_f &= L_v V_p + \Gamma_N^h L_x^h X_p^h + L_e E_f \\ &= L_v V_p + S \end{aligned} \tag{3.59}$$

where

$$L_v = \Gamma_N^h \phi_v$$

To prove the consistency in estimation of L_v, one may substitute V_f in Equation(3.53) from (3.59) and $V_p^\dagger$ by its definition, which gives

$$\begin{aligned} \hat{L}_v =& V_f V_p^\dagger \\ =& L_v V_p V_p^T (V_p V_p^T)^{-1} + \Gamma_N^h L_x^h X_p^h V_p^T (V_p V_p^T)^{-1} + L_e E_f V_p^T (V_p V_p^T)^{-1} \\ =& L_v + \Gamma_N^h L_x^h X_p^h V_p^T (V_p V_p^T)^{-1} + L_e E_f V_p^T (V_p V_p^T)^{-1} \\ =& L_v \qquad (where \quad j \to \infty \quad and \quad N \to \infty) \end{aligned} \tag{3.60}$$

The second term in (3.60) goes to zero because L_x^h converges to zero when $N \to \infty$ (stability of Kalman filter) and the third term goes to zero because $\bar{E}(e_f v_p^T) = 0$ and $\bar{E}(v_p v_p^T)$ exists (persistent excitation). So, $\bar{E}(e_f v_p^T)\bar{E}(v_p v_p^T)^{-1}$ converges to zero and similar to (3.50), we can show that $E_f V_p^T (V_p V_p^T) = 0$ when $j \to \infty$, which proves the consistency of $\hat{L}_v$.

Consistency of noise covariance estimation can be proven following the same procedure as in Section 3.2 by replacing Z with V_p and X_p^{CL} with X_p^h.

3.4.2 Example

Consider a 2×2 process with open-loop transfer function matrix G_p and disturbance transfer function matrix G_l given as [45]

$$G_p = \begin{pmatrix} \frac{z^{-1}}{1-0.4z^{-1}} & \frac{0.5z^{-2}}{1-0.1z^{-1}} \\ \frac{0.3z^{-1}}{1-0.4z^{-1}} & \frac{z^{-2}}{1-0.8z^{-1}} \end{pmatrix}$$
$$G_l = \begin{pmatrix} \frac{1}{1-0.4z^{-1}} & \frac{-z^{-1}}{1-0.1z^{-1}} \\ \frac{z^{-1}}{1-0.7z^{-1}} & \frac{1}{1-0.8z^{-1}} \end{pmatrix} \tag{3.61}$$

The following controller is implemented on the process:

$$G_c = \begin{pmatrix} \frac{0.5-0.2z^{-1}}{1-0.5z^{-1}} & 0 \\ 0 & \frac{0.25-0.2z^{-1}}{(1-0.5z^{-1})(1+0.5z^{-1})} \end{pmatrix} \tag{3.62}$$

Input and output data under routine closed-loop operation of the process have

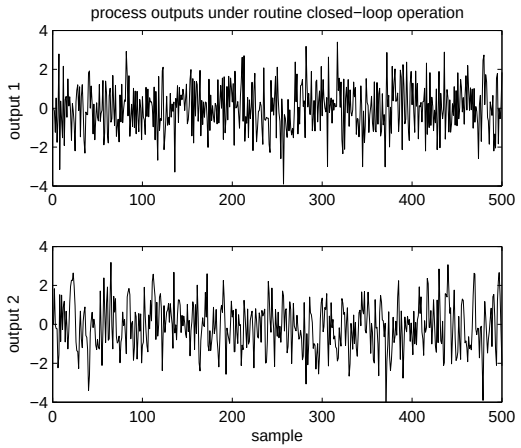


Figure 3.7: Outputs of the process under routine closed-loop operation

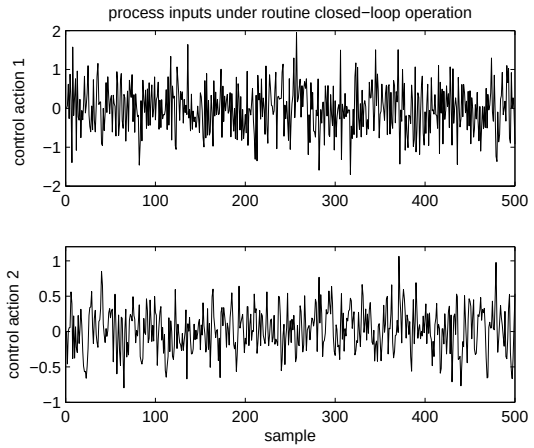


Figure 3.8: Control actions under routine closed-loop operation

been collected and shown in Figures 3.7 and 3.8. Noise contributions in the outputs are estimated based on (3.52). Using process and disturbance models, the true LQG curve for this process can be obtained. The process model has also been used to calculate subspace matrix L_u. Based on the presented procedure, noise subspace matrix, L_e, is estimated from routine closed-loop data. Both the true and estimated trade-off curves are plotted in Figure 3.9 which confirms the ability of this method to estimate the noise model from routine closed-loop data.

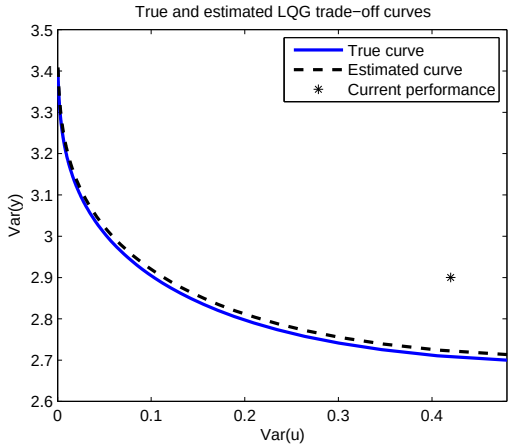


Figure 3.9: True and estimated LQG trade-off curves and actual variances

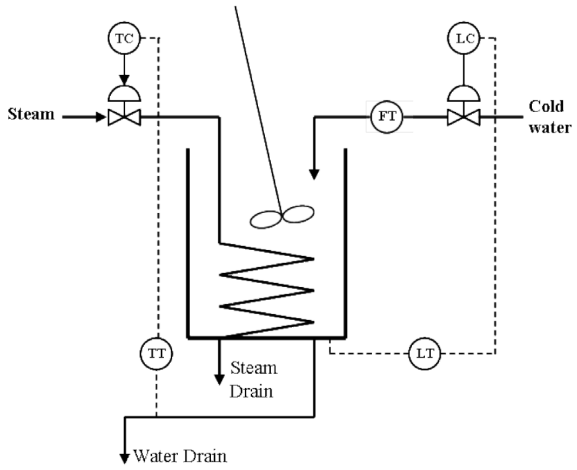


Figure 3.10: A schematic of Continuous Stirred Tank Heater process

3.5 Application on a pilot-scale process

The two proposed methods have been applied for performance assessment of a pilot-scale process. The process is a continuous stirred tank heater (CSTH) and its schematic is shown in Figure 3.10.

There are two controlled variables in the CSTH, the water level inside the tank and outlet water temperature. Manipulated variables are the cold water flowrate and steam flowrate. The head of the water in the inlet pipe as well as the steam supply pressure and temperature can be considered as disturbances. This process is under PID control by *DeltaV* control system. Tuning parameters of both level and temperature controllers are given in Table 3.1.

After some preliminary tests, two '*RBS*' signals are designed using MATLAB for testing the process under closed-loop conditions. Signals are applied to the setpoint

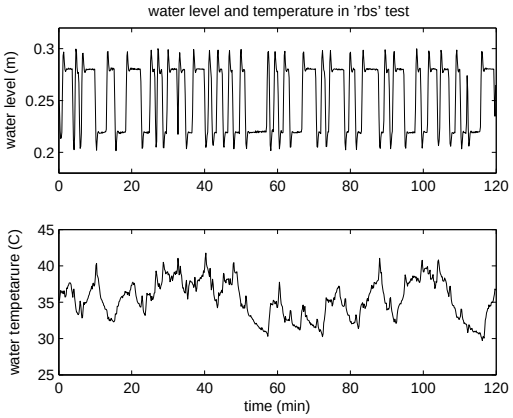


Figure 3.11: CSTH outputs under closed-loop 'RBS' test

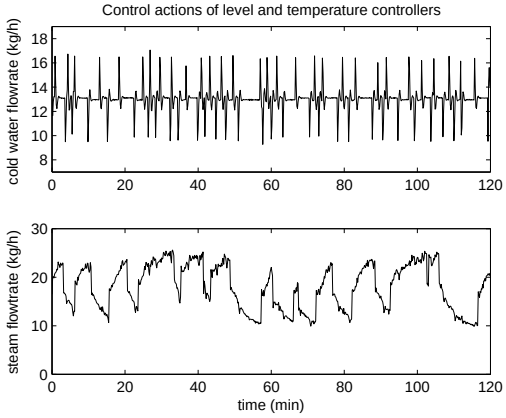


Figure 3.12: PID control actions under closed-loop 'RBS' test

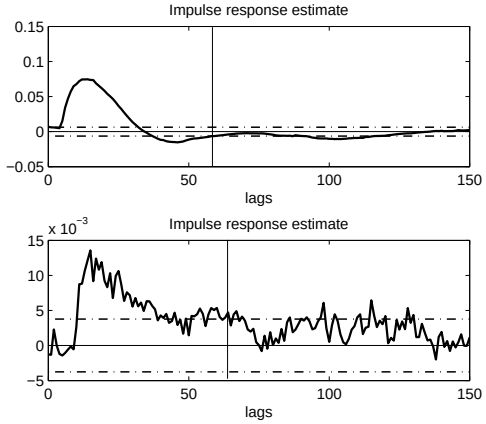


Figure 3.13: Correlation analysis results on closed-loop test data from CSTH process

Table 3.1: Initial tuning parameters of PID controllers in CSTH process

	Level controller	Temp. controller
Gain	1.1	0.9
Reset	75	90
Rate	0	0

Table 3.2: Modified tuning parameters of PID controllers in CSTH process

	Level controller	Temp. controller
Gain	1.5	1.4
Reset	100	80
Rate	0	8

of the two controllers and closed-loop data are collected with 2 seconds sampling time and over 2 hours. The input and output of the process under the identification test are shown in Figures 5.7 and 3.12. Correlation analysis results for the closed-loop data can be seen in Figure 3.13 which suggests the value of N to be around 60. Subspace matrices are identified using the proposed closed-loop identification method and the LQG trade-off curve for the process are estimated.

One hour of routine operating data has also been collected and used as a measure of current performance. Figure 3.17 shows the trade-off curve as well as the current control performance (named as 'PID tuning 1'). The results indicates some

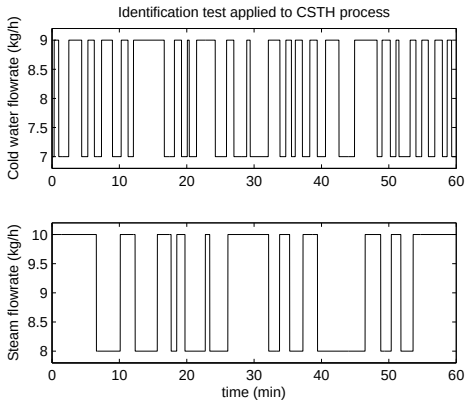


Figure 3.14: RBS test signal used for open-loop identification of the process model

potential for improvement in performance. We have then used the IMC tuning method followed by some fine tuning to get a better set of tuning parameters for the two PIDs given in Table 3.2 and the corresponding performance point is shown in Figure 3.17 by '*' (named as 'PID tuning 2'). From this figure, it can be inferred that improvement in output variance comes with a small increase in the input variance.

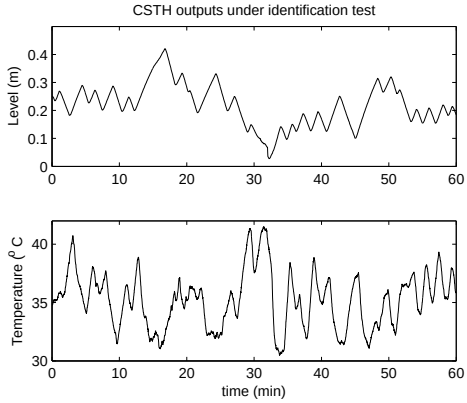


Figure 3.15: Open-loop response of the CSTH process to the test signal

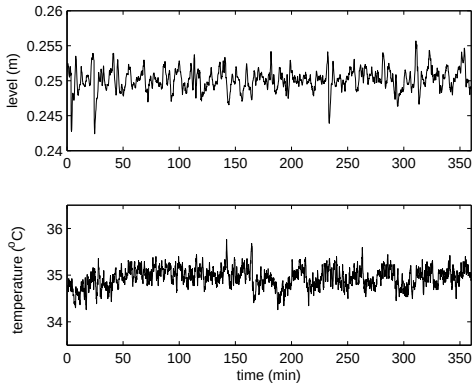


Figure 3.16: Closed-loop routine operating data from CSTH process

We have also implemented an LQG controller on this process based on the subspace LQG law [53]:

$$u_f^{opt} = (L_u^T L_u + \lambda I)^{-1} L_u^T L_w$$

The LQG control algorithm was run in MATALB which was connected to *DeltaV* control system through OPC. Performance of the controller for $\lambda = 7$ is shown in Figure 3.17. It shows that the performance point of this LQG controller is closer but does not lie exactly on the curve owing to the process nonlinearity and experiment errors such as changes in the steam pressure and drift of steam temperature, etc.

We have also applied the estimation procedure of Section 4 on the CSTH process for estimating noise subspace matrix from routine operating data. For this purpose, we have performed an open-loop identification test on the process to identify the process subspace matrices L_w and L_u. Process inputs and outputs under the identification test are shown in Figures 3.14 and 3.15, respectively. Subspace matrices L_u and L_w are identified using the regression analysis method [57] and the corresponding disturbance matrix L_e is estimated by the method of Section 4 using

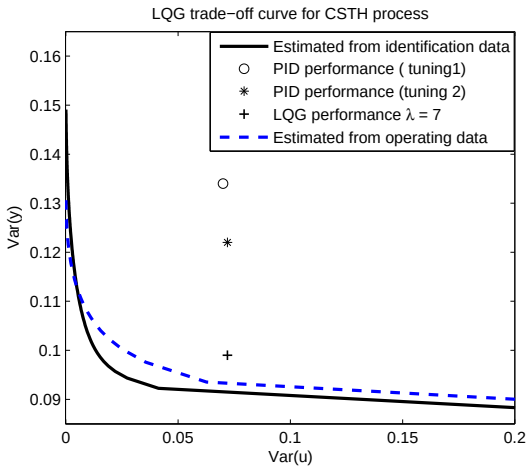


Figure 3.17: LQG trade-off curves and actual performance points

the routine operating data set shown in Figure 3.16. The estimated LQG trade-off curve is shown in Figure 3.17 with dashed line. The result shows an estimation of the curve close to the one from experiment data.

3.6 Concluding remarks

In this chapter, a direct formulation of the joint input-output closed-loop subspace identification method was proposed. A lack of consistency in the noise covariance estimation was observed in many of the previous methods when used for the LQG trade-off curve estimation. As a result, the direct formulation based on two separate open-loop identification problems was proposed which provides consistency of the noise variance estimation. The possible limitation of the proposed method is the requirement of a linear controller. The results of the LQG benchmark estimation for a simulation example and a pilot-scale continuous stirred tank heater process were provided which verified the effectiveness of the presented method for estimation of the LQG trade-off curve.

A procedure for estimating noise subspace matrix and noise variance from routine closed-loop operating data was also proposed along with the consistency analysis. Simulation and pilot-scale application results show the effectiveness of the presented method for estimation of the LQG trade-off curve.

Chapter 4

Performance Assessment of Advanced Supervisory-regulatory Control Systems with Subspace LQG

4.1 Introduction

Automation systems are an essential part of almost any industrial process and controller is the heart of an automation system. High performance control systems require healthy controllers. However, surveys show that about sixty percents of industrial controllers have some kind of performance problem [86]. Controller performance assessment has been one of the most interesting areas of research in the field of control engineering during past decades. Performance assessment of advanced process control (APC) systems is receiving increasing attention because of their high design and implementation cost. Considerable academic as well as commercial interests have been devoted to the monitoring of both univariate and multivariate control systems [13, 31, 39, 47, 61, 59, 75, 76, 80, 89, 101].

Minimum variance control (MVC) benchmark by Harris (1989) [36] provided the fundamental step to measure the performance and presents valuable information about a lower bound on the process variance. However, the *a priori* requirements of this approach are not easily obtainable in the case of multivariate process [27, 37, 38]. Many researchers modified and improved this method by means of different estimation and identification methods to make it more applicable, especially for MIMO processes [37, 46, 55].

The LQG benchmark approach considers variances of both input and output and provides a 'trade-off' curve which represents a feasible range of performance for linear controllers [14, 45]. In other words, this curve provides optimum values of the output variance for a range of the variance of the manipulated variable. This

method is based on a model of the process which has to be obtained through process identification.

Process identification is a necessary step in many control applications. Subspace identification methods provide an alternative approach to classic system identification methods. Various methods of subspace identifications have been developed in the past two decades [57, 64, 77, 91, 92, 96, 90]. Many studies have been devoted to the closed-loop identification as well because of its extensive use in control-relevant identification and controller performance assessment [64, 70, 88, 94, 95].

A method for the LQG trade-off curve estimation in the subspace framework was presented in Chapter 3 which does not need an explicit process model. The work to be presented in this chapter will also be established under the subspace framework.

Implementation of advanced supervisory controllers has been one of the most demanded in control engineering, because of the resulted economic benefits. On the other hand, these economic expectations and the cost of APC implementation always raise questions about performance of these controllers. Enormous amount of studies has been dedicated to this issue in both academia and industry using many different approaches [1, 2, 3, 34, 44, 52, 67, 79, 86, 99, 100]. APC is normally implemented on the process as a supervisory controller at the top of the existing regulatory control system. This control structure is named '*Cascade*' implementation by Lee (2000) [66] and has been discussed in details for the case in which the advanced controller is model predictive control (MPC). Advantages and disadvantages of this type of APC implementation were elaborated in comparison to the so called '*Direct*' implementation. We will review both these cases briefly in the next section. Performance assessment of the cascade control structure is considered in this chapter which has not been studied in the above mentioned references. Ko and Edgar (2000) [58] studied the performance assessment of classic cascade control loops which have a different structure comparing to the supervisory-regulatory cascade control structure addressed in this chapter.

Comparing to our previous work on subspace LQG benchmarking for conventional feedback control in Chapter 3, the advanced supervisory-regulatory control of this chapter constitutes a special structure and results in different mathematical derivations in the control law, the benchmark, and performance indices. The cascade structure provides us a much richer alternative in performance assessment.

In this chapter, we employ the concept of LQG benchmark and use subspace framework as a tool to derive a method of performance assessment for the cascade control system. Three possible scenarios for performance assessment in a cascade supervisory-regulatory control structure are described and the LQG control design is provided for each scenario. We also illustrate how to obtain the minimum input and output variances under the LQG control in a model-free subspace framework. The

results lead us to the LQG trade-off curves for performance assessment using certain subspace matrices. A closed-loop subspace identification procedure for estimation of the required subspace matrices is also provided based on the joint input-output identification approach.

This chapter is organized as follows: In Section 2, a review on the available options for implementation of supervisory advanced control on existing regulatory control layer is presented. Section 3 describes how to use the LQG benchmark for the performance assessment of the cascade control system. Procedures of designing the LQG control and obtaining the trade-off curve are discussed in this section. A closed-loop identification method for estimation of the required subspace matrices and the input noise variance is provided in Section 4. In Section 5, two simulation examples in MATALB and HYSYS are provided. Section 6 provides some concluding remarks.

4.2 Implementation of advanced supervisory control on the regulatory control layer

This section is a review on the issue of interfacing advanced model-based control systems with low-level regulatory controllers. Lee (2000) [66] provided a detailed study on this issue with MPC as the supervisory control.

Consider a process which is under feedback control, e.g. PID control. An advanced process controller, e.g. a model-based controller, is designed to improve performance of the control system. There are two main options for implementing the advanced model-based controller on the existing regulatory control level. One is called *Direct* implementation, which means breaking the regulatory loops, identifying a model of the open-loop process and applying the model-based controller directly to the process (see Figure 4.1). The second option is called cascade control, where the regulatory control level is kept in place and APC provides the setpoints to this controller (see Figure 4.2). Process identification, handling the constraints and disturbance rejection are the major sources of difference between these two approaches [66].

Since the inputs of the model in the *Direct* approach are the valve positions, valve limits can be entered directly into the APC algorithm as the input constraints. Therefore, handling of input constraints can be done in a straightforward manner. However, because the regulatory loops are taken out, one may lose the efficiency in disturbance rejection. Note that the disturbance must propagate through the process to affect the output before any control action can take place.

The cascade control structure, on the other hand, helps to eliminate the disturbances that may occur inside the loop more efficiently and quickly. In addition,

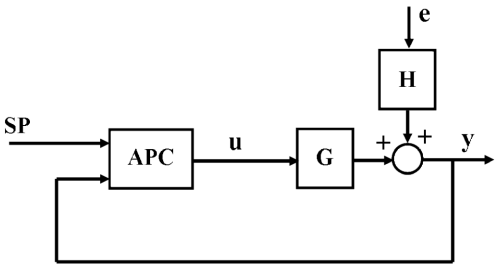


Figure 4.1: A schematic of *Direct* implementation.

