**University of Alberta**

# GENERATING INFORMATION FOR REAL-TIME OPTIMIZATION

by

**George Charles Pfaff** ©

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

Edmonton, Alberta
Fall 2001

0-612-69574-3

**University of Alberta**

**Library Release Form**

**Name of Author**: George Charles Pfaff

**Title of Thesis**: Generating Information for Real-Time Optimization

**Degree**: Master of Science

**Year this Degree Granted**: 2001



George Charles Pfaff
CME 536
University of Alberta
Edmonton, AB
Canada, T6G 2G6

**Date:** SEPT 27/01

**University of Alberta**

**Faculty of Graduate Studies and Research**

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Generating Information for Real-Time Optimization** submitted by George Charles Pfaff in partial fulfillment of the requirements for the degree of **Master of Science** in *Process Control*.

. . . . . . . . . . . . . . . . . . . .
J. Fraser Forbes

. . . . . . . . . . . . . . . . . . . .
P. Jim McLellan

. . . . . . . . . . . . . . . . . . . .
Sirish L. Shah

. . . . . . . . . . . . . . . . . . . .
Jim V. Kresta

**Date:** SEPT. 27/01

# Abstract

Real-time optimization (RTO) is an on-line optimization technique that monitors the behaviour of the process, looking for significant low frequency changes in the true plant optimum, while adjusting the setpoints of the process controllers to track these changes. The performance of the optimizer depends on its ability to track these changes effectively and locate the true plant optimum operating conditions. By incorporating experimental design techniques, this thesis proposes an improvement to RTO performance by integrating information generation into the algorithm to reduce uncertainty in the final optimization results.

Expansion of the command conditioning subsystem evaluates when the predicted result from the economic optimizer will not generate sufficient amount of information. The experimental design calculation uses an A-optimal criterion to reduce uncertainty associated with decision variables by perturbing from the optimal point to another that generates more information. By sacrificing short-term profit, greater profit can be realized in future RTO intervals.

# Acknowledgements

I would like to thank Drs. J. Fraser Forbes and P. James McLellan for their guidance throughout the course of this work. I am indebted to them for their in depth discussions, comments and useful insight. I want to express my gratitude to Dr. P. James McLellan for his effort and devotion to this work even though he was based out of a different university, and for initiating my interest in the fields of process control and optimization. I would also like to thank Dr. J. Fraser Forbes who kept me focused in the end.

I would like to express my appreciation to the support staff in the Department of Chemical and Materials Engineering especially Marilee Barry, AnnMarie Brereton, Jack Gibeau and Bob Barton who never hesitated to offer their help when needed. I would like to thank the other members in my research group who always provided helpful discussions and insight.

The financial support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

As global competition increases, industries are forced to try and improve their competitive advantage and often resort to improving their process efficiency through economic optimization. Real-time optimization (RTO) is a steady-state, model-based, on-line optimization technique that monitors the behaviour of the process, looking for significant low frequency changes in the true plant optimum, while adjusting the setpoints of the process controllers to track these changes. The plant optimum could vary as a result of changes in ambient conditions, equipment fouling, and changes in the inputs to the process. The efficiency of the optimizer is dependent on its ability to track these changes effectively and locate the true optimum operating conditions for the plant.

The two main causes of the RTO system not converging to the plant optimum are: 1) plant/model mismatch in the model structure and fixed parameters; and 2) uncertainty in the adjustable parameter estimates. These factors contribute to the bias cost, which is defined as the loss of profit as a result of the deviation of the expected value of the predicted optimum from the true plant optimum (Forbes and Marlin, 1996). Figure 1.1 illustrates the different contributions to the bias cost. The bias associated with plant/model mismatch occurs from errors in the model structure and/or fixed parameters, for which the model updating system may not be able to compensate (Zhang and Forbes, 2000). The bias that occurs from uncertainty in the adjustable parameters is a result of the optimizer being unable to converge closer to the true plant optimum because additional optimization moves are considered statistically insignificant, and are not implemented. It is the uncertainty in the adjustable parameters, which is forwarded to the optimization calculations, that results in the new moves being rejected. By reducing this uncertainty, future RTO iterations will do a better of locating the plant optimum.

The objective of this thesis is to develop a decision criterion that provides insight into the number of historical steady-steady points that should be used in the model updater subsystem, and also to expand the role of the command conditioning subsystem in the RTO loop to: 1) ensure that each implemented result generates information and secondly; 2) provide a diagnostic check to evaluate model adequacy.

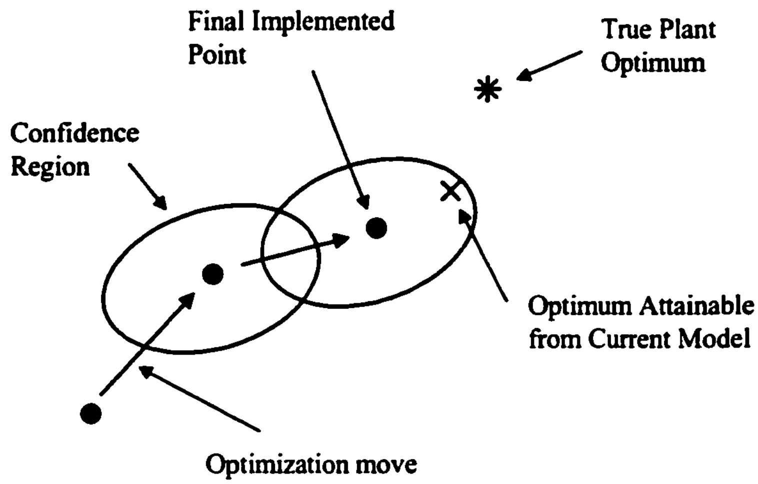



Figure 1.1: Structural mismatch and decision variable uncertainty contribute to the bias cost.

This chapter provides an overview of process optimization and experimental design. It also discusses the purpose of the different units in the RTO system, and provides a motivation for the basis of this thesis. The chapter concludes with a discussion of the conventions and assumptions used in subsequent chapters.

## 1.1 Process Optimization

Many advances have been made to improve the efficiency of RTO algorithms, such as including first-principles information to better describe a wider range of the process behaviour, and attempting to compensate for plant/model mismatch. However, little attention has been provided towards moving in directions that ensure information is generated to allow future RTO iterations to do a better job of locating the plant optimum.

Steady-state process optimization can be separated into two categories: direct-search and model-based methods. Direct-search methods use plant experimentation to identify directions to improve plant performance. Model-based methods typically rely on models that are developed off-line to describe the process behaviour, and only make use of plant data to update certain parameters in the model that are unknown or have values that change frequently during operation.

The act of performing experiments to generate information on how a process can be improved began receiving attention in the late 1950's with the development of Evolutionary Operation (Bacon, 1992). Evolutionary Operation (EVOP), developed by Box (1957), uses 2-level factorial experimental designs to systematically guide changes in operating conditions to more profitable conditions. The factor effects are identified by making many small changes in the operating conditions. Once a significant effect is identified, the operation moves to a new point in the direction of the optimal factor. This is an iterative process that repeats itself once a new point is selected. The advantages of EVOP include:

- the use of small changes in the factors so there is little upset in process operation,
- continual optimization of the process,
- ability to track changes in the operating conditions.

The drawback of EVOP is that it requires 2-level factorial designs, possibly augmented with center points, at steady-state conditions. Since 2-level factorial designs without center points require $2^n$ points (where $n$ is the number of factors), it can take a long time before an optimum direction can be found, especially for processes with long settling times. Since EVOP requires a long time to develop any conclusions, it is only useful for testing a small number of factors for processes with small settling times relative to the frequency of occurrences of process disturbances.

Response Surface Methodology (RSM), as an optimization tool, is similar to EVOP. It was originally developed by Box and Wilson (1951) for the purpose of identifying complex nonlinear processes using simple equations that adequately describe the local behaviour of the process. It was later applied to optimization problems. This method uses a 2-level factorial, or higher, order design to formulate an empirical model, with a predefined structure, that describes a response surface within the region of interest (Myers and Montgomery, 1995). The gradient of this equation provides the direction of steepest improvement of the response variable. The next point is selected by performing a line search along the direction defined by the gradient. Once the next point is selected, the routine starts over with the past information discarded and a new experimental design constructed. The same concerns exist with this algorithm as for EVOP, since the time it takes to perform an experimental design with many factors may be unreasonable given that the points are steady-state points.

Bamberger and Isermann (1978), Garcia and Morari (1981), and McFarlane and Bacon (1989) developed optimization algorithms that evolved from EVOP and adaptive control techniques to extract steady-state information from empirical models, with a predefined structure, that are identified using an experimental design. The experimental design for these algorithms consist of on-line perturbations of the manipulated variables. Edwards and Jutan (1997) combine the ideas of RSM and adaptive optimization algorithms to produce Dynamic Response Surface Methodology (DRSM). DRSM is a modification of RSM that tracks dynamic optima.

The drawback of direct search methods is that substantial plant experimentation is required to make up for the absence of first-principles information. By adding first principle information the time required for experimentation is reduced. White (1997) states that over time, more rigorous first-principles models are being used for optimization because of their ability to accurately cover a wider operating region than the empirical model based methods.

Model-based optimization uses first-principles information but is limited by its inability to compensate for plant/model mismatch. Figure 1.1 shows how the optimizer may converge to a point that is not the true plant optimum, because the model is unable to accurately describe the process behaviour. Many researchers have investigated the problem of improving the efficiency of the optimization routine by reducing the plant/model mismatch. The Two-Phase design, studied by Chen and Joseph (1987), creates two subsystems in the RTO loop where the parameter estimation and optimization problems are solved separately. The Two-Phase approach is the most widely used design for the model updating and model-based optimization subsystems in RTO (Zhang and Forbes, 2001). This approach updates the adjustable parameters in the model; however, it does not compensate for the inevitable mismatch in the model structure or the fixed parameters. Golden and Ydstie (1989) also attempted to compensate for plant/model mismatch by expanding the work of Bamberger and Isermann (1978) to combine the use of theoretical models and an adaptive approach. Roberts (1979) developed a model-based method that compensated for plant/model mismatch by adding a term to the original objective function. This work was extended to produce an iterative technique that solves dynamic optimal control problems (Becerra and Roberts, 1996).

This thesis is only concerned with steady-state model based optimizers, which account for almost all commercial implementations of RTO (Darby and White, 1988).

## 1.2 RTO Design Issues

This section will discuss the role of the individual RTO subsystems, and how errors and uncertainty enter and propagate through the closed-loop RTO system. A discussion of the issues addressing the effect parameter and variable uncertainty has on the performance of the RTO system, and the state of model adequacy tests for optimization purposes is presented at the end of this section.

A simplified schematic of an RTO system is shown in Figure 1.2. The following is a description of the different units in the closed-loop RTO system:

**Measurements**: The measurements of the process are taken from sensors positioned in the plant. These measurements are used to update the adjustable parameters in the model updating subsystem, in the process of which the uncertainty in the measurements is forwarded to the parameter estimates.

**Data Validation:** The data validation subsystem consists of three subsections: steady-state detection, gross error detection and data reconciliation. Since steady-state models are used for most RTO systems it is important to ensure the plant is operating at steady-state before measurements are collected for model updating purposes (White, 1997). Some tests for steady-state detection include: comparing the mean values of a given set of measurements between two time periods, constancy of time series coefficients, small rate of change of a variable over a period, or measurements remaining within prescribed bounds about a mean

(Marlin and Hrymak, 1996). Gross error detection is responsible for removing data that are not self-consistent with respect to material and energy balances and therefore should not be used by the model updater (Reilly and Carponi, 1963; Mah and Tamhane, 1982; Rosenberg, 1987; Crowe, 1988; Tong and Crowe, 1995). This false information would lead to biases and reduce the effectiveness of the RTO algorithm. Gross errors could occur from a series of factors including poor sampling, instrument malfunction, leaks or data from units off-line (Marlin and Hrymak, 1996). The data reconciliation stage is responsible for allocating errors, which still exist after removing the gross errors, across the material and energy material balances to attempt to obtain the best possible estimation of the true plant operation (Crowe *et al.*, 1983; Crowe, 1986; Crowe, 1996).

**Model Updater**: Model uncertainty is due to uncertainty in the structural form of the model or parameter values when compared to the true process. Structural uncertainty exists because of the approximations made to develop the model, and parameter uncertainty is present because of changes in the process such as catalyst deactivation, equipment fouling, reaction kinetics, and so forth. The model updater subsystem attempts to compensate for model uncertainty by using the processed data from previous subsystems to update the model parameters. The efficiency of the optimizer is sensitive to the quality of the model as it determines the rate at which the estimated optimum is reached and the amount of offset from optimal operation that will exist.

**Optimizer**: The optimizer performs a steady-state model-based optimization to determine the point that will provide the most profit within the boundaries specified by the constraints. The two most common nonlinear programming algorithms used for optimization problems are augmented Lagrangian methods or sequential quadratic programs (Marlin and Hrymak, 1996). The uncertainty in the adjustable parameter estimates transmits through the optimization subsystem to the predicted results of the optimizer.

**Command Conditioning**: The command conditioning subsystem acts as a check of the optimization results before they are transmitted. These checks include: determining if the plant operation is not currently upset, the optimization variables remain available for manipulation, and the bounds of the optimization variables remain unchanged (Marlin and Hrymak, 1996). Another role designated to this section is to determine if the calculated change in operation is significant. If the result is not significant, the point should not be implemented. By reducing the amount of unnecessary changes, plant profits have been shown to increase (Miletic and Marlin, 1998). It is this subsystem that this thesis will propose to expand to ensure a level of information generation after each optimization move is implemented.

**Plant and Controllers**: The control system implements the results from the RTO calculations. Advanced control systems, such as the model predictive controller (MPC), are usually layered under the RTO system to ensure the solution can be reliably implemented (Marlin and Hrymak, 1996).

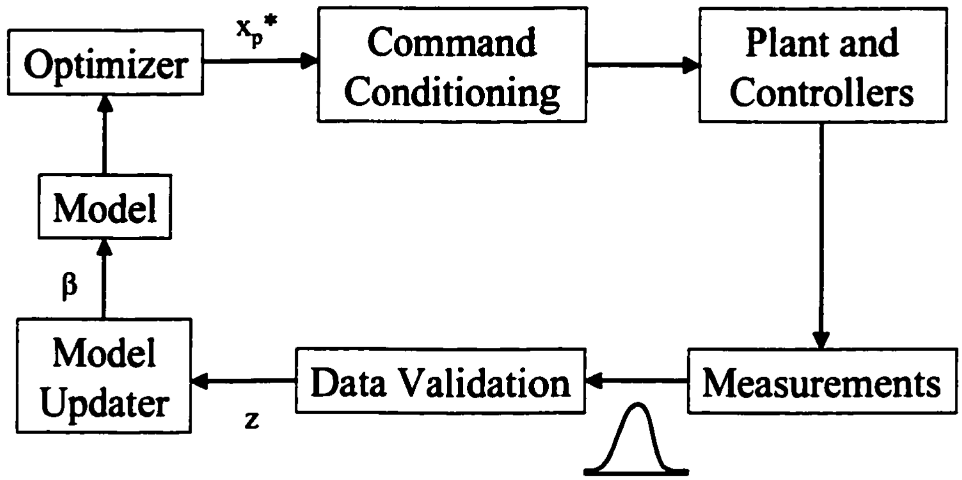



Figure 1.2: The closed-loop RTO system.

### 1.2.1 Decision Variable Uncertainty

There has been some attention directed to the effect parameter uncertainty has on RTO efficiency. Referring to Figure 1.2, RTO is a closed-loop system; therefore, whatever uncertainty is associated with the parameter estimates obtained from the model updater, will be forwarded on to become uncertainty in the result obtained from the economic optimizer. Koninckx (1988) examined how the uncertainty in the decision variables affected the evaluation of the expected gains in profit, and the accuracy of the predicted profit and decision variables. Krishnan *et al.* (1992) investigated how uncertainty from the measurements was forwarded to the adjustable parameter estimates, in order to select which measurements have the greatest influence on the updated parameters. Forbes and Marlin (1996) developed an expression that describes the transmission of variability in the plant data from the measurements to the model parameters and then the propagation of the variability of the adjustable model parameters to the calculated optimization variables. Fraleigh (1999) examined how to select the optimal sensor system that will minimize uncertainty in the decision variables.

It is clear that uncertainty in the adjustable parameters lead to model and optimization results that also contain significant uncertainties. Pinto (1998) identified some of the costs that result from parametric uncertainty and was able to assign an economic value to measure the effect. Zhang *et al.* (2001) showed that for RTO there is a further cost that might occur when the size of the confidence region surrounding the optimization calculation inhibits the progress of the optimizer, by not implementing results that do improve profit. It is possible that the optimizer could get 'stuck' and not converge closer to the true plant optimum creating a bias cost with respect to profit. Reducing parameter uncertainty means better precision in the parameter estimates and ultimately, estimates of optimal operation. This improved precision will enable the optimizer to locate the plant optimum more effectively.

The uncertainty in the decision variables could be reduced by ensuring that the RTO system will always implement setpoints that work to reduce the magnitude of some scalar measure (*i.e.* determinant or trace) of the decision variable covariance matrix.

### 1.2.2 Diagnostics

The models developed for RTO are usually based on first-principles. While developing models to describe the behaviour of the process, assumptions and approximations are made to simplify the system to either reduce the computational load or because some effects are considered insignificant. However, there may be a time when these assumptions are no longer reasonable and the model is unable to adequately represent the process behaviour. The error caused from this plant/model mismatch could result in a bias cost in the RTO system reducing the amount of profit that could be attained.

For any regression analysis it is important that tests be performed to assess the fit of the model to the data once the adjustable parameter estimates have been determined (Bard, 1974; Bates and Watts, 1988). Currently, RTO applications only compensate for the inevitable mismatch by updating the adjustable parameters at the current steady-state operating point (Marlin and Hrymak, 1996). Tests may be performed off-line to determine if the model system is no longer effective, but the RTO system may operate for a long period of time before acknowledgment of a poor model system is identified. Forbes *et al.* (1994) developed a model adequacy check that is performed off-line to investigate the suitability of a candidate process model for use in an optimization routine. They define model adequacy as the ability of the process model to have an optimum coinciding with the true plant optimum by altering the adjustable parameters.

The need for a set of diagnostic checks will become clear in Chapter 3, where the case study will show that the experimental design performs poorly if there is significant mismatch between the model and the true process. A set of on-line diagnostic checks to evaluate model adequacy will inform the optimizer when the model is not able to represent the process behaviour, indicating the conditions for when an experimental design should not be performed.

## 1.3 Experimental Design for Optimization

This section provides a discussion of the past involvements of experimental design in optimization and how other design of experiment (DOE) techniques could be used for optimization applications.

The selection of an experimental design depends on the criteria required by the application. Featherstone (1997) lists some of the following criteria that may or may not be relevant when selecting an experimental design technique. For the purpose of this thesis the first six points will be considered to be of primary importance.

- Provide a satisfactory distribution of information throughout the region of interest.
- ensure a good fit between the model and true process,
- allow for a minimum number of experimental runs,
- minimize complexity of the DOE calculation,
- allow for a model adequacy determination,
- be robust to errors occurring in the settings of the experimental variables,
- provide simple data patterns that allow for ready visual appreciation,
- use a minimal number of levels of the experimental factors,
- allow for estimation of transforms of both the response and the quantitative experimental factors,
- provide replicate data in order to estimate error,
- be robust to outliers,
- provide a diagnostic check on the constancy of variance assumption.

The alphabetic optimal experimental designs are common methods of reducing parameter uncertainty through planned experimentation (Shirt *et al.*, 1994; Myers and Montgomery, 1995). The four most common designs are A, D, E, and G optimal designs. The D-optimal design was the original design whose basis was provided by Wald (1943), and was later fully developed by Kiefer and Wolfowitz (1959). A, D, and E designs were formulated to minimize the variance of the parameter estimates by reducing the volume and/or shape of the parameter estimate's confidence intervals, while the G-optimal design minimizes the variance of model predictions. Please note that this thesis will focus on minimizing the uncertainty in the decision variables instead of the parameter estimates. For the purpose of this thesis, model prediction is not as important as minimizing the variance of the decision variables because optimization is only concerned with finding the optimal direction of increasing profit, and less interested in using the model to predict the values of the response variables. For these same reasons, this thesis is primarily concerned with reducing the volume of the decision variable confidence region, but also needs to address shape to ensure that uncertainty in some variables do not remain large at the expense of further volume reduction. Definitions of the four designs for the purpose of this thesis are provided (Bacon, 1992),

**D-Optimal:** Minimizes the determinant of the decision variable covariance matrix.

**A-Optimal:** Minimizes the sum of the variances of the individual decision variables.

**E-Optimal:** Minimizes the largest eigenvalue of the decision variable covariance matrix.

**G-Optimal:** Minimizes the maximum variance of the predicted responses.

A significant advantage of the alphabet designs is their ability to accommodate the addition of any number of extra points to a current data set. Many authors have used the alphabet criteria as a basis for sequential experimental design purposes, a requirement for use in a RTO application (Box and Hunter, 1965; Draper and Hunter, 1966; Draper and Hunter,1967 a and b; Dykstra, 1971; Evans, 1979; Pinto *et al.*, 1990; Myers and Montgomery, 1995; Featherstone, 1997). Heiberger *et al.* (1992,1993) proposed a strategy called U-optimal designs which use a combination of A, D, and E-optimal designs to augment an experimental set.

A disadvantage of the alphabet designs is that they are developed with the assumption that the model form is correct. Although this may be a reasonable assumption, there is always some structural mismatch that cannot be accounted for in the assumed RTO model. All-bias experimental designs focus on reducing the bias in the current model and ignore the effect of parameter variation. Therefore, all-bias designs are theoretically optimum when variance in the parameter estimates or decision variables are not present ( Box and Draper, 1987). However, for the general case when there is variance and bias error present, several authors have concluded that the all-bias design is a poor choice (Box and Draper, 1959; Box and Draper, 1963; Galil and Kiefer, 1977). Box and Draper (1987) provide a good reference for the theoretical framework of all-bias designs.

The ideal experimental design would distribute the experimental effort to reduce both variance in the decision variables and bias in the RTO model. A few researchers have developed experimental designs to produce this effect. Draper and Guttman (1992) compensate for bias by adding an extra term to the variance function incorporating bias errors as random effects instead of fixed effects. Welch (1983) and Steinberg (1985) have also developed designs that use this assumption. Welch (1983) bases his design on a mean square criterion, and Steinberg (1985) applies a Bayesian approach. DuMouchel and Jones (1994) use a Bayesian approach to modify the D-optimal design and reduce its dependence on the assumed model. It should be noted that these authors examine simple low-order models that are derived solely through empirical methods, and not models that incorporate first principle information. The disadvantages of these designs for consideration in RTO are they require some knowledge of the form of the bias that may be present and most of them are computationally intensive.

Previously, it was stated that although it is advantageous to reduce bias in the RTO model, the experimental designs that exist to accomplish this are not well suited for the application. However, there are designs that exist that would enable the optimization routine to discriminate between two rival models in the effort to reduce structural mismatch, although they do not work to reduce bias in the current model (Hunter and Reiner, 1965; Box and Hill, 1967; Atkinson and Cox, 1974). T-optimal experimental designs were developed by Atkinson and Fedorov (1975) to provide a design that creates a maximum discrimination between two models. For sequential designs, the next experimental point is selected to provide the largest increase in the expected value of the sum of squares of differences between the responses of the two models (Atkinson and Fedorov, 1975). Hill *et al.* (1968) studied the dual problem of model discrimination and parameter estimation as a joint design criterion. Fraleigh (1999) also studied the dual problem, although not for the purpose of selecting an experimental point, but to select the optimal sensor system that provides the smallest level of uncertainty in the calculated setpoints and bias between the nominal calculated setpoints. The drawback to these designs is the amount of trial and error that would have to be performed to select one out of a possible set of candidate models for RTO. These designs will also only compare the prediction ability of the models, when for an optimization application it is the ability of the model to locate the true plant optimum that is of priority.

The experimental design required for this thesis will be required to augment the current data set with one more point that will reduce decision variable uncertainty in a steady-state multiple response nonlinear model. It will focus solely on reducing variance in the decision variables, and will ignore the effects of bias in the assumed RTO model.

## 1.4 Scope and Objectives

This thesis focuses on methods to reduce uncertainty in the decision variables, and the application of on-line diagnostic tests to determine the adequacy of the model. As previously discussed, work has been directed at measuring the level of uncertainty in the optimization calculations for RTO, but little attention has been focused on its reduction. It is important to reduce this uncertainty to allow for tighter convergence to the true plant optimum. Figure 1.3 illustrates how reducing the uncertainty in the optimization variables, by implementing an experimental design calculation, could reduce the confidence region of the decision variables and allow for a closer convergence to the true plant optimum. Ideally, each RTO interval would generate some level of information thus minimizing the confidence region for each move. Since the purpose of RTO is to improve plant profitability, the cost associated with reducing the confidence region surrounding the decision variables must be considered. Therefore an experimental design calculation would only be implemented if the result from the economic optimizer is not expected to generate information. When a DOE is needed, the experimental window will consist of a constraint that limits the amount of potential profit that would be lost.

Chapter 2 investigates the impact of including historical steady-state information in the model updater subsystem of the RTO system to reduce uncertainty in the parameter estimates and decision variables. It will also take into account the effects of increased computational requirements and the weakened ability of the RTO to track changes in the process behaviour that result from including older information. Chapter 2 will also examine the ability of an experimental design to reduce the confidence region of the decision variables.

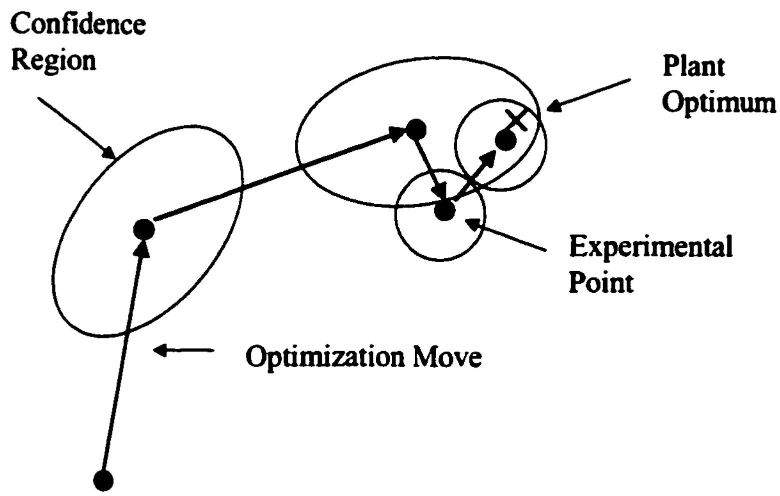



Figure 1.3: Reducing uncertainty allows for tighter convergence to the plant optimum.

Chapter 3 will focus on expanding the role of the command conditioning subsystem to ensure that each implemented point generates information that will reduce the confidence region surrounding the decision variables. In the event that the proposed operating point does not generate a significant amount of information, the experimental design calculations developed in Chapter 2 are performed. The experimental window will be defined as a constraint that will control the amount of potential profit that would be lost.

Chapter 4 will present the Williams-Otto plant case study that investigates how the information generation approach can be applied to a specific RTO problem.

## 1.5 Conventions and Assumptions

The following assumptions have been made throughout this thesis:

- Process measurements are taken at steady-state.
- Measurements are corrupted with normally distributed random noise only.
- Sensors are available to measure all the desired variables.
- The control system is able to implement the setpoints specified by the optimizer.

This thesis uses the following terms and conventions.

**Process Model**: refers to the set of equations that represent mass and energy balances for each process unit, and physical phenomena. This term does not include operational constraints.

**Process Variables**: The **decision** variables are the **independent**, or **manipulated** variables that are adjusted to optimize the plant profit. The **dependent** variables are determined from the process model once the **independent** variables have been specified.

**Uncertainty**: refers to the level of confidence in the numerical value of the variable. Specifically, **measurement** uncertainty is related to the quality of the sensors in terms of measuring the true value of a variable, and uncertainty in the **parameters** and **decision variables** is a result of how the **measurement** uncertainty propagates through the RTO system calculations to the parameter estimates and optimization calculations, respectively.

# Chapter 2

# Decision Variable Uncertainty Reduction

The role of the command conditioning subsystem, as introduced in the previous chapter, was expanded by Koninckx (1988) to use estimates of the decision variable covariance matrix to evaluate the expected economic gains of the RTO system and test for significant changes in the decision variables. Miletic and Marlin (1998) followed this work to determine if the predicted operating conditions are significantly different than the current operating conditions. Practically, it is undesirable to upset the process and change the setpoints to move to an operating point that may not improve the profitability of the process. Reducing the uncertainty of the decision variables improves the ability of the optimizer to locate and converge closer to the true plant optimum.