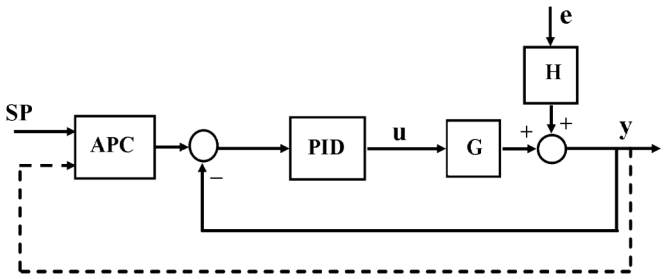


Figure 4.2: A schematic of cascade implementation.

identification can be made easier since unstable or excessively slow dynamics are stabilized or made faster by regulatory loops. However, since the valves are not directly manipulated, handling of valve constraints becomes more complicated and also performance of the control system depends on both supervisory and regulatory controllers.

From a practical point of view, in most cases the regulatory control system should be kept working all time, because usually performing identification tests on the open loop system (without regulatory controllers) is not preferred by the process engineers. Most of the APC systems have a shutdown option for the cases where their computational algorithms (such as linear/nonlinear dynamic programming) fail in computing the next step of the control action [79]. For such cases, having a backup control system is necessary. Therefore, in most of the APC applications, the regulatory control system, or at least some of the regulatory loops, are not broken [66] which results in having cascade control structures. Optimal LQG design and performance assessment of these cascade control systems are discussed in this chapter.

4.3 Performance assessment in a cascade control structure using LQG benchmark

The question to be answered in this section is how to design a LQG controller in a two-layer control structure and how to find optimal input and output variances which leads us to the trade-off curve.

4.3.1 Classic design of the LQG controller

We consider a more general form of LQG formulation than the classical one: controlled variables can be different from the measured variables as shown in the block diagram of Figure 4.3. In this figure γ_t represents the external inputs, u_t is the control signal and y_t and ξ_t represent the measured and controlled variables, respectively. The process P is given by the following state space presentation:

$$P : \begin{cases} x_{t+1} &= Ax_t + B_1\gamma_t + B_2u_t \\ y_t &= C_1x_t + D_{11}\gamma_t + D_{12}u_t \\ \xi_t &= C_1x_t + D_{21}\gamma_t + D_{22}u_t \end{cases}$$

The LQG control aim is to minimize the following objective function:

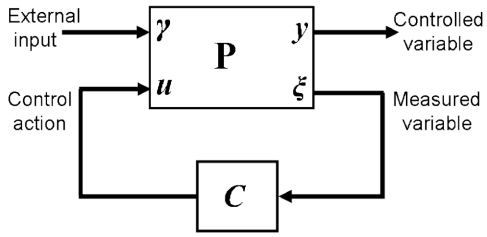


Figure 4.3: A schematic of classic LQG control configuration.

$$J = \lim_{t\to\infty} E\{[y_t^T y_t] + \lambda[u_t^T u_t]\}$$

The two main steps of the LQG control design, state estimation and optimal state-feedback design, are given by [62]

$$\hat{x}_{t+1} = A\hat{x}_t + B_2u_t + K_{KAL}(y_t - C_2\hat{x}_t + D_{22}u_t) \tag{4.1}$$

$$u_t = -K_{SF}\hat{x}_t \tag{4.2}$$

where

$$\begin{cases} K_{KAL} &= (P_1C_2^T + B_1D_{21}^T)[D_{21}D_{21}^T]^{-1} \\ K_{SF} &= [D_{12}^TD_{12}]^{-1}(B_2^TP_1 + D_{12}^TC_1) \end{cases}$$

P_1 and P_2 are the solutions of two algebraic Ricatti equations. Favoreel *et. al.* [32, 33] showed that, for the special case where $\xi_t = y_t$, the LQG control law can be expressed in terms of process and disturbance subspace matrices, L_u and L_e, assuming a finite-horizon LQG control objective. No explicit model of the process is needed and no Ricatti equation is required to be solved in this method. Employing this idea, Kadali and Huang [43, 53] further provided the expressions of input and output variances under this LQG control law which is then used to obtain the trade-off curve for performance assessment.

4.3.2 Subspace-based design of the LQG controller

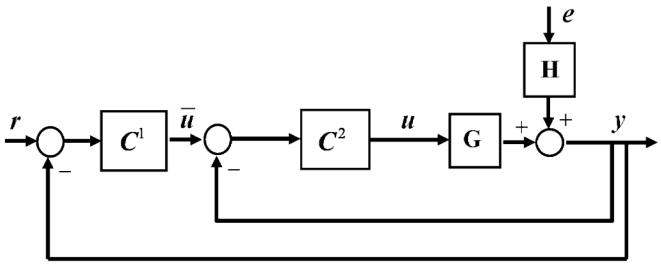


Figure 4.4: A schematic of cascade implementation with general linear controllers.

Let's start the derivation with the block diagram shown in Figure 4.4. In this diagram, the real input to the process (regulatory level control action) is named u while the setpoint of the regulatory controller, coming from the supervisory controller, is named $\bar{u}$. In the LQG control objective, variance of the input and output should be minimized which means minimizing the variances of u and y in Figure 4.4. Note that $\bar{u}$ does not represent a physical control variable, but a setpoint to the regulatory control loop.

For the cascade control structure of Figure 4.4, three possible scenarios can be considered as follows:

1- The controller C^2 (regulatory controller) is considered to be fixed and one is looking for the best achievable performance in terms of input and output variances by designing the supervisory controller C^1.

2- The controller C^1 (supervisory controller) is considered to be fixed and one is looking for the best achievable performance in terms of input and output variances by designing the regulatory controller C^2.

3- None of the two controllers C^1 or C^2 are known and the goal is to design both controllers in a way that optimal input and output variances are obtained.

It should be noted that in all three cases, input variance means variance of the regulatory controller action u.

Before starting the derivations, a very brief review on the subspace equations under the closed-loop is given below.

Subspace definitions in a closed-loop system

In a closed-loop system such as Figure 5.1, two separate open-loop models can be defined: a model from setpoint (r_t) to the output (y_t) and the other one from (r_t) to the controller output (u_t). Similar to Equation (2.19), these two systems can be presented by the following input-output relations [26]:

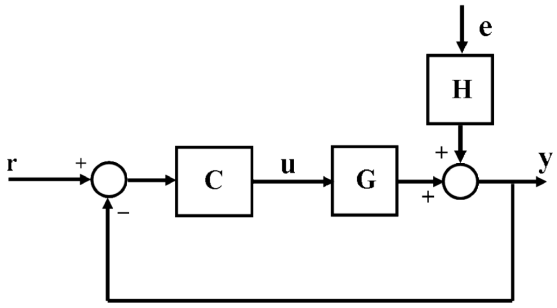


Figure 4.5: A typical closed-loop process.

$$Y_f = \Gamma_N^Y X_f^{CL} + L_{YR} R_f + L_{YE} E_f \tag{4.3}$$
$$= L_Y W_p^{yr} + L_{YR} R_f + L_{YE} E_f \tag{4.4}$$

$$U_f = \Gamma_N^U X_f^{CL} + L_{UR} R_f + L_{UE} E_f \tag{4.5}$$
$$= L_U W_p^{ur} + L_{UR} R_f + L_{UE} E_f \tag{4.6}$$

where:

$$X_f^{CL} = \begin{pmatrix} X_f \\ X_f^c \end{pmatrix} \ , \ W_p^{yr} = \begin{pmatrix} Y_p \\ R_p \end{pmatrix} \ , \ W_p^{ur} = \begin{pmatrix} U_p \\ R_p \end{pmatrix} \tag{4.7}$$

and X_f^c is the future state subspace matrix for the controller. The extended observability matrices, Γ_N^Y and Γ_N^U, are defined similar to (2.12). The closed-loop subspace matrices are given by

$$L_{YR} = (I + L_u L_y^c)^{-1} L_u L_y^c$$
$$L_{UR} = (I + L_y^c L_u)^{-1} L_y^c$$
$$L_{YE} = (I + L_u L_y^c)^{-1} L_e$$
$$L_{UE} = -(I + L_y^c L_u)^{-1} L_y^c L_e$$

where L_y^c contains Markov parameters of the controller.

Scenario 1

Subspace equations for the input and output in a closed-loop system are provided by Equations (4.3) and (4.5). Similar relations for the lower loop of Figure 4.4 are as follows:

$$\begin{aligned} Y_f =& \Gamma_N^{y2} X_f^{cc2} + L_{Y\bar{U}} \bar{U}_f + L_{YE}^{cs} E_f && (4.8)\\ =& L_{Y2} W_p^{y\bar{u}} + L_{Y\bar{U}} \bar{U}_f + L_{YE}^{cs} E_f && (4.9)\\ U_f =& \Gamma_N^{u2} X_f^{cc2} + L_{U\bar{U}} \bar{U}_f + L_{UE}^{cs} E_f && (4.10)\\ =& L_{U2} W_p^{u\bar{u}} + L_{U\bar{U}} \bar{U}_f + L_{UE}^{cs} E_f && (4.11) \end{aligned}$$

where Γ_N^{y2} and Γ_N^{u2} are extended observability matrices defined similar to (2.12) and

$$X_f^{cc2} = \begin{pmatrix} X_f \\ X_f^{c2} \end{pmatrix} \quad , \quad W_p^{y\bar{u}} = \begin{pmatrix} Y_p \\ \bar{U}_p \end{pmatrix} \quad , \quad W_p^{u\bar{u}} = \begin{pmatrix} U_p \\ \bar{U}_p \end{pmatrix} \tag{4.12}$$

and similar to L_{YR} and L_{UR} we have

$$\begin{aligned} L_{Y\bar{U}} &= (I + L_u L_y^{c2})^{-1} L_u L_y^{c2} \\ &= L_u (I + L_y^{c2} L_u)^{-1} L_y^{c2} \end{aligned} \tag{4.13}$$

$$L_{U\bar{U}} = (I + L_y^{c2} L_u)^{-1} L_y^{c2} \tag{4.14}$$

L_y^{c2} contains Markov parameters of the controller C^2 and X_f^{c2} is the subspace matrix for the future states of the regulatory controller C^2. L_{YE}^{cs} and L_{UE}^{cs} represent the closed-loop relation from the input noise e_t to the output y_t and C^2 control action u_t, respectively. These notations will be explained in detail later in Section 5.

Using Equations (4.8) and (4.10), the finite-horizon LQG objective function can be written as follows:

$$\begin{aligned} J =& \; E\{[y_f^T y_f] + \lambda [u_f^T u_f]\} \\ \\ =& (\Gamma_N^{y2} x_f^{cc2} + L_{Y\bar{U}} \bar{u}_f)^T (\Gamma_N^{y2} x_f^{cc2} + L_{Y\bar{U}} \bar{u}_f) \\ & + \lambda (\Gamma_N^{u2} x_f^{cc2} + L_{U\bar{U}} \bar{u}_f)^T (\Gamma_N^{u2} x_f^{cc2} + L_{U\bar{U}} \bar{u}_f) \\ \\ =& (x_f^{cc2})^T (\Gamma_N^{y2})^T \Gamma_N^{y2} x_f^{cc2} + \bar{u}_f^T L_{Y\bar{U}}^T \Gamma_N^{y2} x_f^{cc2} \\ & + (x_f^{cc2})^T (\Gamma_N^{y2})^T L_{Y\bar{U}} \bar{u}_f + \bar{u}_f^T L_{Y\bar{U}}^T L_{Y\bar{U}} \bar{u}_f \\ & + \lambda (x_f^{cc2})^T (\Gamma_N^{u2})^T \Gamma_N^{u2} x_f^{cc2} + \lambda \bar{u}_f L_{U\bar{U}}^T \Gamma_N^{u2} x_f^{cc2} \\ & + \lambda (x_f^{cc2})^T (\Gamma_N^{u2})^T L_{U\bar{U}} \bar{u}_f + \lambda \bar{u}_f^T L_{U\bar{U}}^T L_{U\bar{U}} \bar{u}_f \end{aligned} \tag{4.15}$$

Taking derivative of (4.15) with respect to $\bar{u}_f$ gives

$$\begin{aligned}\frac{\partial J}{\partial \bar{u}_f} =& 2(x_f^{cc2})^T(\Gamma_N^{y2})^T L_{Y\bar{U}} + 2\bar{u}_f^T L_{Y\bar{U}}^T L_{Y\bar{U}} + \\ &+ 2\lambda(x_f^{cc2})^T(\Gamma_N^{u2})^T L_{U\bar{U}} + 2\lambda\bar{u}_f L_{U\bar{U}}^T L_{U\bar{U}} = 0\end{aligned} \tag{4.16}$$

which results in

$$\bar{u}_f = -[L_{Y\bar{U}}^T L_{Y\bar{U}} + \lambda L_{U\bar{U}}^T L_{U\bar{U}}]^{-1}(L_{Y\bar{U}}^T \Gamma_N^{y2} + L_{U\bar{U}}^T \Gamma_N^{u2}) x_f^{cc2} \tag{4.17}$$

The state-feedback control in Equation (4.17) can also be written in terms of past outputs and control actions using Equations (4.9) and (4.11) as follows:

$$\bar{u}_f = -[L_{Y\bar{U}}^T L_{Y\bar{U}} + \lambda L_{U\bar{U}}^T L_{U\bar{U}}]^{-1}(L_{Y\bar{U}}^T L_{Y2} w_p^{y\bar{u}} + L_{U\bar{U}}^T L_{U2} w_p^{u\bar{u}}) \tag{4.18}$$

When implementing the controller, at each sampling interval, only the first control move is applied to the process and calculation procedure is repeated for the next sampling interval.

Scenario 2

In this case, the measured (feedback) variable of the regulatory controller C^2 (to be designed) is different from its controlled variable, so the problem configuration would be similar to the LQG design presented at the beginning of this section. Figure 4.6 shows the reconfiguration of Figure 4.4 and definition of the measured variable, ξ. State space presentation of the process G_T in Figure 4.6 is obtained in the following. The block $\bar{C}^1$ in the figure can be presented by the following state-space model:

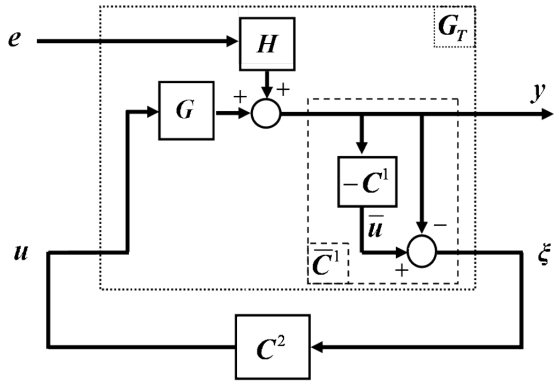


Figure 4.6: Reconfiguration of Figure 4.4 for Scenario 2.

$$\bar{C}^1 : \begin{cases} x_{t+1}^{c1} & = A_{c1}x_t^{c1} + B_{c1}y_t \\ \xi_t & = -C_{c1}x_t^{c1} - (D_{c1} + I_m)y_t \end{cases} \tag{4.19}$$

where $x_t^{c1} \in \mathbb{R}^{n_{c1}}$, $\xi_t \in \mathbb{R}^m$ and $\bar{u}_t \in \mathbb{R}^m$. Substituting y_t from Equation (3.1) in the above equation followed by some matrix manipulations results in the following presentation for G_T:

$$G_T : \begin{cases} \begin{pmatrix} x_{t+1} \\ x_{t+1}^{c1} \end{pmatrix} = & \begin{pmatrix} A & 0 \\ B_{c1} & A_{c1} \end{pmatrix} \begin{pmatrix} x_t \\ x_t^{c1} \end{pmatrix} + \begin{pmatrix} K \\ B_{c1} \end{pmatrix} e_t + \begin{pmatrix} B \\ B_{c1}D \end{pmatrix} u_t \\ y_t = & -\begin{pmatrix} (D_{c1}+I_m)C & C_{c1} \end{pmatrix} \begin{pmatrix} x_t \\ x_t^{c1} \end{pmatrix} - (D_{c1}+I_m)e_t \\ & -(D_{c1}+I_m)Du_t \\ \xi_t = & -\begin{pmatrix} C & 0 \end{pmatrix} \begin{pmatrix} x_t \\ x_t^{c1} \end{pmatrix} + e_t + Du_t \end{cases} \tag{4.20}$$

Using Equation (4.20), both controlled and measured variables, y_t and ξ_t, can be presented in terms of the control action, u_t, and the input noise by the following subspace equations:

$$Y_f = \Gamma_N^{y1} X_f^{cc1} + L_{yU} U_f + L_{yE} E_f \tag{4.21}$$

$$\Xi_f = \Gamma_N^{\xi} X_f^{cc1} + L_{\xi U} U_f + L_{\xi E} E_f \tag{4.22}$$

$$= L_{\Xi} W_p^{\xi u} + L_{\xi U} U_f + L_{\xi E} E_f \tag{4.23}$$

where $X_f^{cc1} = \begin{pmatrix} X_f \\ X_f^{c1} \end{pmatrix}$ and $W_p^{\xi u} = \begin{pmatrix} \Xi_p \\ U_p \end{pmatrix}$. Other subspace matrices can be easily defined using Equation (4.20), similar to the definitions of Section 4.3 for open-loop system.

Now, the two steps of the LQG control design can be performed in the subspace framework. First, the state vector x_f^{cc1} can be estimated from (4.22) as

$$\hat{x}_f^{cc1} = [\Gamma_N^{\xi}]^{-1}(\xi_f - L_{\xi U} u_f)$$

and then controlled variable can be estimated using the estimated state as follows:

$$\begin{aligned} \hat{y}_f &= \Gamma_N^{y1} \hat{x}_f^{cc1} + L_{yU} u_f \\ &= \Gamma_N^{y1} [\Gamma_N^{\xi}]^{-1} \xi_f + (L_{yU} - \Gamma_N^{y1} [\Gamma_N^{\xi}]^{-1} L_{\xi U}) u_f \end{aligned} \tag{4.24}$$

For simplicity in notation, we define

$$\begin{aligned} \Gamma_{y\xi} &\triangleq \Gamma_N^{y1} [\Gamma_N^{\xi}]^{-1} \\ \Gamma_{yu} &\triangleq L_{yU} - \Gamma_N^{y1} [\Gamma_N^{\xi}]^{-1} L_{\xi U} \end{aligned}$$

Now, the LQG objective function can be expressed as

$$\begin{aligned} J =& \hat{y}_f^T \hat{y}_f + \lambda u_f^T u_f \\ =& (\Gamma_{y\xi}\xi_f + \Gamma_{yu} u_f)^T (\Gamma_{y\xi}\xi_f + \Gamma_{yu} u_f) + \lambda u_f^T u_f \\ =& \xi_f^T \Gamma_{y\xi}^T \Gamma_{y\xi} \xi_f + u_f^T \Gamma_{yu}^T \Gamma_{y\xi} \xi_f + \xi_f^T \Gamma_{y\xi}^T \Gamma_{yu} u_f + u_f^T (\Gamma_{yu}^T \Gamma_{yu} + \lambda I) u_f \end{aligned} \tag{4.25}$$

Taking derivative of J with respect to u_f and substituting ξ_f from (4.22) yields

$$\begin{aligned}\frac{\partial J}{\partial u_f} =&2L_{\xi U}^T\Gamma_{y\xi}^T\Gamma_{y\xi}\xi_f + 2\Gamma_{yu}^T\Gamma_{y\xi}\xi_f + 2(\Gamma_{yu}^T\Gamma_{yu} + \lambda I)u_f \\ =&2(L_{\xi U}^T\Gamma_{y\xi}^T\Gamma_{y\xi} + \Gamma_{yu}^T\Gamma_{y\xi})\xi_f + 2(\Gamma_{yu}^T\Gamma_{yu} + \lambda I)u_f \\ =&2(L_{\xi U}^T\Gamma_{y\xi}^T\Gamma_{y\xi} + \Gamma_{yu}^T\Gamma_{y\xi})(\Gamma_N^\xi\hat{x}_f^{cc1} + L_{\xi U}u_f) + 2(\Gamma_{yu}^T\Gamma_{yu} + \lambda I)u_f \\ =&0 \end{aligned} \tag{4.26}$$

which results in the following state-feedback control law:

$$u_f = [L_{\xi U}^T\Gamma_{y\xi}\Gamma_{y\xi}^T L_{\xi U} + \Gamma_{yu}^T\Gamma_{y\xi}L_{\xi U} + \Gamma_{yu}^T\Gamma_{yu} + \lambda I]^{-1}(L_{\xi U}^T\Gamma_{y\xi}\Gamma_{y\xi}^T + L_{\xi U}^T\Gamma_{y\xi})\Gamma_N^\xi\hat{x}_f^{cc1} \tag{4.27}$$

Similar to *Scenario* 1, this control action can also be presented in term of past inputs and measured outputs using (4.23).

Scenario 3

In this case, the objective is to design both controllers C^1 and C^2 at the same time to obtain the minimum input and output variances. A reconfiguration of Figure 4.4 is shown in Figure 4.7 where block C is the combination of two sub-controllers. Clearly, controller C is the same as the conventional LQG feedback controller.

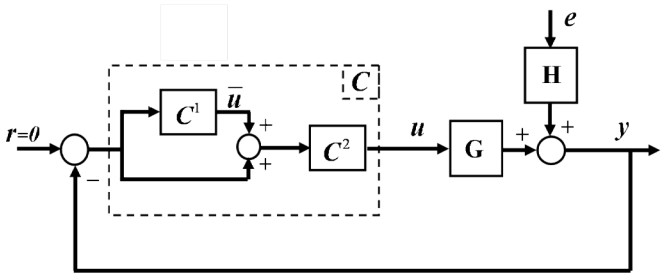


Figure 4.7: Reconfiguration of Figure 4.4 for Scenario 3.

The LQG design for control block C in this figure has been developed in [32, 33, 53] as

$$\begin{aligned} u_f &= -[L_u^T L_u + \lambda I]^{-1}L_u^T\Gamma_N\hat{x}_f \\ &= -[L_u^T L_u + \lambda I]^{-1}L_u^T L_w w_p \end{aligned} \tag{4.28}$$

which provides the global optimum solution to the finite-horizon LQG problem. It should be noted that none of the control designs in first two scenarios can result in a better performance comparing to (4.28). However, there might be many combinations of controllers C^1 and C^2 that can provide the above control action. As an

example, C^2 can be chosen as a simple unit gain and C^1 be given by

$$u_f = y_f - [L_u^T L_u + \lambda I]^{-1} L_u^T L_w w_p \tag{4.29}$$

In other words, different controller sets may result in the optimal variance for the output and the regulatory control action. Therefore, no unique solution can be provided for C^1 and C^2 at the same time. However, the optimal values of the input and output variances are unique and can be used for the purpose of performance assessment.

It should be noted that we assumed the same sampling interval for both supervisory and regulatory controllers. Extension of this study for the case of different sampling times is not considered in this thesis.

In the next section, we show how to drive the optimal input and output variances in terms of the input noise variance which is the key step to obtain the LQG trade-off curve for cascade supervisory-regulatory control.

4.3.3 Estimation of the LQG benchmark from closed-loop data

In this section, the approach of [43] is extended to the calculation of input and output variances under the LQG control for each scenario of the cascade supervisory-regulatory control structure. For better understanding of the derivation, we shall start with the last scenario.

Scenario 3: the global optimum

To derive closed-loop variance, we follow the approach of Ko and Edgar (2000) [60]. In the closed-loop process of Figure 4.4, assume that at time $t = 0$ a single random noise e_0 enters the process when the system is at the steady state. Then the following propagation of e_0 results:

$$\begin{pmatrix} y_0 \\ y_1 \\ \cdots \\ y_{N-1} \end{pmatrix} = \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}B & CA^{N-3}B & \cdots & D \end{pmatrix} \begin{pmatrix} u_0 \\ u_1 \\ \cdots \\ u_{N-1} \end{pmatrix} + \begin{pmatrix} I \\ CK \\ \cdots \\ CA^{N-2}K \end{pmatrix} e_0 \tag{4.30}$$

or in a compact form

$$y_{0|N-1} = L_u u_{0|N-1} + L_{e,1} e_0 \tag{4.31}$$

and similarly for $u_{0|N-1}$

$$u_{0|N-1} = L_{U\bar{U}} \bar{u}_{0|N-1} + L_{UE,1}^{cs} e_0 \tag{4.32}$$

where $L_{U\bar{U}}$ is given by (4.14) and $L_{e,1}$ is the first block column of L_e. $L^{cs}_{UE,1}$ is the first block column of L^{cs}_{UE}. L^{cs}_{UE} is the subspace matrix representing the relation of input noise to the regulatory control action and its detailed expression will be given in Section 4.

Combining Equations (4.31) and (4.32), one can get

$$y_{0|N-1} = L_u L_{U\bar{U}} \bar{u}_{0|N-1} + L_u L^{cs}_{UE,1} e_0 + L_{e,1} e_0 \tag{4.33}$$

Recall the finite-horizon LQG control objective function

$$J = E(\sum_{i=0}^{N-1} y_i^T y_i + \lambda u_i^T u_i) \tag{4.34}$$

This objective function can be written as

$$\begin{aligned} J =& E(y^T_{0|N-1} y_{0|N-1} + \lambda u^T_{0|N-1} u_{0|N-1}) \\ =& E\{[L_u L_{U\bar{U}} \bar{u}_{0|N-1} + L_u L^{cs}_{UE,1} e_0 + L_{e,1} e_0]^T [L_u L_{U\bar{U}} \bar{u}_{0|N-1} + L_u L^{cs}_{UE,1} e_0 + L_{e,1} e_0] \\ &+ \lambda (L_{U\bar{U}} \bar{u}_{0|N-1} + L^{cs}_{UE,1} e_0)^T (L_{U\bar{U}} \bar{u}_{0|N-1} + L^{cs}_{UE,1} e_0)\} \end{aligned} \tag{4.35}$$

Expanding Equation (4.35), omitting the subscript $(0|N-1)$ and taking derivative with respective to $\bar{u}$ yields

$$\begin{aligned} \frac{\partial J}{\partial \bar{u}} =& 2L^T_{U\bar{U}} L^T_u L_u L_{U\bar{U}} \bar{u} + 2L^T_{U\bar{U}} L^T_u L_u L^{cs}_{UE,1} e_0 \\ &+ 2L^T_{U\bar{U}} L^T_u L_{e,1} e_0 + 2\lambda L^T_{U\bar{U}} L_{U\bar{U}} \bar{u} + 2\lambda L^T_{U\bar{U}} L^{cs}_{UE,1} e_0 = 0 \end{aligned}$$

which gives

$$\begin{aligned} \bar{u}^{opt}_{0|N-1} =& -[L^T_{U\bar{U}} L^T_u L_u L_{U\bar{U}} + \lambda L^T_{U\bar{U}} L_{U\bar{U}}]^{-1} \{L^T_{U\bar{U}} L^T_u (L_u L^{cs}_{UE,1} + L_{e,1}) + \lambda L^T_{U\bar{U}} L^{cs}_{UE,1}\} e_0 \\ =& -[L^T_{U\bar{U}} (L^T_u L_u + \lambda I) L_{U\bar{U}}]^{-1} \{L^T_{U\bar{U}} (L^T_u [L_u L^{cs}_{UE,1} + L_{e,1}] + \lambda L^{cs}_{UE,1})\} e_0 \\ =& -L^{-1}_{U\bar{U}} [L^T_u L_u + \lambda I]^{-1} \underbrace{[L^T_{U\bar{U}}]^{-1} (L^T_{U\bar{U}})}_{I} \{L^T_u [L_u L_{UE,1} + L_{e,1}] + \lambda L^{cs}_{UE,1}\} e_0 \\ =& -L^{-1}_{U\bar{U}} [L^T_u L_u + \lambda I]^{-1} \{L^T_u [L_u L^{cs}_{UE,1} + L_{e,1}] + \lambda L^{cs}_{UE,1}\} e_0 \\ =& -L^{-1}_{U\bar{U}} \{[L^T_u L_u + \lambda I]^{-1} L^T_u L_{e,1}\} e_0 - L^{-1}_{U\bar{U}} \{\underbrace{[L^T_u L_u + \lambda I]^{-1} (L^T_u L_u + \lambda I)}_{I} L^{cs}_{UE,1}\} e_0 \\ =& -L^{-1}_{U\bar{U}} \{[L^T_u L_u + \lambda I]^{-1} L^T_u L_{e,1} + L^{cs}_{UE,1}\} e_0 \end{aligned} \tag{4.36}$$

Using Equations (4.32) and (4.36), $u^{opt}_{0|N-1}$ can be obtained by

$$\begin{aligned} u^{opt}_{0|N-1} &= L_{U\bar{U}}\bar{u}^{opt}_{0|N-1} + L^{cs}_{UE,1}e_0 \\ &= -\{[L_u^T L_u + \lambda I]^{-1} L_u^T L_{e,1}\}e_0 \end{aligned} \tag{4.37}$$

Now, Equation (4.31), gives $y^{opt}_{0|N-1}$ as

$$y^{opt}_{0|N-1} = \{I - L_u[L_u^T L_u + \lambda I]^{-1} L_u^T\} L_{e,1} e_0 \tag{4.38}$$

Similar to the approach of [43], one may write $u^{opt}_{0|N-1}$ as a vector form

$$\begin{pmatrix} \psi_0 \\ \psi_1 \\ \vdots \\ \psi_{N-1} \end{pmatrix} = -[L_u^T L_u + \lambda I]^{-1} L_u^T L_{e,1} \tag{4.39}$$

and $y^{opt}_{0|N-1}$ as a vector form

$$\begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{N-1} \end{pmatrix} = \{I - L_u[L_u^T L_u + \lambda I]^{-1} L_u^T\} L_{e,1} \tag{4.40}$$

When the random noises occur at every sampling instant, we can apply the principle of superposition to get the optimal sequence of control inputs as

$$\begin{aligned} u_0^{opt} &= \psi_0 e_0 \\ u_1^{opt} &= \psi_0 e_1 + \psi_1 e_0 \\ u_2^{opt} &= \psi_0 e_2 + \psi_1 e_1 + \psi_2 e_0 \\ &\vdots \\ u_{N-1}^{opt} &= \psi_0 e_{N-1} + \psi_1 e_{N-2} + \cdots + \psi_{N-1} e_0 \\ u_N^{opt} &= \psi_0 e_N + \psi_1 e_{N-1} + \cdots + \psi_{N-1} e_0 \\ &\vdots \\ u_t^{opt} &= \sum_{i=0}^{N-1} \psi_i e_{t-i} \end{aligned}$$

and, similarly for y_t^{opt} we get

$$y_t^{opt} = \sum_{i=0}^{N-1} \gamma_i e_{t-i}$$

Output variance and control actions variances can now be written as

$$Var[u_t] = \sum_{i=0}^{N-1} \psi_i \;\; Var[e_t] \;\; \psi_i^T \tag{4.41}$$

$$Var[y_t] = \sum_{i=0}^{N-1} \gamma_i \;\; Var[e_t] \;\; \gamma_i^T \tag{4.42}$$

Remark. The optimal variance of the regulatory control action and the output in Equations (4.41) and (4.42) are consistent with the previous results of Kadali and Huang [43, 53].

Scenario 1

A procedure similar to the past section can be performed starting from the following equations, assuming a single random noise e_0 has entered the process at time $t = 0$:

$$u_{0|N-1} = L_y^{c2}(\bar{u}_{0|N-1} - y_{0|N-1}) + L_{UE,1}^{cs} e_0 \tag{4.43}$$

$$y_{0|N-1} = L_u u_{0|N-1} + L_{e,1} e_0 \tag{4.44}$$

Substituting (4.43) in (4.44), one can calculate $y_{0|N-1}$ as

$$y_{0|N-1} = [I + L_u L_y^{c2}]^{-1}\{L_u L_y^{c2} \bar{u}_{0|N-1} + L_u L_{UE,1} e_0 + L_{e,1} e_0\} \tag{4.45}$$

Similarly for $u_{0|N-1}$

$$u_{0|N-1} = [I + L_y^{c2} L_u]^{-1}\{L_y^{c2} \bar{u}_{0|N-1} + L_y^{c2} L_{e,1} e_0 + L_{UE,1} e_0\} \tag{4.46}$$

It is shown in the Appendix A that

$$[I + L_u L_y^{c2}]^{-1} = [I + L_y^{c2} L_u]^{-1} \triangleq \Pi \tag{4.47}$$

The LQG objective can now be written as

$$\begin{aligned} J =& E(y_{0|N-1}^T y_{0|N-1} + \lambda u_{0|N-1}^T u_{0|N-1}) \\ =& E\{(L_u L_y^{c2} \bar{u}_{0|N-1} + L_u L_{UE,1} e_0 + L_{e,1} e_0)^T \Pi^T \Pi (L_u L_y^{c2} \bar{u}_{0|N-1} + L_u L_{UE,1} e_0 + L_{e,1} e_0) \\ &+ \lambda (L_y^{c2} \bar{u}_{0|N-1} + L_y^{c2} L_{e,1} e_0 + L_{UE,1} e_0) \Pi^T \; \Pi (L_y^{c2} \bar{u}_{0|N-1} + L_y^{c2} L_{e,1} e_0 + L_{UE,1} e_0)\} \end{aligned} \tag{4.48}$$

Taking derivative of the above objective function, solving for $\bar{u}_{0|N-1}$ and some simplifications results in

$$\bar{u}_{0|N-1}^{opt} = -\,[\Pi L_y^{c2}]^{-1}[L_u^T L_u + \lambda I]^{-1}\{(L_u^T L_u + \lambda I)\Pi L_{UE,1} + (L_u^T + L_y^{c2})\Pi L_{e,1}\} e_0 \tag{4.49}$$

which gives the $u^{opt}_{0|N-1}$ as

$$\begin{aligned} u^{opt}_{0|N-1} =& (\Pi L^{c2}_y)\bar{u}^{opt}_{0|N-1} + \Pi L_{UE,1}e_0 + \Pi L^{c2}_y L_{e,1}e_0 \\ =& -\{[L^T_u L_u + \lambda I]^{-1}(L^T_u + L^{c2}_y)\Pi - \Pi L^{c2}_y\}L_{e,1}e_0 \end{aligned} \tag{4.50}$$

This control action results in the following output:

$$y^{opt}_{0|N-1} = \{I - L_u[L^T_u L_u + \lambda I]^{-1}(L^T_u + L^{c2}_y)\Pi + L_u \Pi L^{c2}_y\}L_{e,1}e_0 \tag{4.51}$$

Now, the input and output variances can be obtained in a similar way as the derivations of Equations (4.41) and (4.42).

Scenario 2

Starting from the following equations, the expressions for $u^{opt}_{0|N-1}$ and $y^{opt}_{0|N-1}$ in this case can be obtained using the same procedure as in the past two subsections:

$$\begin{aligned} \xi_{0|N-1} &= L^{\bar{c}1}_y L_u u_{0|N-1} + L^{\bar{c}1}_y L_{e,1}e_0 \\ \bar{u}_{0|N-1} &= -L^{c1}_y y_{0|N-1} + L_{\bar{U}E,1}e_0 \\ y_{0|N-1} &= \bar{u}_{0|N-1} - \xi_{0|N-1} \end{aligned} \tag{4.52}$$

where $L^{\bar{c}1}_y$ contains the Markov parameters of $\bar{C}^1$ controller presented in (4.19) and $L_{\bar{U}E,1}$ is the first block column of $L_{\bar{U}E}$ which will be explained in detail by (4.71) in Section 5.

The final results for the optimal input and output are as follows:

$$u^{opt}_{0|N-1} = \{L_u[L^T_u L_u + \lambda I]^{-1}L^T_u - [L^{\bar{c}1}_y]^{-1}L^{c1}_y[I + L_u L^{\bar{c}_1}_y]^{-1}\}L_{e,1}e_0 \tag{4.53}$$

$$\begin{aligned} y^{opt}_{0|N-1} =& \{(I - L_u[L^T_u L_u + \lambda I]^{-1}L^T_u) \\ & - L_u[L^T_u L_u + \lambda I]^{-1}L^T_u[L^{\bar{c}1}_y]^{-1}L^{c1}_y[I + L_u L^{\bar{c}_1}_y]^{-1}\}L_{e,1}e_0 \end{aligned} \tag{4.54}$$

4.3.4 Interpretation of the trade-off curves in cascade supervisory-regulatory structure

Based on the results of the past subsection, one trade-off curve can be obtained for each of the three scenarios defined in Section 4.2. The one calculated for scenario 3 represents the best achievable performance which is the conventional feedback LQG performance. Similar to Section 3.3, the following performance indices can be defined to evaluate the current control system performance relative to the optimal performance:

$$\eta^3_y \triangleq \frac{(V^{opt}_y)}{(V_y)} \qquad , \qquad \eta^3_u \triangleq \frac{(V^{opt}_u)}{(V_u)}$$

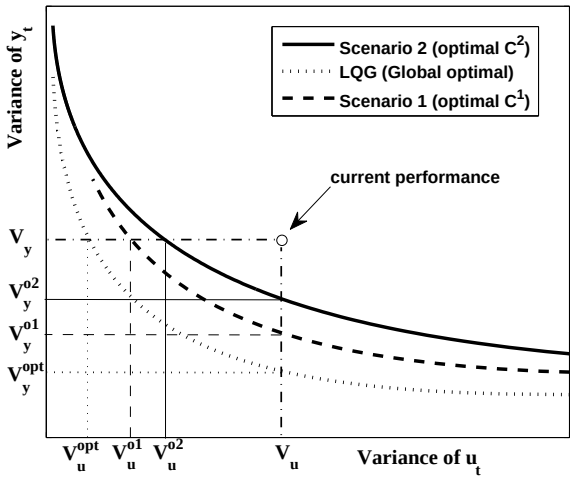


Figure 4.8: Typical LQG trade-off curves for a *cascade* structure.

where V_y^{opt}, V_y, V_u^{opt} and V_u are defined similarly in Figure 4.8.

The trade-off curve obtained for scenario 1 can be used to calculate the possible performance improvement by supervisory controller (C^1) without tuning the existing regulatory controller (C^2). Similar to the above definitions, η^1 and E^1 can be defined for this case by

$$\eta_y^1 \triangleq \frac{(V_y^{o1})}{(V_y)} \qquad , \qquad \eta_u^1 \triangleq \frac{(V_u^{o1})}{(V_u)}$$

Similarly, the trade-off curve of scenario 2 indicates the improvement in performance which can be obtained by re-tuning/re-designing the regulatory controller without changing supervisory control. η^2 and E^2 indices can be defined and used for quantification of the performance as

$$\eta_y^2 \triangleq \frac{(V_y^{o2})}{(V_y)} \qquad , \qquad \eta_u^2 \triangleq \frac{(V_u^{o2})}{(V_u)}$$

Also, comparing these curves with each other and with the LQG curve provides us with the insight to decide which control layer(s) to be used for performance enhancement. This issue will be elaborated more by two illustrative examples later.

4.4 Identification of the required subspace matrices from closed-loop data

Based on the results of Section 3, certain subspace matrices and the noise variance are required to calculate the input and output variances and obtain the LQG trade-off curves. In the following, we present a closed-loop subspace identification method

for estimation of the required subspace matrices and the input noise variance. This method is an extension to the joint input-output identification approach presented in Chapter 3. The problem is divided into three separate open-loop identification problems, one from the setpoint, r_t to the output, y_t, one from r_t, to u_t and the other one from r_t to $\bar{u}_t$. The following represents the details of the algorithm.

For the cascade structure of Figure 4.4, three basic subspace equations for the process, controller C^1 and controller C^2 can be derived as follows, respectively (see Equation(2.9)):

$$Y_f = \Gamma_N X_f + L_u U_f + L_e E_f \tag{4.55}$$

$$\bar{U}_f = \Gamma_N^{c1} X_f^{c1} + L_y^{c1}(R_f - Y_f) \tag{4.56}$$

$$U_f = \Gamma_N^{c2} X_f^{c2} + L_y^{c2}(\bar{U}_f - Y_f) \tag{4.57}$$

X_f^{c1} and X_f^{c2} are the future state matrices for the controllers C^1 and C^2, and Γ_N^{c1} and Γ_N^{c2} are the related extended observability matrices. L_y^{c1} and L_y^{c2} contain the Markov parameters of the controllers C^1 and C^2.

Substituting Equation (4.57) in (4.55) yields

$$Y_f = \Gamma_N X_f + L_u \Gamma_N^{c2} X_f^{c2} + L_u L_y^{c2} \bar{U}_f - L_u L_y^{c2} Y_f + L_e E_f$$

Using $\bar{U}_f$ from Equation (4.56), we get

$$\begin{aligned} Y_f =& \Gamma_N X_f + L_u \Gamma_N^{c2} X_f^{c2} + L_u L_y^{c2} \Gamma_N^{c1} X_f^{c1} + L_u L_y^{c2} L_y^{c1} R_f - L_u L_y^{c2} L_y^{c1} Y_f - L_u L_y^{c2} Y_f + L_e E_f \\ =& \begin{pmatrix} \Gamma_N & L_u L_y^{c2} \Gamma_N^{c1} & L_u \Gamma_N^{c2} \end{pmatrix} \begin{pmatrix} X_f \\ X_f^{c1} \\ X_f^{c2} \end{pmatrix} + L_u L_y^{c2} L_y^{c1} R_f - L_u (L_y^{c2} L_y^{c1} + L_y^{c2}) Y_f + L_e E_f \end{aligned}$$

which gives Y_f as

$$Y_f = [I + L_u L_y^{cc}]^{-1} (\Gamma_N^{cs} X_f^{cs} + L_u L_y^{c2} L_y^{c1} R_f + L_e E_f) \tag{4.58}$$

where

$$\begin{aligned} \Gamma_N^{cs} &= \begin{pmatrix} \Gamma_N & L_u L_y^{c2} \Gamma_N^{c1} & L_u \Gamma_N^{c2} \end{pmatrix} \\ L_y^{cc} &= L_y^{c2} L_y^{c1} + L_y^{c2} \\ X_f^{cs} &= \begin{pmatrix} X_f \\ X_f^{c1} \\ X_f^{c2} \end{pmatrix} \end{aligned} \tag{4.59}$$

In a compact form

$$Y_f = \Gamma_N^y X_f^{cs} + L_{YR}^{cs} R_f + L_{YE}^{cs} E_f \tag{4.60}$$

where

$$\Gamma_N^y = [I + L_u L_y^{cc}]^{-1}\Gamma_N^{cs}$$
$$L_{YR}^{cs} = [I + L_u L_y^{cc}]^{-1} L_u L_y^{c2} L_y^{c1} \tag{4.61}$$
$$L_{YE}^{cs} = [I + L_u L_y^{cc}]^{-1} L_e \tag{4.62}$$

Using the regression analysis method [57], the first term in (4.60) can be represented based on the past setpoints and outputs which yields

$$Y_f = L_Y^{cs} W_p^{yr} + L_{YR}^{cs} R_f + L_{YE}^{cs} E_f \tag{4.63}$$

in which W_p^{yr} is given by (4.7).