The purpose of this chapter is to investigate different approaches for reducing uncertainty in the optimization calculations by including historical steady-state information and applying experimental design techniques to generate information. Section 2.1 examines the issues surrounding the introduction of historical steady-state information to the model updater subsystem which includes: reducing uncertainty in the optimization calculations, increasing computational requirements, and reducing the ability of the model updater to track changes in the process behaviour. Section 2.2 discusses the selection of the decision variable covariance matrix approximation. Section 2.3 investigates the selection of the experimental design criterion for the RTO application. The chapter concludes with a case study that examines the advantages and disadvantages of including historical steady-state information and implementing experimental design optimization problems in the RTO algorithm.

## 2.1 Parameter Estimation Schemes to Reduce Uncertainty

The purpose of the model updater in the closed-loop RTO system is to update the adjustable parameters in the process model based on steady-state plant measurements received from sensors located in the plant. The uncertainty in the adjustable parameters can be reduced by implementing a planned experimental design or using a more robust parameter estimation scheme. Due to the serial nature of the RTO calculation (see Figure 1.2) reducing uncertainty in the adjustable parameters will reduce uncertainty in the decision variable values obtained from the optimization calculations.

Fraleigh (1999) compared three methods used for parameter estimation: back-substitution, least squares regression, and error-in-variables estimation (EVM). She concluded that least squares regression was the best method when considering the combined effects of parameter error and uncertainty, system reliability and computational requirements. Parameter error is defined as the difference between the estimates and true values of the parameters, and parameter uncertainty is the measure of uncertainty in the estimates.

EVM considers that all variables are subject to error and formally recognizes no distinction between dependent and independent variables (Fraleigh, 1999). Fraleigh (1999) concluded that although EVM is the most robust of the three techniques studied, it requires a computational load that is too intensive to be considered for most RTO applications. The algorithm performs simultaneous parameter estimation along with solving for the true values of the measured variables, resulting in a larger number of variables relative to the back-substitution or least-squares regression methods. This substantial increase in the number variables can create an unreasonably large problem for RTO applications that may involve models that contain thousands of variables. Marlin (1997) presented a model of the Sunoco Hydrocracker unit that involves 76,000 variables.

Back-substitution is the most common technique used for parameter estimation in RTO applications because of its simplicity and low computational cost. This method requires solving a system of nonlinear model equations where the number of equations is equal to the number of adjustable parameters. Fraleigh (1999) showed that least squares regression outperformed back-substitution, producing a substantially smaller confidence region for the parameter estimates for all tests including: nominal operation (all measurements are unbiased and available), a 10% bias introduced to one of the response variables, and a complete failure of one of the response variables. Back-substitution actually failed for the last case where one of the response variables failed.

In summary, Fraleigh (1999) was able to show that although least squares regression is computationally more expensive than back-substitution, it is able to significantly reduce parameter error and uncertainty while the computational load is not significantly burdensome. Based on this conclusion this thesis will use the least squares regression method in the model updater.

### 2.1.1 Selecting the Optimal Amount of Historical Information

The need to include historical steady-state information in the model updater becomes necessary if the final goal is to construct an experimental design to calculate the next operating point that will generate the maximum amount of information. In considering implementation of experimental design or regression techniques, it does not make sense to base the calculations on a single point, from either an experimental design point of view or from the perspective of the quality of estimates obtained through regression.

The implicit least squares regression problem for the multiple response case, is formulated with a structure similar to that defined by Box (1970),

$$\begin{aligned} \min \phi &= \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} && (2.1) \\ &= \mathbf{f}^T\left(\boldsymbol{\beta}^{(p)}, \mathbf{z}_u\right) \mathbf{f}\left(\boldsymbol{\beta}^{(p)}, \mathbf{z}_u\right) \\ &= \mathbf{f}^T\left(\boldsymbol{\beta}^{(p)}, \boldsymbol{\xi}_u^{(q)}, \mathbf{y}_u^{(v)}\right) \mathbf{f}\left(\boldsymbol{\beta}^{(p)}, \boldsymbol{\xi}_u^{(q)}, \mathbf{y}_u^{(v)}\right) && (2.2) \end{aligned}$$

where:

- $u = 1, 2, \ldots, N$ represents a collection of measurements at a specific operating point.
- $\mathbf{f}(\boldsymbol{\beta}, \mathbf{z}_u)$ is a stacked vector of $r$ implicit model equations with respect to the measured responses, at $N$ steady-state data points and has a length of $N \times r$,
- $\boldsymbol{\varepsilon}$ is a vector of residuals from the steady-state equations,
- $\boldsymbol{\beta}$ is the vector of $p$ adjustable parameters,
- and $\mathbf{z}_u = \left[\boldsymbol{\xi}_u^{(q)} | \mathbf{y}_u^{(v)}\right]$ is the vector of measurements at the $u^{th}$ steady-state operating point where $\boldsymbol{\xi}$ is the vector of observations of the $q$ decision variables, and $\mathbf{y}$ is the vector of $v$ measured dependent variables.

Using historical steady-state information will provide the additional advantage of utilizing more data points to increase the degrees of freedom to the least squares regression optimization problem, further reducing the uncertainty in the parameters and decision variables, providing that the data is linearly independent (Dahlquist and Bjorck, 1974). However, adding more data will increase the computational requirement to solve the optimization problem and will reduce the ability of the model updater to quickly track changes in the process behaviour. An expression is needed to determine the optimal number of historical data points to incorporate in the model updater, by balancing the reduction in the decision variable covariance matrix with the increased computational load and reduced ability to track process changes.

A similar situation has been identified in the literature for the model structure problem, where the trade-off is made between model precision and complexity. The goal is to identify the transition from relevant model fit to overfit. Expressions have been proposed to penalize the decrease in the loss function associated with introducing extra terms, with a cost to the increase in the number of parameters (Soderstrom and Stoica, 1989),

$$W_N = [1 + \psi(N, p)] \sum_{t=1}^{N} \varepsilon^2(t, \theta) \quad (2.3)$$

or

$$W_N = N \log \left[ \sum_{t=1}^{N} \varepsilon^2 (t, \theta) \right] + \psi (N, p) \tag{2.4}$$

where $\psi (N, p)$ is a function of the number of data points ($N$) and the number of parameters in the model ($p$). Two popular criteria are Akaike's Information Criterion and the Final Prediction Error (Soderstrom and Stoica, 1989).

**Akaike's Information Criterion:**

$$AIC = \log \left[ \sum_{t=1}^{N} \varepsilon^2 (t, \theta) \right] + \frac{2p}{N} \tag{2.5}$$

**Final Prediction Error:**

$$FPE = \left[ \sum_{t=1}^{N} \varepsilon^2 (t, \theta) \right] \frac{1 + \frac{p}{N}}{1 - \frac{p}{N}} \tag{2.6}$$

The developments presented in the remainder of this section are meant only as an introduction to the topic of selecting the amount of historical data to use in the model updater, and it is not within the scope of thesis to find an analytical solution. The structure of the equation and penalty functions will be justified through considerations that are important to this thesis and the small case study of the Williams-Otto reactor, discussed further in this chapter (Williams and Otto, 1960).

Using an expression similar to Equation (2.3), the optimal number of historical data points could be determined by minimizing an expression that in some way penalizes the decrease in uncertainty of the decision variables (as measured by the volume of the confidence region: $\det \mathbf{Q_x}$) that results from increasing the computational load ($\psi_1$) and decreasing the ability of the RTO system to track changes in the process operation ($\psi_2$).

Although A-optimality is selected for the experimental design criteria (as discussed later in this chapter), the determinant of the decision variable covariance matrix is used here to measure the change in the volume of the confidence region. The determinant of a covariance matrix is a common method, that has been used by many previous authors, to measure the volume of the confidence region (Box and Lucas, 1959; Kiefer and Wolfowitz, 1959; Box and Hunter, 1965; Pinto *et al.*, 1990 and 1991). The inconsistency is justified because the experimental design for this thesis is focused on the volume and shape of the confidence interval; however, the application of selecting the number of historical points should only focus on the volume since it cannot guarantee that $N$ points chosen at random will have an impact on shape. To ensure the separability of the various effects, the penalty terms were introduced in an additive form

$$W = [1 + \psi_1 (N, r, p) + \psi_2 (N)] [\det \mathbf{Q_x}] \tag{2 7}$$

where $\mathbf{Q_x}$ is the estimated covariance matrix of the decision variables.

The computational load required for the least squares regression algorithm increases proportionally to the number of equations used to solve for the parameter estimates.

$$Comp.Load \propto N \times r \tag{2.8}$$

This thesis will decrease this penalty as the number of estimated parameters increases. However, please note that this decision criterion was not used for selecting the number of adjustable parameters, only $N$. A larger value for the ratio,

$$\frac{\text{\# of Equations}}{\text{\# of Adjustable Parameters}}$$

helps reduce the amount of uncertainty transmitted from the measurement noise to the adjustable parameter estimates. To accommodate for this, the term $\frac{1}{p}$ will be included in the penalty function (see Proportionality (2.9)). Note other possible expressions include $N \times r$ or $p(N \times r)$, which is a viable choice since more adjustable parameters will increase the computational load.

$$\psi_1 \propto \frac{N \times r}{p} \tag{2.9}$$

If the proportionality constant is set to one,

$$\psi_1(N, r, p) = \frac{N \times r}{p} \tag{2.10}$$

The selection of a linear form for the penalty function, with respect to $N$, is justified by examining the results of the Williams-Otto reactor case study of section 2.4. Figure 2.1, shows that as the number of historical data points is increased, the volume of the confidence region for the decision variables decreases, as expected, and the computational load increases linearly over the range of $N$ investigated. To evaluate the computational load required of the model updater, the number of floating-point operations (FLOPS) was counted using the *flops* function in the computer software package MATLAB ver. 5.3.0 (The Math Works, 1997).

The reduction in the ability of the RTO system to track changes in the process behaviour is solely dependent on the number of historical data points used by the model updater. As the number of historical data points increases, the model updater is put to 'sleep' as its sensitivity to new points is reduced. This thesis will make the case that points further in the past should be penalized more heavily than recent information because it becomes more probable that the operating conditions have changed, and the information provided is not well-suited to describe the behaviour at the current operating state. A possible proportionality expression is described by Inequality (2.11). Ensuring that $b > 1$, will provide the desired effect of increasing the value of $\psi_2$ exponentially as $N$ gets larger. It is noted that other functions that produce the same result include: $b^N$ and $e^N$:

$$\psi_2 \propto N^b \tag{2.11}$$

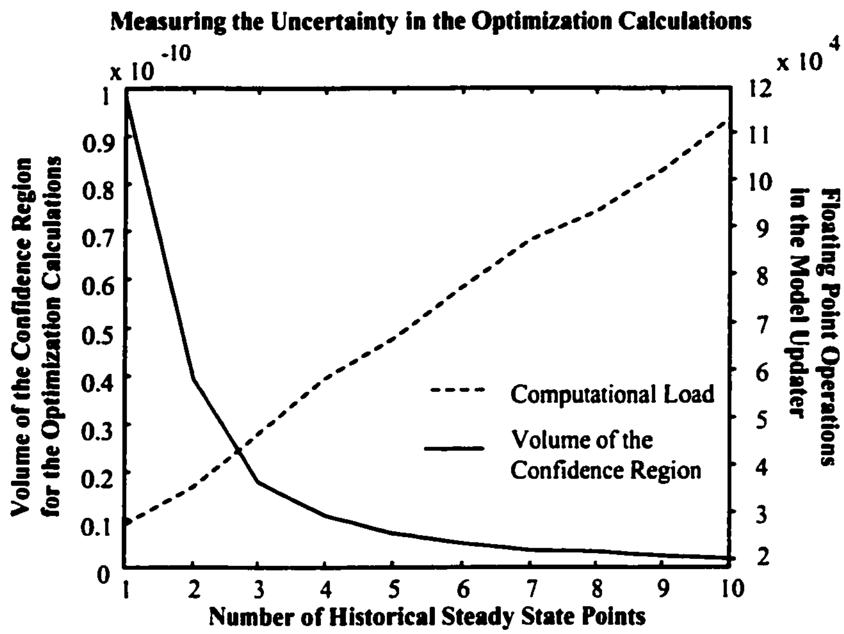



Figure 2.1: The uncertainty in the optimization calculations decreases, and the computational load increases linearly as the number of historical steady-state points used in the model updater increases.

If the process behaviour does not change relatively often, then the penalty function described by Inequality (2.11) should be less restrictive and it would be ideal to use more historical points. Using the Williams-Otto reactor, values of $b = 3$ and a proportionality constant equal to $\frac{1}{3}$, provided results where the number of points is reasonable,

$$\psi_2(N) = \frac{N^3}{3} \tag{2.12}$$

A reasonable range is considered to be two - ten historical points, since exceeding ten would be unreasonable for an RTO system to track changes in the process behaviour, since the data would be describing the process behaviour from several days before the calculation is made.

The calculation to determine the optimal number of historical points may have to be repeated if the operation of the process changes significantly, or if the set of equations in the model updater subsystem is changed, particularly if a simplified model structure is used to describe the local behaviour of the process.

Although a least squares regression method was used for this case study, the consideration of the developments in this section are not specific to this parameter estimation algorithm.

## 2.2 Parameter Covariance Matrix Estimation

This section discusses methods for estimating the adjustable parameter covariance matrix and to combine it with the sensitivity of the optimization variables to the adjustable parameters, to produce an estimate of the decision variable covariance matrix. The selected method will be used in an experimental design to select the next point that is expected to generate the most information possible within a specified experimental window.

Two methods will be discussed in this section that take different approaches in estimating the uncertainty for the adjustable parameter estimates. One method was developed for RTO applications and the other for experimental design purposes. The approach by Forbes and Marlin (1996) can be extended to estimating the decision variable covariance matrix ($\mathbf{Q}_x$), by using the sensitivity of the optimization variables to the adjustable parameters, $\left(\frac{d\mathbf{x}_p^*}{d\boldsymbol{\beta}}\right)$, as a weighting matrix to the parameter covariance matrix ($\mathbf{Q}_{\beta}$). Their method will be discussed further in this section:

$$\mathbf{Q}_{\mathbf{x}} = \left(\frac{d\mathbf{x}_p^*}{d\boldsymbol{\beta}}\right) \mathbf{Q}_{\beta} \left(\frac{d\mathbf{x}_p^*}{d\boldsymbol{\beta}}\right)^T \tag{2.13}$$

The use of the estimate of the parameter covariance matrix in the experimental design literature was originally developed by Box and Lucas (1959). This result was also achieved by Bard (1974), who took a different approach. The estimate of the parameter covariance matrix found in the experimental design literature has not been used for any RTO applications, but has been applied to the development of many sequential experimental designs (Box and Hunter, 1965; Draper and Hunter, 1966; Draper and Hunter, 1967 a and b; Pinto *et al.*, 1990). This approach will be discussed in detail later in this section.

Forbes and Marlin (1996) developed their estimate of the decision variable and parameter covariance matrix using sensitivity analysis to track how the process measurement noise propagated through the closed-loop RTO system. Their approximation has been used for the following RTO applications: design cost, results analysis and optimal sensor selection (Forbes and Marlin, 1996; Miletic and Marlin, 1998; Fraleigh, 1999; Zhang and Forbes, 2000; Zhang *et al.*, 2001). Other optimization researchers have used sensitivity analysis for their applications as well. Koninckx (1988) investigated how the uncertainty in the decision variables affects the evaluation of the expected gains in profit and the accuracy of the predicted profit and decision variables. Krishnan *et al.* (1992) examined how uncertainty from the measurements is forwarded to the adjustable parameter estimates, in order to select which measurements have the greatest influence on the updated parameters. Of significance to this thesis, Fraleigh (1999) used this estimate with D and T-optimal criteria to find an optimal sensor system, and Miletic and Marlin (1998) used this result to estimate the confidence region of the decision variables.

A practical advantage to the approach proposed by Forbes and Marlin (1996) is that it estimates the parameter covariance matrix from the measurement noise, where the development in the experimental design literature uses the residuals of the model equations, which are not as accessible. This advantage along with the desired consistency of using linear sensitivity methods to calculate the entire decision variable covariance matrix and the documented success of the approximation developed by Forbes and Marlin (1996) for RTO applications, make it the preferred approach to estimate the covariance matrix of the adjustable parameters and decision variables for this thesis.

This section will conclude with a detailed discussion of the two approximations mentioned.

### 2.2.1 Covariance Estimates (RTO Approach)

In an RTO application, the variation in the calculated optimization variables and the adjustable parameter estimates is due to uncertainty in the process measurements propagating through the closed-loop RTO system (Forbes and Marlin, 1996). Figure 2.2 shows the sensitivities governing the propagation of uncertainty through the RTO subsystems.

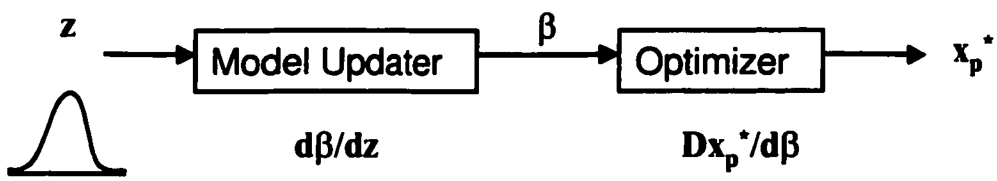



Figure 2.2: Tracking the uncertainty through the RTO subsystems.

Forbes and Marlin (1996) showed that a closed-loop RTO system, as shown in Figure 1.2, can be represented by a linear approximation which is written in terms of the deviations in the calculated setpoints from the applied setpoints and the subsystem sensitivities. However, since this thesis is limited to a discussion about reducing uncertainty in the optimization calculations caused by sensor noise and model uncertainty, it is assumed that only measurement variance is propagating through the RTO system. Therefore the closed-loop system can be reduced to an open-loop approximation using a one-step ahead approximation of the setpoint covariance matrix as shown in Equation (2.14):

$$\mathbf{Q_x} = \frac{\partial \mathbf{x}_p^*}{\partial \mathbf{z}} \mathbf{U} \frac{\partial \mathbf{x}_p^*}{\partial \mathbf{z}}^T \tag{2.14}$$

which is rewritten as,

$$\mathbf{Q_x} = \frac{\partial \mathbf{x}_p^*}{\partial \beta}\frac{\partial \beta}{\partial \mathbf{z}} \mathbf{U} \left(\frac{\partial \mathbf{x}_p^*}{\partial \beta}\frac{\partial \beta}{\partial \mathbf{z}}\right)^T \tag{2.15}$$

where **U** is the covariance matrix of the measurement noise sampled from a Gaussian distribution and having zero mean.

This assumption has been made previously by other researchers (Miletic and Marlin, 1998; Fraleigh, 1999; Zhang *et al.*, 2001).

If only the uncertainty in the adjustable parameters is of interest, then the sensitivity of the decision variables to the parameter estimates can be removed from Equation (2.15) to produce the approximation,

$$\mathbf{Q}_{\beta} = \frac{\partial \boldsymbol{\beta}}{\partial \mathbf{z}} \mathbf{U} \left( \frac{\partial \boldsymbol{\beta}}{\partial \mathbf{z}} \right)^{T} \tag{2.16}$$

Equation (2.16) describes how the uncertainty in the process measurements propagates through the model updater subsystem, to contribute to uncertainty in the adjustable parameters.

Linear sensitivity analysis is used to represent the RTO subsystem sensitivities. It is a tool that provides an approximation of the variability of the RTO results based on the local sensitivity of the parameter updater and the economic optimizer. These sensitivities depend on the definition of the optimization problems (*i.e.*, parameter estimation and model-based optimization) and the resulting Karush-Kuhn-Tucker conditions (Ganesh and Biegler,1987). A description of the sensitivities is given in *Appendix A*.

### 2.2.2 Covariance Estimates (Experimental Design Approach)

The estimate of the covariance matrix for the adjustable parameters, originally developed by Box and Lucas (1959), is based on the posterior distribution of the parameter estimates. Draper and Hunter (1966) expanded on the original work to account for multiple-response models. Draper and Hunter (1967 a and b) adapted their estimate to include, when available, *a priori* information of the parameters in the form of a multivariable normal distribution for single-response nonlinear models ( Draper and Hunter, 1967a) and multiple-response nonlinear models (Draper and Hunter, 1967b).

The expression used to estimate the covariance matrix of the adjustable parameter estimates in the sequential multiple response experimental designs developed by Draper and Hunter (1967 b) and Pinto *et al.* (1990) is,

$$\mathbf{Q}_{\beta} = \sum_{i=1}^{r} \sum_{j=1}^{r} \sigma^{ij} \mathbf{X}_i' \mathbf{X}_j + \mathbf{V}_o^{-1} \tag{2.17}$$

where

$$x_{iu}^{(t)} = \left[ \frac{\partial f_i \left( y_{iu}^{(v)}, \xi_{iu}^{(q)}, \theta_p \right)}{\partial \theta_t} \right]_{\theta=\hat{\theta}} \tag{2.18}$$

and

$$\mathbf{X}_i = \begin{bmatrix} x_{i1}^{(1)} & x_{i1}^{(2)} & \cdots & x_{i1}^{(p)} \\ x_{i2}^{(1)} & x_{i2}^{(2)} & \cdots & x_{i2}^{(p)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{im}^{(1)} & x_{im}^{(2)} & \cdots & x_{im}^{(p)} \end{bmatrix} \tag{2.19}$$

and $\mathbf{X}$ is the incremental design matrix of partial derivatives and $\mathbf{V}_o$ is the past estimate of the parameter covariance matrix. Since the result will not be used in this thesis, a full description of the development of this approximation is presented in *Appendix B*.

## 2.3 Experimental Design Algorithms

This section will discuss the different alphabetic experimental design criteria, and will provide justification for the selection of the A-optimal criterion for use in this thesis.

The final form of the experimental design is defined by Problem (2.20). The experimental design is required to select the next point that will minimize the trace of the estimate of the decision variable covariance matrix, with the vector of decision variables at observation $N+1$ as the optimization variables.

$$\min_{\xi_{N+1}^{(q)}} \quad trace\left(\mathbf{Q_x}\right) \tag{2.20}$$

$$\begin{aligned} s.t. \quad \mathbf{f}\left(\mathbf{y}_{N+1}^{(v)}, \boldsymbol{\xi}_{N+1}^{(q)}, \widehat{\boldsymbol{\theta}}_p\right) &= 0 \\ \mathbf{g}\left(\boldsymbol{\xi}_{N}^{(q)}, \boldsymbol{\xi}_{N+1}^{(q)}\right) &\leq a_q \end{aligned}$$

where the first set of constraints represents the steady-state model equations and the second set defines the experimental window.

The inequality constraints are present to define the experimental window (or trust regions), which prevent the system from being subjected to drastic process changes. The solution of the optimization problem will provide a design point, contained within the experimental window, that will produce high quality steady-state information to reduce uncertainty in the decision variables.

### 2.3.1 Variance-Optimal Designs

Although this thesis is not concerned with linear experimental designs, since the models used for RTO applications are mostly nonlinear in the parameters, an introduction explaining the different optimal design criteria is best provided with linear models.

The objective for this thesis is to generate a sequential experimental design. Several design criteria based on the information matrix and their experimental objectives are described below (Box and Draper, 1987; Myers and Montgomery, 1995). For clarity purposes, first consider the model,

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{2.21}$$

where $\mathbf{Y}=\left[\mathbf{Y}_1^T\,\mathbf{Y}_2^T\right]^T$, $\mathbf{X}=\left[\mathbf{X}_1^T\,\mathbf{X}_2^T\right]^T$, $\boldsymbol{\varepsilon}=\left[\boldsymbol{\varepsilon}_1^T\,\boldsymbol{\varepsilon}_2^T\right]^T$ and $(\mathbf{Y}_2, \mathbf{X}_2)$ is the new observation.

**D-Optimality**: D-optimal designs, originally developed by Kiefer and Wolfowitz (1959), are the most common and widely studied of the alphabetic optimal designs (Myers and Montgomery, 1995). Their objective is to minimize the generalized variance of the parameter estimates by maximizing the determinant of the information matrix:

$$\max \left|(\mathbf{X}^T\mathbf{X})\right| \tag{2.22}$$

Geometrically, the volume of the confidence ellipsoid of the parameter estimates is proportional to $\left|\mathbf{X}^T\mathbf{X}\right|^{-1/2}$. Therefore the D-optimal design is considered to minimize the volume of the confidence ellipsoid.

This design is augmented by calculating the next point, within the design region, which maximizes Equation (2.22).

**A-Optimality**: A-optimality deals strictly with the individual variances of the parameter estimates. Whereas the D-optimal design focuses on minimizing the volume of the confidence region, the A-optimal design focuses on both minimizing the volume and creating a more symmetrical shape by equalizing the uncertainty in all the parameter estimates:

$$\min \quad \text{trace } \left(\mathbf{X}^T\mathbf{X}\right)^{-1} \tag{2.23}$$

or from Myers and Montgomery (1995),

$$\max \quad \text{trace } \left(\mathbf{X}^T\mathbf{X}\right) \tag{2.24}$$

To augment the optimal design with an extra point, all that is required is to maximize the trace of the information matrix of the new observation,

$$\max \quad \text{trace } \left(\mathbf{X}_2^T\mathbf{X}_2\right) \tag{2.25}$$

This concept is derived based on the following property,

$$\text{trace } \left(\mathbf{X}^T\mathbf{X}\right) = \text{trace } \left(\mathbf{X}_1^T\mathbf{X}_1\right) + \text{trace } \left(\mathbf{X}_2^T\mathbf{X}_2\right) \tag{2.26}$$

**E-Optimality**: The objective of E-optimality is to minimize the variance of the most poorly estimated coefficient. This design is often called the shape design because of its focus on equalizing the amount of uncertainty for each parameter,

$$\min \left[\max \lambda \left\{\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\right\}\right] \tag{2.27}$$

Heiberger *et al.* (1993) indicated that once all the eigenvalues are equal, the design is orthogonal.

This design is augmented by calculating the next point, within the design region, which maximizes Equation (2.27).

**G-optimality**: The primary focus of G-optimality is to improve the prediction of the response, $E(y)$. It operates by minimizing the maximum prediction variance over the design region. For the non-sequential application, $n$ experimental conditions are chosen to minimize the maximum prediction variance,

$$\min\left[\max\left\{\mathbf{x}^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}\right\}\right] \tag{2.28}$$

where $\mathbf{X}$ is manipulated and has row dimension $n$, and $\mathbf{x}$ is fixed as the location in the design space that is being predicted. For experimental design purposes $\mathbf{x}$ is the center of the design space. For sequential designs, the data set used to define $\mathbf{X}$ is augmented by the next experimental point. Dette and Wong (1995) suggest that this design is a good choice when it is desired to estimate the entire response surface, as it provides global protection against unreliable estimates at points in the design space.

**U-optimal**: U-optimal designs were developed by Heiberger *et al.* (1993) as an improvement to the other more traditional A, D, E and G optimal experimental designs. They have shown through formal definition and proofs that U optimality implies D, A, and E optimalities (Heiberger *et al.*, 1992).

The authors show the best strategy is to begin augmenting the design, with an approach similar to E-optimality, with points that increase the smaller eigenvalues of the information matrix until all the eigenvalues are equal. Once this is the case, the design will be orthogonal and any further new points should be selected from any other balanced design. However, U-optimal designs have not received much attention in the literature so their efficiency compared to other experimental designs is unknown. Since the objective of this thesis is to only augment the design with one point, this design is essentially reduced to an E-optimal design.

Although D-optimal designs, and to a lesser extent A and E optimal designs, have received the most attention in the literature, some authors believe that experimenters do not make use of the prediction variance used in G-optimality as much they should (Welch, 1983; Myers and Montgomery, 1995). However, as previously discussed, the purpose of the experimental design for this thesis is to develop a function that is able to measure the amount of information that will reduce uncertainty in the parameter estimates, and to a lesser extent to reduce the prediction variance over the entire response surface, therefore making the G-optimal design less desirable when comparing it to other experimental designs.

The next section will discuss the justification for selecting the A-optimal criterion for this application.

### 2.3.2 Comparing the Alphabetic Design Criteria

There have been some comparisons to distinguish which of the alphabet designs is generally the most efficient. Considering linear models, Wong (1994) examined the robustness properties of the A, D, E, and G-optimal designs by comparing the efficiencies of each design under various model assumptions. Wong (1994) stresses the importance of selecting a design that is deemed adequate for several optimality criteria. Kiefer (1975) shares this same opinion in writing that it does not make sense to select a design based on precise calculations of one particular criterion and not examine how it performs in other respects.

Wong (1994) showed that the experimenter should be cautious when using an E-optimal design, since a small amount of structural error in the model can result in a severe loss in efficiency of the design. The A-optimal designs, although not to the same degree, also showed problems in handling model mismatch. The G and D-optimal designs were able to provide protection from deviations of the true model.

The alphabet efficiencies are measures of how the results of an experimental design compare to the theoretical optimal design that would be produced by the criterion specified by the letter efficiency. Therefore a design created using an E-optimal criterion would have an E-efficiency of 1. Fedorov (1972) provides the background for the computation of D and G-efficiencies, while Pukelsheim and Torsney (1991) and Wong (1994) discuss how to calculate the A and E-efficiencies, respectively. Table 2.1 provides a summary of the results reached by Wong (1994), where a design is considered very efficient if it has an efficiency above 0.9, and efficient if its efficiency is between 0.8 and 0.9.

Table 2.1: Comparing Optimal Designs

| **Optimal Design** | **A-Efficiency** | **D-Efficiency** | **E-Efficiency** | **G-Efficiency** |
|---|---|---|---|---|
| A | | v. efficient | v. efficient | not efficient |
| D | efficient | | not efficient | v. efficient |
| E | v. efficient | efficient | | not efficient |
| G | not efficient | efficient | not efficient | |

Wong's results showed that G-optimal designs should be used sparingly since they are only compatible with D-efficiency, although they do posses good protection when there are departures from the true model. E-optimal results showed the opposite effect in that they are compatible with all criteria except G-efficiency, but they do not perform well when the assumed model and the true model are not similar.

In summary, Wong (1994) showed the best designs were A and D-optimal designs. The A-optimal design showed slightly better results when comparing the efficiencies between other designs. However, the D-optimal design also performs well and is stated to be slightly better equipped to handle larger amounts of plant/model mismatch (Wong, 1994).

Pinto *et al.* (1990,1991) compared A, D, and E-optimal design criteria for experimental designs where a prior distribution of the model parameters was known and unknown for multiple response nonlinear models. Pinto *et al.* (1990,1991) concluded that the A-optimal design provided the best parameter estimates with the smallest parameter uncertainty in the examples studied. They noted that using the D-optimal criterion was good for prediction-orientated problems but was not well-suited for general parameter estimation, which is more consistent with the objectives of this thesis. They found that the D-optimality criterion reduced the volume of the confidence region faster, but at the expense of reducing the uncertainty of one parameter while the uncertainty of the other parameters remained relatively large (Pinto *et al.*, 1990). As mentioned previously, the volume as well as the shape of the confidence region is important to this thesis because the primary need of the optimizer is to have a good estimate of the direction of steepest improvement, and is less concerned with the prediction abilities of the model.

This thesis will use an A-optimal criterion for the experimental design, because it focuses on the volume and shape of the confidence region, and previous authors have shown it to be an effective approach for applications focusing on parameter estimation and uncertainty problems (Pinto *et al.*, 1990; Pinto *et al.*, 1991; Wong, 1994).

## 2.4 Williams-Otto Reactor Case Study

This series of case studies will examine the following issues:

- the effect of including historical steady-state information in the model updater of the RTO system;
- the ability of an experimental design to generate information to reduce uncertainty in the optimization calculations.

The first case study will show how to select the optimal number of past steady-state data points to balance the effects of reducing variable uncertainty with increased computational load and the weakened ability of the model updater to track changes in the process behaviour. The second case study will show how generating and implementing the result from an experimental design is able to reduce the decision variable uncertainty further than implementing the result predicted by the economic optimizer.

### 2.4.1 Process Description

The Williams-Otto reactor, shown in Figure 2.3, is one unit of the Williams-Otto plant model, which also contains a heat exchanger, decanter and a distillation column (Williams and Otto, 1960). The Williams-Otto reactor has been widely used by a number of researchers to test their applications (Roberts, 1979; McFarlane and Bacon, 1989; Forbes *et al.*, 1994; Forbes and Marlin, 1996; Miletic and Marlin, 1998; Fraleigh, 1999; Zhang and Forbes, 2000; Zhang *et al.*, 2001).

The reactor is a continuously stirred tank reactor (CSTR), with mass holdup $V_R = 4640$ lb and temperature $T_R$. The reactor is fed with two pure component reactant streams, $FA$ and $FB$. Stream flow rate $FA$ is held constant at $14,500$ lb/hr. There are three simultaneous reactions that occur involving six components,

$$\begin{aligned} A + B &\rightarrow C \\ B + C &\rightarrow P + E \\ C + P &\rightarrow G \end{aligned}$$

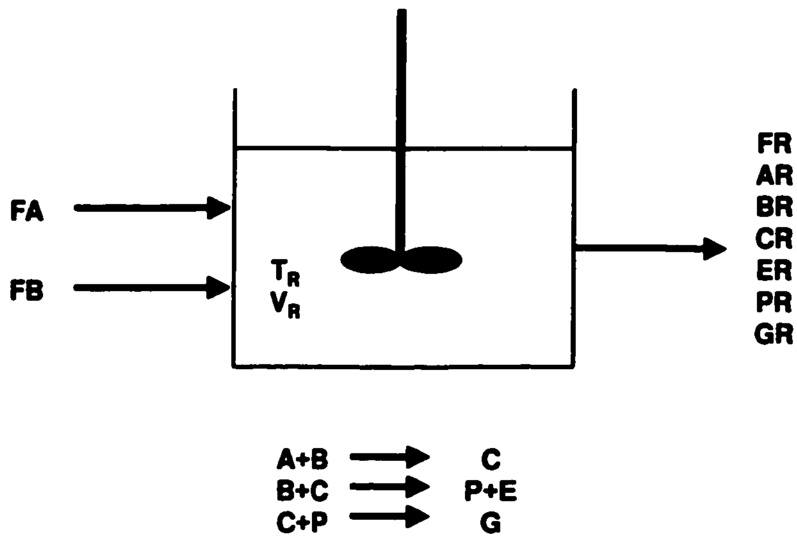



Figure 2.3: Williams-Otto reactor (Williams and Otto, 1960).

The kinetics of each reaction are represented by the Arrhenius expression,

$$k_i = A_i \exp\left(-B_i/T_R\right) \tag{2.29}$$

where $i$ represents the reaction equation, $k$ is the rate of reaction (hr$^{-1}$), $A$ is the frequency factor (hr$^{-1}$), and $B$ is the activation energy ($R$). The Arrhenius constants for the reactions are presented in Table 2.2.

Table 2.2: Reaction Constants

| Reaction Number | Frequency Factors (hr$^{-1}$) | Activation Energy (R) |
|---|---|---|
| 1 | $5.9755 \times 10^{9}$ | 12,000 |
| 2 | $2.5962 \times 10^{12}$ | 15,000 |
| 3 | $9.6283 \times 10^{15}$ | 20,000 |

For this case study, all ten variables are considered measurable (flows, temperatures, component compositions). The frequency factors ($A_i$) will be the parameters updated on-line and for the purposes of this case study the activation energy constants ($B_i$) are assumed to be known. The decision variables for optimization will be the flow of reactant $B$ into the reactor ($FB$) and the temperature of the reactor ($T_R$).

The steady-state model includes the overall mass balance and the six component mass balances. For simplicity, the energy balances have not been included.

$$
\begin{aligned}
FA + FB - FR &= 0 \quad (2.30)\\
\frac{FA}{V_R} - \frac{FR}{V_R}(AR) - k_1(AR)(BR) &= 0 \quad (2.31)\\
\frac{FB}{V_R} - \frac{FR}{V_R}(BR) - k_1(AR)(BR) - k_2(BR)(CR) &= 0 \quad (2.32)\\
-\frac{FR}{V_R}(CR) + 2k_1(AR)(BR) - 2k_2(BR)(CR) - k_3(CR)(PR) &= 0 \quad (2.33)\\
-\frac{FR}{V_R}(ER) + 2k_2(BR)(CR) &= 0 \quad (2.34)\\
-\frac{FR}{V_R}(GR) + 1.5k_3(CR)(PR) &= 0 \quad (2.35)\\
-\frac{FR}{V_R}(PR) + k_2(BR)(CR) - 0.5k_3(CR)(PR) &= 0 \quad (2.36)
\end{aligned}
$$

where $R$ represents the stream exiting the reactor. This model uses the assumption that the tank is well mixed so the composition in the exit stream is equal to the composition in the tank.

The instantaneous profit function is defined in Equation (2.37) and is the same as the one used by Miletic and Marlin (1998).

$$
P(x,u) = 70000.3(FA + FB)PR + 1586.9(FA + FB)ER - 4667.0FA - 7000.2FB \quad (2.37)
$$

The profit response surface and contours are shown in Figures 2.4 and 2.5. The optimal point occurs at $[FB, TR] = [0.3799, 0.6531]$ for a maximum profit of \$928 /s.

### 2.4.2 Method

The problem was scaled to ensure a more reliable solution for the optimization and equation solving routines. This transformation also helps the statistical estimation of parameters in the kinetic expression by reducing correlation (Watts, 1994). The kinetic expressions were manipulated as follows,

$$
\begin{aligned}
k_i &= A_i \exp\left[-B_i / \left(T_R - (T_{Ro})_i\right)\right] \quad (2.38)\\
&= A_i \exp\left(B_i / (T_{Ro})_i\right) \exp\left(-B_i / T_R\right) \quad (2.39)
\end{aligned}
$$

with the scaled reference temperatures,

$$
\begin{aligned}
(T_{Ro})_1 &= 0.9022 \quad R\\
(T_{Ro})_2 &= 0.7907 \quad R\\
(T_{Ro})_3 &= 0.7434 \quad R
\end{aligned}
$$

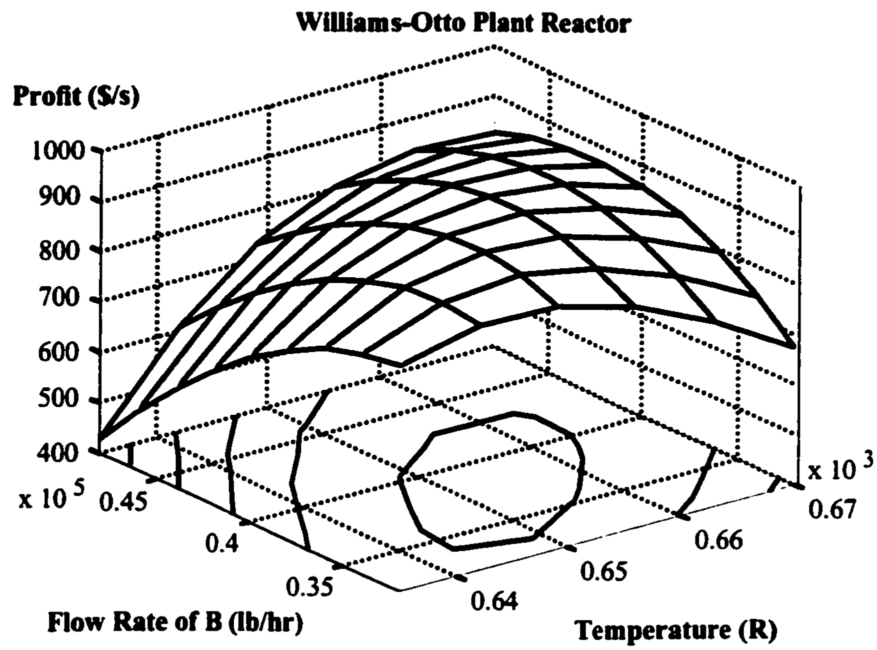



Figure 2.4: Surface plot of the Williams-Otto reactor.

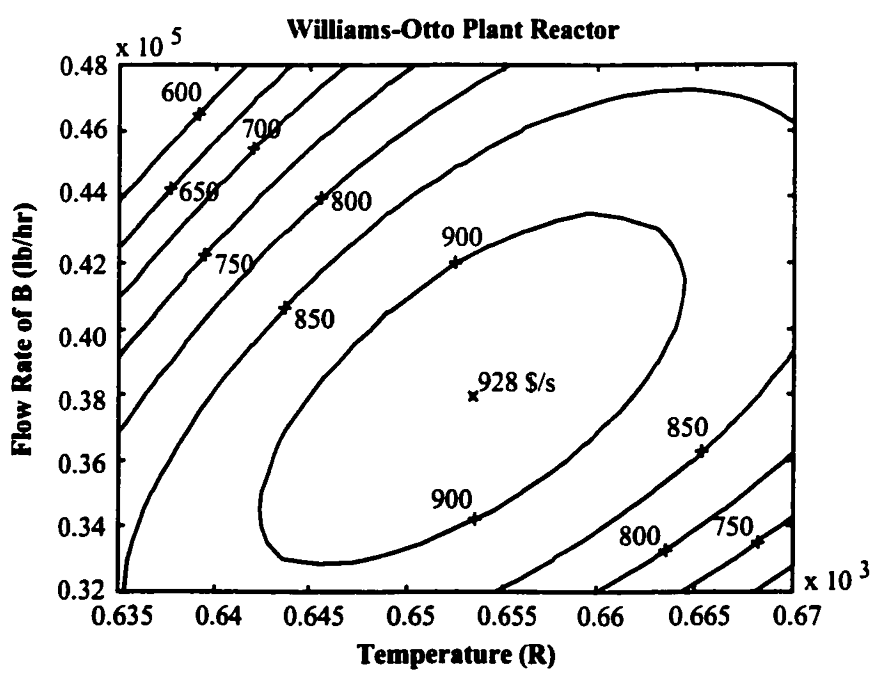



Figure 2.5: Contour plot of the Williams-Otto reactor.

The frequency factors were scaled by dividing by $1,000$, so the true values of the adjustable parameters become,

$$\begin{aligned} A_1 &= 10 \\ A_2 &= 15 \\ A_3 &= 20 \end{aligned}$$

The following variables were further scaled as shown in Table 2.3.

Table 2.3: Variable Scaling for Williams-Otto Reactor

| Variable | Symbol | Scaling Factor |
|---|---|---|
| Flows | FA, FB | divided by 100,000 |
| Temperatures | $T_R$ | divided by 1,000 |
| Reactor Volume | $V_R$ | divided by 100 |
| Arrhenius Exponent | $B_1$, $B_2$, $B_3$ | divided by 1,000 |

In order to generate the measurement covariance, it is assumed that *a priori* knowledge of the covariance structure for the measurements is unavailable. The measurement covariance matrix for the compositions of materials exiting the reactor was found by generating a set of $1,000$ process data points around a nominal value with a white noise distribution added to each data point for each measurement. Table 2.4 shows the true applied standard deviations of the measurements, expressed as a percentage of the nominal case. The nominal point and measurement covariance matrix can be found in *Appendix C.* Although this is not strictly correct for this case, it creates a study of mismatch in the covariance structure. This thesis allows for future discussion to evaluate the benefits of integrating knowledge of the covariance structure into the RTO system.

Table 2.4: Standard Deviations of the Measurements

| Measurement | Standard Deviation |
|---|---|
| Flows | 1.0 % |
| Temperatures | 0.5 % |
| Compositions | 3.0 % |

The initial data set, which represented historical operation before the case studies were initiated, consisted of ten steady-state operating points selected randomly from a uniform distribution in the range $[FB, TR] = [40,000, 42,000] \times [640, 645]$. This data set is found in *Appendix C.*

The first case study investigated the advantages and disadvantages of using information from past steady-state operating points in the model updater stage. A maximum of ten historical points were considered, since exceeding ten would be unreasonable for an RTO system to track changes in the process behaviour, as the data would be describing the process behaviour from several days before the calculation was made. For each set of data, the model updater used a least squares regression routine to calculate the adjustable parameter estimates. The updated model was used by the economic optimizer to calculate the new optimal point. The estimated covariance matrices of the adjustable parameters and the optimization calculations were evaluated using linear sensitivity analysis (see *Appendix A*). The volumes of the confidence regions were compared to determine the transmission of noise from the measurements to the adjustable parameters and the decision variables. The criterion used in Equations (2.7), (2.10), and (2.12) was used to determine the optimal number of historical points that should be used for this case study.

The result from the first case study was used in the second case study to compare the uncertainty of a result of an economic optimization calculation after the initial data set has been augmented with a new point from: 1) an economic optimization calculation; and 2) an experimental design optimization calculation. Similar to the first case study, the determinant of the covariance matrix was used to measure the volume of the confidence region of the decision variables.

The experimental design problem requires the value of the matrix representing the sensitivity of the decision variables to the adjustable parameters, $\frac{\partial \mathbf{x}}{\partial \beta}$, which is only obtained from the results of an economic optimization calculation. Therefore the matrix, $\frac{\partial \mathbf{x}}{\partial \beta}$, used in the computation of the experimental design problem is estimated from a previous calculation. This approximation is adequate, since unless significant changes occur in the process operation, the matrix $\frac{\partial \mathbf{x}}{\partial \beta}$ does not need to be recalculated (Miletic and Marlin, 1998). Therefore the sensitivity of the optimization variables to the adjustable parameters can be considered as a weighting matrix for the experimental design calculation. For the case study, an experimental window of $[\pm 0.05, \pm 0.020]$ surrounding the current operating point was defined. A large experimental window was selected for this case study to effectively compare the results to the unconstrained economic optimization problem. A more practical problem with a smaller experimental window is discussed in Chapter *3*.

### 2.4.3 Results and Discussion

This section concludes with the presentation and discussion of the results for the case studies involving the Williams-Otto reactor. The first case study is performed to determine the number of historical steady-state points that should be used in the model updater subsystem. The second case study investigates the use of DOE in the RTO system. Please see *Appendix C* for the results of the optimization routines and the covariance matrices.

#### Historical Data

The results of the first case study show that, for the criterion used in Equation (2.7), the optimal number of historical steady-state points that should be used for updating the adjustable parameter estimates is three. The structure of the decision criterion is held constant while the number of historical data points is adjusted. The decision criterion is based on

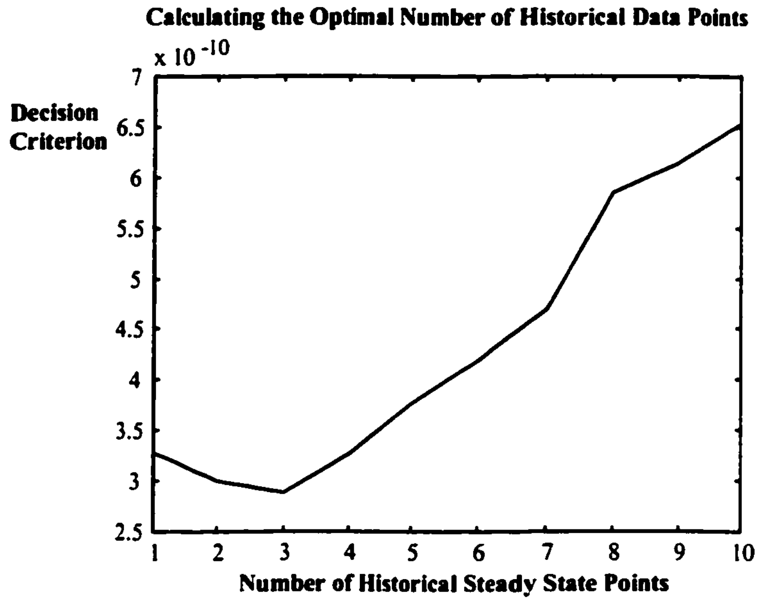



Figure 2.6: The optimal number of historical steady-state date points to be used by the model updater subsystem for the Williams-Otto reactor is three.

the following three considerations: reducing uncertainty in the optimization calculation, limiting the computational load, and ensuring that the RTO loop is able to effectively track changes in the process behaviour. Figure 2.1 shows that by using more historical steady-state information, the uncertainty associated in the optimization calculations is reduced. Figure 2.1 also shows how the computational load of the model updater increases linearly as more historical points are used. This result is consistent with the structure of the penalty function, represented by Equation (2.10), which predicted the computational load as a linear function of $N$ (number of steady-state points). However, Figure 2.6 illustrates that the benefit of this further reduction in the decision variable confidence region is minimal as the influence of the penalty functions become more significant.

Table 2.5 compiles the data illustrated in Figures 2.1 and 2.6, as well as devoting an extra column to show the changes in the volume of the confidence region of the adjustable parameters.

There are many methods available to solve the nonlinear economic optimization for this case study (Edgar and Himmelblau, 1988). The method applied for this thesis was the Sequential Quadratic Programming (SQP) algorithm in the optimization toolbox for the computer software package MATLAB ver. 5.3.0 (Branch and Grace, 1996).

### DOE Use

The second case study introduces the potential benefits of using experimental design for RTO applications, by showing that a DOE calculation does reduce uncertainty in future

Table 2.5: Results Summary for Optimal Historical Point Selection for the Williams-Otto Reactor

| Historical Data Points | det ($\mathbf{Q}_\beta$) | det ($\mathbf{Q}_x$) ($\times 10^{-10}$) | Regression (FLOPS) | Decision Criterion (Equation 2.7) ($\times 10^{-9}$) |
|---|---|---|---|---|
| 1 | 0.8026 | 0.9835 | 27158 | 0.3278 |
| 2 | 0.1815 | 0.3922 | 34917 | 0.3007 |
| 3 | 0.0632 | 0.1798 | 46037 | 0.2876 |
| 4 | 0.0242 | 0.1075 | 57838 | 0.3261 |
| 5 | 0.0120 | 0.0716 | 66032 | 0.3773 |
| 6 | 0.0073 | 0.0492 | 76845 | 0.4185 |
| 7 | 0.0053 | 0.0363 | 86947 | 0.4693 |
| 8 | 0.0038 | 0.0312 | 92669 | 0.5851 |
| 9 | 0.0027 | 0.0234 | 101634 | 0.6139 |
| 10 | 0.0020 | 0.0184 | 112299 | 0.6537 |

economic optimization calculations. Two tests were performed using the same initial data set as the first case study. The first test consisted of performing an economic optimization calculation and the second completed an experimental design calculation. Each test implemented the result from the first RTO cycle, and the RTO loop was run a second time with an economic optimization completed in both cases. Table 2.6 shows the volume of the confidence region of the economic optimization calculation for the second RTO cycle is 29% smaller in the second test where the result of the experimental design was implemented in the first RTO cycle. However, Figure 2.7 shows that the result from the experimental design is farther from the true optimum than the result from the economic optimizer, indicating that a trade-off needs to be made between information generation and profit realization.

Table 2.6: Results Summary for Benefits of Experimental Design

| Economic Optimization Calculation | det ($\mathbf{Q}_\beta$) | det ($\mathbf{Q}_x$) ($\times 10^{-10}$) |
|---|---|---|
| First RTO Interval | 0.0632 | 0.1798 |
| Second RTO Interval (after Economic Optimization) | 0.0858 | 0.2263 |
| Second RTO Interval (after DOE) | 0.0560 | 0.1604 |

The same SQP algorithm used for the economic optimization problem was applied to the experimental design optimization problem.

Figure 2.7 also shows that the experimental design does not produce a point that lies on the border of the experimental window. This occurs because the nonlinearities in the objective function and the process model produce a local minimum, as shown in Figures 2.8 and 2.9. This is in contrast to the more common linear experimental designs, which will always push for points to be selected on the boundary of the experimental window.

Since the primary purpose of the RTO system is to maximize profit, it is not prudent to apply an experimental design without considering the loss in profit that would result.

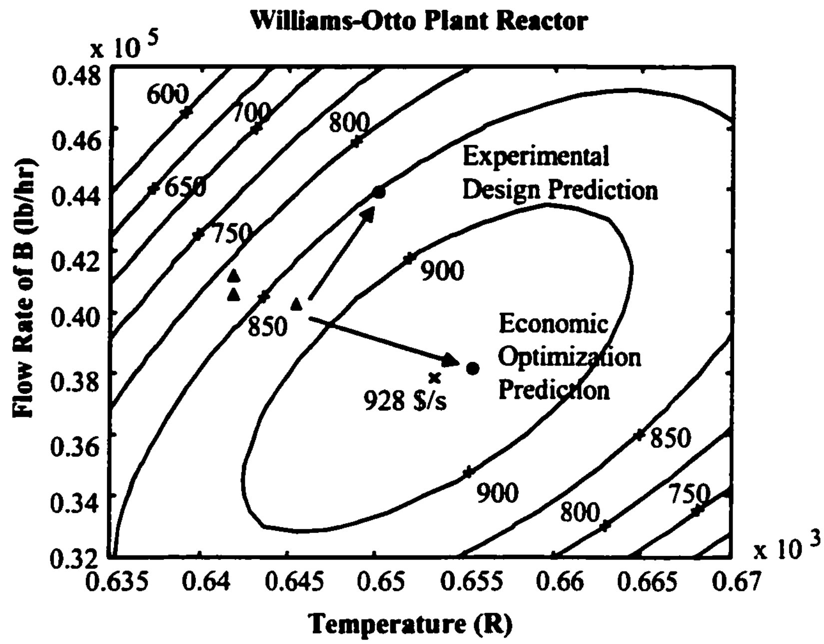



Figure 2.7: The economic optimizer and the experimental design push the operation of the process in different directions.

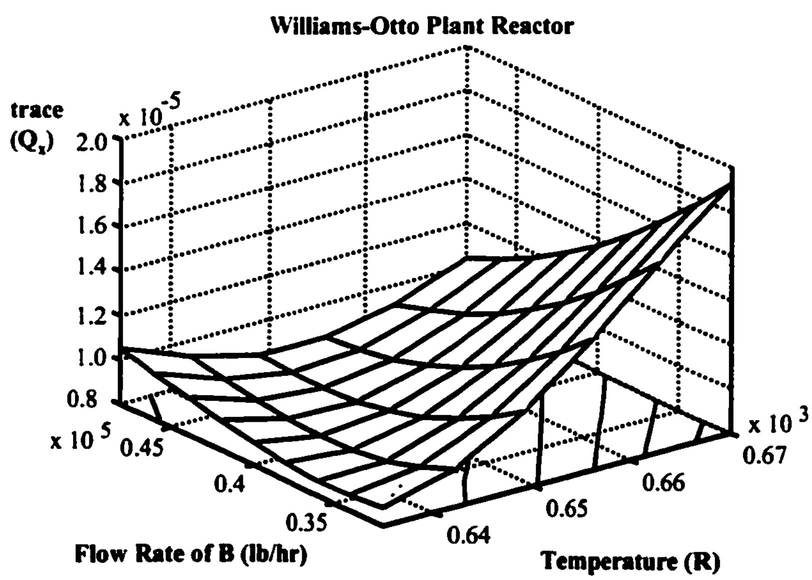



Figure 2.8: The surface plot shows there is a local minimum for the experimental design problem.

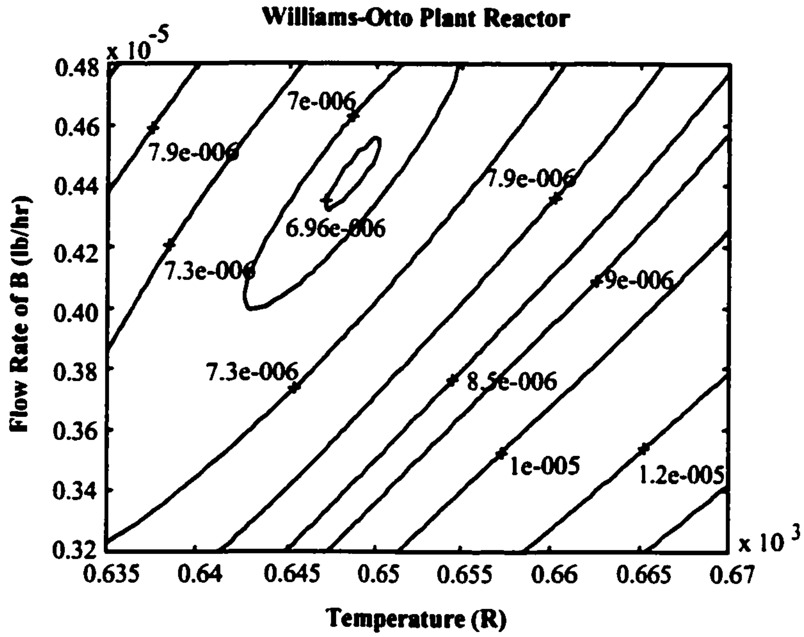



Figure 2.9: The contour plot shows there is a local minimum for the experimental design problem.

Chapter 3 will extend the study initiated in Chapter 2 to develop a more practical usage for experimental design, that involves a trade-off between information generation and producing greater profit. It will consist of expanding the role of the command conditioning unit to judge whether the result from the economic optimizer exceeds a threshold value of information generation. If this value is not met, then an experimental design calculation is performed that consists of an experimental window represented by a constraint in the optimization algorithm. This constraint will limit the amount of profit that the engineer is willing to lose for the purpose of improved information generation.

# Chapter 3

# Enhanced Command Conditioning

The primary purpose of RTO is to maximize profit generated by efficient process operation. Chapter 2 showed that experimental design techniques can be used to reduce uncertainty in the decision variables; however, they can move the process operation into a region of significantly lower profit.