Substituting Equations (4.55) and (4.57) in (4.56) and following the similar procedure as (4.60), one can write U_f as

$$\begin{aligned} U_f =& (-L_y^{c2}\Gamma_N \;\; L_y^{c2}\Gamma_N^{c1} \;\; \Gamma_N^{c2}) X_f^{cs} + L_y^{c2} L_y^{c1} R_f - (L_y^{c2} L_y^{c1} + L_y^{c2}) L_u U_f - (L_y^{c2} L_y^{c1} + L_y^{c2}) L_e E_f \\ =& \Gamma_N^u X_f^{cs} + L_{UR}^{cs} R_f + L_{UE}^{cs} E_f \end{aligned} \tag{4.64}$$

where

$$\Gamma_N^u = [I + L_y^{cc} L_u]^{-1}(-L_y^{c2}\Gamma_N \;\; L_y^{c2}\Gamma_N^{c1} \;\; \Gamma_N^{c2})$$
$$L_{UR}^{cs} = [I + L_y^{cc} L_u]^{-1} L_y^{c2} L_y^{c1} \tag{4.65}$$
$$L_{UE}^{cs} = -[I + L_y^{cc} L_u]^{-1} L_y^{cc} L_e \tag{4.66}$$

Representing the future state term by past data Hankle matrices U_p and R_p using linear regression, we have the following relation for U_f:

$$U_f = L_U^{cs} W_p^{ur} + L_{UR}^{cs} R_f + L_{UE}^{cs} E_f \tag{4.67}$$

where W_p^{ur} is given by (4.7).

Similarly, for $\bar{U}_f$ the following expression can be obtained:

$$\bar{U}_f = L_{\bar{U}}^{cs} W_p^{\bar{u}r} + L_{\bar{U}R}^{cs} R_f + L_{\bar{U}E}^{cs} E_f \tag{4.68}$$

where

$$W_p^{\bar{u}r} = \begin{pmatrix} \bar{U}_p \\ R_p \end{pmatrix}$$
$$L_{\bar{U}R}^{cs} = L_y^{c1}\{I - [I + L_u L_y^{cc}]^{-1} L_u L_y^{c2} L_y^{c1}\} \tag{4.69}$$
$$= \{I - L_y^{c1}[I + L_u L_y^{cc}]^{-1} L_u L_y^{c2}\} L_y^{c1} \tag{4.70}$$

$$L_{\bar{U}E}^{cs} = -L_y^{c1}[I + L_u L_y^{cc}]^{-1} L_e \tag{4.71}$$

R_f can be chosen as a random binary signal uncorrelated with W_p^{yr}, W_p^{ur}, $W_p^{\bar{u}r}$ and E_f (past inputs and outputs and future noise), and then least squares estimation can be used to estimate closed-loop subspace matrices in (4.60) , (4.67) and (4.68) as follows:

$$(\hat{L}_Y^{cs} \quad \hat{L}_{YR}^{cs}) = Y_f \begin{pmatrix} W_p^{yr} \\ R_f \end{pmatrix}^{\dagger} \tag{4.72}$$

$$(\hat{L}_U^{cs} \quad \hat{L}_{UR}^{cs}) = U_f \begin{pmatrix} W_p^{ur} \\ R_f \end{pmatrix}^{\dagger} \tag{4.73}$$

$$(\hat{L}_{\bar{U}}^{cs} \quad \hat{L}_{\bar{U}R}^{cs}) = \bar{U}_f \begin{pmatrix} W_p^{\bar{u}r} \\ R_f \end{pmatrix}^{\dagger} \tag{4.74}$$

The first row of $\hat{Y}_f$ is the one step ahead prediction of the output. So the innovation sequence can be estimated by

$$\begin{aligned} \hat{e}_f &= (e_N \quad e_{N+1} \quad \cdots \quad e_{N+j-1})^T \\ &= Y_f(1:m,:) - \hat{Y}_f(1:m,:) \end{aligned} \tag{4.75}$$

where $(1:m,:)$ represents first m rows and all columns of the matrix. Using $\hat{e}_f$, the block Hankel matrix for noise, $\hat{E}_f$, can be built. Now define

$$\Xi_f^u \triangleq U_f - \hat{U}_f = L_{UE}^{cs} \hat{E}_f$$

So, for L_{UE}^{cs} we have

$$\hat{L}_{UE}^{cs} = \Xi_f^u \hat{E}_f^{\dagger} \tag{4.76}$$

Similarly, one may define

$$\Xi_f^y \triangleq Y_f - \hat{Y}_f = L_{YE}^{cs} \hat{E}_f$$

which gives $\hat{L}_{YE}^{cs}$ as follows:

$$\hat{L}_{YE}^{cs} = \Xi_f^y \hat{E}_f^{\dagger} \tag{4.77}$$

$L_{\bar{U}E}^{cs}$ can also be estimated in a similar way.

Now, open-loop subspace matrices of the process and controllers should be estimated using the identified closed-loop matrices. Based on Equations (4.61) and (4.65), the following derivation can be provided using the matrix inverse lemma:

$$\begin{aligned} L_{YR}^{cs}[L_{UR}^{cs}]^{-1} =& [I + L_u L_y^{cc}]^{-1} L_u (L_y^{c2} L_y^{c1})[L_y^{c2} L_y^{c1}]^{-1} (I + L_y^{cc} L_u) \\ =& L_u [I + L_y^{cc} L_u]^{-1} (I + L_y^{cc} L_u) \\ =& L_u \end{aligned} \tag{4.78}$$

which can be used to estimate L_u. To estimate L_e, we have to first estimate L_y^{cc} from closed-loop matrices by

$$\begin{aligned} -L_{UE}^{cs}[L_{YE}^{cs}]^{-1} &= [I + L_y^{cc} L_u]^{-1} L_y^{cc} L_e [L_e]^{-1} (I + L_u L_y^{cc}) \\ &= L_y^{cc} [I + L_u L_y^{cc}]^{-1} (I + L_u L_y^{cc}) \\ &= L_y^{cc} \end{aligned} \tag{4.79}$$

Substituting (4.79) in (4.62) gives

$$\begin{aligned} L_{YE}^{cs} &= [I + L_u L_y^{cc}]^{-1} L_e \\ &= [I - L_u L_{UE}^{cs} [L_{YE}^{cs}]^{-1}]^{-1} L_e \end{aligned} \tag{4.80}$$

Finally,

$$\begin{aligned} L_e &= \{I - L_u L_{UE}^{cs} [L_{YE}^{cs}]^{-1}\} L_{YE}^{cs} \\ &= L_{YE}^{cs} - L_u L_{UE}^{cs} \end{aligned} \tag{4.81}$$

The subspace matrices representing the two controllers should also be estimated. Note that $L_{\bar{U}E}^{cs}$ can be identified in a similar way to the estimation of L_{UE}^{cs}. It is straightforward to show that

$$L_y^{c1} = -L_{\bar{U}E}^{cs} [L_{YE}^{cs}]^{-1} \tag{4.82}$$

Using Equations (4.59) and (4.82) we get

$$\begin{aligned} L_y^{cc} &= L_y^{c2} + L_y^{c2} L_y^{c1} = L_y^{c2} (I + L_y^{c1}) \\ &= L_y^{c2} \{I - L_{\bar{U}E}^{cs} [L_{YE}^{cs}]^{-1}\} \end{aligned} \tag{4.83}$$

Substituting L_y^{cc} from (4.79) in (4.83) yields

$$-L_{UE}^{cs} [L_{YE}^{cs}]^{-1} = L_y^{c2} \{I - L_{\bar{U}E}^{cs} [L_{YE}^{cs}]^{-1}\}$$

which gives the following estimation of L_y^{c2}:

$$\begin{aligned} L_y^{c2} &= -L_{UE}^{cs} [L_{YE}^{cs}]^{-1} \{I - L_{\bar{U}E}^{cs} [L_{YE}^{cs}]^{-1}\}^{-1} \\ &= -L_{UE}^{cs} \left[\{I - L_{\bar{U}E}^{cs} [L_{YE}^{cs}]^{-1}\} L_{YE}^{cs} \right]^{-1} \\ &= -L_{UE}^{cs} [L_{YE}^{cs} - L_{\bar{U}E}^{cs}]^{-1} \end{aligned} \tag{4.84}$$

Remark. Note that the proposed closed-loop identification procedure consists of three separate open-loop identifications. Consistency analysis of the regression analysis approach is provided in [57] for open-loop identification and in Chapter 3 for joint input-output closed-loop identification. Consistency analysis similar to Chapter 3 can be performed here to show that for a sufficiently large value of N, asymptotic consistency is obtained in the estimation of the subspace matrices and the input noise variance.

4.5 Simulation

In the first example, two linear controllers are used as C^1 and C^2 in a cascade structure to control a SISO process. This allows us to evaluate the proposed closed-loop identification method. Next example provides more realistic simulation where C^1 is replaced with a linear MIMO LQG controller running in MATLAB® and C^2 is PID control running in HYSYS.

4.5.1 Example 1

The process to be controlled is described by the following state-space representation [92]:

$$x_{t+1} = \begin{pmatrix} 0.6 & 0.6 & 0 \\ -0.6 & 0.6 & 0 \\ 0 & 0 & 0.7 \end{pmatrix} x_t + \begin{pmatrix} 1.616 \\ -0.348 \\ 2.631 \end{pmatrix} u_t + \begin{pmatrix} -1.15 \\ -1.52 \\ -3.20 \end{pmatrix} e_t$$
$$y_t = \begin{pmatrix} -0.437 & -0.504 & 0.093 \end{pmatrix} x_t - 0.775 u_t + e_t$$

The low-level controller is a PI controller, [0.1+0.05/s], and the high-level controller is given by

$$x_{t+1} = \begin{pmatrix} 1 & 0 \\ 1 & 0.01 \end{pmatrix} x_t + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u_t$$
$$y_t = \begin{pmatrix} 0.02 & 0.01 \end{pmatrix} x_t + 0.2 u_t$$

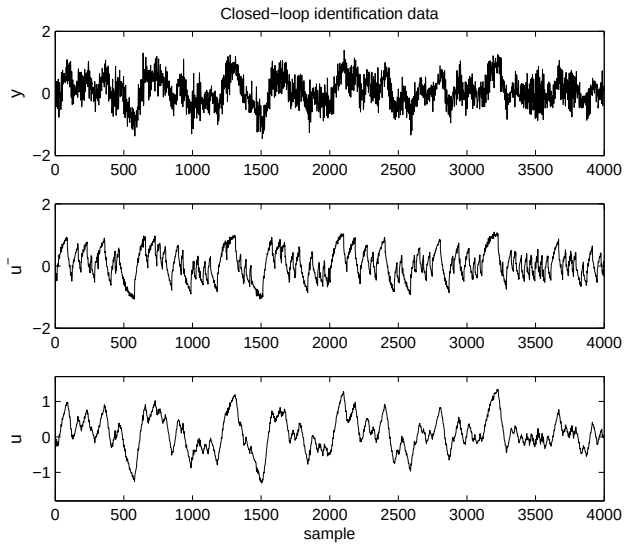


Figure 4.9: Identification data for Example 1 collected under cascade control.

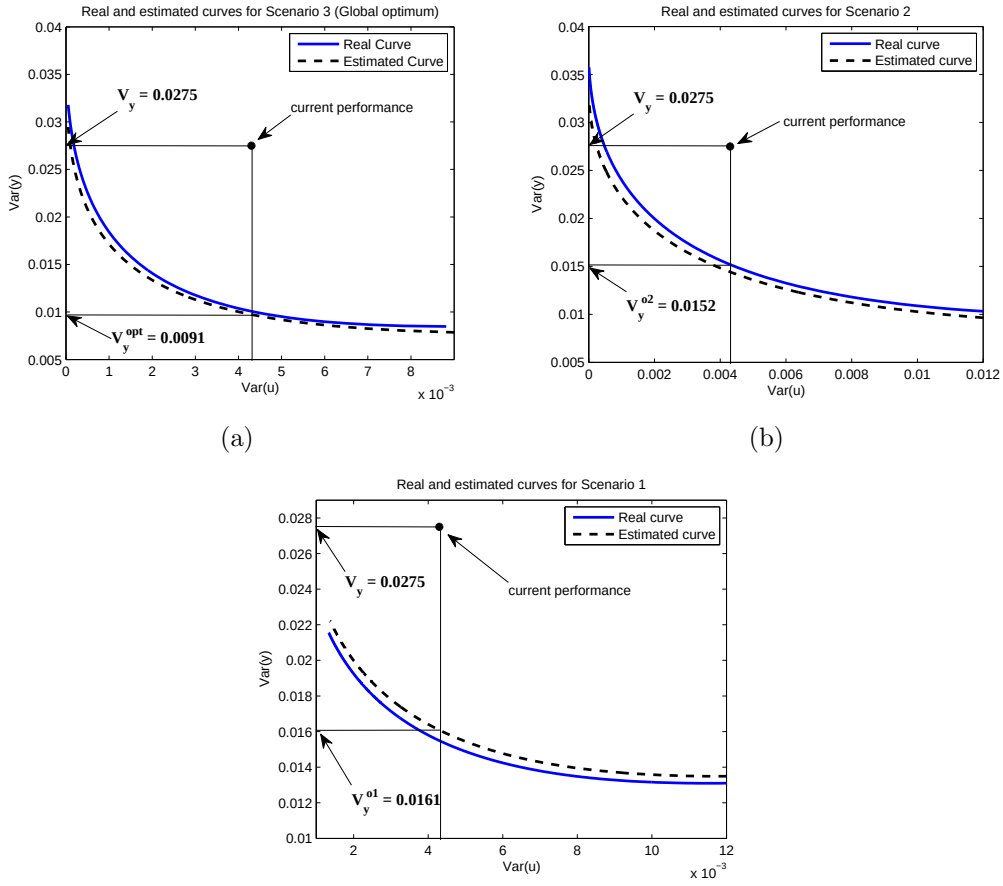


Figure 4.10: Real and estimated LQG trade-off curves for *Scenario 3* (a), *Scenario 2* (b) and *Scenario 1* (c).

The first step in this example is to evaluate the presented closed-loop identification procedure. For this purpose, 4000 points of identification data are collected from the closed-loop system with a RBS test signal, generated by MATLAB® function '*idinput*', injected through the system setpoint. A sequence of white noise disturbance with variance 0.01 is applied to the process. Simulation data are plotted in Figure 4.9. Using the proposed joint input-output identification method for the cascade structure, subspace matrices are estimated and used to obtain three trade-off curves shown in Figure 4.10. In the figures, the true trade-off curves, obtained from the original model, are also shown for the sake of comparison. These figures verify the proposed closed-loop subspace identification method.

A set of routine closed-loop operating data is also collected to obtain the current performance of the control system.

Figures 4.10 presents three possibilities for improving the performance. If one can re-design or re-tune both of the regulatory and supervisory controllers, the curve in Figure 4.10(a) would be the limit of achievable performance. For instance,

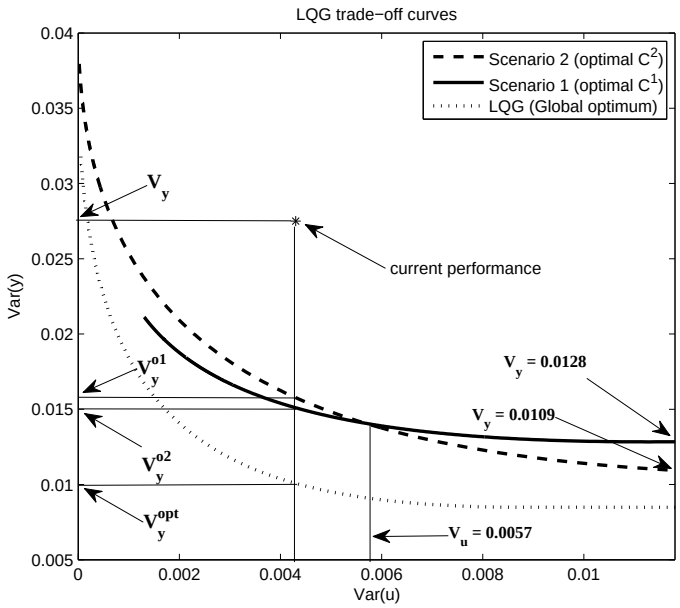


Figure 4.11: LQG trade-off curves and actual performance point.

this curve shows the possibility of decreasing the output variance by 67% without increasing the control effort relative to the existing controller.

If only changing or re-tuning of the supervisory controller is possible, then the estimated curve in Figure 4.10(b) shows how much improvement may be achieved. For the same input variance, re-tuning only the supervisory controller can result in $(1 - 0.0152/0.0275) \times 100 = 45\%$ reduction in the output variations. A similar statement can be proposed for the regulatory controller according to Figure 4.10(c) which shows that regulatory controller re-tuning may lead to $(1 - 0.0161/0.0275) \times 100 = 42\%$ variance reduction in the output.

Figure 4.11 shows all three trade-off curves and the current performance point. This figure provides more insight for comparing different strategies of performance enhancement. It shows that for this example, the minimum achievable output variance by increasing the C^2 control effort is $(1 - 0.0109/\ 0.0275) \times 100 = 61\%$ where as the one from manipulating C^1 is $(1 - 0.0128/\ 0.0275) \times 100 = 54\%$. Also, it is evident that if control action variations more than $V_u = 0.0057$ is acceptable, then regulatory controller (C^2) re-tuning can be more effective to reduce the output variance.

In the next example we provide a more realistic simulation using a MIMO process simulated in HYSYS and controlled by a cascade LQG-PID structure.

4.5.2 Example 2

The process used in this simulation is a plug-flow reactor (PFR) which is used in many industrial processes particularly those in which a solid catalyst is required. This reactor has a vessel packed with solid catalyst. In a tubular reactor, temperature and composition vary along the length of the reactor which makes the process dynamics complicated. A schematic of this process is shown in Figure 4.12 along with the HYSYS folwsheet. The reaction considered is the chlorination of propylene. The reaction rate and operating data are given in [72]. There are two parallel gas-phase reactions. The first forms allyl chloride and HCl as follows:

$$C_3H_6 + Cl_2 \rightarrow CH = CH_2Cl + HCl \tag{4.85}$$

The second forms 1,2 dichloro propane

$$C_3H_6 + Cl_2 \rightarrow CH_2Cl - CHCl + HC_3 \tag{4.86}$$

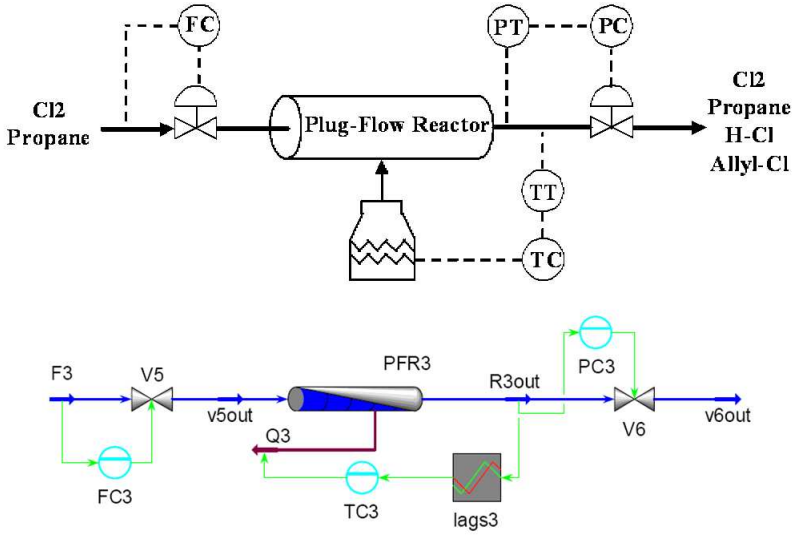


Figure 4.12: A schematic of Plug-Flow Reactor and Hysys flowsheet.

The reaction rate is a first order dependent on the partial pressures of the reactants. The reaction takes place in a pipe which is 2 inches in diameter and 15 feet in length. The inlet gas fed to the reactor is 0.85 $lb - mol/hr$ at 392 $^\circ F$ and 29.4 $psia$ with a composition 80 $mol\%$ propylene and 20 $mol\%$ chlorine. Pressure drop through the reactor is 3.7 psi at design conditions. If the reactor is operated adiabatically, the temperature of the gas leaving the reactor is predicted by HYSYS as 272.4 $^\circ C$ and the chlorine concentration is 9.89 $mol\%$.

A control valve on the gas feeding the reactor is designed for a 20 psi drop when 50 % open at design flowrate. A flow controller manipulates this valve to control

feed flow. Feed flowrate is considered as a disturbance variable. Outlet pressure is controlled by valve $V6$ and reactor temperature is controlled by coolant water flow rate.

Two PI controllers have been designed for temperature and pressure control (see Figure 4.12), and tuned based on a relay-feedback test [72]. Proportional gain and integral time are set to be 5 and 40 for TC3 and 1 and 2 for PC3, respectively. Temperature should be maintained at its setpoint of 272.4 $^{\circ}C$ and pressure at 177.2 $Kpas$.

The first step to design a LQG is to identify a linear model of the lower-level closed-loop process (including current PID controllers). The MOESP identification method from MATLAB® system identification toolbox is used to identify a model for the LQG controller. Weighting factors are chosen 15 for both control actions. LQG algorithm is implemented in a MATLAB® code and MATLAB® is connected to HYSYS in a real-time manner to apply LQG control. For this purpose, the '*hysyslib*' toolbox [12] has been used after some required modifications for the current version of HYSYS. Therefore, HYSYS is used for process simulation and PID control, and MATLAB® is used for running the LQG algorithm.