This chapter will continue to investigate the benefits of experimental design and will propose methods that enable it to be a practical addition to the RTO system. The command conditioning (CC) subsystem is expanded to measure the level of potential information that could be generated if the result predicted by the economic optimizer is implemented. If this measure is too low, then a constrained experimental design calculation is performed that limits the size of the experimental window to control the amount of profit that could be lost if the result is implemented.

This chapter begins with a discussion concerning the current structure of the command conditioning subsystem and how results analysis is used to evaluate the significance of the result predicted by the economic optimizer. Section 3.2 discusses the issues involved in incorporating experimental design into an RTO algorithm. This chapter concludes with a case study based on the Williams-Otto reactor (Williams and Otto, 1960) that demonstrates how the concepts presented in Sections 3.2 perform as a part of the RTO system.

## 3.1 Results Analysis for Low Frequency Disturbances

Miletic and Marlin (1998) developed a method that applies multivariable statistical hypothesis tests based on control charts to distinguish between high frequency disturbances transmitted through the calculations and significant low frequency changes in the plant optimization variables. By making this distinction, they were able to show that they could reduce the number of unnecessary changes in the manipulated variables and improve profits. From an operating benefit stand point, the process is not subjected to as many setpoint changes since only significant results are implemented.

The hypothesis test used has the form,

$$\begin{aligned} H_o &: \boldsymbol{\mu}_p = \boldsymbol{\mu}_o \\ H_1 &: \boldsymbol{\mu}_p \neq \boldsymbol{\mu}_o \end{aligned} \tag{3.1}$$

where $\boldsymbol{\mu}$ represents a vector of expected values of the optimal operating conditions, and $p$ and $o$ represent the model prediction and applied plant values, respectively. If the hypothesis test is not rejected, it is considered that the difference in $\boldsymbol{\mu}$'s does not exceed a significant threshold which accounts for the inherent background statistical variation caused by measurement noise, therefore the computed change is not significant and the new point should not be implemented.

The hypothesis test in Problem (3.1) is conducted by comparing the calculated $T^2$ statistic to a control limit value. Wierda (1994) identified several different control limit values. A single multivariate hypothesis, such as the Hotelling $T^2$ statistic, is selected over a set of individual tests for each decision variable, because of the possible high degree of interaction among the decision variables in practical applications (Marlin and Hrymak, 1997). If the $T^2$ statistic is larger than the control limit, then $H_o$ is rejected and the new point should be implemented. Assuming the distribution of the calculated setpoints is multinormal, which is true if the distribution of the plant measurements is multinormal and if the transformations applied to the measurement covariance matrix are linear, the Hotelling $T^2$ statistic has the form (Miletic and Marlin, 1998),

$$T^2 = (\mathbf{x}_p - \mathbf{x}_o)^T \mathbf{Q}_x^{-1} (\mathbf{x}_p - \mathbf{x}_o) \tag{3.2}$$

where $\mathbf{Q}_x$ is an approximation of the decision variable covariance matrix, see Equation (2.15). Please note that the variable values, $\mathbf{x}$, are estimated values of the expected values $\boldsymbol{\mu}$. The control limits for the hypothesis test in Problem (3.1) can be based on the most recent optimization calculation or on successive averages of the most recent calculations:

$$UCL_1 = \frac{p(N+1)(N-1)}{N(N-p)} F_\alpha(p, N-p) \tag{3.3}$$

$$UCL_2 = \frac{p(m+1)(mn-1)}{m(mn-p)} F_\alpha(p, mn-p) \tag{3.4}$$

where $p$ is the row dimension of $\mathbf{x}$, $N$ is the number of data points used to calculate the measurement covariance matrix, $n$ is the number of optimization results tested in $UCL_2$, and $m$ is the number of $n$-sized groups used to compute the measurement covariance matrix. Both of the control limits above were used by Miletic and Marlin (1998), and Equation (3.3) was used by Zhang *et al.* (2001) for their work.

Zhang *et al.* (2001) extended these developments to analyze steady-state RTO results in the presence of inequality constraints that may exist due to operating constraints or trust-region limits. They expanded the statistical testing to include the dual variables of the optimization problem, as well as the primal variables.

## 3.2 Incorporating DOE into RTO

Chapter 2 concluded by showing that experimental design techniques can be used to generate information to reduce uncertainty in future optimization calculations. However, the case study in Section 2.4 illustrated the potential drawback of the experimental design was selecting an operating point in a region of significant potential profit loss. Of further concern includes the computational burden that DOE calculations can have on the RTO system. This section will discuss two approaches to reduce the influence of these concerns.

### 3.2.1 Command Conditioning Structure

As discussed in Section 3.1, one of the purposes of the CC subsystem is to evaluate whether the point predicted by the economic optimizer should be implemented if it is not significantly different than the current operating point. This thesis builds on the work originally developed by Miletic and Marlin (1998), by expanding the structure of the CC subsystem to also determine if the predicted point is expected to generate information. By selecting instances when the DOE calculation is most needed, the CC subsystem can reduce the computational burden that would be involved in performing an extra optimization calculation for each RTO cycle. The DOE calculation cannot be performed between RTO intervals, to alleviate the computational burden, because information is needed at steady-state to complete the necessary calculations. It cannot be performed after the result from the economic optimizer is implemented either, because this exercise is performed to avoid implementing results that do not generate information. The modified CC subsystem continues to focus the RTO loop on maximizing profit by only deviating from the predicted result when necessary. The proposed structure for the CC subsystem is shown in Figure 3.1, where the expansion occurs after the first layer to ensure information is generated prior to implementation.

The criterion that measures the potential level of information generation has a form similar to the criteria used for the experimental design:

$$trace\ (\mathbf{Q}_x)_{predicted} \leq \alpha\ trace\ (\mathbf{Q}_x)_{current} \tag{3.5}$$

Inequality (3.5) states that the volume and shape of the confidence region for the decision variable covariance matrix will be improved if the predicted point is implemented. If the experimenter wants the decision on whether a DOE should be performed to be based solely on the volume of the confidence interval, without considering shape, then the trace can be replaced with the determinant function.

The selection of $\alpha$ as a tuning parameter, is dependent on the importance of information generation to the application. Setting $\alpha$ to one, states that the experimental design calculation is performed only if the point predicted by the economic optimizer is expected to produce a poorer estimate of the confidence region than what currently exists. Values of $\alpha > 1$ are not good selections, because it allows for points to be implemented that will

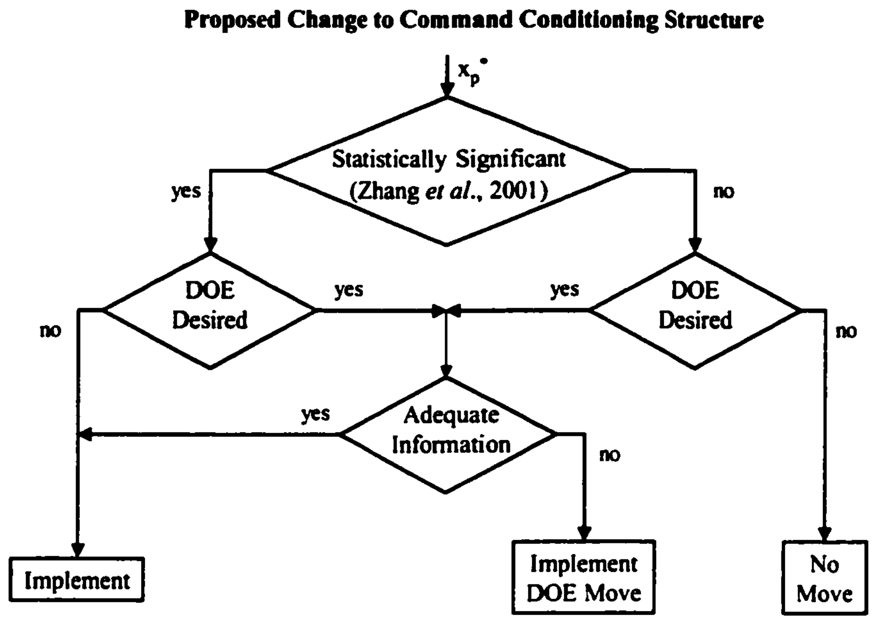



Figure 3.1: Proposed adjustment to the command conditioning structure.

potentially increase the uncertainty in the decision variables. However, if $\alpha \ll 1$, the criterion will too frequently conclude that an experimental design should be performed when it is not needed. This is not desirable since the primary goal of RTO is to improve profit, so deviations from the result predicted by the economic optimizer should be performed selectively. Therefore, it is recommended to select a value for $\alpha$ close to one or slightly less.

Defining the command conditioning structure in Figure 3.1 identifies when the DOE should be performed. The next part of this section will discuss how to formulate the DOE problem so the amount of profit lost by implementing the result from the experimental design is controlled.

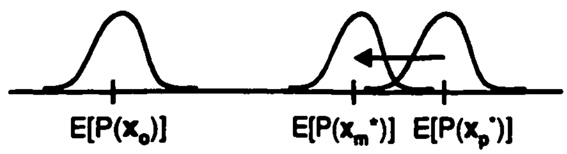
### 3.2.2 Defining the Experimental Window

The case study in Section 2.4 showed there is a cost associated with performing experimentation. That cost results from moving to an operating point that generates more information but less profit than the point predicted by the economic optimizer. This problem is reduced by adding a constraint to the experimental design problem that limits the experimental window to a region of acceptable profit loss determined by the engineer.

In defining the CC structure in Figure 3.1, two scenarios are considered for when a DOE might need to be performed. One scenario occurs when the results from the economic optimizer are significant, and the other occurs when they are not significant. Therefore, the formulation of the constraint describing the experimental window should depend on the conclusions from the results analysis work.

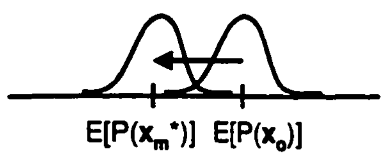
For the case when the result from the economic optimizer is statistically significant, the experimental design is centered around the predicted point. Therefore, the experimental design will select a point that deviates from the predicted result, determined by the economic optimizer, with an allowable amount of predicted profit loss defined by Inequality (3.6). Where $E[P(x)]$ is the expected profit produced at the values specified by the decision variables, and $\mathbf{x}_o$, $\mathbf{x}_p^*$, and $\mathbf{x}_m^*$ are the current operating point, the point predicted by the economic optimizer and the implemented point predicted by the DOE, respectively:

$$E\left[P\left(\mathbf{x}_m^*\right)\right] \geq E\left[P\left(\mathbf{x}_o\right)\right] + c_1\left\{E\left[P\left(\mathbf{x}_p^*\right)\right] - E\left[P\left(\mathbf{x}_o\right)\right]\right\} \tag{3.6}$$




When the result from the economic optimizer is found to be statistically insignificant, the DOE problem is focused on the region surrounding the current operating point. If the value of $c_2$ in Inequality (3.7) is set to one, then the DOE problem is limited to an experimental window that will potentially only implement a result that is still expected to produce more profit than the current operating point:

$$E\left[P\left(\mathbf{x}_m^*\right)\right] \geq c_2\left\{E\left[P\left(\mathbf{x}_o\right)\right]\right\} \tag{3.7}$$




Please note that unlike $c_2$ in Inequality (3.7), $c_1$ is a fraction of the difference in profit between the result from the economic optimizer and the current operating point. The constraint is structured in this manner to ensure that the point resulting from the DOE is expected to at least produce a result that generates more profit than the current operating point. If Inequality (3.6) was given a structure similar to Inequality (3.7) such as,

$$E\left[P\left(\mathbf{x}_m^*\right)\right] \geq c_1\left\{E\left[P\left(\mathbf{x}_p^*\right)\right]\right\} \tag{3.8}$$

this statement could not be made. This is important because, if the economic optimizer found a point that is significant, the implemented result should be at least expected to generate more profit than the current operating point.

The selection of the tuning parameters $c_1$ and $c_2$ in Inequalities (3.6) and (3.7), respectively, can be dependent on one of three criteria. The first method is to fix the values of $c_1$ and $c_2$, so the size of the experimental window is consistently set to sacrifice a specified percentage of expected profit every time the DOE is performed.

The second method would base the selection of the tuning parameters on the result of the significance tests. The result from Inequality (3.5), would dictate the values of $c_1$ or $c_2$.

As the difference between the two scalar values widen, the values for the tuning parameters should get smaller to allow for a larger experimental window, which would provide an increased level of information generation. Care must be taken when selecting the size of $c_1$ or $c_2$, since values $\ll 1$ will give the DOE too much freedom, resulting in deviations far from the optimum and significant profit loss, as shown in the Chapter 2 case study.

The third approach would define the values of $c_1$ and $c_2$ as functions of the profit distribution, assuming it is known. The experimental window would then be restricted to a size of one or two standard deviations of the profit distribution. The constraints for the two scenarios would then have a form similar to the following,

Result from economic optimizer is statistically significant:

$$E\left[P\left(\mathbf{x}_m^*\right)\right] \geq \left(1 - c_1\right) E\left[P\left(\mathbf{x}_p^*\right)\right] \tag{3.9}$$

Result from economic optimizer is statistically insignificant:

$$E\left[P\left(\mathbf{x}_m^*\right)\right] \geq \left(1 - c_2\right) E\left[P\left(\mathbf{x}_o\right)\right] \tag{3.10}$$

This thesis will use the first method, because it is considered important to define the constraints in terms of the loss in expected profit.

## 3.3 Williams-Otto Reactor Case Study

Chapter 2 showed that implementing the result from an experimental design will reduce the uncertainty in future optimization calculations. However, it also showed that the cost associated with this result is the potential move of the process operation to a region of significant profit loss.

The purpose of this case study is to identify the benefits and weaknesses of adding a DOE component to the RTO system.

### 3.3.1 Method

To add parametric mismatch in one of the case studies, the activation energy constants were given a 1, 3, 5 or 10% additive bias in the working model, while the original values remained the same in the true plant model. The fixed values for the activation energies are shown in Table 3.1.

The same initial data set and measurement covariance matrix used in Section 2.4 will also apply to the case studies in this section. Please see *Appendix C* for their values. The first case study will examine whether the performance of the RTO system can be improved by including an information generation component. Table 3.2 shows the different tests that will be performed to determine which sets of tuning parameters produce an RTO system that out-performs the base case. The base case is defined as the RTO system without an information generation component incorporated into the RTO structure (this reduces to the work developed by Miletic and Marlin (1998)). The tuning parameters are defined as:

Table 3.1: Assumed Values of Activation Energy Constants for Mismatch Case Study

| Test # | Bias (%) | Activation Energies (R) $B_1$ (R) | $B_2$ (R) | $B_3$ (R) |
|---|---|---|---|---|
| 1 | 1 | 12,120 | 15,150 | 20,200 |
| 2 | 3 | 12,360 | 15,450 | 20,600 |
| 3 | 5 | 12,600 | 15,750 | 21,000 |
| 4 | 10 | 13,200 | 16,500 | 22,000 |

Table 3.2: Case Study Tuning Parameters

| Test # | $\alpha$ | $c_1$ | $c_2$ | Test # | $\alpha$ | $c_1$ | $c_2$ |
|---|---|---|---|---|---|---|---|
| 1 | 0.90 | 0.90 | 0.990 | 19 | 0.95 | 0.95 | 0.997 |
| 2 | 0.90 | 0.90 | 0.995 | 20 | 0.95 | 0.95 | 0.999 |
| 3 | 0.90 | 0.90 | 0.997 | 21 | 0.95 | 0.98 | 0.990 |
| 4 | 0.90 | 0.90 | 0.999 | 22 | 0.95 | 0.98 | 0.995 |
| 5 | 0.90 | 0.95 | 0.990 | 23 | 0.95 | 0.98 | 0.997 |
| 6 | 0.90 | 0.95 | 0.995 | 24 | 0.95 | 0.98 | 0.999 |
| 7 | 0.90 | 0.95 | 0.997 | 25 | 1.00 | 0.90 | 0.990 |
| 8 | 0.90 | 0.95 | 0.999 | 26 | 1.00 | 0.90 | 0.995 |
| 9 | 0.90 | 0.98 | 0.990 | 27 | 1.00 | 0.90 | 0.997 |
| 10 | 0.90 | 0.98 | 0.995 | 28 | 1.00 | 0.90 | 0.999 |
| 11 | 0.90 | 0.98 | 0.997 | 29 | 1.00 | 0.95 | 0.990 |
| 12 | 0.90 | 0.98 | 0.999 | 30 | 1.00 | 0.95 | 0.995 |
| 13 | 0.95 | 0.90 | 0.990 | 31 | 1.00 | 0.95 | 0.997 |
| 14 | 0.95 | 0.90 | 0.995 | 32 | 1.00 | 0.95 | 0.999 |
| 15 | 0.95 | 0.90 | 0.997 | 33 | 1.00 | 0.98 | 0.990 |
| 16 | 0.95 | 0.90 | 0.999 | 34 | 1.00 | 0.98 | 0.995 |
| 17 | 0.95 | 0.95 | 0.990 | 35 | 1.00 | 0.98 | 0.997 |
| 18 | 0.95 | 0.95 | 0.995 | 36 | 1.00 | 0.98 | 0.999 |

- $\alpha$ weighs the importance of information generation in Inequality (3.5);
- $c_1$ represents the maximum amount of expected profit that can be sacrificed when the result from the economic optimizer is significant (see Inequality (3.6));
- $c_2$ represents the maximum amount of expected profit that can be sacrificed when the result from the economic optimizer is not significant (see Inequality (3.7)).

The second case study will investigate whether the RTO system that generates information, continues to perform well if plant/model mismatch is introduced. The structural mismatch will take the form of bias in the fixed parameters of the activation energies between the working model and the true plant model of: 1, 3, 5 and 10%. The tuning parameters that produce the best economic results in the first case study will be used in the second.

The extended design cost criterion, developed by Zhang and Forbes (2000), will be used to compare the performance of the different RTO designs. They defined extended design cost as the total loss of performance relative to perfect optimization within a pre-specified performance evaluation period.

$$C = \int_{t_o}^{t_f} E\left[P\left(x^{**}\right) - P\left(x_m^{*}\left(t\right)\right)\right] dt \tag{3.11}$$

Where $C$ is the amount of profit the RTO system is not able to attain, $P$ is the amount of profit at the true plant optimum ($x^{**}$) and the values of the implemented decision variables ($x_m^*$), and $E$ is the expected value. The variation around the expected value of the predicted optimum was also investigated to supplement the performance evaluation, for the base case and the test that produced the most optimal results. This variation is caused by propagation of common cause variation around the closed-loop RTO system (Zhang and Forbes, 2000). It will be assumed that each RTO interval is the same length.

In order to analyze the final results, ten separate runs consisting of 25 RTO intervals were performed at each set of test conditions, with the means of these results compared.

One-sided hypothesis tests were used to compare the means of the extended design cost measure for the 36 tests to the base case, and to determine if the standard deviations of the data sets used to find the sample means have the same population variances. The statistic used to compare the sample means, is dependent on whether the population variances were found to be equal. The hypothesis tests were also used to evaluate whether the variation around the expected value of the predicted optimum improved, from the base case to the RTO system with DOE. The form of the statistics are shown in *Appendix D.* The structure of the hypothesis tests are:

Testing population variance - standard deviation of test case $(\sigma_t^2)$ is larger than the base case$(\sigma_b^2)$:

$$\begin{aligned} H_o &: \frac{\sigma_t^2}{\sigma_b^2} = 1 \\ H_1 &: \frac{\sigma_t^2}{\sigma_b^2} > 1 \end{aligned} \tag{3.12}$$

Testing population variance - standard deviation of test case is smaller than the base case:

$$\begin{aligned} H_o &: \frac{\sigma_t^2}{\sigma_b^2} = 1 \\ H_1 &: \frac{\sigma_t^2}{\sigma_b^2} < 1 \end{aligned} \tag{3.13}$$

Testing population mean - sample mean of test case ($\mu_t$) is larger than the base case ($\mu_b$):

$$\begin{aligned} H_o &: \mu_t - \mu_b = 0 \\ H_1 &: \mu_t - \mu_b > 0 \end{aligned} \tag{3.14}$$

Testing population mean - sample mean of test case is smaller than the base case:

$$\begin{aligned} H_o &: \mu_t - \mu_b = 0 \\ H_1 &: \mu_t - \mu_b < 0 \end{aligned} \tag{3.15}$$

One-sided hypothesis tests were used because once the sample values are known, the interval can only be violated on one side. For example, a test that had a sample mean smaller than the base case, is not susceptible to having the hypothesis test discover that $\mu_t - \mu_b > 0$, therefore it is only necessary to investigate the alternate hypothesis of $\mu_t - \mu_b < 0$.

The following algorithm outlines the steps used in the RTO loop.

1. Complete the parameter estimation step using a least squares regression optimization routine, see Problem (2.1), with the optimal number of historical steady-state points predefined.

2. The sensitivity matrix of the parameters to the measurements, $\frac{\mathbf{d}\boldsymbol{\beta}}{\mathbf{dz}}$, is found using the optimal number of historical steady-state measured points.

3. Using the updated model, perform the optimization step using the SQP algorithm in the optimization toolbox of the computer software package MATLAB ver. 5.3.0 (Branch and Grace, 1996).

4. The measured values of the manipulated variables combined with the process model, are used to solve for the expected values of the response variables based on those measurements. This result is later used to develop the profit constraints that define the experimental window.

5. The results analysis step is performed to determine whether the result from the economic optimizer is significant. Please see the hypothesis test represented by Problem (3.1) (Miletic and Marlin, 1998). The testing is performed for a hypothesis test at a 5% significance level.

6. The sensitivity matrix of the optimization variables to the parameters, $\frac{\mathbf{dx_p}}{\mathbf{d}\boldsymbol{\beta}}$, is found in this step.

7. The command conditioning subsystem determines if the point predicted by the economic optimizer is expected to generate information with respect to Inequality (3.5).

8. If the result from the economic optimizer is expected to generate information, the result is implemented despite its status of being significant. If it is not expected to generate information then an experimental design (see Problem (2.20)) is performed with an added constraint to limit the amount of profit lost. The form of the constraint is dependent on whether the prediction from the economic optimizer is significant (see Inequalities (3.6) and (3.7)). The sensitivity matrix, $\frac{\mathbf{d}\boldsymbol{\beta}}{\mathbf{dz}}$, used in the experimental

design objective function is based on the most recent optimal number of steady-state data points subtract one, which is replaced by the next design point to augment the data set. The sensitivity matrix, $\frac{dx_p}{d\beta}$, is found from a previous calculation since it can only be estimated from the result of an economic optimization calculation. If the experimental design is centered around the predicted optimum, then the result from step #7 is used. However, if the experimental design is centered around the current operating point then the result from the previous RTO cycle is used.

9. The actual amount of profit generated is based on the setpoints returned from the optimizer since it is assumed the controllers are able to implement them.

10. Steps #1-#9 are repeated.

### 3.3.2 Results and Discussion

This section concludes with the presentation and discussion of the results for the case studies involving the Williams-Otto reactor. The first case study is performed to show that DOE can improve the performance of the RTO system. The second case study investigates how the RTO with DOE performs when parametric mismatch between the model and plant exist.

#### RTO with DOE

The results for the first case study are shown in Table 3.2. The data is arranged in ascending order of the sample means of the extended design cost measure for ten runs calculated over 25 RTO intervals. The column representing the standard deviation, refers to the variation in the extended design cost data set used to calculate the sample mean.

Using one-sided hypothesis tests (see Equations (3.12) to (3.15)) the sample means of the extended design costs from the tests shown in Table 3.3, were compared to the sample mean of the base case, which was comprised of the RTO system running without the information generation component (this reduces to the results analysis case developed by Miletic and Marlin (1998)). The tests show, which sets of fixed tuning parameters result in an RTO system that produces significantly lower or greater profit than the base case. Please see *Appendix C* for the results of the statistical tests. Note that the following tests were found to have sample variances different from the base case, using a one-sided hypothesis test with a 5% significance level: 1-5, 7, 9, 10, 13-18, 21, 24, 25, 34, 35

The results of the first case study show that when plant/model mismatch is absent, incorporating DOE into the RTO system does allow for more profit to be realized. Figure 3.2 compares the profit profile, over 25 RTO intervals, of the first runs for the base case test and test #36. It illustrates how incorporating DOE can allow the RTO system to converge closer to the true plant optimum. However, the results also show that the selection of the tuning parameters is important.

Table 3.3: Design Cost of RTO Systems

| Test # | Sample Mean ($/s) | Standard Deviation ($/s) | Sig. Mean | Test # | Sample Mean ($/s) | Standard Deviation ($/s) | Sig. Mean |
|---|---|---|---|---|---|---|---|
| 36 | 34.98 | 11.09 | * | 26 | 69.57 | 16.67 | |
| 35 | 40.30 | 9.26 | * | 18 | 70.61 | 25.19 | |
| 27 | 41.82 | 11.48 | * | 29 | 71.67 | 12..22 | * |
| 28 | 43.71 | 14.94 | * | 25 | 79.00 | 31.07 | |
| 32 | 43.83 | 12.86 | * | 4 | 79.69 | 25.34 | * |
| 31 | 45.33 | 15.43 | * | 15 | 81.38 | 20.94 | * |
| 20 | 48.30 | 17.60 | * | 10 | 82.66 | 23.71 | * |
| 12 | 52.29 | 12.61 | * | 33 | 82.82 | 17.73 | * |
| 24 | 52.40 | 6.95 | * | 17 | 83.73 | 21.34 | * |
| 34 | 53.57 | 19.09 | | 21 | 83.76 | 19.19 | * |
| 8 | 55.18 | 12.21 | | 7 | 86.12 | 26.40 | * |
| 23 | 56.22 | 16.45 | | 9 | 88.59 | 17.96 | * |
| 11 | 57.44 | 15.87 | | 14 | 90.26 | 28.26 | * |
| 19 | 58.13 | 14.91 | | 6 | 91.53 | 12.46 | * |
| 16 | 61.18 | 22.03 | | 13 | 95.22 | 32.03 | * |
| **Base** | **62.45** | **10.01** | | 5 | 99.50 | 23.01 | * |
| 30 | 62.57 | 13.58 | | 2 | 105.14 | 19.80 | * |
| 22 | 67.32 | 12.17 | | 1 | 138.42 | 30.94 | * |
| 3 | 68.80 | 20.99 | | | | | |

Table 3.3 shows that the RTO system performs better when larger settings for $c_1$ and $c_2$ are applied. The larger values for these tuning parameters creates a smaller experimental window for the DOE that lessens the amount of profit that is willing to be traded-off for improved information generation. This case study shows that deviating from the predicted optimum to generate information can improve profit, although if the experimental window is too large the moves will not converge to the plant optimum but instead move in directions of significant profit loss. Table 3.3 also shows that RTO systems allowing large experimental windows, produced substantially poorer results than the base case, determined from one-sided hypothesis tests of 5% significance.

Table 3.3 also illustrates that the RTO system performs more effectively with a tightened decision criterion (see $\alpha$ in Inequality (3.5)). Tightening the decision criterion reduced the number of times the DOE is implemented. This indicates that it is beneficial for the optimizer to deviate from the predicted economic optimum only when there is a significant need for generating information.

Table 3.4 displays how the variation around the expected value of the predicted optimum is reduced when DOE (for tuning constants defined by test #36) is added to the RTO system. The results were found to be significant using a one-sided hypothesis test at a significance level of 5%. This is attributed to better optimization results, that occur because the information generation step is present to ensure high quality information is produced to improve the predictability of the optimizer.

Table 3.4: Variance of RTO system

| Case Study | Variance ($/s) |
|---|---|
| RTO without DOE | 6.95 |
| RTO with DOE | 5.21 |

The reduction in variance is also visualized in Figures 3.2, 3.3 and 3.4, for a comparison of the first runs from the base case and test #36. Figure 3.2 illustrates how the profit profile of the RTO system with DOE, was consistently closer to the expected value of the predicted optimum Figures 3.3 and 3.4 illustrate the moves made by the optimizers with and without DOE incorporated in the RTO design, respectively. Figure 3.4, shows a more scattered behaviour because of the greater uncertainty in the decision variables. The RTO system that used DOE, shown in Figure 3.3, illustrates how the process operation is kept closer to the plant optimum even though deviations were made from the predicted economic optimum to perform the experimental design. These figures show that the optimizer consistently predicts points closer to the true plant optimum when using experimentation, even though deviations are made from the predicted economic optimum to perform the experimentation.

Figure 3.5 displays a set histograms for eight RTO intervals, that compare the ten runs performed for each RTO system. The eight selected intervals, out of a possible twenty-five, were chosen because they provided the most informative observations. The histograms reinforce the results that showed the RTO system with DOE out-performs the base case. All eight figures illustrate how the profit distribution shifts to higher profit levels and the variance, or spread of the distribution, is reduced.

### RTO with DOE: Parametric Mismatch

The first case study showed that DOE can improve the performance of the RTO system based on the extended design cost measure for cases where plant/model mismatch does not exist. The second case study examines the change in performance of the RTO systems when parametric mismatch is present by creating a bias in the activation energies, which act as fixed parameters, according to Table 3.1.

Similar to the first case study, the statistical tests shown in *Appendix D* were used to determine if the RTO system with a DOE component still out-performed the base case, when plant/model mismatch was present. Table 3.5 presents the sample means of the extended design cost, and the variances around the expected optimum for the results of the second case study. The RTO system which incorporates DOE out-performed the base case only for the 1% bias test, and produced significantly worse results for the 5% and 10% bias tests. Please see *Appendix C* for the results of the hypothesis tests. However, please note that the variance was reduced until the 10% bias case. Since the objective of DOE is to reduce

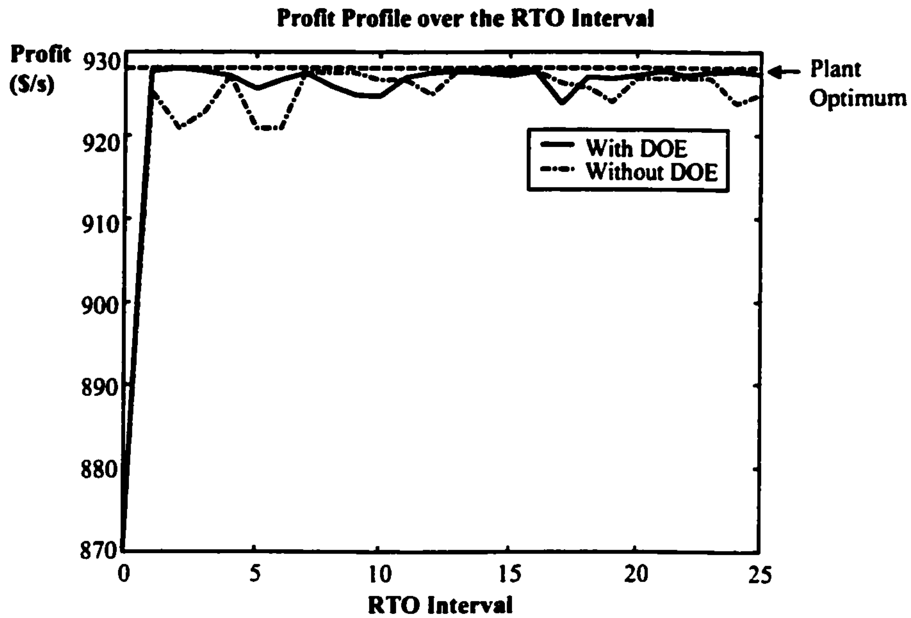



Figure 3.2: The RTO system performs better with an information generation component added to the algorithm.

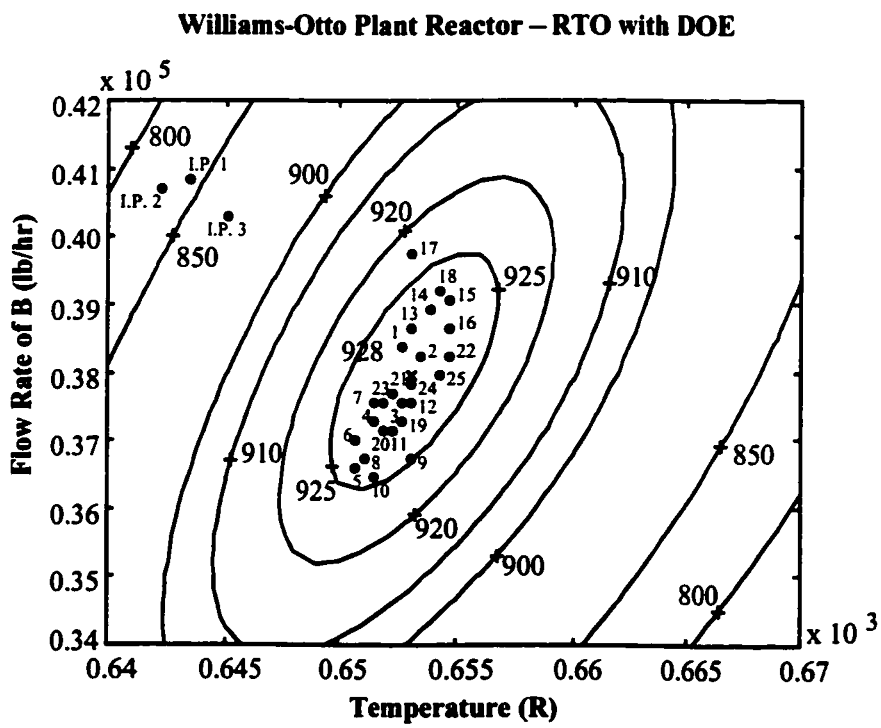



Figure 3.3: DOE can be implemeted in the RTO system without implementing moves far from the true plant optimum.

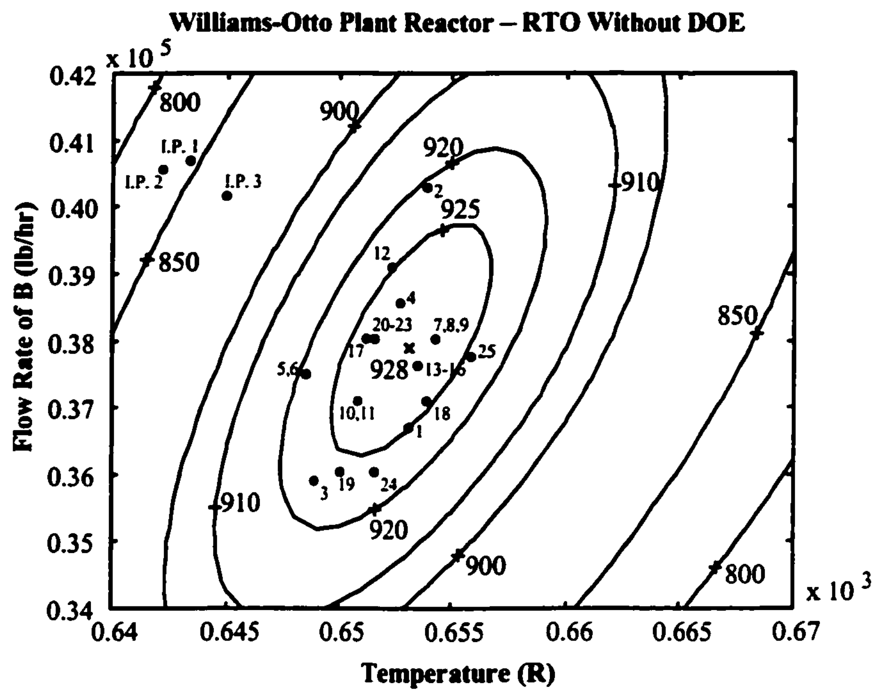



Figure 3.4: Running the RTO system without DOE causes more significant deviations from the true plant optimum because of a reduced understanding of the process behaviour.

the variance, the DOE performed well until the mismatch became too large. The failure of the DOE to effectively reduce variance in the presence of large structural mismatch is due to the assumption that states the defined process model is expected to represent the true process, which is also used for least squares regression. Therefore, the DOE is calculating points that do not generate as much information as expected.

Table 3.5: Design Cost of RTO Systems in the Presence of Structural Mismatch

| Bias (%) | DOE | Sample Mean ($/s) | Variance ($/s) | Means Significant | Variances Significant |
|---|---|---|---|---|---|
| 1 | Yes | 34.7809 | 4.4722 | Y | Y |
| | No | 60.6389 | 6.6205 | | |
| 3 | Yes | 93.3772 | 8.8720 | N | Y |
| | No | 87.5840 | 6.9034 | | |
| 5 | Yes | 156.0228 | 14.6544 | Y | Y |
| | No | 130.8010 | 18.2053 | | |
| 10 | Yes | 425.9288 | 78.0358 | Y | Y |
| | No | 348.5020 | 19.7630 | | |

Figure 3.6 shows the profit profiles for the first runs of the two RTO systems for the 5% bias case. The RTO system using information generation produces lower profit because the experimentation pushes the operation of the process into regions of significant profit loss, as shown in Figure 3.7. Figure 3.8 shows the path of operation calculated by the RTO system

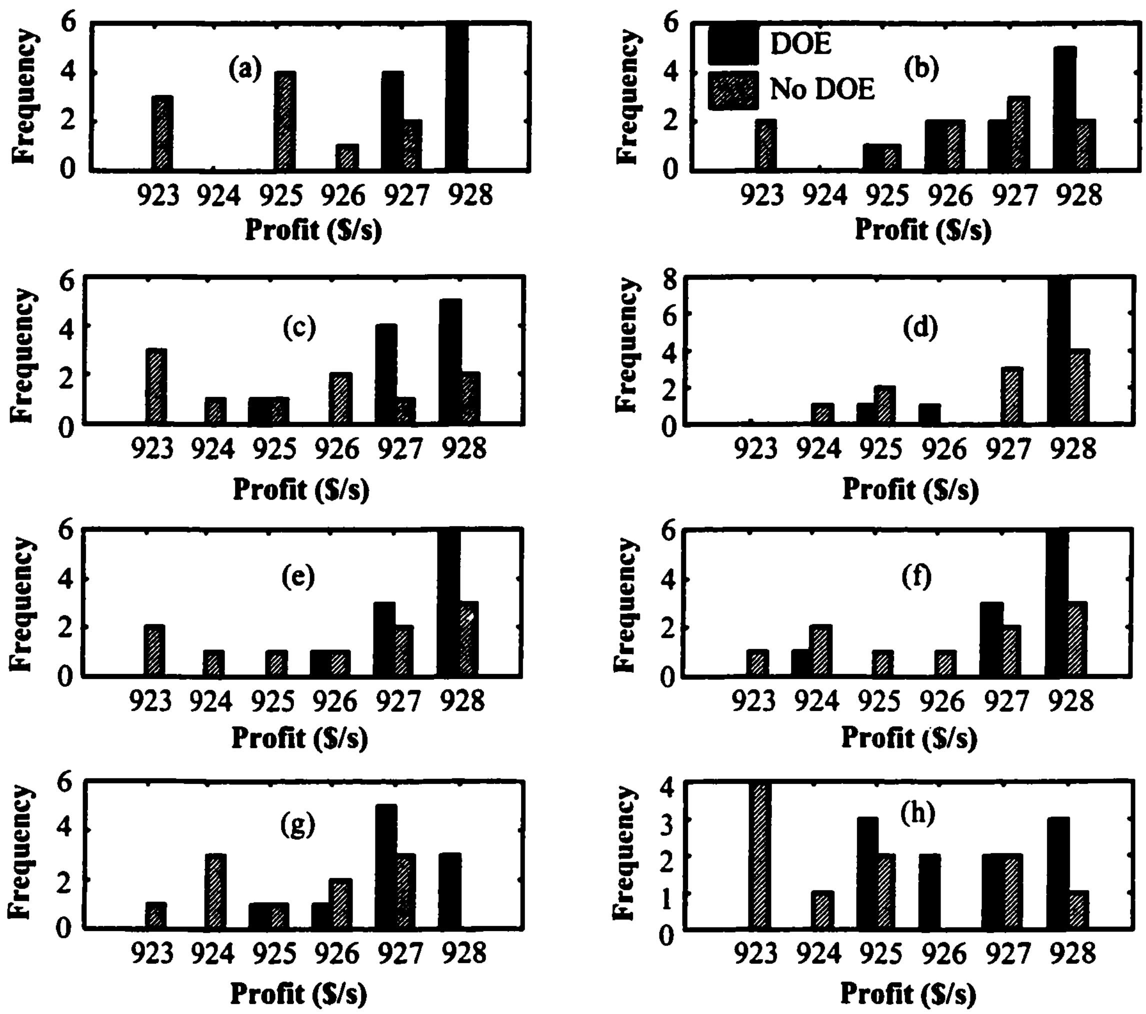



Figure 3.5: The set of histograms provides a visualization of the ten runs for each RTO system for the Williams-Otto Reactor: a) 2nd Interval; b) 5th Interval; c) 6th Interval; d) 10th Interval; e) 12th Interval; f) 16th Interval; g) 21st Interval; h) 24th Interval.

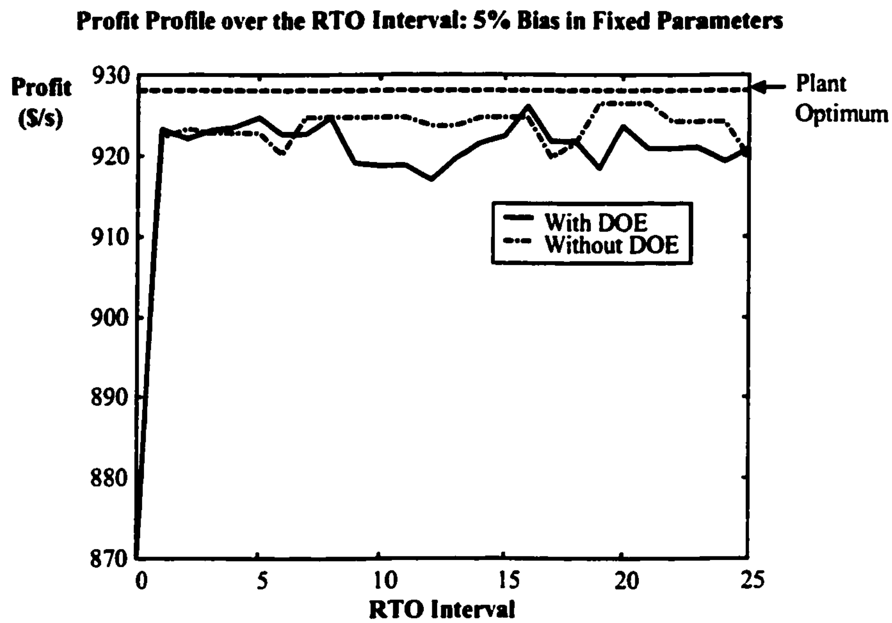



Figure 3.6: The DOE moves operation to regions of significant profit loss because the profit constraint defining the experimental window is not as well defined when structural mismatch exists.

without information generation avoids implementing any costly moves.

These poor results occur because the experimental window used in the experimental design is composed of a set of profit constraints that limit the amount of profit that is willing to be lost in exchange for improved information generation. The profit constraints are dependent on the form of the assumed model, so any structural errors that exist will lead to poor estimates of the behaviour of the process with respect to profit. This will lead to generation of experimental windows that extend to operating regions where, unexpectedly, profit loss exceeds the desired amounts.

Figure 3.9 displays a set histograms for eight RTO intervals, that compare the ten runs performed for each RTO system with a 5% bias added to the activation energies. The histograms do not show a clear difference in the variances that result from the two systems, but they do illustrate that the RTO system with DOE consistently operates in regions of lower profit.

The case study did show that for minimal amounts of bias, adding a DOE component to the RTO system is beneficial in terms of improving the profitability. To overcome the poor results identified in the case studies of Chapter 3, the following possible solutions exist: 1) if significant structural mismatch is suspected to exist, adjust the tuning parameters to restrict the size of the experimental window to prevent the DOE from selecting poor operating moves and 2) implement a set of on-line diagnostic checks to provide a warning that the model is not considered able to adequately predict the process behaviour, and a

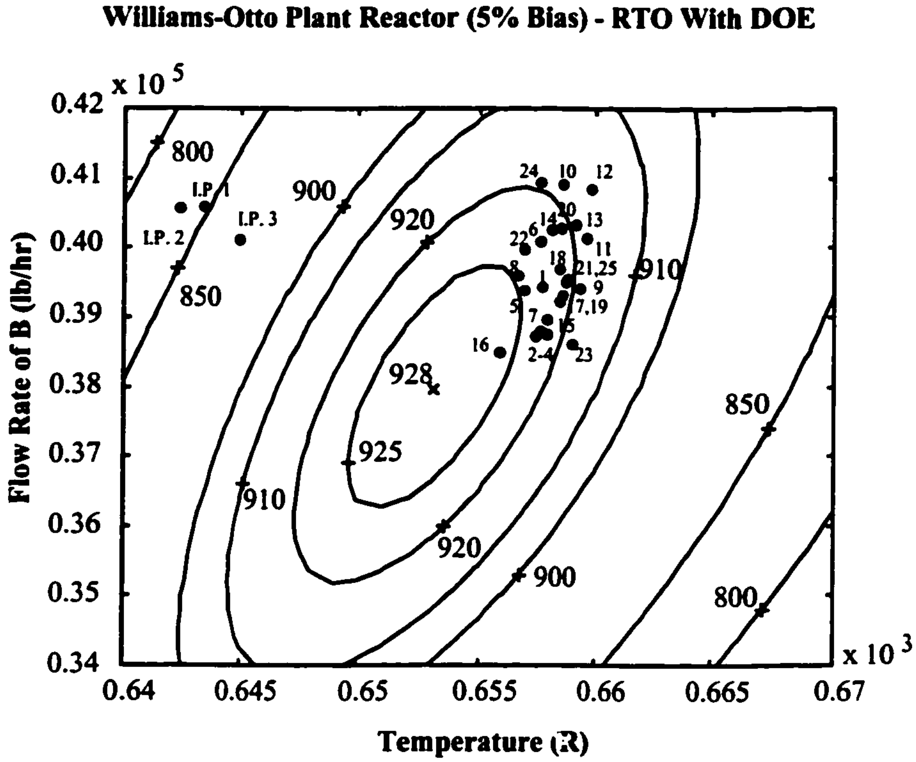



Figure 3.7: Poor move selections lower the profitability of the DOE in the RTO system.

DOE should not be implemented.

Chapter 4 will implement the results developed in Chapter 3 for the full Williams-Otto plant model (Williams and Otto, 1960). It will also investigate the use of implementing a series of on-line diagnostic checks that will evaluate model adequacy, and determine if the DOE should be performed.

**Williams-Otto Plant Reactor (5% Bias) - RTO Without DOE**

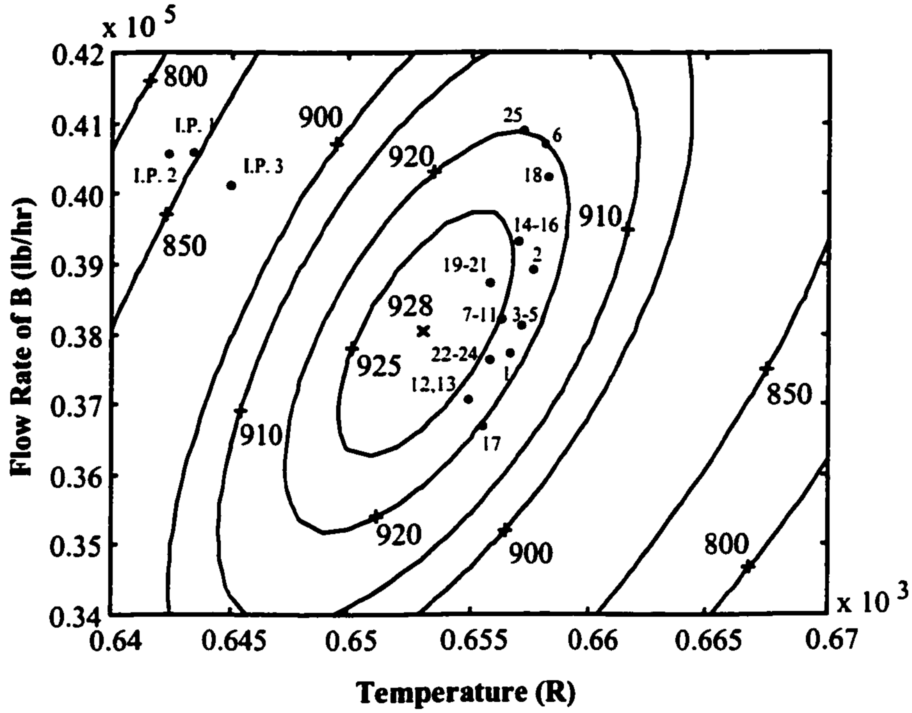



Figure 3.8: The RTO system converges to a point that is not the plant optimum, but it does not implement any points that result in significant profit loss.

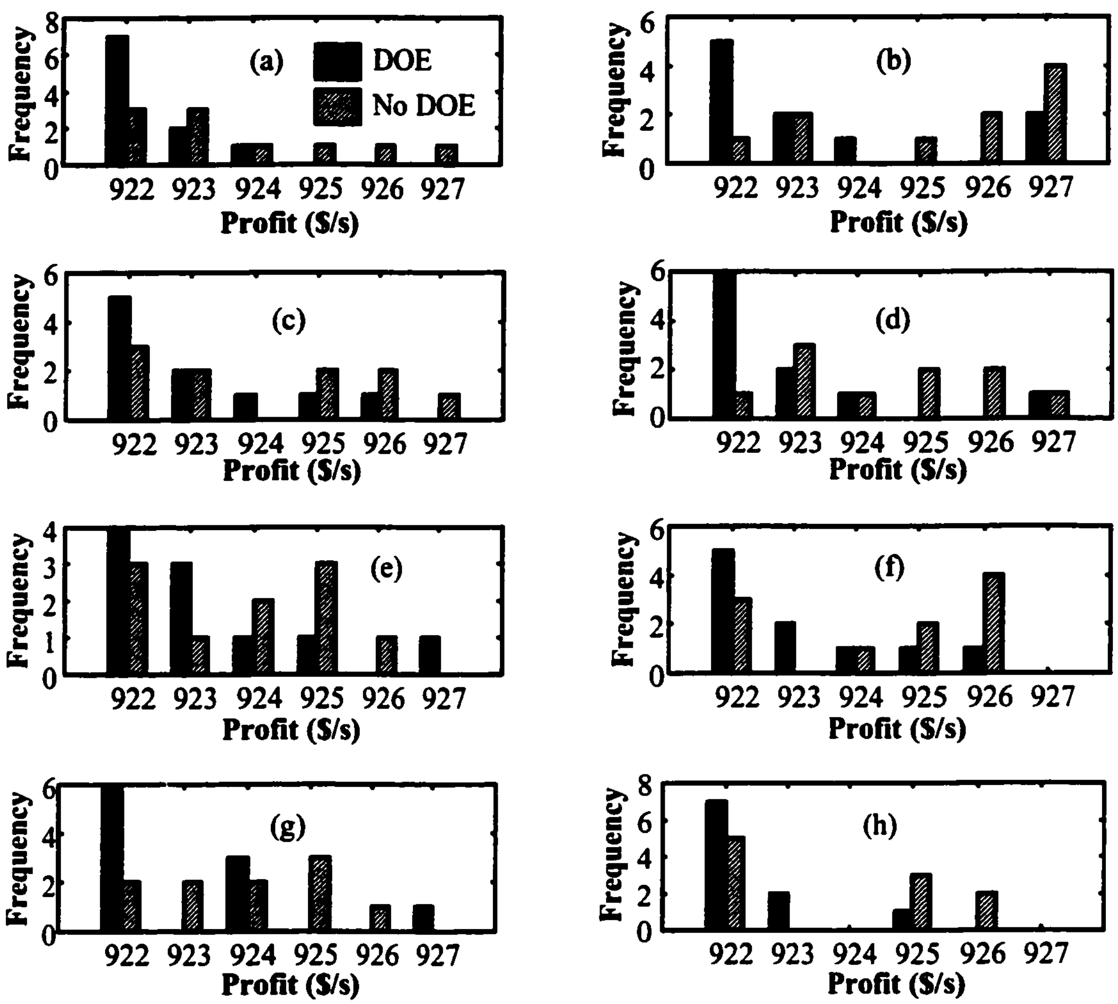



Figure 3.9: The set of histograms provides a visualization of the ten runs for each RTO system for the Williams-Otto Reactor in the presence of parametric mismatch: a) 3rd Interval; b) 6th Interval; c) 8th Interval; d) 11th Interval; e) 15th Interval; f) 19th Interval; g) 22nd Interval; h) 25th Interval.

# Chapter 4

# Case Study: Williams-Otto Plant

Chapter 3 introduced the DOE tool for use in RTO to reduce uncertainty in the decision variables and subsequently improve profit generation. Case studies in Chapter 3 case studies showed that implementing moves that deviate from the result predicted by the economic optimum, with the intention of generating information, can improve the profitability of the RTO system, providing plant/model mismatch is minimal. However, the Williams-Otto reactor is not a realistic application since the model consists of only one unit, and a small number of model equations. The objective of this chapter is to show how DOE is applied in a RTO system for a more realistic application, consisting of a larger model with interacting units.

## 4.1 Diagnostic Check for Model Adequacy

Chapter 3 showed how significant amounts of plant/model mismatch can cause the RTO system with DOE to perform poorly compared to an RTO system without DOE. One contributing factor is the inability of the constraints that define the experimental window to prevent the process from moving to areas of profit loss that are more significant than expected. This occurs because the constraints (see Equations (3.6) to (3.10)) are attempting to predict the amount of profit generation at the points calculated by the economic optimizer and DOE problems, with the intention of ensuring that the DOE does not implement a point located in a region of significant profit loss, relative to the current operating point or the point predicted by the economic optimizer. If the model developed for optimization is not able to adequately predict profit generation, then the ability of the constraints to avoid implementation of poor results is weakened. Performing diagnostic checks to ensure the model is able to adequately predict profit generation would provide an indication of when the DOE should not be performed.

The four diagnostic tests, summarized below, are performed after the least squares regression to investigate model adequacy and determine if the constraints in the DOE will be effective. If the tests fail, the DOE will not be performed and the decision of whether to implement the result from the economic optimizer will be dependent on the result of the significance tests. Please note that these tests do not guarantee that model prediction at new points will be good, because the tests are performed based on the current operating point. It is possible that the model could be considered adequate under the current set of operating conditions; however, due to the nonlinear nature of the underlying true process behaviour, the model structure may not be adequate for the extrapolated points.

The four tests to evaluate model adequacy are listed below, where the first two tests are defined as approximate, because of the nonlinearity in the equations used for least squares regression.

1. the approximate t ratio;
2. approximate parameter estimate correlation matrix;
3. autocorrelation test;
4. and the cross correlation test.

The approximate $t$ ratio is calculated from the $t$ distribution for the parameter estimates (Bates and Watts, 1988):

$$t \sim \frac{\widehat{\beta} - \beta_o}{se_{\widehat{\theta}}} \tag{4.1}$$

The $t$ ratio is considered significant if it exceeds the $t$ statistic of 1.96, for a 5% level of significance and 999 degrees of freedom, which results from using 1000 data points to construct the measurement covariance matrix:

$$\text{t ratio} = \frac{\text{parameter estimate}}{\text{approximate standard error}} \tag{4.2}$$

where the approximate standard error is the standard deviation of the parameter estimate, found using linear sensitivity analysis (see *Appendix A*). If the $t$ ratio is not considered significant, then it is an indication that the respective parameter should be removed and the assumed model may be incorrect (Bates and Watts, 1988). This approximate $t$ test is essentially testing the hypothesis test:

$$\begin{aligned} H_o &: \widehat{\beta} = 0 \\ H_1 &: \widehat{\beta} \neq 0 \end{aligned} \tag{4.3}$$

The approximate parameter correlation matrix provides an indication of whether any of the parameters are highly correlated, which could be a result of the model being over parameterized. Bates and Watts (1988) discuss that, in general, correlations above 0.99 in absolute value should be examined.

The autocorrelation test is used to evaluate the presence of serial correlations in the residuals ($\varepsilon$) from the model equations. Although RTO primarily uses steady-state equations, these tests still provide a tool to use for diagnostic tests. Concerning the model updater subsystem, an assumption of least squares regression is that the disturbances are independent. Therefore if there is significant autocorrelation, then it gives an indication that plant/model mismatch is present and/or the disturbances have serial correlation which violates the previously stated assumption. This is true if the sample autocorrelation function ($p_\varepsilon(\tau)$) is not zero, except at $\tau = 0$ (Ljung, 1987),

$$p_\varepsilon(\tau) = \frac{\widehat{r}_\varepsilon(\tau)}{\widehat{r}_\varepsilon(0)} \tag{4.4}$$

where $\tau$ is the lag, and the sample autocovariance function is defined as (Ljung, 1987):

$$\widehat{r}_\varepsilon(\tau) = \frac{1}{N}\sum_{t=1}^{N-\tau}\varepsilon(t+\tau)\,\varepsilon(t), \qquad (t \geq 0) \tag{4.5}$$

where $N$ is the number of data points tested. If the residuals were white noise, then the distribution exists (Box and Jenkins, 1976),

$$\sqrt{N}p_\varepsilon(\tau) \sim N(0,1) \tag{4.6}$$

The hypothesis test is defined for $|\tau| \geq 1$:

$$\begin{aligned} H_0 &: p_\varepsilon(\tau) = 0 \\ H_1 &: p_\varepsilon(\tau) \neq 0 \end{aligned} \tag{4.7}$$

where the Normal test is used to determine if $p_\varepsilon(\tau) \neq 0$. For a significance level of 5%, a threshold value of 1.96 is selected (Ljung, 1987),

- if $|p_\varepsilon(\tau)| > \frac{1.96}{\sqrt{N}}$ then invalidate the model;
- if $|p_\varepsilon(\tau)| < \frac{1.96}{\sqrt{N}}$ then validate the model.