Now, for the purpose of performance assessment, a set of data should be collected from the process under LQG-PID control to estimate the LQG trade-off curves. For this purpose, two RBS test signals are generated in MATLAB® and are applied to the setpoints of the LQG controller which manipulates the HYSYS PID controllers. LQG controller outputs (sepoints to the PID level), low-level PI control actions and process outputs under the test are shown in Figure 4.13. During the test, a white noise signal with valiance of 0.03 is applied to FC3 setpoint which is treated as a disturbance to the process. Note that FC3 controller is used in this simulation only to apply changes to the feed flowrate as a disturbance.

Using the proposed joint input-output identification method, the required subspace matrices are identified and trade-off curves for the three scenarios are obtained and shown in Figure 4.14. A set of routine operating data has also been collected which is used to determine the performance of the current cascade LQG-PID control system as shown in the figure by the '*' symbol. Based on the results, $(1 - 0.7/1.39) \times 100 = 75\%$ decrease in the output variance is possible with the same control effort if LQG control is used, while re-tuning only the LQG controller may provide up to $(1 - 0.7/1.39) \times 100 = 50\%$ improvement and re-tuning PID controllers may provide $(1 - 0.59/1.39) \times 100 = 48\%$ improvement. The curves also show that at the current working point, increasing the control effort does not have much effect on reducing the output variance.

To verify that the optimal performance as suggested by the LQG trade-off curves can be achieved, the two optimal designed controllers for scenario 1 and 2 and the

LQG control are also applied on the process using the identified subspace matrices. The resulting performances are shown in the figure by 'o', '×' and '+' symbols, respectively. These points are very close to the optimal curve although do not exactly lie on the corresponding trade-off curves.

4.6 Concluding remarks

In this chapter, we investigated the problem of subspace LQG design and performance assessment for control systems with a supervisory-regulatory structure usually resulted from applying advanced controllers on the regulatory control systems. This type of cascade control implementation was briefly reviewed in this chapter and compared to the *Direct* method of implementation. LQG control was employed as the benchmark for performance assessment in this study. Three possible LQG designs in a cascade control structure were studied. For each case, we proposed the controller design and provided the expression of input and output variances which led us to obtain the LQG trade-off curves. As a result, three trade-off curves could be obtained which provided three possibilities for control performance improvement depending on which controller (supervisory, regulatory or both) is chosen to be redesigned or re-tuned. The derivations of the LQG and constructions of the trade off curves were provided in the subspace framework. It was shown that the trade-off curves can be obtained from certain subspace matrices without the need of an explicit model. A closed-loop subspace identification method was provided based on the joint input-output identification approach to estimate the required subspace matrices from closed-loop data. A simulation on a SISO example in MATLAB® was used to elaborate the proposed method of performance assessment and verify the proposed closed-loop identification method. Another simulation study using MATLAB® and HYSYS was also performed on a multivariate process which provided more realistic application results.

4.7 Appendix A

Lemma. The following relation holds for L_u and L_y^{c2}:

$$[I + L_u L_y^{c2}]^{-1} = [I + L_y^{c2} L_u]^{-1}$$

Proof. We show that the above relation holds for any two matrices which have the same structure as L_u. Consider the following 4×4 generally defined matrices:

$$M_1 = \begin{pmatrix} a & 0 & 0 & 0 \\ b & a & 0 & 0 \\ c & b & a & 0 \\ d & c & b & a \end{pmatrix} \qquad M_2 = \begin{pmatrix} e & 0 & 0 & 0 \\ f & e & 0 & 0 \\ g & f & e & 0 \\ h & g & f & e \end{pmatrix}$$

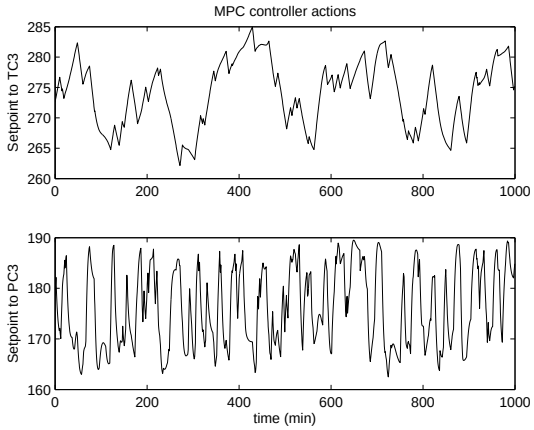
Now, simple matrix multiplication shows

$$\begin{aligned} M_1 M_2 &= M_2 M_1 \\ &= \begin{pmatrix} ae & 0 & 0 & 0 \\ be + af & ae & 0 & 0 \\ ec + fb + ag & be + af & ae & 0 \\ de + fc + bg + ah & ec + fb + ag & be + af & ae \end{pmatrix} \end{aligned}$$

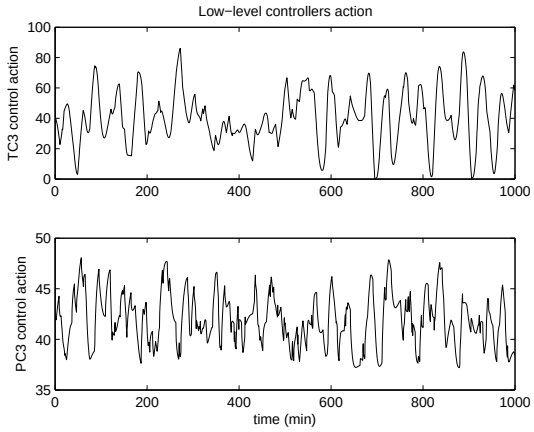
Therefore, we have

$$[I + M_1 M_2]^{-1} = [I + M_2 M_1]^{-1}$$

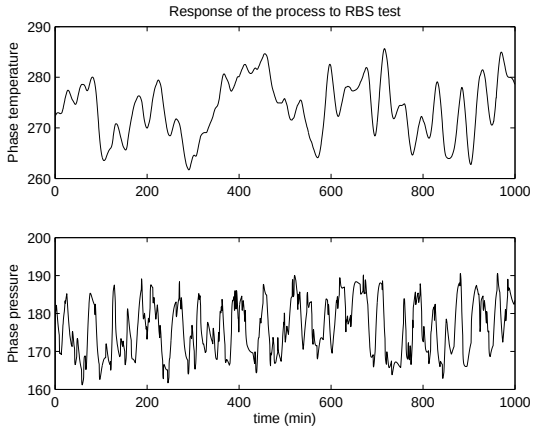
This result can be easily extended for any other size of the matrices.



(a)



(b)



(c)

Figure 4.13: Supervisory and regulatory control actions and process outputs under RBS test.

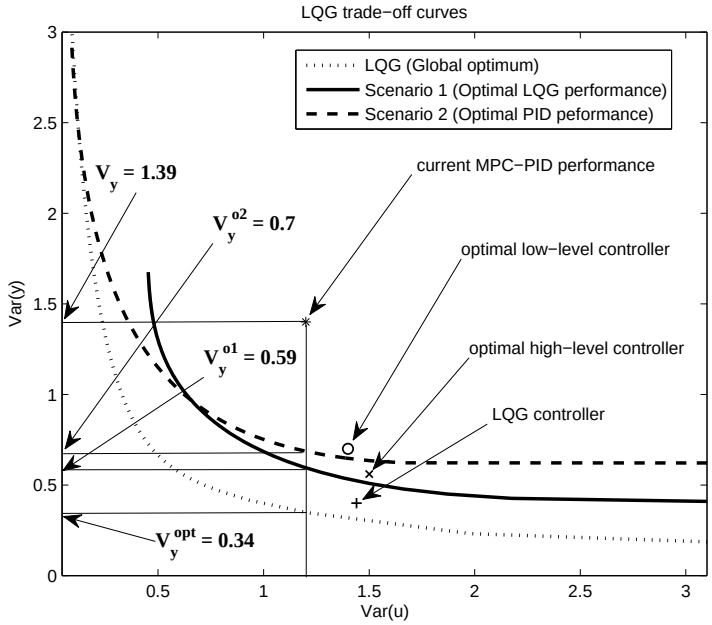


Figure 4.14: LQG trade-off curves and actual performance point.

Chapter 5

Subspace approach to identification of step response model from closed-loop data

5.1 Introduction

Motivations: Most system identification methods, particularly the subspace identification method, estimate state space models or equivalent parametric models. Practical process control applications, such as tuning of PID or design of MPC, however require nonparametric models particularly the step responses. It is often necessary to convert the parametric models to the nonparametric ones. Practical experiences have shown such a conversion can often result in unexpected results including even a wrong sign of the process gain. It is well known that one of the subspace matrices obtained from the intermediate step of subspace identification contains process impulse response coefficients. Can we directly estimate the step response model from this matrix? Can this direct extraction of the step response models be more reliable than the conventional approach? With these questions in mind, this chapter explores a practical solution to estimation of step response models from closed-loop data directly.

Subspace identification approach has found its applications not only for process modeling but also in other areas of control engineering such as predictive control [29, 54, 102] and controller performance assessment [43, 53]. In many recent novel applications of subspace identification methods, the complete procedure of subspace identification is not required, but only the first step which is the estimation of the intermediate subspace matrices is needed. In other words, these applications only need the process impulse response coefficients embedded in these subspace matrices. This approach is often called as "model-free" approach in the literature.

In this regard, direct estimation of the process step responses from closed-loop data can be considered as a useful application of subspace identification in practice.

This application has a close relation to the subspace predictive control. The results are specially useful for MPC controller design or model validation form closed-loop data.

The step response model of the process can be easily obtained from the estimated impulse response coefficients by integration. However, more than one set of the impulse response coefficients are contained in the estimated subspace matrices. Therefore, it is important to make the best use of all the estimated parameters for the step response calculation. A proper estimation of the step response can be obtained through weighted average of estimated impulse response coefficients if their variances are available.

The particular subspace matrix of interest in this study, which is denoted by L_u, contains a series of impulse response coefficients in a Toeplitz structure

$$L_u = \begin{pmatrix} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ CA^2B & CAB & CB & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ CA^{N-2}B & CA^{N-3}B & CA^{N-4}B & \cdots & D \end{pmatrix} \tag{5.1}$$

The process step responses can be obtained by integration over these impulse response coefficients. Normally, because of the noise effect, the estimated $\hat{L}_u$ does not have the above Toeplitz structure. Each column consists of a finite number of impulse response coefficients and the matrix provides N different sets of the impulse response coefficients with decreasing lengths. In fact, the matrix contains N estimates for the first impulse response coefficient, $N-1$ estimates for the second coefficient and so on. The question which motivated this study is how to properly use all the coefficients provided in this matrix.

A simple answer is to take average over the diagonals for each coefficient. In this paper, we show that there is a more efficient approach to solve this problem. This approach requires the element-by-element variance calculation of the above subspace matrix. We shall discuss this issue with details in the following sections.

The asymptotic properties of different subspace identification methods have been studied in the past decade. Open-loop identification methods have received more attention in this regard [5, 6, 7, 8, 9, 11, 17, 19, 23, 35, 49, 51, 81]. The asymptotic distribution for the MOESP-type methods has been established in Bauer and Jansson (2000) [11] and extended to more general cases in Jansson (2000) [49]. Discussion on the influence of past and future horizons as well as the weighting matrices on the asymptotic variance can be found in Gustafsson (1999) [35], Jansson (1997) [48], Bauer and Jansson (2000) [11] and Chiuso and Picci (2003) [17]. The Asymptotic analyses for state approaches such as Canonical Variate Analysis (CVA) [63, 64]

and CCA [78] were presented by Bauer (1998, 1999, 2000) [5, 9, 10], Buaer and Ljung (2002) [7] and Chiuso (2007) [16]. Similar discussion on N4SID can be found in Bauer (1998) [5] and Chiuso and Picci (2004) [19]. A detailed review on the asymptotic properties of the open-loop methods is provided in Bauer (2005) [8].

Some authors have provided the results of statistical analysis on the closed-loop subspace identification. In particular, Chiuso and Picci have explored this issue in a series of papers for different closed-loop methods [15, 16, 18, 20, 22, 23, 29] including the SSARX method by Jansson (2003) [50], the innovation estimation method of Qin and Ljung (2003) [83] and the Whitening filter or predictor-based approach [20].

The main focus in most of the above mentioned studies is on the asymptotic properties of the estimated system matrices. For this purpose, the framework of stochastic systems and stochastic realization theory have been employed extensively. In this paper, we are concerned with the variance of the estimated impulse response coefficients for practical applications, so we will avoid using the complex mathematical notations. Furthermore, the closed-loop method of interest in our work is the joint input-output identification method of Chapter 3 which divides the closed-loop problem into two open-loop identification problems and makes the analysis more straightforward.

It should be noted that the previous studies do not provide the element-by-element variance for the subspace or system matrices, as provided in this paper. For this purpose, we present a sequential version of the joint input-output identification method based on the idea of enforced casual modeling [81].

This chapter is organized as follows: in Section 2, a brief review on the joint input-output subspace identification of Chapter 3 is provided. Section 3 presents the main results the joint input-output method and the variance calculation for the impulse response coefficients. Section 4 provides the results of two Monte-Carlo simulations and two experimental application of the proposed method. Concluding remarks are given in Section 5.

5.2 A review of the joint input-output identification

This section provides a review of the joint input-output closed-loop identification method of Chapter 3 for which a modified version will be presented in the next section.

This method provides a direct estimation of process and disturbance Markov parameters in subspace matrices L_u and L_e. The following is a brief review on the method. In the closed-loop system of Figure 5.1, two separate open-loop models can be defined: a model from setpoint (r_t) to the output (y_t) and the other one from

(r_t) to the controller output (u_t). Similar to Equation (2.19), these two systems can be presented by the following input-output relations:

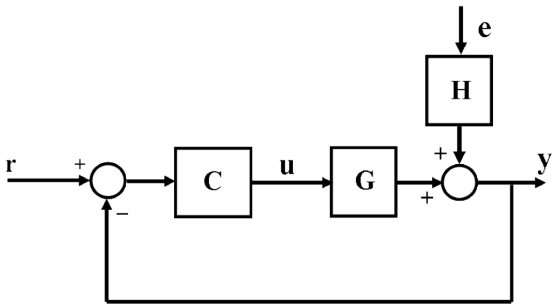


Figure 5.1: A typical closed-loop process.

$$Y_f = L_Y W_p^{yr} + L_{YR} R_f + L_{YE} E_f \tag{5.2}$$

$$U_f = L_U W_p^{ur} + L_{UR} R_f + L_{UE} E_f \tag{5.3}$$

where:

$$W_p^{yr} = \begin{pmatrix} Y_p \\ R_p \end{pmatrix} \quad , \quad W_p^{ur} = \begin{pmatrix} U_p \\ R_p \end{pmatrix} \tag{5.4}$$

Based on this method, two lease squares estimation should be performed

$$(\hat{L}_U \quad \hat{L}_{UR}) = U_f \begin{pmatrix} W_p^U \\ R_f \end{pmatrix}^{\dagger} \tag{5.5}$$

$$(\hat{L}_Y \quad \hat{L}_{YR}) = Y_f \begin{pmatrix} W_p^Y \\ R_f \end{pmatrix}^{\dagger} \tag{5.6}$$

$\hat{U}_f$ and $\hat{Y}_f$ are found by the orthogonal projection of the row space of U_f and Y_f. Then, the estimation of the innovation sequence is performed as follows:

$$\hat{e}_f = Y_f(1:m,:) - \hat{Y}_f(1:m,:) \tag{5.7}$$

After constructing $\hat{E}_f$ by the use of $\hat{e}_f$, L_{UE} can be estimated as:

$$\hat{L}_{UE} = (U_f - \hat{U}_f)\hat{E}_f^{\dagger} \tag{5.8}$$

Finally, $\hat{L}_u$ and $\hat{L}_e$ are estimated by

$$\hat{L}_u = \hat{L}_{YR}(\hat{L}_{UR})^{-1} \tag{5.9}$$

$$\hat{L}_e = -(\hat{L}_{UR})^{-1}\hat{L}_{UE} \tag{5.10}$$

In the next section, we present a modified version of this joint input-output identification method that breaks the problem into N separate least squares for which the variance evaluation is provided.

5.3 A modified method and variance calculation

The two least squares in (5.5) and (5.6) are actually being used to solve two open-loop identification problems, so the variance of estimations can be obtained for the elements of $\hat{L}_{UR}$ and $\hat{L}_{YR}$ (to be discussed shortly). However, because of the matrix inversion and matrix product in equation (5.9), the element-by-element variance evaluation for $\hat{L}_u$ is too complex for practical use.

The purpose of this study is to provide a proper method to use the estimated impulse response coefficients, to reduce the influence of the noise. Because of the above mentioned complication in the variance evaluation of $\hat{L}_u$, we take an alternative approach to solve the problem. Assuming the variance of estimation for all elements in $\hat{L}_{UR}$ and $\hat{L}_{YR}$ is available, one may first refine these matrices by proper variance-based weighted averaging and then use (5.9) to calculate $\hat{L}_u$. In this way, $\hat{L}_{UR}$ and $\hat{L}_{YR}$ have the Toeplitz structure and the final $\hat{L}_u$ estimated from Equation (5.9) will also have the Toeplitz structure. Therefore, the variance calculation for $\hat{L}_u$ is no longer required. Note that this approach is in fact more effective, since it attenuates the noise effect before it propagates to $\hat{L}_u$ through Equation (5.9).

For this purpose, we employ the idea of casual subspace model identification of [81] to present a sequential version of the above joint input-output identification method for which the element-by-element variance evaluation can be performed.

5.3.1 Sequential joint input-output identification with enforced casual modeling

As explained, the problem of interest narrows down to variance calculation for all the elements in $\hat{L}_{YR}$ and $\hat{L}_{UR}$. For this purpose, we break Equations (5.2) and (5.3) into N row equations to estimate each row of L_{YR} and L_{UR} by a separate least squares. Taking this approach not only makes it possible to calculate element-by-element variance of $\hat{L}_{YR}$ and $\hat{L}_{UR}$, but also provides the possibility to remove the non-casual input terms from the estimation. These terms are conveniently included in the second term of the right hand side of Equations (5.2) and (5.3) for performing subspace projections. The coefficients of these non-casual terms should be zeros theoretically. However, their presence increases the number of model parameters to be estimated and thus results in higher variance. This idea is elaborated in the open-loop subspace identification approach with enforced causal models by Qin *et. al.* (2005) [81]. The authors showed that enforcing casual models decreases the variance in the estimation of subspace matrices. However, they did not provide the variance for each element in the subspace matrices.

In the following, we provide the element-by-element variance for $\hat{L}_{YR}$ and $\hat{L}_{UR}$ which are then used to improve the step response estimation. This analysis can also

provide the variance of estimation for $\hat{L}_Y$ and $\hat{L}_U$. Since the variance expression for $\hat{L}_{YR}$ and $\hat{L}_{UR}$ follow the same procedure, we provide the detailed analysis only for L_{YR}. First, we consider the case of single-input single-output process. Based on the subspace definitions given in Section 2, Equation (5.6) can be expressed as

$$\begin{pmatrix} y_N & \cdots & y_{N+j-1} \\ y_{N+1} & \cdots & y_{N+j} \\ \vdots & \vdots & \vdots \\ y_{2N-1} & \cdots & y_{2N+j-2} \end{pmatrix} = \begin{pmatrix} l_{1,0} & \cdots & l_{1,N-1} & l_{1,N} & \cdots & l_{1,2N-1} \\ l_{2,0} & \cdots & l_{2,N-1} & l_{2,N} & \cdots & l_{2,2N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{N,0} & \cdots & l_{N,N-1} & l_{N,N} & \cdots & l_{N,2N-1} \end{pmatrix} \begin{pmatrix} y_0 & \cdots & y_{j-1} \\ \vdots & \vdots & \vdots \\ y_{N-1} & \cdots & y_{N+j-2} \\ u_0 & \cdots & u_{j-1} \\ \vdots & \vdots & \vdots \\ u_{N-1} & \cdots & u_{N+j-2} \end{pmatrix} +$$

$$+ \begin{pmatrix} h_1^{(1)} & 0 & \cdots & 0 \\ h_2^{(1)} & h_1^{(2)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ h_N^{(1)} & h_{N-1}^{(2)} & \cdots & h_1^{(N)} \end{pmatrix} \begin{pmatrix} u_N & \cdots & u_{N+j-1} \\ u_{N+1} & \cdots & u_{N+j} \\ \vdots & \vdots & \vdots \\ u_{2N-1} & \cdots & u_{2N+j-2} \end{pmatrix} + \begin{pmatrix} g_1^{(1)} & 0 & \cdots & 0 \\ g_2^{(1)} & g_1^{(2)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ g_N^{(1)} & g_{N-1}^{(2)} & \cdots & g_1^{(N)} \end{pmatrix} \begin{pmatrix} e_N & \cdots & e_{N+j-1} \\ e_{N+1} & \cdots & e_{N+j} \\ \vdots & \vdots & \vdots \\ e_{2N-1} & \cdots & e_{2N+j-2} \end{pmatrix} \tag{5.11}$$

where $h_i^{(j)} \in \mathbb{R}^{1\times 1}$ $(i, j = 1, 2, \cdots, N)$, represents the j^{th} estimate of the i^{th} impulse response coefficient in L_{YR}. Our purpose is to find the variance of estimation for $\hat{h}_i^{(j)}$. Using a similar notation, the elements of L_{YE} are denoted by $g_i^{(j)} \in \mathbb{R}^{1\times 1}$. Note that $g_1^{(i)} = 1$.