The cross correlation test evaluates whether there are any lagged dependencies between the residuals and inputs ($u$). If there is significant cross correlation, then it is an indication of the presence of plant/model mismatch. The sample cross correlation function is defined as (Ljung, 1987),

$$p_{\varepsilon u}(\tau) = \frac{\widehat{r}_{\varepsilon u}(\tau)}{\sqrt{\widehat{r}_\varepsilon(0)\,\widehat{r}_u(0)}} \tag{4.8}$$

where the sample cross covariance between one residual and one input is (Ljung, 1987),

$$\widehat{r}_{\varepsilon u}(\tau) = \frac{1}{N}\sum_{N-\min(0,\tau)}^{N-\max(\tau,0)} \varepsilon(t+\tau)\,u(t) \tag{4.9}$$

If the correlation between the inputs and the residuals were white, then the distribution exists (Box and Jenkins, 1976),

$$\sqrt{N}p_{\varepsilon u}(\tau) \sim N(0,1) \tag{4.10}$$

The hypothesis test is defined as:

$$
\begin{aligned}
H_0 &: \; p_{\varepsilon u}(\tau) = 0 \\
H_1 &: \; p_{\varepsilon u}(\tau) \neq 0
\end{aligned}
\tag{4.11}
$$

Similarly to the autocorrelation test, the normal test with a significance level of 5%, yields a threshold value of 1.96 where,

- if $|p_{\varepsilon u}(\tau)| > \frac{1.96}{\sqrt{N}}$ then invalidate the model;
- if $|p_{\varepsilon u}(\tau)| < \frac{1.96}{\sqrt{N}}$ then validate the model.

## 4.2 Williams-Otto Plant

The Williams-Otto plant, shown in Figure 4.1, is chosen as the case study for this chapter (Williams and Otto, 1960). The plant consists of the following four operating units: a reactor (investigated in Chapters 2 and 3), heat exchanger, decanter and distillation column. The Williams-Otto plant was developed to provide a realistic model that would allow for the evaluation of the applicability of proposed process control schemes before they are implemented in a real plant. Several researchers have also used the Williams-Otto plant to test the validity of their own developments (Krishnan *et al.*, 1992; Forbes, 1994; Lin *et al.*, 1994; Fraleigh, 1999).

The stream labels shown in Figure 4.1 conform to those in the original paper (Williams and Otto, 1960).

The three streams that enter the reactor, of mass holdup $V_R = 4640$ lb, include the pure reactant streams $FA$ and $FB$, and the recycle stream from the distillation tower ($FL$). The reactor temperature is controlled at $T_R$ by the cooling water flow rate, $FUr$. Similar to the reactor case studies in Chapters 2 and 3, the following three exothermic reactions occur:

$$
\begin{aligned}
A + B &\rightarrow C \\
B + C &\rightarrow P + E \\
C + P &\rightarrow G
\end{aligned}
$$

The kinetics of the reactions are temperature dependent as described by the Arrhenius expression in Equation (2.29). The reaction data is presented in Table 4.1.

Table 4.1: Reaction Data

| Reaction Number | Frequency Factors ($\text{hr}^{-1}$) | Activation Energy (R) | Heat of Reaction (BTU/lb of reactant) |
|---|---|---|---|
| 1 | $5.9755 \times 10^{9}$ | 12,000 | −125 |
| 2 | $2.5962 \times 10^{12}$ | 15,000 | −50 |
| 3 | $9.6283 \times 10^{15}$ | 20,000 | −143 |

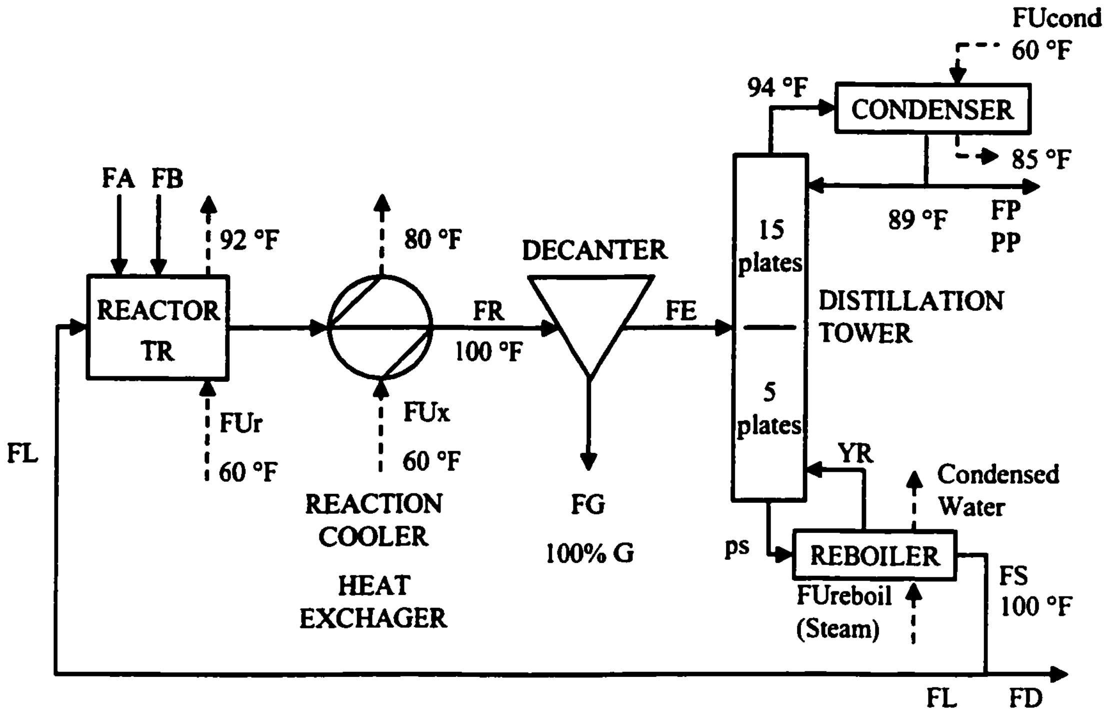



Figure 4.1: Williams-Otto Plant (Williams and Otto, 1960).

The purpose of the reaction cooler is to lower the temperature of the product stream from the reactor to $100^oF$, where material by-product $G$ becomes insoluble in the reaction mixture. The exit temperature of the reaction mixture is controlled by the cooling water flow rate, $FUx$. The decanter is used to remove the by-product, where it is assumed that all of the material $G$ is removed from the reaction mixture (Williams and Otto, 1960).

Once material $G$ is removed, the remaining mixture is forwarded to a distillation tower where component $P$ is separated from the remaining components. Product P has a relative volatility ($\alpha$) of 2.8 with the remaining liquid in stream $FE$, including $P$ trapped in the azeotrope with material $E$,

$$w_P = 0.1 w_E \tag{4.12}$$

where the components are measured in weight fractions.

The feed enters the column as 100% liquid and is separated using 20 bubble cap trays, where the enriching and stripping sections contain 15 and 5 trays, respectively. A total condenser is used to condense the product vapour, and a partial reboiler acts as an extra equilibrium stage to vapourize the column bottoms. The liquid stream from the reboiler ($FS$) is split into two streams, with $FD$ sold for minimal profit and $FL$ recycled back to the reactor.

The design specifications for the various operating units are presented in Table 4.2.

Table 4.2: Unit Design Specifications

| Unit | Specification | Value |
|---|---|---|
| Reactor Cooling Coils | Area | 100 ($ft^2$) |
| | Heat Transfer Coefficient | 0.0195 $F$ ($Btu/hr/R/ft^2$) |
| Heat Exchanger | Area | 569 ($ft^2$) |
| | Heat Transfer Coefficient | 0.0059 $F$ ($Btu/hr/R/ft^2$) |
| Reboiler | Area | 2770 ($ft^2$) |
| | Heat Transfer Coefficient | 0.0041 $F$ ($Btu/hr/R/ft^2$) |
| Condenser | Area | 4960 ($ft^2$) |
| | Heat Transfer Coefficient | 0.0054 $F$ ($Btu/hr/R/ft^2$) |

The plant objective function is defined to maximize the percent return on investment and has the form:

$$\% \text{ Return} = 100 \times \frac{\left(\begin{array}{c}\text{Price of Products FP and FD - Cost of Feeds FA and FB} \\ \text{- Cost of Disposal for FG - Utility Costs} \\ \text{- Charges for Sales, Administration, Research and Engineering}\end{array}\right)}{\text{Total Investment}} \tag{4.13}$$

with the generated revenues and costs for the process streams entered, the objective function becomes,

$$\% \text{ Return} = \frac{1}{0.278} \times \left(\begin{array}{c} 2207.52FP + 50.03712FD - 168FA - 252FB - 84FG \\ -1.27173FR - \frac{8400 \cdot .25 \cdot 10^{-3}}{62.4}(FUr + FUx + FUcond) \\ -8400 \cdot 1.00 \cdot 10^{-3} \cdot FUreboil - 2.76 \end{array}\right) \tag{4.14}$$

where $FUcond$ is the cooling water flow through the condenser and $FUreboil$ is the steam flow through the reboiler.

In this case study the plant is modelled based on the original paper with the following exceptions (Forbes, 1994; Fraleigh, 1999),

- steady-state simulations were performed;
- a log-mean temperature difference was used to model the heat transfer dynamics;
- all heat transfer coefficients were made flow dependent with their exponents set to one (Incropera and DeWitt, 1996);
- the distillation tower was modelled using a tray-by-tray equilibrium relationship and assuming constant molal overflow (King, 1980);
- the relative volatility was set to 2.8 to match the nominal case presented by Williams and Otto (1960);
- the minimum concentration of $P$ in the product stream ($FP$) was set at 95 wt%.

The following operating constraints are implemented,

- the product flow rate must be less than 4,763 lb/hr, because any extra exceeds the available sales and cannot be sold;
- the concentration of $P$ in the product stream must exceed 95 wt.% to satisfy product specifications.

Similar to the case studies presented in Chapters 3 and 4, the reactant stream $FA$ is fixed at 14,500 lb/hr. This leaves five degrees of freedom available for optimization. The five selected decision variables are:

1. the flow of reactant $B$ into the reactor ($FB$);
2. reactor temperature ($T_R$);
3. recycle flow into the reactor ($FL$);
4. distillate flow from the distillation tower ($FP$);
5. weight fraction of material P in the distillate flow ($PP$).

Using this set of decision variables the true plant optimum is located at the point shown in Table 4.3. It is expected that the optimum should be located at the intersection of the two operating constraints, since maximizing $FP$ generates the most revenue and minimizing $PP$ lowers utilities usage. Therefore this case study will focus on setting $FB$, $T_R$ and $FL$ as the decision variables. Fraleigh (1999) used a similar method for her case study of the Williams-Otto plant.

Table 4.3: Williams-Otto Plant Optimum

| Variable | Optimal Value |
|---|---|
| FB | 29,216 lb/hr |
| TR | 635.4 R |
| FL | 63,132 lb/hr |
| FP | 4,763 lb/hr |
| PP | 0.95 |
| % Return | 64.3372 % |

## 4.3 Experimental Design Problem

Consider a scenario in which the Williams-Otto plant described in section 4.2 has been operating for an extended period of time with a RTO system supervising the control layer. A cost estimate has been performed and concluded that a significant increase in profit could be realized if more precise estimates of the relative volatility and frequency factors in the reaction equations could be developed. It is thought that adding an information generation component to the RTO system will improve parameter estimation and increase profit.

This section will investigate the feasibility of adding a DOE component to a RTO system operating for a large plant with interacting units.

### 4.3.1 Method

The problem was scaled to ensure a more reliable solution for the optimization and equation solving routines. The reaction equations were scaled in a similar format to that discussed in section 2.4.2, so in general the variable values lie in the range 0.01-10. The following variables were further scaled as shown in Table 4.4.

Table 4.4: Variable Scaling for Williams-Otto Plant

| Variable | Notes | Scaling Factor |
|---|---|---|
| Flows | All flows excluding *FUreboil* | divided by 100,000 |
| | *FUreboil* | divided by 1,000 |
| Molar Flow Rates | liquid and vapour flow in tower | divided by 1,000 |
| Molecular Weights | | divided by 100 |
| Temperatures | | divided by 1,000 |
| Arrhenius Exponent | ($B_1$, $B_2$, $B_3$) | divided by 1,000 |
| Reactor Volume | ($V_R$) | divided by 100 |
| Heats of Reaction | | divided by 10 |
| Heat Capacities | | multiplied by 100 |
| Heat Duties | | divided by 1,000,000 |
| Heat Transfer Coefficients | All except Reboiler (HTC) | multiplied by 100 |

For this case study the set of adjustable parameters consists of:

- the frequency factors from the reaction equations: $A_1$, $A_2$, $A_3$;
- and the volatility of $P$ relative to the remainder of the feed mixture to the distillation tower: $\alpha$.

To update these parameters, a least squares regression optimization routine was applied using a steady-state model represented by the following set of equations:

$$\frac{FL}{V_R}AS + \frac{FA}{V_R} - \frac{FR}{V_R}AR - k_1\,(AR)\,(BR) = 0 \quad (4.15)$$

$$\frac{FL}{V_R}BS + \frac{FB}{V_R} - \frac{FR}{V_R}BR - k_1\,(AR)\,(BR) - k_2\,(BR)\,(CR) = 0 \quad (4.16)$$

$$\frac{FL}{V_R}CS - \frac{FR}{V_R}CR + 2k_1\,(AR)\,(BR) - 2k_2\,(BR)\,(CR) - k_3\,(CR)\,(PR) = 0 \quad (4.17)$$

$$\frac{FL}{V_R}ES - \frac{FR}{V_R}ER + 2k_2\,(BR)\,(CR) = 0 \quad (4.18)$$

$$-\frac{FR}{V_R}GR + 1.5k_3\,(CR)\,(PR) = 0 \quad (4.19)$$

$$\frac{FL}{V_R}PS - \frac{FR}{V_R}PR + k_2\,(BR)\,(CR) - 0.5k_3\,(CR)\,(PR) = 0 \quad (4.20)$$

$$(1 + (\alpha - 1)\cdot ps)\cdot YR - \alpha \cdot ps = 0 \quad (4.21)$$

where

$$FR = FA + FB + FL$$

and $k_i$ is represented by the Arrhenius expression in Equation (2.38), and $YR$ and $ps$ are the mole compositions of $P$ in the vapour and liquid phases leaving the reboiler (excluding $P$ caught in the azeotrope), respectively. Equations (4.15) to (4.20) represent the component material balances of the reactor and Equation (4.21) represents the vapour-liquid equilibrium (VLE) relationship around the reboiler unit. The least squares regression is formulated as an implicit problem as discussed in Chapter 2. This formulation allows for the residuals to reflect the lack of satisfaction of the steady-state balance equations.

Equations (4.15) to (4.21) define a reduced model that is used to estimate the parameters. The reduced model focuses on the two pieces of equipment (reactor and column reboiler) that are directly characterized and affected by the parameters. This is accomplished by assuming that available measurements are located close to the process equipment *i.e.* reaction outlet compositions, and compositions around the reboiler. It should be noted that the estimates of the parameters might change somewhat if the full model were used; however, since the selected equations and measurements were strongly linked to the parameters, the change would likely be relatively small. The advantage of using a reduced model is to improve the computational simplicity. The topic of incorporating reduced models for parameter estimation is left for future research and is not within the scope of this thesis.

Similar to the case studies involving the reactor in Chapters 2 and 3, it was assumed that *a priori* knowledge of the covariance structure for the measurements is unavailable. The measurement covariances of the material compositions for streams FP and FS were found by generating a set of $1,000$ independent process data points around a nominal value with a white noise distribution added to each data point for each measurement. This nominal value along with the measurement variances and covariances may be found in *Appendix E*. Table

2.4 shows the applied standard deviations of the measurements, expressed as a percentage of the nominal case. The same concerns as discussed in Chapter 2 are also considered for this case study by stating that although this is not strictly correct for this case, it does create a study of mismatch in the covariance structure.

The initial data set consisted of ten steady-state operating points selected randomly from a uniform distribution in the range $[FB, TR, FL] = [35,000, 39,000] \times [630, 635] \times [40,000, 45,000]$. This data set is found in *Appendix E*.

Referring to Chapter 2, the first step was to evaluate the optimal number of historical steady-state data points that should be used in the model updater. Up to ten historical points were tested using the decision criterion defined in Equation (2.7). The estimated covariance matrices of the adjustable parameters and the optimization calculations were evaluated using linear sensitivity analysis (see *Appendix A*). The results of the case study from Chapter 2 (see *Appendix C*), showed a minimal difference in the sensitivity matrix of the decision variables to the parameter estimates between tests; therefore, the matrix was evaluated once and kept constant for the ten tests. Miletic and Marlin (1998) discuss that it does not need to be recalculated unless significant changes in measurement accuracy or process operation occur. The volume of the confidence regions were compared to determine the extent of transmission of noise from the measurements to the adjustable parameters and the optimization calculations.

The algorithm used to implement the DOE in the RTO system is listed in the following numbered list. Similar to the optimal result found in the Chapter 3 case study, the decision criterion tuning parameter ($\alpha$) was set to one, and it was assumed that the allowed amount of sacrificed profit for DOE was defined by Equations (3.6) and (3.7) with,

$$c_1 = 0.980$$
$$c_2 = 0.999$$

In order to compare the performance of the RTO systems with and without DOE, data from ten runs consisting of 25 RTO intervals was collected for each system. The extended design cost (EDC) measure, developed by Zhang and Forbes (2000), was used to measure the performance of the RTO system (see Equation (3.11)).

The case study was repeated to investigate the comparison of the RTO systems when parametric mismatch exists. Similar to one of the case studies in Chapter 3, a 5% increase was applied to the assumed values of the activation energies. Referring to Table 3.1, the assumed values became,

$$B_1 = 12,600 \quad R$$
$$B_2 = 15,750 \quad R$$
$$B_3 = 21,000 \quad R$$

Similar to the case study in Chapter 3, one-sided hypothesis tests were used to evaluate whether the performance of the RTO system improves with DOE incorporated in the algorithm.

The following algorithm outlines the steps used in the RTO loop.

1. Complete the parameter estimation step using an implicit least squares regression optimization routine, see Problem (2.1), with the optimal number of historical steady-state points predefined.

2. Perform the set of diagnostic tests discussed in section 3.1. If any of the tests fail, the experimental design is not performed (Steps #5, #8 and #9 are ignored). The structure of the tests are shown at the conclusion of this list.

3. The sensitivity matrix of the parameters to the measurements, $\frac{d\boldsymbol{\beta}}{d\mathbf{z}}$, is found using the optimal number of historical steady-state measured points.

4. Using the updated model, perform the optimization step using the SQP algorithm in the optimization toolbox of the computer software package MATLAB ver. 5.3.0 (Branch and Grace, 1996).

5. The measured values of the manipulated variables combined with the process model, are used to solve for the expected values of the response variables based on those measurements. This result is later used to develop the profit constraints that define the experimental window.

6. The results analysis step is performed to determine whether the result from the economic optimizer is significant. Please see the hypothesis test represented by Problem (3.1) (Miletic and Marlin, 1998). The testing is performed for a hypothesis test at a 5% significance level.

7. The sensitivity matrix of the optimization variables to the parameters, $\frac{d\mathbf{x}_p}{d\boldsymbol{\beta}}$, is found in this step.

8. The command conditioning subsystem determines if the point predicted by the economic optimizer is expected to generate information with respect to Inequality (3.5).

9. If the result from the economic optimizer is expected to generate information, the result is implemented despite its status of being significant. If it is not expected to generate information then an experimental design (see Problem (2.20)) is performed with an added constraint to limit the amount of profit lost. The form of the constraint is dependent on whether the prediction from the economic optimizer is significant (see Inequalities (3.6) and (3.7)). The sensitivity matrix, $\frac{d\boldsymbol{\beta}}{d\mathbf{z}}$, used in the experimental design objective function is based on the most recent optimal number of steady-state data points subtract one, which is replaced by the next design point to augment the data set. The sensitivity matrix, $\frac{d\mathbf{x}_p}{d\boldsymbol{\beta}}$, is found from a previous calculation since it can only be estimated from the result of an economic optimization calculation. If the experimental design is centered around the predicted optimum, then the result from

step #7 is used. However, if the experimental design is centered around the current operating point then the result from the previous RTO cycle is used.

10. The actual amount of profit generated is based on the setpoints returned from the optimizer since it is assumed the controllers are able to implement them.

11. Steps #1-#10 are repeated.

The following displays the structure of the diagnostic tests performed in step #2. If the test is rejected, then the diagnostic test is failed.

**Approximate $t$ Ratio:**

$$\text{if } \min |\text{t ratio}| < 1.96, \text{ reject}$$

**Approximate Parameter Estimate Correlation Matrix:**

$$\text{if } \max_{\substack{ij \\ i \neq j}} |v_{ij}| > 0.99, \text{ reject}$$

where $v_{ij}$ represents the $ij^{th}$ coefficient of the approximate parameter estimate correlation matrix.

**Autocorrelation Test:**

$$\text{if } \max |p_{\varepsilon}(\tau)| > \frac{1.96}{\sqrt{N}}, \text{ reject}$$

**Cross Correlation Test:**

$$\text{if } \max |p_{\varepsilon u}(\tau)| > \frac{1.96}{\sqrt{N}}, \text{ reject}$$

### 4.3.2 Results and Discussion

This section concludes with the presentation and discussion of the results for the case studies involving the Williams-Otto plant. The first case study is performed to determine the number of historical steady-state points that should be used in the model updater subsystem. The second case study investigates the improvement in profit that can be realized by incorporating DOE into the RTO system when plant/model mismatch is absent. The final case study is similar to the second; however, it investigates the performance of the RTO systems with and without DOE in the presence of parametric mismatch.

#### Historical Points

The results for this case study are tabulated in Table 4.5, and illustrated in Figure 4.2. Please see *Appendix E* for the results of the model updater and the values of the sensitivity matrices.

Table 4.5: Results Summary for Optimal Historical Point Selection for the Williams-Otto Plant

| Historical Data Points | det ($\mathbf{Q}_\beta$) ($\times 10^{-3}$) | det ($\mathbf{Q}_\mathbf{x}$) ($\times 10^{-15}$) | Decision Criterion ($\times 10^{-13}$) |
|---|---|---|---|
| 1 | 15.2524 | 23.1367 | 0.7134 |
| 2 | 0.7076 | 2.6354 | 0.1889 |
| 3 | 0.1424 | 0.6550 | 0.0999 |
| 4 | 0.0595 | 0.3127 | 0.0917 |
| 5 | 0.0212 | 0.1504 | 0.0773 |
| 6 | 0.0112 | 0.0957 | 0.0799 |
| 7 | 0.0066 | 0.0622 | 0.0793 |
| 8 | 0.0042 | 0.0427 | 0.0793 |
| 9 | 0.0028 | 0.0337 | 0.0876 |
| 10 | 0.0019 | 0.0248 | 0.0873 |

Table 4.5 shows, as anticipated, how the volumes of the confidence regions for both the parameter estimates and the decision variables decrease as more historical points are used. Figure 4.2 displays that the decision criterion, defined by Equation (2.7), selects five steady-state points as the optimal amount of historical data. Please note that the horizontal axis of Figure 4.2 begins at 2, to better illustrate the change in the design criterion as more historical points are added.

The decision criterion plot for the Williams-Otto plant, unlike the reactor case study in Chapter 2, is not unimodal because of noise effects. The determinant, similar to any statistic, has confidence regions that surround its mean value. This uncertainty in its value could lead to a plot such as Figure 4.2 that does define a clear minimum. The larger amount of uncertainty in the plant case study compared to the reactor case study is attributed to a more significant trade-off that results from including more points. It is advantageous to include more points as more information is obtained; however, the drawback is that more noise is also introduced. The extra information is limited as the placement of the past historical points is dictated by earlier considerations, so if the process operation changes then the criteria defining the DOE will change. In this case, the noise that is introduced from the measurements is not offset by the information obtained.

### 4.3.3 RTO Use

This case study assesses performance improvement of the RTO system with the additions proposed by this thesis, which include adding a DOE component and incorporating historical data, when structural mismatch does not exist. The original RTO system was built to implement points that were predicted by the optimizer, only if they were found to be statistically significant (Miletic and Marlin,1998). The second set of simulations involves the implementation of the RTO with DOE as outlined in Chapter 3.

The results for the Williams-Otto plant case study are shown in Table 4.6. The average

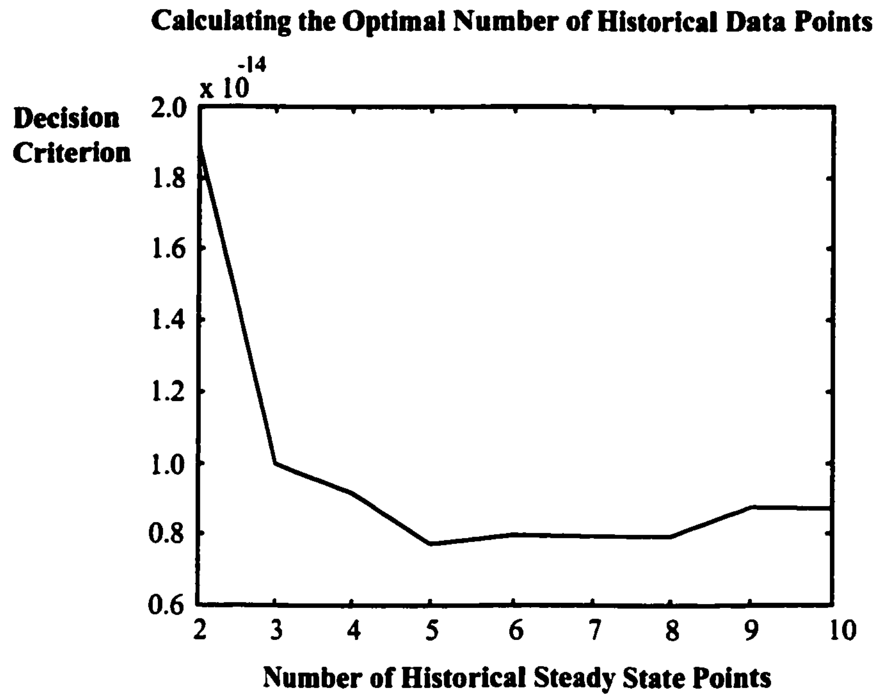



Figure 4.2: The optimal number of historical steady-state date points to be used by the model updater subsystem for the Williams-Otto plant is five.

EDC for the two systems is defined as the average of the EDC measure, developed by Zhang and Forbes (2000), over ten runs. The variance column refers to the variation around the expected value of the predicted optimum. Please see *Appendix E* for the profit profiles and the design costs for the individual tests.

Table 4.6: Williams-Otto Plant Case Study Results

| **Case Study** | **Average EDC** (% Return) | **Variance** (% Return) |
|---|---|---|
| RTO without DOE | 90.1605 | 8.94 |
| RTO with DOE | 72.8142 | 3.55 |

Both the results for the average EDC and variance were found to improve significantly based on the results from one-sided hypothesis tests, with a 5% significance level (see Problems (3.12) to (3.15)). Please see *Appendix E* for the results of the hypothesis tests.

It is also noted that the diagnostic tests were passed for all the runs. This result is expected since structural mismatch did not exist for this case study. To visualize the results submitted by the diagnostic tests, specific details are shown as an example for the first interval of the first run for the RTO system with DOE.

The results of the approximate $t$ ratio are shown in Table 4.7. Since min $|t \text{ ratio}| > 1.96$, the test was passed.

Table 4.7: Approximate t Test

| Parameter | t Ratio |
|---|---|
| $A_1$ | 22.62 |
| $A_2$ | 22.23 |
| $A_3$ | 18.59 |
| $\alpha$ | 150.93 |

With the maximum off-diagonal of the approximate parameter estimate correlation matrix, shown here, less than 0.99 the test was passed.

$$v = \begin{bmatrix} 1 & 0.90 & 0.89 & 0 \\ 0.90 & 1 & 0.82 & 0 \\ 0.89 & 0.82 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The sample autocorrelation function for the model updating equations, defined as Equations (4.15) to (4.21), are illustrated in Figures 4.3 a) to g), respectively. The threshold limit of $\frac{1.96}{\sqrt{N}}$ was not exceeded by the sample autocorrelation function for any of the model equations, so the test was passed.

For the sample cross correlation function, the max $|p_{\varepsilon u}(\tau)|$ for each set of residuals belonging to Equations (4.15) to (4.21) is displayed in Table 4.8. To show the entire sample cross correlation function for each set of residuals and measured variables would present too much information. The test was passed since the threshold limit of $\frac{1.96}{\sqrt{N}}$ was not exceeded by the sample cross correlation function for any of the model equations.

Table 4.8: Cross Covariance Data

| Residuals | max $\lvert p_{\varepsilon u}(\tau)\rvert$ |
|---|---|
| Equation (4.15) | 0.0306 |
| Equation (4.16) | 0.0380 |
| Equation (4.17) | 0.0960 |
| Equation (4.18) | 0.0887 |
| Equation (4.19) | 0.0817 |
| Equation (4.20) | 0.0571 |
| Equation (4.21) | 0.0293 |

Figure 4.4 displays histograms for eight RTO intervals, comparing the ten runs performed for each RTO system. The eight selected intervals, out of a possible twenty-five, were chosen because they provided the most informative observations. Please note that the histograms show the % Return and not the EDC measure. Figures 4.4 a), b), c), and g) show that reducing uncertainty in the decision variables can move the profit distribution closer to the true plant optimum. Figures 4.4 a), b), d), e), f), and h) illustrate how the RTO system with an information generation component, is able to avoid implementing points in regions of significant profit loss. This is attributed to better optimization results, that occur

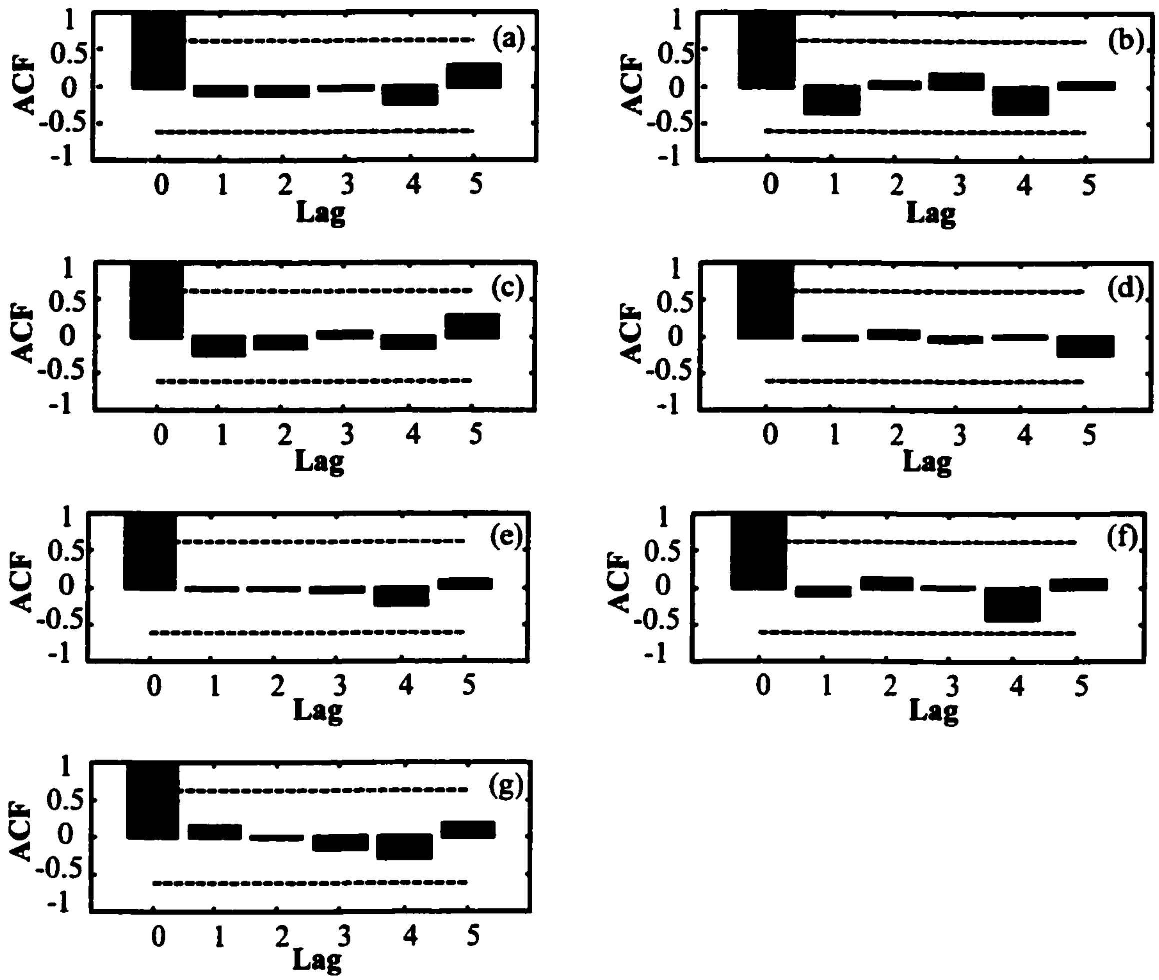



Figure 4.3: The sample autocorrelation function was not violated by Equations (4.15) to (4.21) used for parameter estimation (Data removed from the first interval of the first run for the RTO system without DOE).

because the information generation step is present, to ensure high quality information is produced to improve the predictability of the optimizer.

The results of this case study are consistent with those generated from the case study for Chapter 3 showing that for a more realistic model, incorporating DOE and reducing uncertainty in the decision variables can improve the performance of the RTO system.

### 4.3.4 RTO Use with Parametric Mismatch

This case study compares the performance, in the presence of parametric mismatch, of the two RTO systems tested in the previous case study. The mismatch was created with a 5% bias in the activation energies, as defined in Table 3.1, for the assumed model.

The results for the Williams-Otto plant case study, with mismatch, are presented in Table 4.9. Please see *Appendix E* for the profit profiles and the extended design costs for

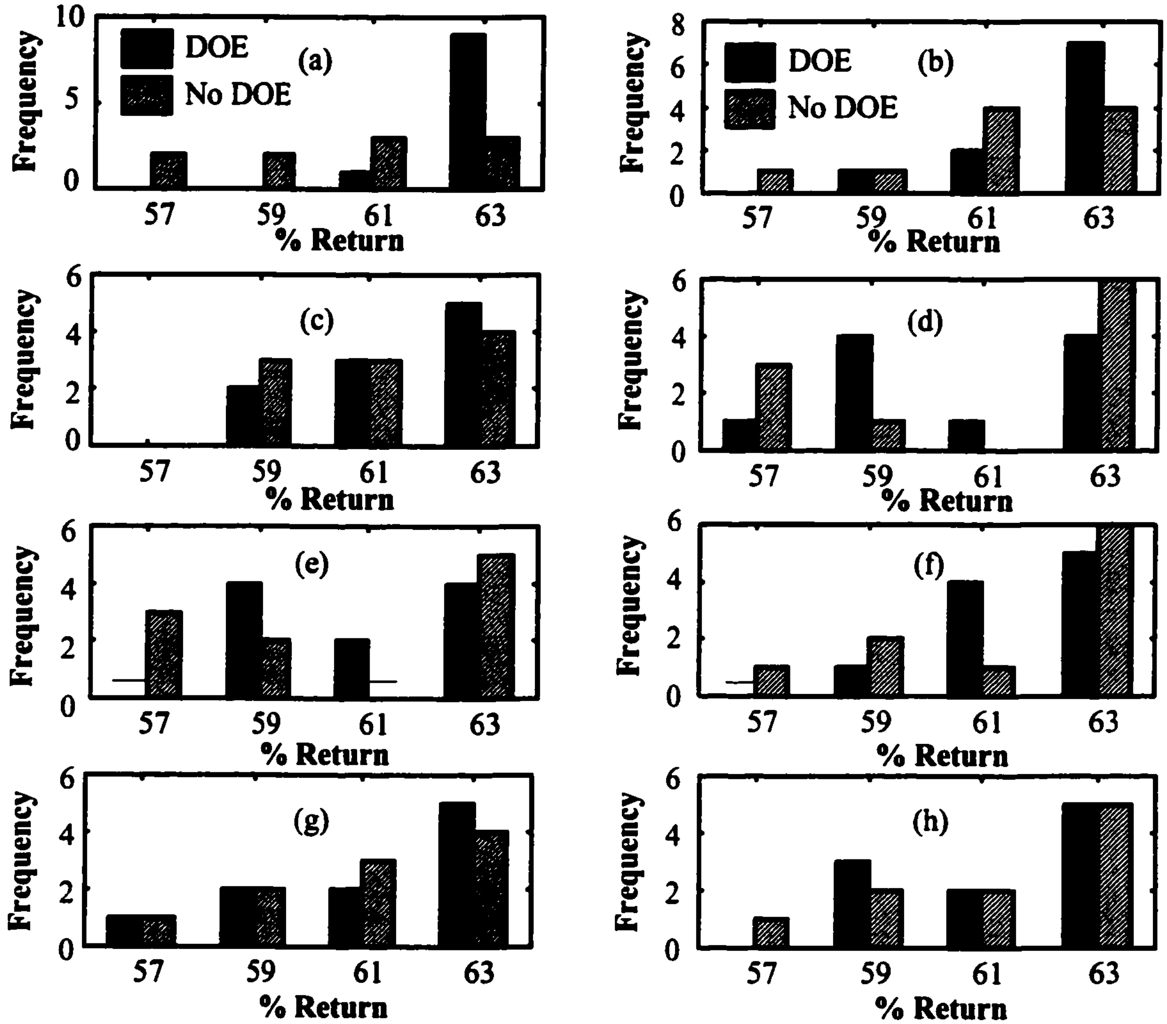



Figure 4.4: The set of histograms provides a visualization of the ten runs for each RTO system for the Williams-Otto Plant: a) 4th Interval; b) 5th Interval; c) 11th Interval; d) 17th Interval; e) 18th Interval; f) 19th Interval; g) 21st Interval; h) 23rd Interval.

the individual tests.

Table 4.9: Williams-Otto Plant Case Study Results: With Mismatch

| Case Study | Average EDC (% Return) | Variance (% Return) |
|---|---|---|
| RTO without DOE | 147.4623 | 17.08 |
| RTO with DOE | 118.7382 | 2.03 |

Both the results for the average EDC and variance were found to improve significantly based on the results from one-sided hypothesis tests, with a 5% significance level (see Problems (3.12) to (3.15)). This is in contrast to the case study performed in Chapter 3, where the 5% bias case performed poorly for the RTO with DOE system. Please see *Appendix E* for the results of the hypothesis tests.

The improvement in the mismatch case studies from Chapter 3 to Chapter 4 is credited to the addition of diagnostic tests to evaluate model adequacy. If the constraints representing the experimental window in the DOE are not expected to be able to adequately predict the process behaviour, then DOE is not considered. The RTO system is reduced to the base case, where predictions from the economic optimizer are implemented only if they are considered statistically significant.

Figure 4.5 displays histograms showing % Return for eight RTO intervals, comparing the ten runs performed for each RTO system. The eight selected intervals were chosen because they provided the most informative observations. Please note that the histograms show the % Return and not the EDC measure. Figures 4.5 a) to h) show that reducing uncertainty in the decision variables reduces variation around the expected value of the predicted optimum at the RTO intervals displayed. This is important to avoid implementing relatively costly results that lie in regions of significant profit loss. However, Figure 4.5 also shows that the RTO system with DOE does not produce as many highly profitable results as the system without DOE (points on the right side of the histograms). This occurs because the DOE is implementing results that deviate from the true plant optimum to generate information. This drawback is considered acceptable, since the improved information generation reduces the number of poor predictions leading to a better overall RTO system performance, specifically a significantly lower EDC measure.

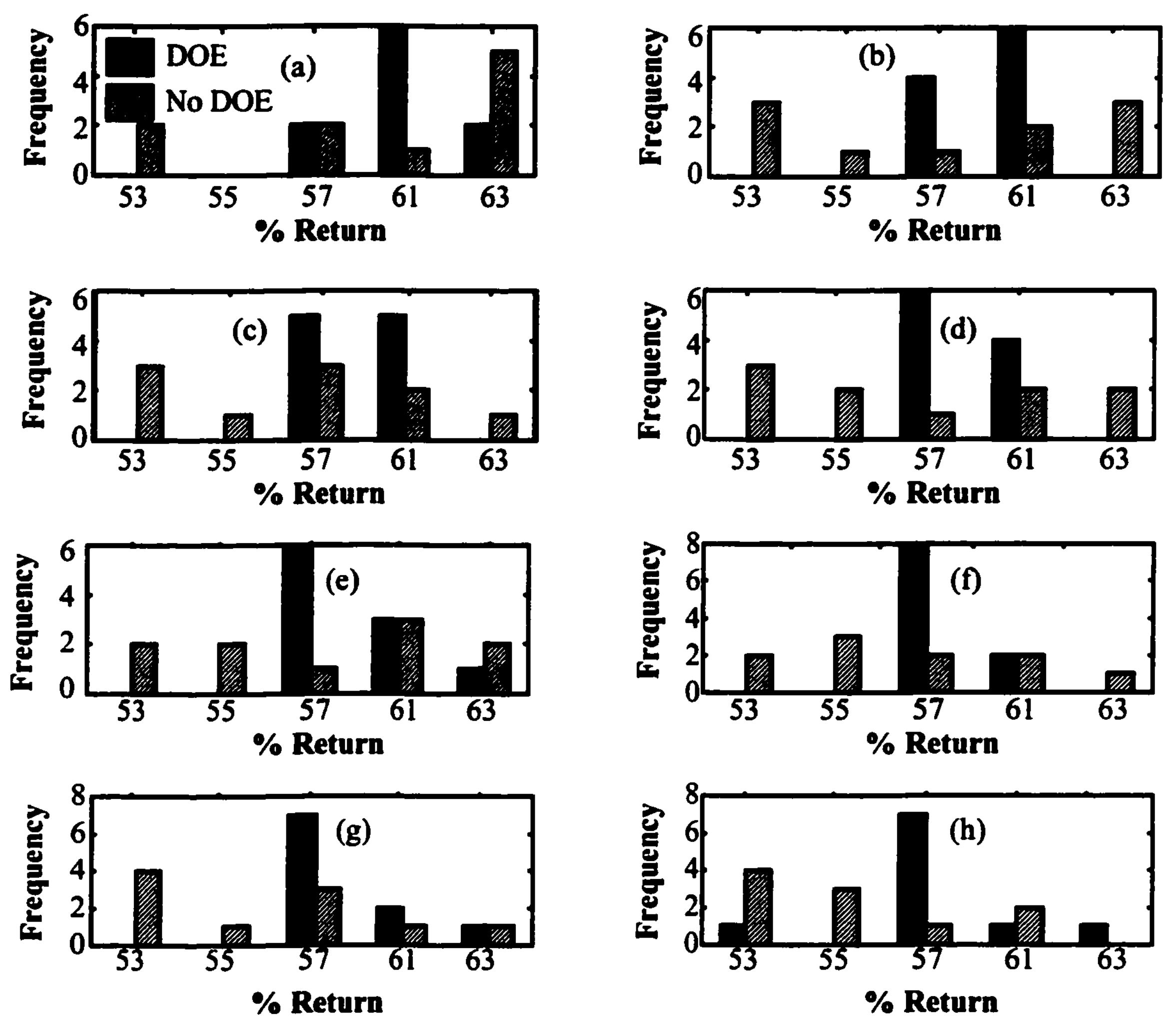



Figure 4.5: The set of histograms provides a visualization of the ten runs for each RTO system for the Williams-Otto Plant in the presence of parametric mismatch: a) 7th Interval; b) 9th Interval; c) 11th Interval; d) 13th Interval; e) 17th Interval; f) 19th Interval; g) 21st Interval; h) 25th Interval.

# Chapter 5

# Summary and Conclusions

With competition increasing as the global market place continues to open, many industries are moving to become the most efficient producer of their product by operating their facilities as optimal as possible. On-line model-based optimization approaches, such as the type investigated by this thesis, have garnered significant interest to help achieve these goals. By integrating experimental design techniques, this thesis has proposed a new approach for incorporating information generation into the RTO algorithm to reduce uncertainty in the final optimization results.

## 5.1 Summary

Currently, most RTO algorithms only use the current operating point to update the adjustable parameters in the model used for optimization. The majority of applications do not even make use of a least squares regression optimization routine, instead selecting a nonlinear equation solver, or back-substitution. In considering implementation of DOE or regression techniques, it does not make sense to base the calculations on a single point, from either an experimental design point of view or from the perspective of the quality of estimates obtained through regression. Therefore, it becomes necessary to alter the RTO system to include more than one historical point for calculation in the model updater and command conditioning subsystems. Chapter 2 introduced the topic of including historical steady-state information, and identified the trade-offs as: reducing uncertainty in the decision variables, increasing the computational load, and putting the model updater to 'sleep' or reducing the optimizer's ability to track changes in the process behaviour.

Chapter 2 also investigated the advantages of DOE for the purposes of optimization. It showed that moving in a different direction through planned experimentation, instead of the direction specified by the economic optimizer, can reduce uncertainty in the decision variables for future RTO intervals. However, the result from the DOE can predict points that lie in regions of significant profit loss. Since the primary purpose of RTO is to increase profit, a trade-off is required that will sacrifice profit in the short term for improved profit generation in future RTO intervals.

Chapter 3 introduced a series of alterations to the command conditioning subsystem (see Figure 3.1). The adjustments included expanding the results analysis component to determine whether the point predicted by the optimizer will generate a sufficient amount of information. If the predicted point was not expected to produce a level of information above a threshold value, then the DOE was performed with constraints that defined an experimental window based on acceptable profit loss. When structural mismatch was not present, the case study showed for tuning parameters that defined a small experimental window, that the proposed RTO system was able to out-perform the more conventional RTO system. This indicated that small deviations from the predicted result of the economic optimum to generate information can improve the performance of the RTO system. If the deviations are allowed to be too large, then the sacrificed profit at the current RTO interval is not made up for at future intervals. A second case study in the chapter showed that for minimal amounts of parametric mismatch, the RTO system with DOE still out-performed the traditional RTO algorithm. However, as the mismatch became more significant, the results reversed with respect to the extended design cost measure. This was attributed to the inability of the constraints, defining the experimental window, to predict the behaviour of the profit function. The constraints were established to prevent the DOE from predicting points that lie in regions of significant profit loss; however, if the model was unable to adequately predict the behaviour of the process, it resulted in points being predicted that lead to more than expected profit loss. The DOE was able to reduce variance, until the 10% bias case, which was expected since that is the objective of the experimentation. The additional loss in the extended design cost is associated with bias.

The case study in Chapter 4 demonstrated how the RTO system with DOE performs for a larger system with interacting operating units. Chapter 4 also introduced the concept of incorporating a series of diagnostic checks to test model adequacy, with the purposes of identifying when DOE should not be performed, because of its reliance on the model to predict process behaviour for the experimental window constraints. The case study results were positive to this addition, as the new RTO system out-performed the more conventional system for the same level of mismatch that it performed poorly for in Chapter 3.

## 5.2 Contributions

This thesis proposed a new approach for incorporating information generation into the RTO algorithm to reduce uncertainty in the final optimization results. It was shown, for the scenarios when parametric mismatch is either absent or minimal, that the performance of the RTO system improved with the proposed adjustments incorporated in the algorithm. Where performance was measured by the average extended design cost, and also considered the variance around the expected value of the optimum point. However, the case study in Chapter 3 also showed that as parametric mismatch was made more significant, the performance of the proposed RTO system deteriorated relative to the RTO system without

DOE.

As mentioned earlier, this deteriorating effect was hypothesized to be a result of a reduced ability in the constraints, defining the experimental windows, to predict the process behaviour. The mismatch contributed to the experimental windows defining regions of profit loss more significant than desired, which enabled the DOE to implement poor economical points. This thesis proposed employing a set of diagnostic tests to track model prediction inadequacies to determine when the experimental windows might not be able to predict process behaviour effectively; therefore, identifying instances when the DOE should not be implemented. The case study in Chapter 4, showed that these adjustments improved the performance of the proposed RTO system in the presence of parametric mismatch.

## 5.3 Future Work

Although the results in this thesis were positive, there are some concerns for implementing DOE in RTO systems. The primary issue was the effect of plant/model mismatch. The DOE developed in this thesis was labeled as a variance-optimal design. This means that the experimental design focuses on minimizing the variance of the decision variables estimated, assuming that the model structure is correct. Inevitably, every practical RTO application will contain some degree of plant/model mismatch, and therefore, some amount of bias in the assumed model. However, the current state of development for experimental designs that consider reducing bias and variance together, is not at a level that could be considered for an RTO application for the reasons of heavy computational load, difficulty in implementing for complicated implicit models, and the requirement to know the structure of the sources of possible biases.

Chapter 4 introduced the concept of using reduced models for the purpose of model updating in the RTO loop. The selected equations in the reduced model describe the process equipment that are directly characterized and affected by the parameters. To use a reduced model it is beneficial to position the sensors that are used close to the process equipment. This ensures that the selected equations and measurements are strongly linked to the parameters to limit the error in the estimates of the parameters compared to the estimated calculations if the full model was used. The advantage of using a reduced model is to improve the computational simplicity. It is recommended to address the topic of using reduced models to evaluate the trade-off between reduced precision in the parameter estimates and reduced computational work.

This thesis did not incorporate *a priori* knowledge into the structure of the measurement covariance matrix that was used for the DOE or the diagnostic tests. It is left for future work to investigate the incorporation of knowledge into the measurement covariance structure. It is expected that it will improve the performance of the proposed RTO system, since adding quality information about the measurement variance will improve the estimates generated by the DOE further, allowing for further reduction in decision variable uncertainty.

An additional point that needs to be reiterated from Chapter 2 is the method used to determine the number of historical points that should be used is not an analytical solution. The expression illustrates what issues needs to be considered, and a possible structure to represent the criteria. It is suggested to focus future work on developing a method that would provide a better indication of the optimal number of historical steady-state data points to use in RTO.

There are several adjustments to the DOE that should be explored. One approach is to perform the economic optimization with an embedded information generation constraint in the optimization problem. Genceli and Nikolaou (1996) proposed a similar method by adding an excitation constraint to the model predictive control algorithm for the purpose of improving model identification. Another formulation consists of performing a vector optimization for the competing objectives of profit and information generation. The drawback to this method is the computational demand that is required. Finally, it may be beneficial to direct some of the experimental effort to evaluate where the plant/model mismatch exists, and attempt to improve the model, thus reducing bias.

# Bibliography


Atkinson, A. C. and D. R. Cox (1974). Planning experiments for discriminating between models. *Journal of the Royal Statistical Society* **36**, 321–348.

Atkinson, A. C. and V. V. Fedorov (1975). The design of experiments for discriminating between two rival models. *Biometrika* **62**(1), 57–70.

Bacon, D. W. (1992). Participant guide: Design of experiments for engineers workshops. Council for Continuous Improvement.

Bamberger, W. and R. Isermann (1978). Adaptive on-line steady-state optimization of slow dynamic processes. *Automatica* **14**, 223–230.

Bard, Y. (1974). *Nonlinear Parameter Estimation.* Academic Press. New York.

Bates, D. M. and D. G. Watts (1988). *Nonlinear Regression Analysis and its Applications.* Wiley. New York.

Becerra, V. M. and P. D. Roberts (1996). Dynamic integrated system optimization and parameter estimation for discrete time optimal control of nonlinear systems. *International Journal of Control* **63**(2), 257–281.

Box, G. E. P. (1957). Evolutionary operation: A method for increasing industrial productivity. *Applied Statistics* **6**(2), 81–101.

Box, G. E. P. (1970). Improved parameter estimation. *Technometrics* **12**(2), 219–229.

Box, G. E. P. and G. M. Jenkins (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day. San Francisco.

Box, G. E. P. and H. L. Lucas (1959). Design of experiments in non-linear situations. *Biometrika* **46**, 77–89.

Box, G. E. P. and K. B. Wilson (1951). On the experimental attainment of optimum conditions. *Journal of the Royal Statistical Society* **13**, 1–45.

Box, G. E. P. and N. R. Draper (1959). A basis for the selection of a response surface design. *Journal of the American Statistical Association* **54**, 622–654.

Box, G. E. P. and N. R. Draper (1963). The choice of a second order rotatable design. *Biometrika* **50**, 335–352.

Box, G. E. P. and N. R. Draper (1987). *Empirical Model-Building and Response Surfaces.* Wiley. New York.

Box, G. E. P. and W. G. Hunter (1965). Sequential design of experiments for nonlinear models. *I.B.M. Scientific Computing Symposium in Statistics* pp. 113–137.

Box, G. E. P. and W. J. Hill (1967). Discrimination among mechanistic models. *Technometrics* **9**(1), 57–71.

Branch, M. A. and A. Grace (1996). *MATLAB Optimization Toolbox User's Guide*. 2nd ed.. The Math Works. Natick, MA.

Chen, C. Y. and B. Joseph (1987). On-line optimization using a two-phase approach: An application study. *Industrial and Engineering Chemistry Research* **26**, 1924–1930.

Crowe, C. M. (1986). Reconciliation of process flow rates by matrix projection. II: The nonlinear case. *AIChE J.* **32**(4), 616–623.

Crowe, C. M. (1988). Recursive identification of gross errors in linear data reconciliation. *AIChE J.* **34**(4), 541–550.

Crowe, C. M. (1996). Formulation of linear data reconciliation using information theory. *Chemical Engineering Science* **51**(12), 3359–3366.

Crowe, C. M., Y. A. Garcia Campos and A. N. Hrymak (1983). Reconciliation of process flow rates by matrix projection. i: The linear case. *AIChE J.* **29**(6), 881–888.

Dahlquist, G. and A. Bjork (1974). *Numerical Methods*. Prentice-Hall. New Jersey.

Darby, M. L. and D. C. White (1988). On-line optimization of complex process units. *Chemical Engineering Progress* pp. 51–59.

Dette, H. and W. K. Wong (1995). On g-efficiency calculation for polynomial models. *The Annals of Statistics* **23**(6), 2081–2101.

Draper, N. R. and I. Guttman (1992). Treating bias as variance for experimental design purposes. *Annals of the Institute of Statistical Mathematics* **44**(4), 659–671.

Draper, N. R. and W. G. Hunter (1966). Design of experiments for parameter estimation in multiresponse situations. *Biometrika* **53**, 525–533.

Draper, N. R. and W. G. Hunter (1967*a*). The use of prior distributions in the design of experiments for parameter estimation in nonlinear situations. *Biometrika* **54**, 147–153.

Draper, N. R. and W. G. Hunter (1967*b*). The use of prior distributions in the design of experiments for parameter estimation in non-linear situations: Multiresponse case. *Biometrika* **54**, 662–665.

DuMouchel, W. and B. Jones (1994). A simple bayesian modification of d-optimal designs to reduce dependence on an assumed model. *Technometrics* **36**(1), 37–47.

Dykstra, O (1971). The augmentation of experimental data to maximize det (X'X). *Technometrics* **13**, 682–688.

Edgar, T. F. and D. M. Himmelblau (1988). *Optimization of Chemical Processes*. McGraw-Hill. New York.

Edwards, I. M. and A. Jutan (1997). Optimization and control using response surface methods. *Computers and Chemical Engineering* **21**(4), 441–453.

Evans, J. W. (1979). Computer augmentation of experimental designs to maximize det (X'X). *Technometrics* **21**, 321–330.

Featherstone, A. P. (1997). Control Relevant Identification of Sheet and Film Processes. PhD thesis. University of Illinois at Urbana-Champaign.

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

Forbes, J. F. (1994). Model Structure and Adjustable Parameter Selection for Operations Optimization. PhD thesis. McMaster University. Hamilton, ON.

Forbes, J. F. and T. E. Marlin (1996). Design cost: A systematic approach to technology selection for model-based real-time optimization. *Computers and Chemical Engineering* **20**, 717–734.

Forbes, J. F., T. E. Marlin and J. F. MacGregor (1994). Model adequacy requirements for optimizing plant operations. *Computers and Chemical Engineering* **18**(6), 497–510.

Fraleigh, L. M. (1999). Optimal sensor selection and parameter estimation for real-time optimization. Master's thesis. University of Alberta. Edmonton, AB.

Galil, Z. and J. Kiefer (1977). Comparison of box-draper and d-optimum designs for experiments with mixtures. *Technometrics* **19**, 441–444.

Ganesh, N. and L. T. Biegler (1987). A reduced hessian strategy for sensitivity analysis of optimal flowsheets. *AIChE Journal* **33**(2), 282–296.

Garcia, C. E. and M. Morari (1981). Optimal operation of integrated processing systems. *AIChE Journal* **27**(6), 960–968.