The first row of Equation 5.11 can be re-arranged into the following standard regression form:

$$\begin{pmatrix} y_N \\ y_{N+1} \\ \vdots \\ y_{N+j-1} \end{pmatrix} = \begin{pmatrix} y_0 & \cdots & y_{N-1} & u_0 & \cdots & u_{N-1} & u_N \\ y_1 & \cdots & y_N & u_1 & \cdots & u_N & u_{N+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{j-1} & \cdots & y_{N+j-2} & u_{j-1} & \cdots & u_{N+j-2} & u_{N+j-1} \end{pmatrix} \begin{pmatrix} l_{1,0} \\ \vdots \\ l_{1,N-1} \\ l_{1,N} \\ \vdots \\ l_{1,2N-1} \\ h_1^{(1)} \end{pmatrix} + \begin{pmatrix} e_N \\ e_{N+1} \\ \vdots \\ e_{N+j-1} \end{pmatrix} \tag{5.12}$$

Using a compact format yields

$$Y_1 = X_1\theta_1 + \epsilon_1 \tag{5.13}$$

where the definitions of $Y_1 \in \mathbb{R}^{j-1,1}$, $X_1 \in \mathbb{R}^{j-1,2N+1}$, $\theta_1 \in \mathbb{R}^{2N+1,1}$ and $\epsilon_1 \in \mathbb{R}^{j-1,1}$ can be simply inferred by comparing (5.12) and (5.13).

Equation (5.12) represents a linear regression problem with $j-1$ equations and $2N+1$ parameters to be estimated. The variance of estimation for $\hat{\theta}_1$ is given by [69]

$$\mathcal{V}_1 \triangleq Cov(\hat{\theta}_1) = \sigma_{\epsilon_1}^2 (X_1^T X_1)^{-1} \tag{5.14}$$

where $\sigma^2_{\epsilon_1}$ is the white noise variance and is estimated by

$$\sigma^2_{\epsilon_1} = \frac{\sum_{i=1}^{j-1}(y_i - \hat{y}_i)^2}{(j-1)-(2N+1)} \tag{5.15}$$

The last element in $Cov(\hat{\theta}_1)$ provides the variance for the first estimation of the first impulse response coefficient, $\hat{h}_1^{(1)}$, as

$$\Lambda_1^{(1)} = \mathcal{V}_1(2mN+1:2mN+m,\ 2lN+1:2lN+l) \tag{5.16}$$

where MATLAB® notation is used in (5.16).

Now, consider the second row in Equation 5.11. This equation can also be re-arranged in the form of a linear regression as in Equation (5.17):

$$\begin{pmatrix} y_{N+1} \\ y_{N+2} \\ \vdots \\ \vdots \\ y_{N+j} \end{pmatrix} = \begin{pmatrix} y_0 & \cdots & y_{N-1} & r_0 & \cdots & r_{N-1} & r_N & r_{N+1} \\ y_1 & \cdots & y_N & r_1 & \cdots & r_N & r_{N+1} & r_{N+2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{j-1} & \cdots & y_{N+j-2} & r_{j-1} & \cdots & r_{N+j-2} & r_{N+j-1} & r_{N+j} \end{pmatrix} \begin{pmatrix} l_{2,0} \\ \vdots \\ l_{2,N-1} \\ l_{2,N} \\ \vdots \\ l_{2,2N-1} \\ h_2^{(1)} \\ h_1^{(2)} \end{pmatrix}$$
$$+ \begin{pmatrix} g_2^{(1)} e_N + e_{N+1} \\ g_2^{(1)} e_{N+1} + e_{N+2} \\ \vdots \\ \vdots \\ g_2^{(1)} e_{N+j-1} + e_{N+j} \end{pmatrix} \tag{5.17}$$

Similar to (5.13), the compact form of Equation (5.17) can be derived as

$$Y_2 = X_2\theta_2 + \epsilon_2 \tag{5.18}$$

where $Y_2 \in \mathbb{R}^{j-1,1}$, $X_2 \in \mathbb{R}^{j-1,2N+2}$, $\theta_2 \in \mathbb{R}^{2N+2,1}$ and $\epsilon_2 \in \mathbb{R}^{j-1,1}$.

Equation (5.18) is a linear regression with $j-1$ equations and $2N+2$ parameters and the last two parameters, $h_2^{(1)}$ (the second estimate of the first coefficient) and $h_1^{(2)}$ (the first estimate of the second coefficient), are of interest here. However, the residual of regression (5.17), ϵ_2, is not white. Therefore, the variance of estimation must be calculated by [87]

$$\mathcal{V}_2 \triangleq Cov(\hat{\theta}_2) = (X_2^T X_2)^{-1} X_2^T R_{\epsilon_2} X_2 (X_2^T X_2)^{-1} \tag{5.19}$$

where $R_{\epsilon_2} \in \mathbb{R}^{j-1,j-1}$ is the covariance matrix of the non-white noise. The variance of estimation for $\hat{h}_2^{(1)}$ and $\hat{h}_1^{(2)}$ is given by

$$\Lambda_2^{(1)} = \mathcal{V}_2(2mN+1:2mN+m,\ 2lN+1:2lN+l) \tag{5.20}$$

$$\Lambda_1^{(2)} = \mathcal{V}_2(2mN+m+1:2mN+2m, 2lN+l+1:2lN+2l) \tag{5.21}$$

Note that calculating R_{ϵ_2} requires the knowledge of $g_2^{(1)}$ from the noise subspace matrix L_{YE}. This issue will be discussed in Remark 1. A simple approximation of R_{ϵ_2} can be obtained directly from the residual of the estimation. The residual ϵ_2 is estimated by

$$\hat{\epsilon}_2 = Y_2 - X_2\hat{\theta}_2 \tag{5.22}$$

which can be used to estimate $\hat{R}_{\epsilon_2}$.

Following the same procedure for the rest of the rows in Equation (5.11) results in another $N-2$ linear regressions, all with non-white residuals. For each regression, the same calculation as in (5.19) provides the variance of the estimated impulse response coefficients. Therefore, performing total of N least squares provides the estimation of all non-zero elements in $\hat{L}_{YR}$ and their corresponding variances as well as $\hat{L}_Y$.

The estimation of $\hat{L}_U$ and $\hat{L}_{UR}$ can also be obtained from Equation (5.3) in the same way along with the variance of all parameters in $\hat{L}_{UR}$.

Remark 1. After the estimation of $\hat{L}_Y$, $\hat{L}_{YR}$, $\hat{L}_U$ and $\hat{L}_{UR}$, Equation (5.8) can be used to estimate $\hat{L}_{UE}$. Similarly, $\hat{L}_{YE}$ can be estimated at this point. Therefore, the required noise parameters, $g_i^{(j)}$, for the calculation of R_{ϵ_i} are available.

Remark 2. Generalization of the presented analysis to the MIMO case does not change the variance calculation procedure. For the case of a multi-input process, the same number of regressions must be solved, but the size of unknown parameters vector (e.g. θ_1 or θ_2) and the regressor increase. For the ith regression, we will have $\theta_i \in \mathbb{R}^{(l+1)N+l\times 1}$, because $h_i^{(j)}$ is no longer scalar but has the dimension of $h_i^{(j)} \in \mathbb{R}^{l\times 1}$ and $l_{k,i} \in \mathbb{R}^{l\times 1}$ for $i = N, \cdots, 2N-1$. The regressor size will also change to $X_i \in \mathbb{R}^{(j-1)\times l(N-1)}$.

In the case of multi-output process, the number of least squares to be solved increases from N to mN. Therefore, the same analysis can be used for the variance estimation in the MIMO case.

Remark 3. Note that except for the first regression, the noise term is not white for which the weighted least squares is required for the optimal solution. In the presented method, these regressions are approximately solved by simple least squares. In fact, this is a common approximation in most of the subspace methods. Based on Remark 1, the knowledge of $g_i^{(j)}$'s and the corresponding noise covariance matrices, R_{ϵ_i}, are available at this point which makes it possible to perform weighted lease squares. This would improve the accuracy in the estimation of the impulse response coefficients. It should be noted that in the improved case, the variance of

estimation should be calculated by [69]

$$Cov(\hat{\theta}_i) = (X_i^T R_{\epsilon_i}^{-1} X_i)^{-1} \tag{5.23}$$

From the computational point of view, this second step of weighted lease squares doubles the number of regression equations to be solved, thus increases the computation cost.

Now, the estimate of $\hat{L}_{YR}$ and $\hat{L}_{UR}$ and their corresponding element-by-element variances are available. Consider the estimated $\hat{L}_{YR}$ and its corresponding variance matrix as

$$\hat{L}_{YR} = \begin{pmatrix} \hat{h}_1^{(1)} & 0 & 0 & \cdots & 0 \\ \hat{h}_2^{(1)} & \hat{h}_1^{(2)} & 0 & \cdots & 0 \\ \hat{h}_3^{(1)} & \hat{h}_2^{(2)} & \hat{h}_1^{(3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{h}_N^{(1)} & \hat{h}_{N-1}^{(2)} & \hat{h}_{N-2}^{(3)} & \cdots & \hat{h}_1^{(N)} \end{pmatrix} \tag{5.24}$$

$$Cov(\hat{L}_{YR}) = \begin{pmatrix} \Lambda_1^{(1)} & 0 & 0 & \cdots & 0 \\ \Lambda_2^{(1)} & \Lambda_1^{(2)} & 0 & \cdots & 0 \\ \Lambda_2^{(1)} & \Lambda_1^{(2)} & \Lambda_1^{(3)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \Lambda_N^{(1)} & \Lambda_{N-1}^{(2)} & \Lambda_{N-2}^{(3)} & \cdots & \Lambda_1^{(N)} \end{pmatrix} \tag{5.25}$$

where each element in $Cov(\hat{L}_{YR})$ represents the variance of estimation for the corresponding element in $\hat{L}_{YR}$.

A weighted average of $\hat{h}_i$'s can be obtained by

$$h_i^{ave} = \left(\sum_{j=1}^{N-i+1} (\Lambda_i^{(j)})^{-1/2} \right)^{-1} \left(\sum_{j=1}^{N-i+1} (\Lambda_i^{(j)})^{-1/2} \hat{h}_i^{(j)} \right) \tag{5.26}$$

This averaged value can be regarded as a better estimation of all h_i's in $\hat{L}_{YR}$. Substituting h_i's with h_i^{ave} results in a Toeplitz structure for $\hat{L}_{YR}$. Performing a same procedure to refine $\hat{L}_{UR}$ forces the final estimation of $\hat{L}_u$ to be Toeplitz as well.

Remark 4. Note that in practical applications, the non-Toeplitz structure of the estimated $\hat{L}_{YR}$ and $\hat{L}_{UR}$ could be not just because the noise effect, but also as a result of time-varying process dynamics.

In the next section, we provide two Monte-Carlo simulations on SISO and MIMO examples followed by the application results from two pilot-scale experiments to evaluate the proposed method.

5.4 Simulations and pilot-scale application

To evaluate the proposed method of this paper, we run Monte-Carlo simulations on a univariate and a multivariate system. The results of application on a pilot-scale Continuous Stirred Tank Heater (CSTH) and a pilot-scale four-tank process are also provided in this section.

5.4.1 SISO Example

This example is the first example of Chapter 3 in which the system description is given by

$$x_{t+1} = \begin{pmatrix} 0.6 & 0.6 & 0 \\ -0.6 & 0.6 & 0 \\ 0 & 0 & 0.7 \end{pmatrix} x_t + \begin{pmatrix} 1.616 \\ -0.348 \\ 2.631 \end{pmatrix} u_t - \begin{pmatrix} 1.147 \\ 1.520 \\ 3.199 \end{pmatrix} e_t$$
$$y_t = \begin{pmatrix} -0.437 & -0.504 & 0.093 \end{pmatrix} x_t - 0.775 u_t + e_t$$

A PI controller, $[0.1 + 0.05/s]$, is used to control the process. The test signal is designed by MATLAB® command '*idinput*' with Nyquist frequency of 0.12 and magnitude of 1. Variance of the input noise is 0.01.

The impulse response coefficients of the system are estimated using the proposed sequential joint input-output identification method followed by a weighted averaging step. The system step response then is calculated by integration. For the sake of comparison, we also implement the original closed-loop subspace identification method of [26] (as described in Section 2) as well as the 'innovation estimation' method of Qin and Ljung (2003) [83, 68] which provide direct estimation of the subspace matrix L_u. Since the variance of impulse response coefficients in these two methods is not available, simple averaging is used for the final step response estimation. For a comparative study, we also simulate the representative closed-loop identification algorithm [93, 94] and the CVA method, for which MATLAB® codes are available, although these latter two methods cannot provide the required subspace matrix directly. For the last two methods, the step response coefficients are *calculated* from the identified state space model. We run 50 Monte-Carlo simulation using each method. The results are shown in Figures 5.2 and 5.3.

Figure 5.2 shows that consistency of the estimation as well as the lowest variance is obtained using the proposed method. Figure 5.3 shows that the original method of [26] and the method of [83] provide consistent estimation too, but with higher variance. It also indicates that both the method of [93, 94] and the CVA method deliver bias, although CVA provides lower variance. This is expected since CVA normally applies to the open-loop identification.

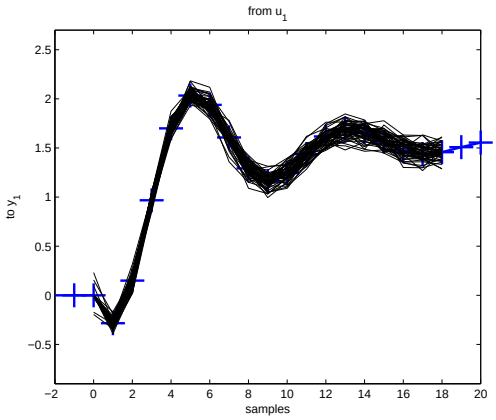


Figure 5.2: Monte-Carlo simulation results using sequential regressions followed by *weighted* averaging of impulse response coefficients.

5.4.2 MIMO Example

Consider a 2×2 process with open-loop transfer function matrix G_p and disturbance transfer function matrix G_l given as [45]:

$$G_p = \begin{pmatrix} \frac{z^{-1}}{1-0.4z^{-1}} & \frac{0.5z^{-2}}{1-0.1z^{-1}} \\ \frac{0.3z^{-1}}{1-0.4z^{-1}} & \frac{z^{-2}}{1-0.8z^{-1}} \end{pmatrix}$$
$$G_l = \begin{pmatrix} \frac{1}{1-0.4z^{-1}} & \frac{-z^{-1}}{1-0.1z^{-1}} \\ \frac{z^{-1}}{1-0.7z^{-1}} & \frac{1}{1-0.8z^{-1}} \end{pmatrix} \tag{5.27}$$

The following controller is implemented on the process:

$$G_c = \begin{pmatrix} \frac{0.5-0.2z^{-1}}{1-0.5z^{-1}} & 0 \\ 0 & \frac{0.25-0.2z^{-1}}{(1-0.5z^{-1})(1+0.5z^{-1})} \end{pmatrix} \tag{5.28}$$

Similar to the previous example, 50 Monte-Carlo simulation runs are performed by each of the methods and results are presented in Figure 5.4 and Figure 5.5, respectively.

Similar to example 1, the proposed method of this paper provides higher performance in the step response estimation over all other methods.

5.4.3 CSTH experiment

The process in this experiment is a continuous stirred tank heater (CSTH) and its schematic is shown in Figure 5.6.

There are two controlled variables in CSTH, the water level inside the tank and outlet water temperature. Manipulated variables are the cold water flowrate and steam flowrate. The head of the water in the inlet pipe as well as the steam supply

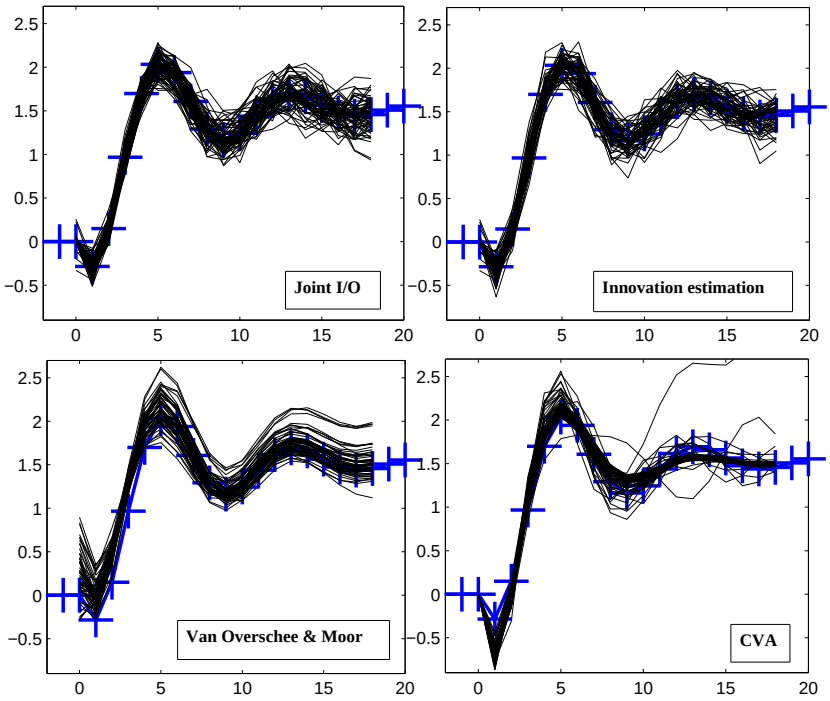


Figure 5.3: Monte-Carlo simulation results using four different methods of closed-loop subspace identification.

pressure and temperature can be considered as disturbances. This process is under PID control by *DeltaV* control system.

After some preliminary tests, two '*RBS*' signals are designed using MATLAB for testing the process under closed-loop conditions. Signals are applied to the setpoint of the two controllers and closed-loop data are collected with 5 seconds sampling time and over 2 hours. The input and output of the process under the identification test are shown in Figure 5.7.

The proposed method of joint input-output identification by sequential least squares and weighted averaging of impulse response coefficients is implemented on the collected data. The value of N is selected to be 50, so it provides the step responses for 250 seconds. The results are presented in Figure 5.8 which verifies the utility of the proposed method for this applications. The step responses between different inputs and outputs have been captured by this method.

5.4.4 Four-Tank experiment

The Four-Tank process consists of four equally-sized transparent tanks with orifices. The system also has two water pumps and split-valves, which determine the distribution of water flow into the tanks. The two left and right flowrates are manipulated using the pumps and the levels of the lower tanks are controlled variables.

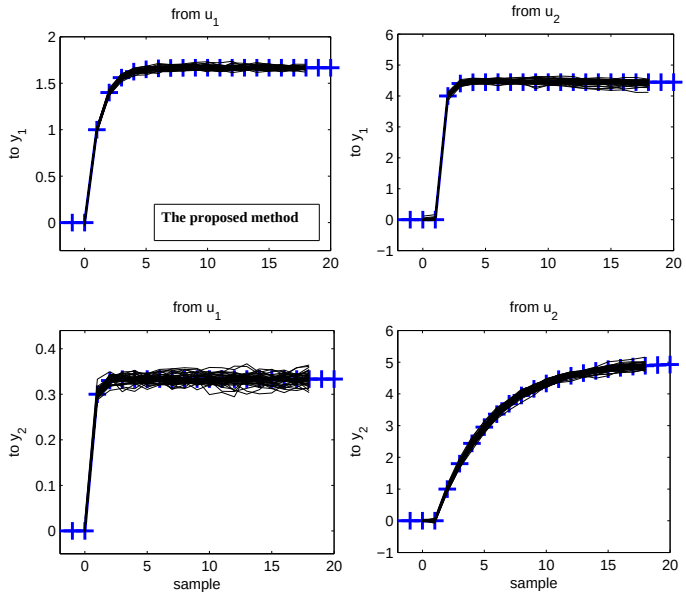


Figure 5.4: Monte-Carlo simulation results using sequential regressions followed by *weighted* averaging of impulse response coefficients.

Split-valves are fixed on 50%. A schematic of the process is shown in Figure 5.9.

Two '*RBS*' signals are designed using MATLAB® to test the process under closed-loop conditions. The process outputs and inputs under the test are shown in Figures 5.10. Closed-loop data are collected over 90 minuets with sampling time of 5 seconds. The choice of $N = 80$ results in the step responses of length 400 seconds. The results are shown in Figure 5.11 which demonstrate again the utility of the proposed method for step response estimation in practical applications.

5.5 Concluding remarks

The problem of direct step response estimation from closed-loop data using subspace approach was considered. The joint input-output identification method was employed. It is known that the intermediate subspace matrix containing impulse response coefficients of the process provides N sets of impulse response coefficients with different lengths. To use all the estimated impulse response coefficients for the step response calculation, one requires the variance of all parameters individually to perform a weighed averaging on them. A solution to this problem was proposed by breaking the high-dimension linear regression step of the subspace identification into a series of N one-dimension least squares from which the variance of estimation can be obtained for each impulse response coefficient. Also, the non-casual input terms which normally appear in the subspace identification algorithms can be avoided by using this approach, which together with the weighted average of the impulse

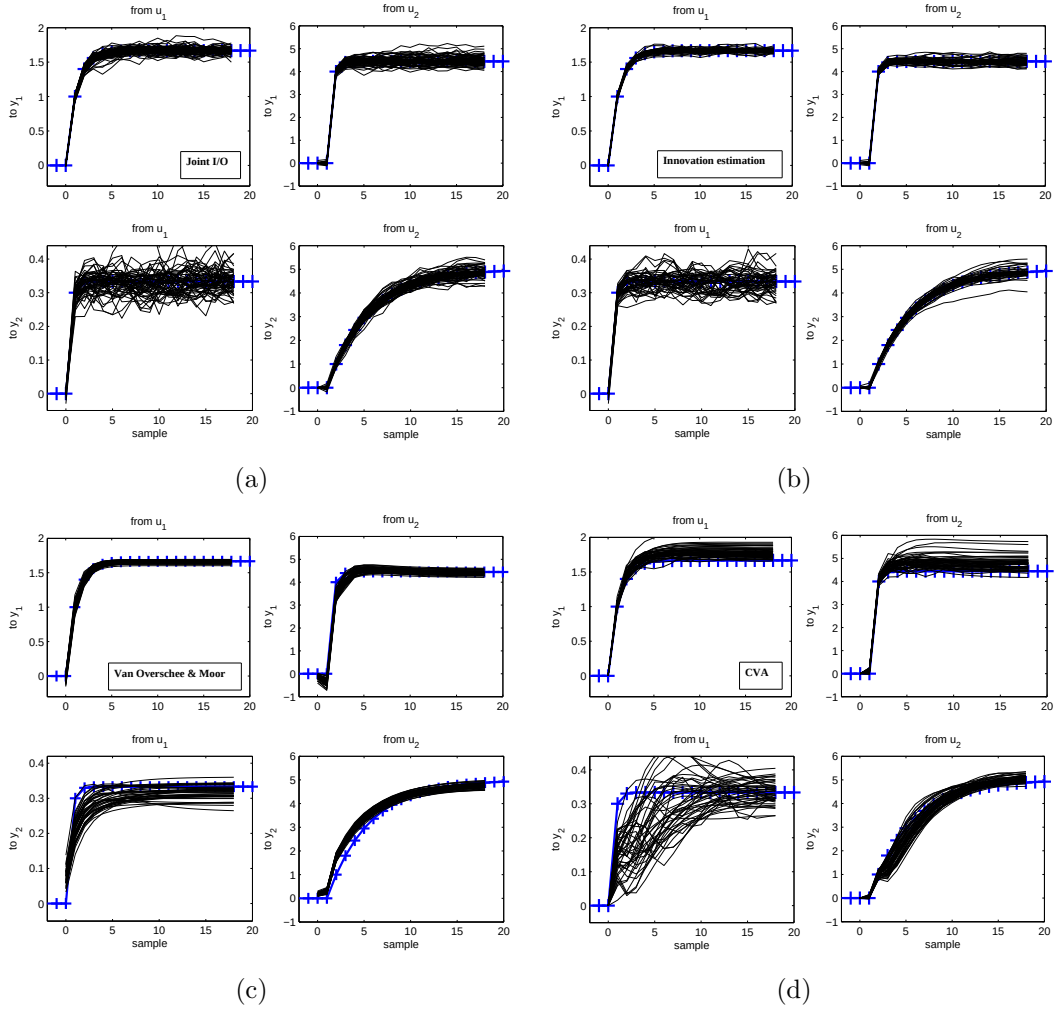


Figure 5.5: Monte-Carlo simulation results using four different methods of closed-loop subspace identification.

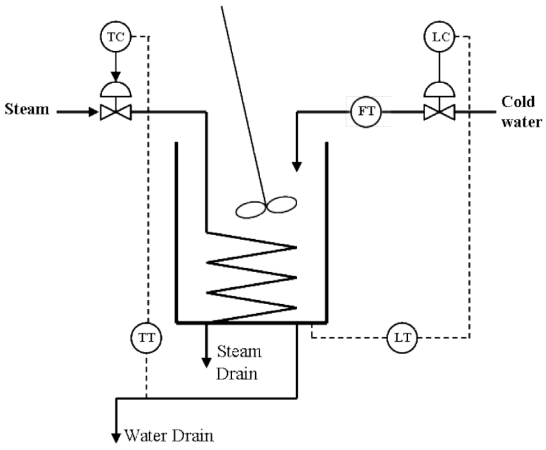
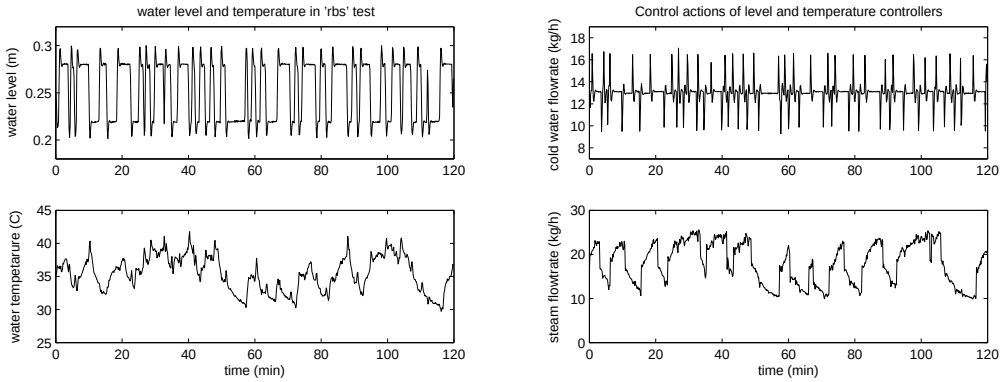


Figure 5.6: A schematic of Continuous Stirred Tank Heater process.



(a) (b)

Figure 5.7: CSTH outputs and control actions under closed-loop 'RBS' test.

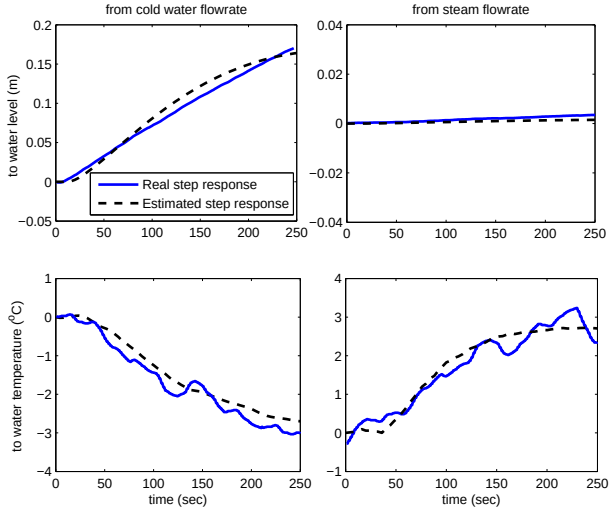


Figure 5.8: Real and estimated step responses of CSTH process.

response coefficients yield in lower variance of the estimates. Two Monte-Carlo simulations were used to compare the performance of the proposed method with four other closed-loop subspace identification methods. The results of two pilot-scale experiment were also provided to verify this method for practical applications.

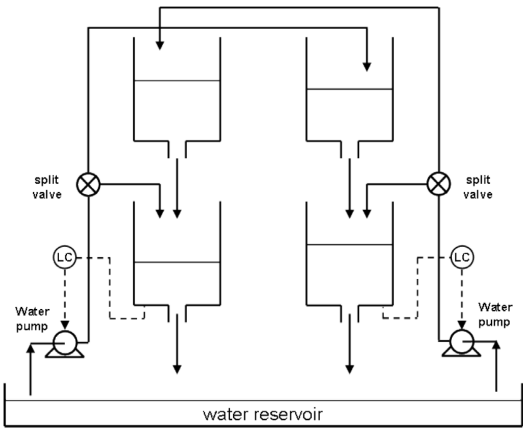
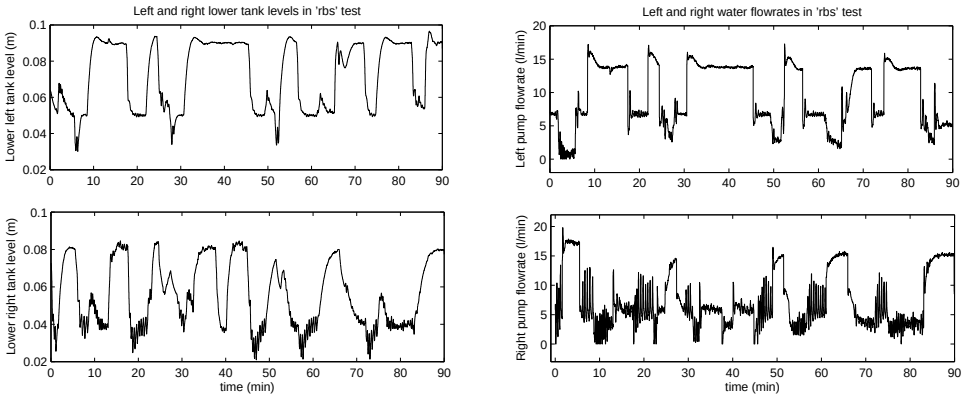


Figure 5.9: Real and estimated step responses of CSTH process.



(a) (b)

Figure 5.10: Four-tank process outputs under closed-loop 'RBS' test.

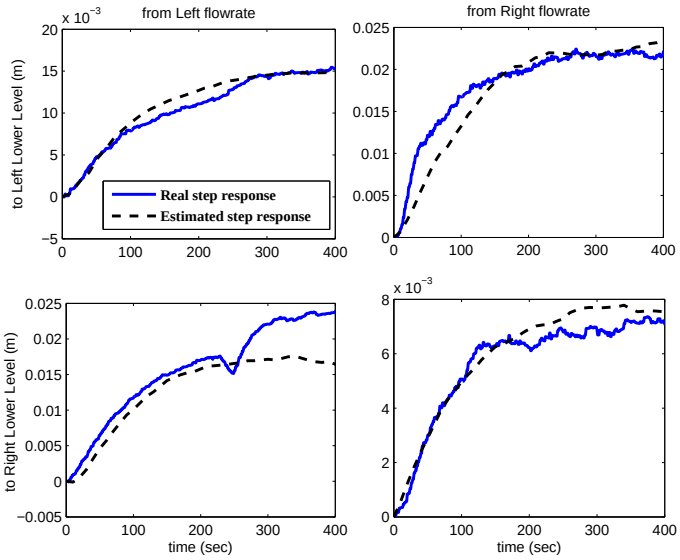


Figure 5.11: Real and estimated step responses of Four-tank process.

Chapter 6

Conclusions

6.1 Concluding Remarks

The main contributions of this thesis can be summarized as follows:

- An insightful study on the mechanism of the open-loop subspace identification method for the noise-free case was provided including a simple numerical example of the regression analysis method which have not been provided in any of the subspace identifications literatures.

- An alternative direct formulation of the joint input-output subspace identification method was proposed for performance assessment based on the LQG benchmark. This direct formulation enabled us to provide consistency analysis for the proposed method. A practical approach for estimation of the required number of rows in the data Hankle matrices was also derived. The result was to offer a data-driven method to obtain consistent estimate of the LQG trade-off curve.

- A subspace approach to obtain the LQG curve from closed-loop routine operating data was provided for the case where a model of the process (not the disturbance model) is available, such as in model-based controllers. Consistency of the estimation was proven as well.

- The proposed method of subspace-based performance assessment was applied on a pilot-scale Continuous Stirred Tank Heater (CSTH) process.

- The subspace LQG control was designed for control systems with supervisory-regulatory structure for the objective of controller performance assessment. Three possible scenarios for performance improvement was considered and the LQG control was provided for all three cases. As well, the LQG trade-off curve for each case was developed using certain subspace matrices without the requirement of explicit models.

- To complete the study, a closed-loop subspace identification method based on the joint input-output identification approach was provided to estimate the required subspace matrices and the noise covariance under the supervisory-regulatory control.

- A sequential implementation of the closed-loop joint input-output subspace identification was proposed which was used to estimate the process step response model directly from closed-loop data. The variance calculation for all elements in the required subspace matrix was provided and used for weighted averaging of the estimated impulse response coefficients resulting in a more accurate estimation of the step response model. Use of the complex mathematical tools was avoided to make the results applicable for practical applications.

- The proposed method of direct step response model estimation was implemented on the pilot-scale CSTH process and a pilot-scale four-tank process.

- Two Graphic User Interfaces (GUI) were generated for closed-loop subspace identification and controller performance assessment based on LQG benchmark.

6.2 Recommendations for future work

Research initiatives on the topic of the current research and the following related field are worthy of future investigations:

- An insightful study on the mechanism of the noise-free subspace identification method was provided in this thesis. Generalization of this study to the case of noisy data can be very useful to improve the performance of the subspace identification methods.

- Since there is some estimation error in any identification method, using the results of identification to obtain the LQG curve leads to some uncertainty in the trade-off curve. Very few available methods for quantification of this uncertainty have relied on very complex mathematical tools. Initiatives can be taken to propose simpler analysis regarding this problem.

- One practical approach to estimate an upper bound for the number of rows in the data Hankle matrices (N) from closed-loop data was provided in this thesis. Since the identification results depend on this value, it is worthwhile to investigate how an optimal value of N can be estimated for closed-loop subspace identification.

- The proposed controller performance assessment method for the supervisory-regulatory control systems is limited to linear controllers, but nonlinear controllers like constrained MPC and PID controllers with limited output are common in industry. Development of a subspace method applicable to the constrained controllers would provide a more practical tool.

- The element-by-element variance calculation was provided for a modified version of the joint input-output identification in this thesis. The results of this study improved the step response model estimation from closed-loop data. This type of analysis may also be performed for some of the direct closed-loop subspace identification methods.

Bibliography

[1] N. Agarwal, B. Huang, and E.C. Tamayo. Assessing mpc performance part i: Probabilistic approach for constraint analysis. *Industrial & Engineering Chemistry Research*, 46:8101–8111, 2007.

[2] N. Agarwal, B. Huang, and E.C. Tamayo. Assessing mpc performance part ii: Probabilistic approach for constraint analysis. *Industrial & Engineering Chemistry Research*, 46:8112–8119, 2007.

[3] A. AlGhazzawi and B. Lennox. Model predictive control monitoring using multivariate statistics. *J. Process Control*, 19(2):314–327, 2009.

[4] K.J. Åström and B. Wittenmark. *Computer-Controlled Systems, Theory and Design, 2nd ed.* Prentice Hall, Englewood Cliffs, NJ, 1990.

[5] D. Bauer. *Some asymptotic theory for the estimation of linear systems using maximum likelihood methods or subspace algorithms.* PhD thesis, Tu Wein, 1998.

[6] D. Bauer. Order estimation for subspace methods. *Automatica*, 37(10):1561–1573, 2001.

[7] D. Bauer. Some facts about the choice of the weighting matrices in larimore type of subspace algorithms. *Automatica*, 38(5):763–773, 2002.

[8] D. Bauer. Asymptotic properties of subspace estimators. *Automatica*, 41(3):359–376, 2005.

[9] D. Bauer, M. Deistler, and W. Scherrer. Consistency and asymptotic normality of some subspace algorithms for systems without observed inputs. *Automatica*, 35(7):1243–1254, 1999.

[10] D. Bauer, M. Deistler, and W. Scherrer. On the impact of weighting matrices in subspace algorithms. In *11th IFAC conference SYSID*, Santa Barbara, California, 2000.

[11] D. Bauer and M. Jansson. Analysis of the asymptotic properties of the moesp type of subspace algorithms. *Automatica*, 36(4):497–509, 2000.

[12] O.T. Berglihn. A toolbox for using matlab as an activex/com controller for hysys. Matlab Central, http://www.pvv.org/olafb/hysyslib/, 1999.

[13] W.L. Bialkowski. Dreams versus reality: a review from both sides of the gap. *Pulp Pap. - Canada*, 94:19–27, 1993.

[14] S.P. Boyd and C.H. Barratt. *Linear Control Design.* Prentice Hall, Eaglewood Cliffs, New Jersy, 1991.

[15] A. Chiuso. Asymptotic equivalence of certain closed-loop subspace identification methods. In *17th IFAC World Congress*, Newcastle, Australia, March 2006.

[16] A. Chiuso. On the relation between cca and predictor-based subspace identification. *IEEE Trans. Automatic Control*, 52:1795–1812, 2007.

[17] A. Chiuso and G. Picci. The asymptotic variance of subspace estimates. *J Econometrics*, 118:257–291, 2003.

[18] A. Chiuso and G. Picci. Consistency analysis of closed-loop subspace identification. In *43rd Conf. on Decision and Control*, Atlantis, Bahamas, Dec. 2004. IEEE.

[19] A. Chiuso and G. Picci. Numerical conditioning and asymptotic variance of subspace estimates. *Automatica*, 40:677–683, 2004.

[20] A. Chiuso and G. Picci. Consistency analysis of some closed-loop subspace identification methods. *Automatica*, 41(3):377–391, 2005.

[21] A. Chiuso and G. Picci. Prediction error vs. subspace methods in closed-loop identification. In *16th IFAC World Congress*, Prague, Czech Republic, July 2005.

[22] A. Chiuso and G. Picci. Asymptotic variance of closed-loop subspace identification methods. *IEEE Trans. Automatic Control*, 51:1299–1314, 2006.

[23] A. Chiuso and G. Picci. Estimating the asymptotic variance of closed-loop subspace estimators. In *17th IFAC World Congress*, Newcastle, Australia, March 2006.

[24] C.T. Chou and M. Verhaegen. Subspace algorithms for the identification of multivariable dynamic error-in-variables models. *Automatica*, 33(10):1857–1869, 1997.

[25] N.LC. Chui. *Subspace methods and informative experiments for system identification*. PhD thesis, Pembroke College Cambridge, 1997.

[26] N. DaneshPour, B. Huang, and S.L. Shah. Consistency in joint input-output subspace identification for lqg benchmark. *submitted to J. Process Control*, 2009.

[27] L. Desborough and T.J. Harris. Performance assessment measure for univariate feedback control. *Cand. J. of Chem. Eng.*, 70:1186–1197, 1992.

[28] L. Desborough and T.J. Harris. Performance assessment measure for univariate feedforward/feedback control. *Cand. J. of Chem. Eng.*, 71:605–616, 1993.

[29] J. Ding, M. Verhaegen, and E. Holweg. Closed-loop subspace predictive control for fault tolerant design. In *17th IFAC World Congress*, pages 3216–3221, Seoul, Korea, July 2008.

[30] J. Dong, M. Verhaegen, and E. Holweg. Closed-loop subspace predictive control for fault tolerant mpc design. In *17th IFAC World Congress*, pages 3216–3221, Seoul, Korea, July 2008.

[31] D. Ender. Process control performance: not as good as you think. *Control Eng.*, 9:180–190, 1993.

[32] W. Favoreel, B. DeMoor, M. Gevers, and P.van Overschee. Closed loop model-free subspace-based lqg-design. Technical Report Tch. rep. 98-108, Katholieke Universiteit Leuven, 1998.

[33] W. Favoreel, B. DeMoor, P. Van Overschee, and M. Gevers. Model-free subspace-based lqg-design. *In Proc. of the American Control Conference*, pages 3372–3376, 1999.

[34] J. Gao, R. Patwardhan, K. Akamatsu, Y. Hashimoto an, G. Emonto, S.L. Shah, and B. Huang. Performance evaluation of two industrial mpc controllers. *Can. J. Chem. Eng.*, 11:1371–1387, 2003.

[35] T. Gustafsson. *Subspace methods for system identification and signal processing.* PhD thesis, Chalmers University, Gothenburg, Sweden, 1999.

[36] T.J. Harris. Assessment of control loop performance. *Can. J. Chem. Eng.*, 67(10):856–861, 1989.

[37] T.J. Harris, F. Boudreau, and J.F. MacGregor. Performance assessment of multivariable feedback controllers. *Automatica*, 32(11):1505–1518, 1996.

[38] T.J. Harris, C.T. Seppala, and L.D. Desborough. A review of performance monitoring and assessment techniques for univariate and multivariate control systems. *J. of Process Control*, 9:1–17, 1999.

[39] K.A. Hoo, M.J. Piovoso, P.D. Schnelle, and D.A. Rowan. Process and controller performance monitoring: overview with industrial applications. *Int. J. Adaptive Control Signal Process*, 17:635–662, 2003.

[40] B. Huang. *Multivariable statistical methods for control loop performance assessment.* PhD thesis, Dep. of Chem. and Mate. Eng., University of Alberta, Edmonton, Alberta, Canada, 1997.

[41] B. Huang. A pragmatic approach towards assessment of control loop performance. *Int. J. Adapt. Control & Signal Process.*, 17:589–608, 2003.

[42] B. Huang, S.X. Ding, and S.J. Qin. Closed-loop subspace identification: an orthogonal projection approach. *J. Process Control*, 15:53–66, 2005.

[43] B. Huang and R. Kadali. *Dynamic Modelling, Predictive Control and Performance Monitoring: A Data-driven Subspace Approach.* Springer Verlag, London, 2008.

[44] B. Huang, R. Kadali, X. Zhao, E.C. Tamayo, and A. Hanafi. An investigation into the poor performance of a model predictive control system on an industrial cgo coker. *Control Engineering Practice*, 8(6):619–631, 2000.

[45] B. Huang and S.L. Shah. *Control Loop Performance Assessment: Theory and Applications.* Springer Verlag, London, 1999.

[46] B. Huang, S.L. Shah, and H. Fujii. The unitary interactor matrix and its estimation using closed-loop data. *J. Process Control*, 7(3):195–207, 1997.

[47] B. Huang, S.L. Shah, and E.K. Kwok. Good, bad or optimal? performance assessment of multivariable processes. *Automatica*, 33(6):1175–1183, 1997.

[48] M. Jansson. *On subspace methods in system identification and sensor arraysignal processing.* PhD thesis, KTH, Stockholm, 1997.

[49] M. Jansson. Asymptotic variance analysis of subspace identification methods. 2000.

[50] M. Jansson. Subspace identification and arx modeling. IFAC Symp. on System Identification, Aug. 2003.

[51] M. Jansson and B. Wahlberg. On consistency of subspace methods for system identification. *Automatica*, 34(12):1507–1519, 1998.

[52] M. Jelali. An overview of control performance assessment technology and industrial applications. *Control Engineering Practice*, 14:441–446, 2006.

[53] R. Kadali and B. Huang. Controller performance assessment using lqg-benchmark obtained under closed loop. *ISA Transactions*, 41:521–537, 2002.

[54] R. Kadali and B. Huang. Estimation of dynamic matrix and noise model for predictive control using closed-loop data. *Ind. Eng. Chem. Res.*, 41:842–852, 2002.

[55] R. Kadali and B. Huang. Multivariate control performance assessment without interactor matrix. In *IFAC Advanced Control of Chemical Processes*, pages 61–66, 2003.

[56] T. Katayama and H. Tanaka. An approach to closed-loop subspace identification by orthogonal decomposition. *Automatica*, 43:1623–1630, 2007.

[57] T. Knudsen. Consistency analysis of subspace identification methods based on linear regression approach. *Automatica*, 47:81–89, 2001.

[58] B.S. Ko and T.F. Edgar. Performance assessment of cascade control loops. *AIChE Journal*, 46:281–291, 2000.

[59] B.S. Ko and T.F. Edgar. Performance assessment of constrained model predictive control systems. *AIChE Journal*, 37:1363–1371, 2001.

[60] B.S. Ko and T.F. Edgar. Performance assessment of multivariable feedback control systems. *Automatica*, 37:899–905, 2001.

[61] D.J. Kozub. Controller performance monitoring and diagnosis: experiences and challenges. In J.C. Kantor, C.E. Garcia, and B.C. Carnahan, editors, *Fifth Int. Conf. on Chemical Process Control*, pages 83–96, Tahoe, CA, 1996. AIChE and CACHE.

[62] H. Kwakernaak. h_2 optimization-theory and application to robust control design. *Anual Reviews in Control*, 26:45–56, 2002.

[63] W.E. Larimore. ystem identification, reduced order filters and modeling via canonical variate analysis. In H.S. Rao and P. Dorato, editors, *American control conference*, volume 2, pages 445–451, Piscataway, NJ, 1983.

[64] W.E. Larimore. Statistical optimality and canonical variate analysis system identification. *Signal Processing*, 52:131–144, 1996.

[65] W.E. Larimore. Large sample efficiency for adaptx subspace system identification with unknown feedback. In *the 7th international symposium on DYCOPS*, Boston, MA, July 2004.

[66] J.H. Lee. On interfacing model predictive controllers with low-level loops. *Ind. Eng. Chem. Res.*, 39(1):92–102, 2000.

[67] K. Lee, B. Huang, and E.C. Tamayo. Sensitivity analysis for constraint and variability tuning in performance assessment of industrial mpc. *Control Engineering Practice*, 16(10):1195–1215, 2008.

[68] W. Lin, S.J. Qin, and L. Ljung. On consistency of closed-loop subspace identification with innovation estimation. In *43rd Conf. on Decision and Control*, Atlantis, Bahamas, Dec. 2004. IEEE.

[69] L. Ljung. *System Identification: Theory for the user*. Prentice-Hall, 1999.

[70] L. Ljung and T. McKelvey. Subspace identification from closed loop data. *Signal Processing*, 52:209–215, 1996.

[71] L. Ljung and T. McKelvey. Interpretation of subspace methods: Consistency analysis. In *11th IFAC Symposium on System Identification*, volume 3, pages 1125–1130, Fukuoka, Japan, July 1997.

[72] W.L. Luyben. *Plantwide Dynamic Simulators in Chemical Processing and Control*. Marcel Dekker Inc., NY, 2002.

[73] C.B. Lynch and G.A. Dumont. Control loop performance monitoring. *IEEE Trans. Control Systems Technol.*, 4:185–192, 1996.

[74] K. Peternell M. Deistler and W. Scherrer. Consistency and relative efficiency of subspace methods. *Automatica*, 31(12):1865–1875, 1993.

[75] C.A. McNabb. *MIMO control performance monitoring based on subspace projections*. PhD thesis, The University of Texas at Austin, 2002.

[76] C.A. McNabb and S.J. Qin. Projection based mimo control performance monitoring: I - covariance monitoring in state space. *J. Process Control*, 13:739–757, 2003.

[77] M. Moonen, B. De Moor, L. Vandenberghe, and J. Vandewalle. On- and off-line identification of linear state-space models. *International J. Control*, 49(1):219–232, 1989.

[78] K. Peternell, W. Scherrer, and M. Deistler. Statistical analysis of novel subspace identification methods. *Signal Processing*, 52:161–177, 1996.

[79] S. J. Qin and T. A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

[80] S.J. Qin. Control performance monitoring a review and assessment. *Comput. Chem. Eng.*, 23:178–186, 1998.

[81] S.J. Qin, W. Lin, and L. Ljung. A novel subspace identification approach with enforced causal models. *Automatica*, 41(12):2043–2053, 2005.

[82] S.J. Qin and L. Ljung. A framework for subspace identification methods. In *American control conference*, Arlington, VA, June 2001.

[83] S.J. Qin and L. Ljung. Closed-loop subspace identification with innovation estimation. In *13th IFAC SYSID Symposium*, Rotterdam, Netherlands, Aug. 2003.

[84] S.J. Qin and L. Ljung. Parallel qr-implementation of subspace identification with parsimonious models. In *13th IFAC SYSID Symposium*, Rotterdam, Netherlands, Aug. 2003.

[85] S.J. Qin and J. Yu. Recent developments in multivariable controller performance monitoring. *J. Process Control*, 17:221–227, 2007.

[86] J. Schäfer and A. Cinar. Multivariable mpc system performance assessment, monitoring, and diagnosis. *J. of Process Control*, 14:113–129, 2004.

[87] T. Söderström and P. Stoica. *System Identification*. Prentice Hall International, 1989.

[88] A.K. Tangirala, S. Lakshminarayanan, and S.L. Shah. Closed-loop identification using canonical variate analysis. Edmonton, Canada, 1997. 47th CSChE Conference.

[89] N.F. Thornhill, M. Oettinger, and B. Surgenor. Refinery-wide control loop performance assessment. *J. Process Control*, 9:109–124, 1999.

[90] P. van Overschee and B. de Moor. *Subspace Identification for Linear Systems; Theory, Implementation, Application*. Katholieke Universiteit, 1996.

[91] P. van Overschee and B. De Moor. N4sid: Subspace algorithm for the identification of combined deterministic-stochastic systems. *Automatica*, 30(1):75–93, 1994.

[92] P. van Overschee and B. De Moor. A unifying theorem for three subspace system identification algorithms. *Automatica*, 31(12):1877–1883, 1995.

[93] P. van Overschee and B. De Moor. Closed-loop subspace system identification. Technical Report ESAT-SISTA/TR 1996-52I, Katholieke Universiteit, Leuven, 1996.

[94] P. van Overschee and B. De Moor. Closed-loop subspace system identification. In *36th conference on Decision and Control*, pages 1834–1853, 1997.

[95] M. Verhaegen. Application of a subspace model identification techniques to identify lti systems operating on closed-loop. *Automatica*, 29:1027–1040, 1993.

[96] M. Verhaegen and P. Dewilde. Subspace model identification part 1. the output-error state-space model identification class of algorithms. *International J. of Control*, 56(5):1187–1210, 1992.

[97] J. Wang and S.J. Qin. A new subspace identification approach based on principal component analysis. *J. Process Control*, 12:841–855, 2002.

[98] J. Wang and S.J. Qin. Closed-loop subspace identification using the parity space. *Automatica*, 42:315–320, 2006.

[99] D. Wei, I.K. Craig, and M. Bauer. Multivariate economic performance assessment of an mpc controlled electric arc furnace. *ISA Transactions*, 46(3):429–436, 2007.

[100] F. Xu, B. Huang, and S. Akande. Performance assessment of model predictive control for variability and constraint tuning. *Industrial & Engineering Chemistry Research*, 46:1208–1219, 2007.

[101] J. Yu and S. J. Qin. Statistical mimo controller performance monitoring. part i: Data-driven covariance benchmark. *J. Process Control*, 18(3):277–296, 2008.

[102] Q. Zhang and S. Li. Enhanced performance assessment of subspace model-based predictive controllers with parameters tuning. *Canadian J. of Chem. Eng.*, 85:537–548, Aug, 2007.

Appendix A

Graphic User Interfaces

For the purpose of generalizing some of the algorithms presented in this thesis, we have generated two Graphic User Interfaces (GUI) for multivariate controller performance assessment and closed-loop subspace identification. In this appendix these two GIUs are briefly introduced.

Graphic user interfaces have provided some degree of simplicity in using MATLAB-coded algorithms. Many new released MATLAB toolboxes provide this facility. Based on this fact and to follow the previous efforts made in the Computer Process Control (CPC) group in this regard, we generated two GUIs based on the material of this thesis. The first one called 'LQGPA', can be used for controller performance assessment using LQG benchmark and the second one is called 'CLsysID' and provides a tool for closed-loop subspace identification. We review the provided options by each GUI in the following.

A.1 LQGPA

LQGPA has been produced to provide different options for multivariate controller performance assessment based on the LQG benchmark. A snapshot of this GUI is shown in Figure A.1. Different modes of operation have been considered in this tool and it can be used both in simulation and industrial studies There are four option available in this GUI in term of the model/data requirements. If a complete model of the process and disturbance dynamics is available, it can be used for to solve the LQG problem and provide the trade-off curve. Identification data collected from the process under open-loop or closed-loop condition can also be used to generate the LQG trade-off curve. In the case where a model-based controller is in use, the available controller model (normally, only the process model) and a set of routing closed-loop operating data (no designed excitations are required) is another option to estimate the LQG curve. This information is used in the GUI to estimate the noise model and the noise covariance matrix for performance assessment.

Different performance indices are calculated based on the current control performance. The overall minimum variance index is calculated for each output as well as the overall least error index, overall least cost index and global cost index [41].

All simulation models or data must be changed to a 'compact' format before being used in the tool. A function is provided with the GUI to perform this transformation. Some SISO and MIMO models and data sets have been included in the toolbox.

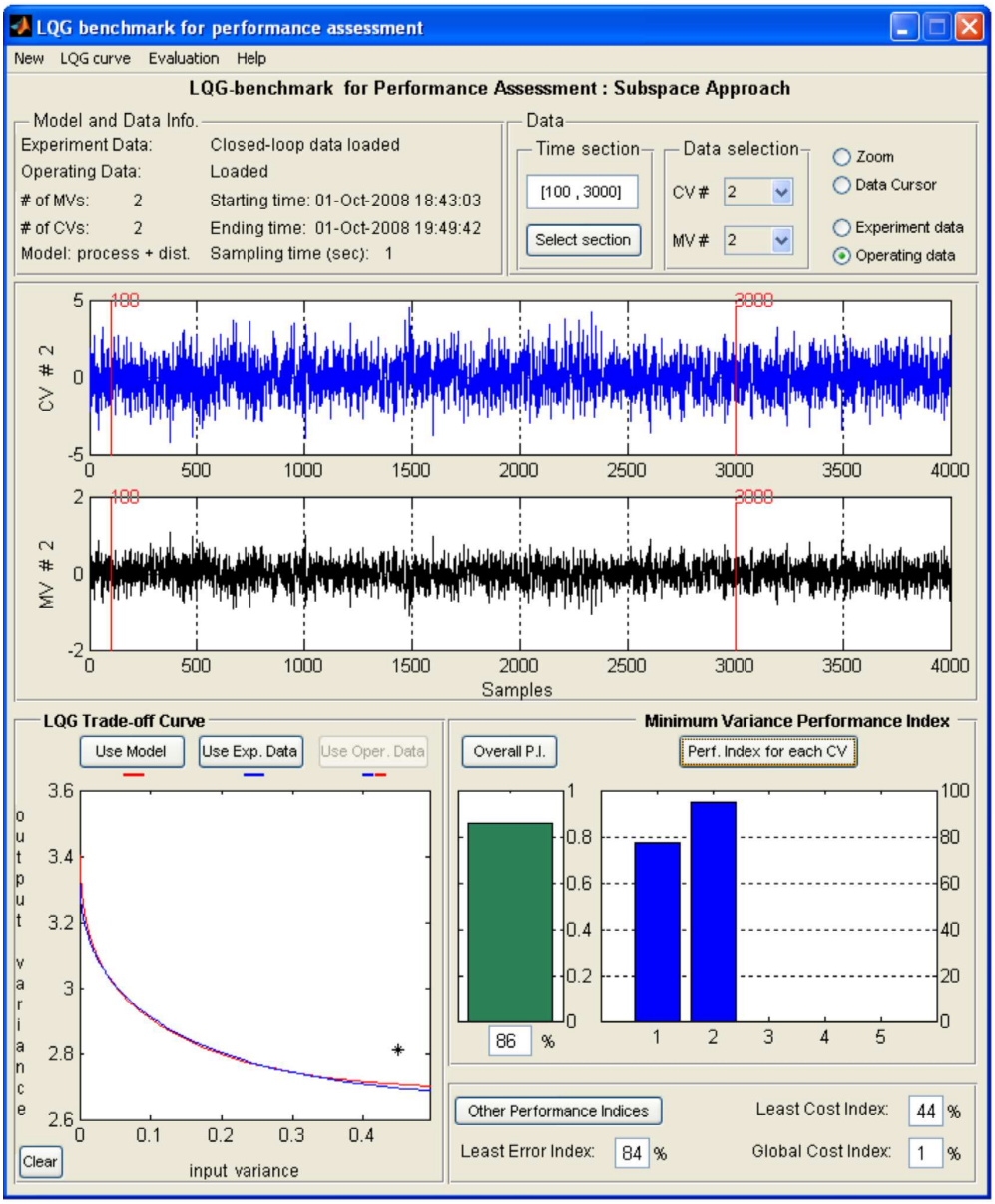


Figure A.1: A snapshot of the LQGPA.

A.2 CLsysID

CLsysID has been produced for the purpose of closed-loop multivariable process identification using subspace identification. A snapshot of this GUI is shown in Figure A.2. The methods of closed-loop joint input-output identification presented in Chapter 3 has been implemented in this GUI. The first step of estimating the subspace matrices has been replaced by the sequential least squares approach proposed in Chapter 5. Final system matrices are estimated using the method of Knudsen (2001) [57] reviewed in Chapter 2. The model order is estimated by the GUI, but there is an option for the user to change the order. An upper bound on the number of rows in the data Hankle matrices, N (The past and future horizons) is also estimated and used in the GUI based on the correlation analysis on closed-loop data (see Chapter 3), but the user is also allowed to make changes in this parameter which may results in some improvements. A reasonable range for this parameter is provided for the the user.

The final estimated model can also be refined by PEM method for possible improvement. Model validation options based on prediction fit and residue test are provided by the GUI.

Direct estimation of the process step responses, as discussed in Chapter 5, is also available. Continuous-time models can also be fitted to the estimated step response coefficients using nonlinear regression method available in the MATLAB statistics toolbox. Continuous-time first-order plus time-delay (FOPTD) or second-order plus time-delay (SOPTD) models are fitted to the step response coefficients based on an automated procedure. This type of continuous-time models are used in some industrial MPC packages. A snapshot of the estimated step response coefficients and the fitted continuous-time models are shown in Figure A.3 for a 2×2 example.

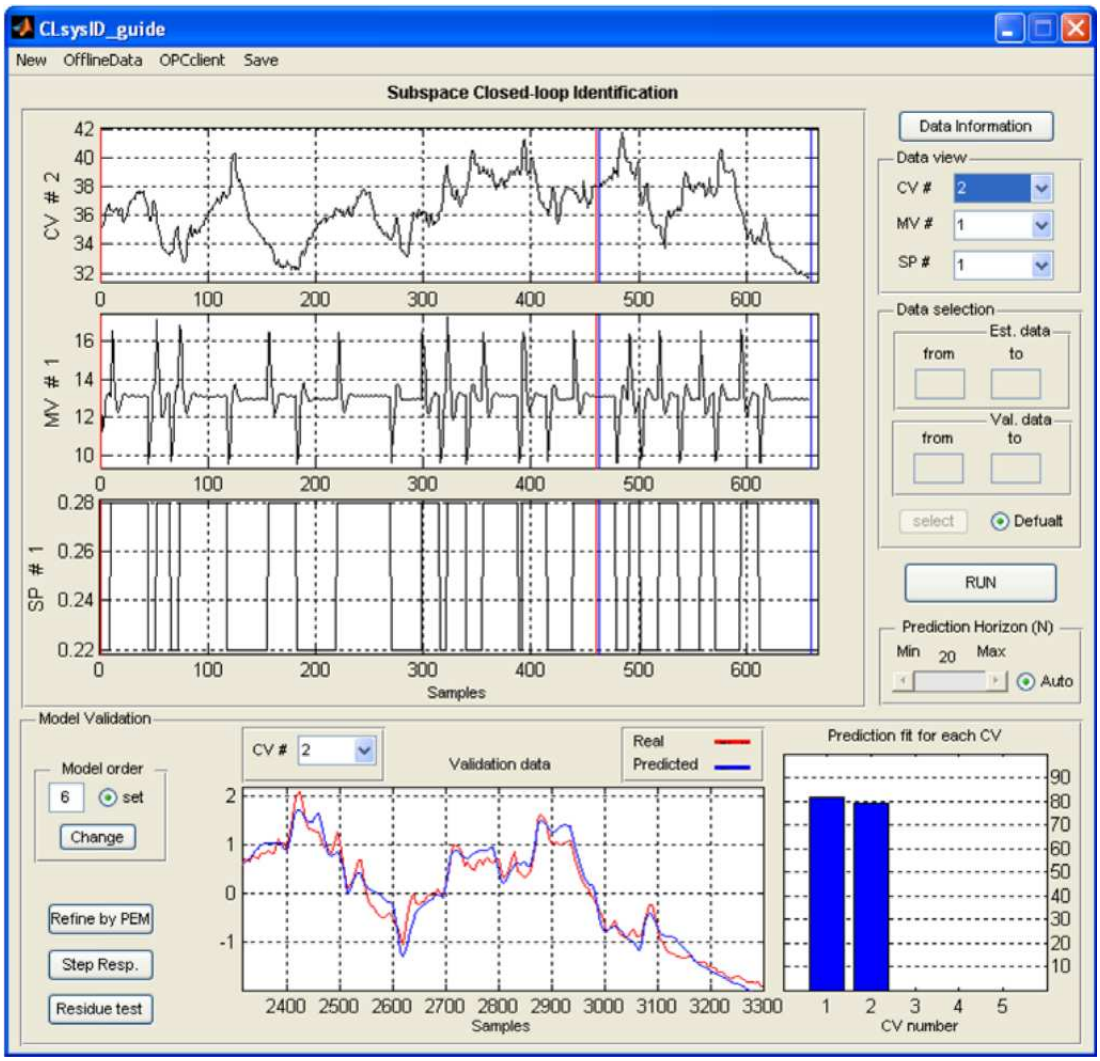


Figure A.2: A snapshot of the CLsysID.

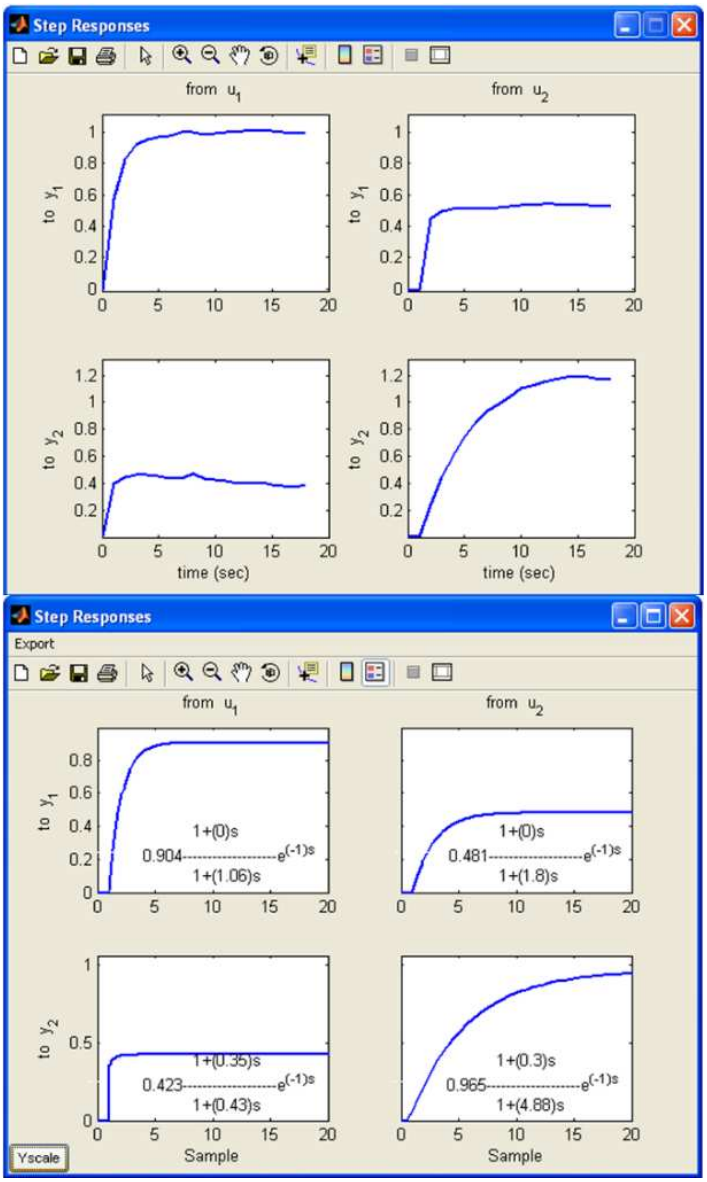


Figure A.3: An example of the continuous-time step response model estimation using CLsysID.