Genceli, H. and M. Nikolaou (1996). New approach to constrained predictive control with simultaneous model identification. *AIChE Journal* **42**(10), 2857–2868.

Golden, M. P. and B. E. Ydstie (1989). Adaptive extremum control using approximate process models. *AIChE Journal* **35**(7), 1157–1169.

Grossman, S. I. (1986). *Multivariable Calculus, Linear Algebra, and Differential Equations.* 2nd ed.. Harcourt Brace Jovanovich. New York.

Heiberger, R. M., D. K. Bhaumik and B. Holland (1992). Universally optimal data augmentation strategies for additive models. Technical report. Temple University. Dept. of Statistics.

Heiberger, R. M., D. K. Bhaumik and B. Holland (1993). Optimal data augmentation strategies for additive models. *Journal of the American Statistical Association* **88**(423), 926–938.

Hill, W. J., W. G. Hunter and D. W. Wichern (1968). A joint design criteria for the dual problem of model discrimination and parameter estimation. *Technometrics* **10**, 145–160.

Hunter, W. G. and A. M. Reiner (1965). Designs for discriminating between two rival models. *Technometrics* **7**, 307–323.

Incropera, F. P. and D. P. DeWitt (1996). *Fundamentals of Heat and Mass Transfer.* 4th ed.. Wiley. New York.

Kiefer, J. (1975). Optimal design: Variation in structure and performance under change of criterion. *Biometrika* **62**(2), 277–288.

Kiefer, J. and J. Wolfowitz (1959). Optimum designs in regression problems. *Annals of Mathematics and Statistics* **30**(2), 271–294.

King, J. C. (1980). *Separation Processes.* 2nd ed.. McGraw-Hill. New York.

Koninckx, J. (1988). Online Optimization of Chemical Plants Using Steady State Models. PhD thesis. University of Maryland.

Krishnan, S., G. W. Barton and J. D. Perkins (1992). Robust parameter estimation in on-line optimization-part i. methodology and simulated case study. *Computers and Chemical Engineering* **16**(6), 545–562.

Lin, X.G, M.O. Tade and R. B. Newell (1994). Structural approach to the synthesis of control-systems. *Chemical Engineering Research and Design* **72**, 26–37.

Ljung, L. (1987). *System Identification: Theory for the User.* Prentice-Hall. New Jersey.

Mah, R. S. H. and A. C. Tamhane (1982). Detection of gross errors in process data. *AIChE J.* **28**(5), 828–830.

Marlin, T. E. (1997). Introduction to real-time optimization. In: *Real-Time Operations Optimization of Continuous Processes.*

Marlin, T. E. and A. N. Hrymak (1996). Real-time operations optimization of continuous processes. In: *Chemical Process Control - V Conference.*

Mat (1997). *MATLAB The Language of Technical Computing.*

McFarlane, R. C. and D. W. Bacon (1989). Empirical strategies for open-loop on-line optimization. *The Canadian Journal of Chemical Engineering* **67**, 665–677.

Miletic, I. P. and T. E. Marlin (1998). On-line statistical results analysis in real-time operations optimization. *Industrial and Engineering Chemistry Research* **37**, 3670–3684.

Montgomery, D. C. and G. C. Runger (1994). *Applied Statistics and Probability for Engineers.* Wiley. New York.

Myers, R. H. and D. C. Montgomery (1995). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments.* Wiley. New York.

Pinto, J. C. (1998). On the costs of parameter uncertainties. effects of parameter uncertainties during optimization and design of experiments. *Chemical Engineering Science* **53**(11), 2029–2040.

Pinto, J. C., M. W. Lobao and J. L. Monteiro (1990). Sequential experimental design for parameter estimation: A different approach. *Chemical Engineering Science* **45**(4), 883–892.

Pinto, J. C., M. W. Lobao and J. L. Monteiro (1991). Sequential experimental design for parameter estimation: Analysis of relative deviations. *Chemical Engineering Science* **46**(12), 3129–3138.

Pukelsheim, F. and B. Torsney (1991). Optimal weights for experimental designs on linearly independent support points. *Annals of Statistics* **19**(3), 1614–1625.

Roberts, P. D. (1979). An algorithm for steady-state system optimization and parameter estimation. *International Journal of Systems Science* **10**(7), 719–734.

Rosenberg, J., R. S. H. Mah and Iordache C. (1987). Evaluation of schemes for detecting and identifying gross errors in process data. *Industrial and Engineering Chemistry Research* **26**, 555.

Shirt, R. W., T. J. Harris and D. W. Bacon (1994). Experimental design considerations for dynamic systems. *Industrial and Engineering Chemistry Research* **33**, 2656–2667.

Soderstrom, T. and P. G. Stoica (1989). *System Identification.* Prentice-Hall. New Jersey.

Steinberg, D. M. (1985). Model robust response surface designs: Scaling two-level factorials. *Biometrika* **72**(3), 513–526.

Tong, H. and C. M. Crowe (1995). Detection of gross errors in data reconciliation by principal component analysis. *AIChE J.* **41**(7), 1712–1722.

Wald, A. (1943). On the efficient design of statistical investigations. *Annals of Mathematical Statistics* **14**, 134–140.

Watts, D. G. (1994). Estimating parameters in nonlinear rate-equations. *Canadian Journal of Chemical Engineering* **72**(4), 701–710.

Welch, W. J. (1983). A mean squared error criterion for the design of experiments. *Biometrika* **70**(1), 205–213.

White, D. C. (1997). Online optimization: What, where and estimating ROI. *Hydrocarbon Processing* pp. 43–51.

Wierda, S. J. (1994). Multivariate statistical process control - recent results and directions for future research. *Statistica Neerlandica* **48**(2), 147–168.

Williams, T. J. and R. E. Otto (1960). A generalized chemical processing model for the investigation of computer control. *AIEE Transactions* **79**, 458–473.

Wong, W. K. (1994). Comparing robust properties of a, d, e and g-optimal designs. *Computational Statistics and Data Analysis* **18**, 441–448.

Zhang, Y. and J. F. Forbes (2000). Extended design cost: A performance criterion for real-time optimization systems. *Computers and Chemical Engineering* **24**, 1829–1841.

Zhang, Y. and J. F. Forbes (2001). Performance analysis of perturbation-based methods for real-time optimization. Technical report. University of Alberta.

Zhang, Y., D. Nadler and J. F. Forbes (2001). Results analysis for trust constrained real-time optimization. *Journal of Process Control* **11**(3), 329–341.

# Appendix A

# Linear Sensitivity Analysis

Linear sensitivity analysis is a tool that provides an approximation of the variability of the RTO results based on the local sensitivity of the parameter updater and the economic optimizer. These sensitivities depend on the definition of the optimization problems and the resulting Karush-Kuhn-Tucker conditions (Ganesh and Biegler, 1987).

Following the development of Ganesh and Biegler (1987), the sensitivity of the economic optimizer, $\frac{\partial \mathbf{x}_p^*}{\partial \boldsymbol{\beta}}$, may be found. Consider the optimization problem,

$$\min P(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta})$$

$$\begin{aligned} s.t. \quad \mathbf{h}(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}) &= 0 \\ \mathbf{g}(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}) &\leq 0 \end{aligned}$$

where $\mathbf{x}$ is the vector of decision variables and $\mathbf{u}$ is the vector of dependent variables. The Lagrangian of the optimization problem is written as,

$$\begin{aligned} L(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}, \mathbf{v}) &= P(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}) + \mathbf{v}^T \mathbf{h}(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}) + \boldsymbol{\mu}_A^T \mathbf{g}_A(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}) \\ \mathbf{h}(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}) &= 0 \\ \mathbf{g}_A(\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}) &= 0 \end{aligned} \tag{A.1}$$

where $\mathbf{g}_A$ is the set of active inequality constraints, and $\mathbf{v}$ and $\boldsymbol{\mu}$ are the vectors of Lagrangian multipliers. Let $\mathbf{x}_v = [\mathbf{x}, \mathbf{u}]$.

At the optimal solution, the Karush-Kuhn-Tucker conditions are satisfied ( Edgar and Himmelblau, 1988),

$$\nabla_{\mathbf{x}_v} L(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}) = \nabla_{\mathbf{x}_v} P(\mathbf{x}_v, \boldsymbol{\beta}) + \mathbf{v}^T \nabla_{\mathbf{x}_v} \mathbf{h}(\mathbf{x}_v, \boldsymbol{\beta}) = 0 \tag{A.2}$$

$$\mathbf{h}(\mathbf{x}_v, \boldsymbol{\beta}) = 0 \tag{A.3}$$

$$\mathbf{g}_A(\mathbf{x}_v, \boldsymbol{\beta}) = 0 \tag{A.4}$$

The sensitivity of the economic optimizer can be formed by evaluating the full differential of Equations (A.2) to (A.4).

$$d\left[\nabla_{\mathbf{x}_v} L\left(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}\right)\right] = \begin{pmatrix} \nabla^2_{\mathbf{x}_v\mathbf{x}_v} L\left(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}\right) d\mathbf{x}_v + \nabla^{2T}_{\boldsymbol{\beta}\mathbf{x}_v} L\left(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}\right) d\boldsymbol{\beta} \\ +\nabla^{2T}_{\mathbf{v}\mathbf{x}_v} L\left(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}\right) d\mathbf{v} \end{pmatrix} = 0 \quad \text{(A.5)}$$

$$d\left[\mathbf{h}\left(\mathbf{x}_v, \boldsymbol{\beta}\right)\right] = \nabla_{\mathbf{x}_v} \mathbf{h}\left(\mathbf{x}_v, \boldsymbol{\beta}\right) d\mathbf{x}_v + \nabla_{\boldsymbol{\beta}} \mathbf{h}\left(\mathbf{x}_v, \boldsymbol{\beta}\right) d\boldsymbol{\beta} = 0 \quad \text{(A.6)}$$

$$d\left[\mathbf{g}_A\left(\mathbf{x}_v, \boldsymbol{\beta}\right)\right] = \nabla_{\mathbf{x}_v} \mathbf{g}_A\left(\mathbf{x}_v, \boldsymbol{\beta}\right) d\mathbf{x}_v + \nabla_{\boldsymbol{\beta}} \mathbf{g}_A\left(\mathbf{x}_v, \boldsymbol{\beta}\right) d\boldsymbol{\beta} = 0 \quad \text{(A.7)}$$

Realizing $\nabla^{2T}_{\mathbf{v}\mathbf{x}_v} L\left(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}\right) = \nabla_{\mathbf{x}_v} \mathbf{h}\left(\mathbf{x}_v, \boldsymbol{\beta}\right) + \nabla_{\mathbf{x}_v} \mathbf{g}_A\left(\mathbf{x}_v, \boldsymbol{\beta}\right)$, and reducing Equations (A.5) to (A.7) yields the following,

$$\begin{bmatrix} \nabla_{\boldsymbol{\beta}} \mathbf{x}_v \cdot \nabla^2_{\mathbf{x}_v\mathbf{x}_v} L\left(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}\right) + \nabla^2_{\mathbf{x}_v\boldsymbol{\beta}} L\left(\mathbf{x}_v, \boldsymbol{\beta}, \mathbf{v}\right) \\ +\nabla^T_{\mathbf{x}_v} \mathbf{h}\left(\mathbf{x}_v, \boldsymbol{\beta}\right) \cdot \nabla_{\boldsymbol{\beta}} \mathbf{v} + \nabla^T_{\mathbf{x}_v} \mathbf{g}_A\left(\mathbf{x}_v, \boldsymbol{\beta}\right) \cdot \nabla_{\boldsymbol{\beta}} \boldsymbol{\mu}_A \end{bmatrix} = 0 \quad \text{(A.8)}$$

$$\nabla_{\mathbf{x}_v} \mathbf{h}\left(\mathbf{x}_v, \boldsymbol{\beta}\right) \cdot \nabla_{\boldsymbol{\beta}} \mathbf{x}_v + \nabla_{\boldsymbol{\beta}} \mathbf{h}\left(\mathbf{x}_v, \boldsymbol{\beta}\right) = 0 \quad \text{(A.9)}$$

$$\nabla_{\mathbf{x}_v} \mathbf{g}_A\left(\mathbf{x}_v, \boldsymbol{\beta}\right) \cdot \nabla_{\boldsymbol{\beta}} \mathbf{x}_v + \nabla_{\boldsymbol{\beta}} \mathbf{g}_A\left(\mathbf{x}_v, \boldsymbol{\beta}\right) = 0 \quad \text{(A.10)}$$

Combining Equations (A.8) to (A.10) yields the following system of linear equations,

$$\begin{bmatrix} \nabla_{\boldsymbol{\beta}\mathbf{x}} L \\ \nabla_{\boldsymbol{\beta}\mathbf{u}} L \\ \nabla_{\boldsymbol{\beta}} \mathbf{g}_A \\ \nabla_{\boldsymbol{\beta}} \mathbf{h} \end{bmatrix} = \begin{bmatrix} \nabla_{\mathbf{xx}} L & \nabla_{\mathbf{xu}} L & \nabla_{\mathbf{x}} \mathbf{g}_A^T & \nabla_{\mathbf{x}} \mathbf{h}^T \\ \nabla_{\mathbf{ux}} L & \nabla_{\mathbf{u,u}} L & \nabla_{\mathbf{u}} \mathbf{g}_A^T & \nabla_{\mathbf{u}} \mathbf{h}^T \\ \nabla_{\mathbf{x}} \mathbf{g}_A & \nabla_{\mathbf{u}} \mathbf{g}_A & \mathbf{0} & \mathbf{0} \\ \nabla_{\mathbf{x}} \mathbf{h} & \nabla_{\mathbf{u}} \mathbf{h} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \nabla_{\boldsymbol{\beta}} \mathbf{x} \\ \nabla_{\boldsymbol{\beta}} \mathbf{u} \\ \nabla_{\boldsymbol{\beta}} \boldsymbol{\mu}_A \\ \nabla_{\boldsymbol{\beta}} \mathbf{v} \end{bmatrix} \quad \text{(A.11)}$$

To calculate the sensitivity, $\frac{\partial \boldsymbol{\beta}}{\partial \mathbf{z}}$, consider the least squares regression parameter estimation problem. The objective function is,

$$\begin{aligned} \phi &= \mathbf{e}^T \mathbf{e} \\ &= \mathbf{f}^T\left(\boldsymbol{\beta}, \mathbf{z}^{(n)}\right) \mathbf{f}\left(\boldsymbol{\beta}, \mathbf{z}^{(n)}\right) \end{aligned} \quad \text{(A.12)}$$

where $\mathbf{f}\left(\boldsymbol{\beta}, \mathbf{z}\right)$ is the vector containing all the model equations at the $nth$ RTO interval, $\boldsymbol{\beta}$ is the vector of $p$ parameters to be evaluated and $\mathbf{z}$ is the vector of process measurements at the $nth$ RTO interval. Although the structure of the equations are the same over the different RTO intervals the operating conditions are different. At the optimum,

$$\nabla_{\boldsymbol{\beta}} \phi = \frac{\partial \phi}{\partial \boldsymbol{\beta}} = 2\mathbf{f}^T\left(\boldsymbol{\beta}, \mathbf{z}^{(n)}\right) \frac{\partial \mathbf{f}}{\partial \boldsymbol{\beta}} = 0 \quad \text{(A.13)}$$

Define,

$$\mathbf{F}\left(\boldsymbol{\beta}, \mathbf{z}^{(n)}\right) = \left[\mathbf{f}^T\left(\boldsymbol{\beta}, \mathbf{z}^{(n)}\right) \frac{\partial \mathbf{f}}{\partial \boldsymbol{\beta}}\right]^T = 0 \quad \text{(A.14)}$$

Using the Implicit Function Theorem (Grossman, 1986),

$$\frac{\partial \mathbf{F}}{\partial \boldsymbol{\beta}} \partial \boldsymbol{\beta} + \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}} \partial \mathbf{z} = 0 \quad \text{(A.15)}$$

Solving Equation (A.15) yields,

$$\frac{\partial \boldsymbol{\beta}}{\partial \mathbf{z}} = -\left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\beta}}\right]^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}} \tag{A.16}$$

The estimate of the covariance matrix of the optimization variables becomes,

$$\mathbf{Q}_{\mathbf{x}} = \frac{\partial \mathbf{x}_p^*}{\partial \boldsymbol{\beta}} \frac{\partial \boldsymbol{\beta}}{\partial \mathbf{z}} \mathbf{U} \left(\frac{\partial \mathbf{x}_p^*}{\partial \boldsymbol{\beta}} \frac{\partial \boldsymbol{\beta}}{\partial \mathbf{z}}\right)^T \tag{A.17}$$

Substituting values for $\frac{\partial \mathbf{x}_p^*}{\partial \boldsymbol{\beta}}$ and $\frac{\partial \boldsymbol{\beta}}{\partial \mathbf{z}}$ into Equation (A.17) yields,

$$\mathbf{Q}_{\mathbf{x}} = \frac{\partial \mathbf{x}_p^*}{\partial \boldsymbol{\beta}} \left[-\frac{\partial \mathbf{F}}{\partial \boldsymbol{\beta}}\right]^{-1} \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}} \mathbf{U} \left(\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}}\right)^T \left(-\left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\beta}}\right]^{-1}\right)^T \left(\frac{\partial \mathbf{x}_p^*}{\partial \boldsymbol{\beta}}\right)^T \tag{A.18}$$

Consider using $n$ historical steady-state points, the term $\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}} \mathbf{U} \left(\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}}\right)^T$ becomes,

$$\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}} \mathbf{U} \left(\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}}\right)^T = \left[\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(1)}} | \cdots | \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}}\right] \begin{bmatrix} \mathbf{U}^{(1)} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{U}^{(2)} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{U}^{(n)} \end{bmatrix} \left[\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(1)}} | \cdots | \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}}\right]^T \tag{A.19}$$

assuming that $\mathbf{U}$ remains constant over the different RTO intervals,

$$\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}} \mathbf{U} \left(\frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}}\right)^T = \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(1)}} \mathbf{U} \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(1)}}^T + \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(2)}} \mathbf{U} \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(2)}}^T + \cdots + \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}} \mathbf{U} \frac{\partial \mathbf{F}}{\partial \mathbf{z}^{(n)}}^T \tag{A.20}$$

For large problems, a numerical approximation may be more appropriate.

# Appendix B

# Multiple Response Nonlinear Experimental Design

The models used to present the results of estimating the uncertainty in the estimates of the adjustable parameters presented by both Bard (1974) and Draper and Hunter (1967b), were written explicitly with respect to the measured response variables. However, for the models involved in most RTO applications, the models will be written with the measured response variables implicitly imbedded in the model equations. This thesis will alter the results from the previous authors to compensate for this by considering the errors to be the response variables. Consider the model that consists of $i = 1, 2...r$ nonlinear equations and $u = 1, 2...m$ sets of observations,

$$\epsilon_{iu} = f_i\left(y_{iu}^{(v)}, \xi_{iu}^{(q)}, \theta_p\right) \approx 0 \tag{B.1}$$

where $\mathbf{y}$ is the vector of $v$ measured dependent variables, $\boldsymbol{\xi}$ is the vector of observations of the $q$ dependent variables, $\boldsymbol{\theta}$ is the vector of $p$ unknown variables, and $\epsilon_i \sim N\left(0, \sigma_i^2\right)$ is the random error associated with the model equation $i$. When the final experimental design optimization problem is defined the values of the future measured response variables will be restricted to values defined by the model equations as constraints. Therefore the manipulated variables will still be the only decision variables in the optimization problem.

If the vector $\boldsymbol{\epsilon}_u' = (\epsilon_{1u}, \epsilon_{2u}, \ldots, \epsilon_{ru})$ is defined to represent the vector of errors for each of the $r$ model equations at the $uth$ observation, the covariance matrix of $\boldsymbol{\epsilon}_u'$ is (Draper and Hunter, 1967a),

$$\{\sigma_{ij}\} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{r1} & \sigma_{r2} & \cdots & \sigma_{rr} \end{bmatrix} = \mathbf{A} \tag{B.2}$$

The values of $\boldsymbol{\epsilon}_u'$ at different observations are considered uncorrelated.

The purpose of experimental design, for this application, is to assume that a set of $N$ observations from the process are available and that the next design point, $\xi_{N+1}^{(q)}$, should be selected to generate the most information possible to improve parameter estimation, while satisfying all the constraints (*i.e.*, $u = 1, 2, \ldots, N+1$). To determine the level of uncertainty of the adjustable parameters only consider the first $N$ points (*i.e.*, $u = 1, 2, \ldots, N$).

The following equation is the sum of squares and sums of products of the deviations from zero of the appropriate model equations.

$$v_{ij} = \sum_{u=1}^{m} \{e_{iu}\}\{e_{ju}\} \tag{B.3}$$

Considering $N$ sets of observations, the likelihood function of the equation errors is (Draper and Hunter, 1966),

$$p\left(\epsilon|\boldsymbol{\theta},\sigma^{ij}\right) = (2\pi)^{-\frac{1}{2}Nr}\left|\mathbf{A}\right|^{\frac{1}{2}Nr}\exp\left\{-\frac{1}{2}\sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}v_{ij}\right\} \tag{B.4}$$

and assuming that prior information is available before $N$ runs are performed, the multinormal distribution for the parameters is (Draper and Hunter, 1967b),

$$(2\pi)^{-\frac{1}{2}p}\left|\mathbf{V}_o\right|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\right\}(\boldsymbol{\theta}-\boldsymbol{\theta}_o)'\mathbf{V}_o^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}_o) \tag{B.5}$$

where $\boldsymbol{\theta}_o = (\boldsymbol{\theta}_{1o}, \boldsymbol{\theta}_{2o}, \ldots, \boldsymbol{\theta}_{po})$ is a vector of preliminary estimates and $\mathbf{V}_o$ is a $pxp$ parameter covariance matrix, all specified.

Applying Bayes' theorem and combining the prior distribution, Equation (B.5), with the likelihood function, Equation (B.4), provides the posterior distribution for $\boldsymbol{\theta}$ after $N$ runs (Draper and Hunter, 1967a).

$$p_N\left(\boldsymbol{\theta}|\boldsymbol{\epsilon},\sigma^{ij},\mathbf{V}_o\right) = C_N(\xi)\exp\left\{-\frac{1}{2}\sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}v_{ij}\right\}\exp\left\{-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}_o)'\mathbf{V}_o^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}_o)\right\} \tag{B.6}$$

where $C_N(\xi)$ is the appropriate normalizing factor.

Bard (1974) and Draper and Hunter (1966, 1967a and b) made the approximation that for a region in the $\boldsymbol{\theta}$ space sufficiently close to the maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$, the following expansion is valid,

$$\epsilon_{iu} = f_i\left(y_{iu}^{(v)},\xi_{iu}^{(q)},\theta_p\right) = f_i\left(y_{iu}^{(v)},\xi_{iu}^{(q)},\widehat{\theta}_p\right) + \sum_{t=1}^{p}\left(\theta_t - \widehat{\theta}_t\right)x_{iu}^{(t)} \tag{B.7}$$

where

$$x_{iu}^{(t)} = \left[\frac{\partial f_i\left(y_{iu}^{(v)},\xi_{iu}^{(q)},\theta_p\right)}{\partial\theta_t}\right]_{\theta=\widehat{\theta}} \tag{B.8}$$

Define the following,

$$\epsilon_{iu} - f_i\left(y_{iu}^{(v)}, \xi_{iu}^{(q)}, \widehat{\theta}_p\right) = d_{iu} \tag{B.9}$$

and

$$\mathbf{X}_i = \begin{bmatrix} x_{i1}^{(1)} & x_{i1}^{(2)} & \cdots & x_{i1}^{(p)} \\ x_{i2}^{(1)} & x_{i2}^{(2)} & \cdots & x_{i2}^{(p)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{im}^{(1)} & x_{im}^{(2)} & \cdots & x_{im}^{(p)} \end{bmatrix} \tag{B.10}$$

Then write

$$\begin{aligned} \sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}v_{ij} &= \sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}\sum_{u=1}^{m}\left[\begin{array}{c} \left\{d_{iu} - \sum_{t=1}^{p}\left(\theta_t - \widehat{\theta}_t\right)x_{iu}^{(t)}\right\} \times \\ \left\{d_{ju} - \sum_{t=1}^{p}\left(\theta_t - \widehat{\theta}_t\right)x_{ju}^{(t)}\right\} \end{array}\right] \\ &= \sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}\sum_{u=1}^{m}d_{iu}d_{ju} + \sum_{i=1}^{r}\sum_{j=1}^{r}\left(\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}\right)'\left\{\sigma^{ij}\mathbf{X}_i'\mathbf{X}_j\right\}\left(\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}\right) \end{aligned} \tag{B.11}$$

The cross product terms vanish because the likelihood function is maximized by $\widehat{\boldsymbol{\theta}}$ if and only if for $q = 1, 2, \ldots, p$ (Draper and Hunter, 1966),

$$\sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}\sum_{u=1}^{m}\left\{d_{iu}x_{ju}^{(q)} + d_{ju}x_{iu}^{(q)}\right\} = 0 \tag{B.12}$$

If Equation (B.11) is inserted into Equation (B.6) the posterior distribution is approximately (Draper and Hunter, 1967 a),

$$p_N\left(\boldsymbol{\theta}|\boldsymbol{\epsilon}, \sigma^{ij}, \mathbf{V}_o\right) = C_N'(\xi)\exp\left\{\begin{array}{c} -\frac{1}{2}\sum_{i=1}^{r}\sum_{j=1}^{r}\left(\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}\right)'\left\{\sigma^{ij}\mathbf{X}_i'\mathbf{X}_j\right\}\left(\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}\right) \\ -\frac{1}{2}\left(\boldsymbol{\theta}-\boldsymbol{\theta}_o\right)'\mathbf{V}_o^{-1}\left(\boldsymbol{\theta}-\boldsymbol{\theta}_o\right) \end{array}\right\} \tag{B.13}$$

This result provides a posterior distribution that can be used to approximate the probability distribution of the parameter estimates. Bard (1974) came to a similar result but took a different approach. Equation (B.13) can be reduced further for implementation in an experimental design.

Draper and Hunter (1967b) developed the other experimental design based on the result of Equation (B.13). Expanding the analysis to include the observation at $N+1$, the posterior distribution is equal to,

$$p_{N+1}\left(\boldsymbol{\theta}|\boldsymbol{\epsilon}, \sigma^{ij}, \mathbf{V}_o\right) = C_{N+n}''(\xi)\exp\left\{\begin{array}{c} -\frac{1}{2}\sum_{i=1}^{r}\sum_{j=1}^{r}\left(\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}_{N+1}\right)'\left\{\sigma^{ij}\mathbf{X}_i'\mathbf{X}_j\right\}\left(\boldsymbol{\theta}-\widehat{\boldsymbol{\theta}}_{N+1}\right) \\ -\frac{1}{2}\left(\boldsymbol{\theta}-\boldsymbol{\theta}_o\right)'\mathbf{V}_o^{-1}\left(\boldsymbol{\theta}-\boldsymbol{\theta}_o\right) \end{array}\right\} \tag{B.14}$$

where, because of the combination of two multinormal densities, $\left(C''_{N+n}(\xi)\right)^2$ is proportional to (Draper and Hunter, 1967 a),

$$\left|\sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}\mathbf{X}_i'\mathbf{X}_j + \mathbf{V}_o^{-1}\right| \tag{B.15}$$

The goal of the experimental design is to maximize Equation (B.14) with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\xi}_{N+1}$. However, the maximum with respect to $\boldsymbol{\theta}$ occurs at a value which depends on both $\widehat{\boldsymbol{\theta}}_{N+1}$ and $\boldsymbol{\theta}_o$ and cannot be determined before the $N+1th$ observation is available. Inconsequential of the value for $\boldsymbol{\theta}$, the exponent now vanishes. For D-optimal criterion, the optimization problem now reduces to,

$$\max\det\left|\sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}\mathbf{X}_i'\mathbf{X}_j + \mathbf{V}_o^{-1}\right| \tag{B.16}$$

$$\begin{aligned} s.t. \quad f_i\left(y_{i,N+1}^{(v)}, \xi_{i,N+1}^{(q)}, \widehat{\theta}_p\right) &= 0 \\ \left|\xi_N^{(q)} - \xi_{N+1}^{(q)}\right| &\leq a_q \end{aligned}$$

Bard (1974) reached the same result for experimental design of parameter estimates, but took a different approach. Pinto (1990) uses this development as the basis for his own experimental design but uses an A-optimal criterion.

$$\max \quad trace\left[\sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}\mathbf{X}_i'\mathbf{X}_j + \mathbf{V}_o^{-1}\right] \tag{B.17}$$

$$\begin{aligned} s.t. \quad f_i\left(y_{i,N+1}^{(v)}, \xi_{i,N+1}^{(q)}, \widehat{\theta}_p\right) &= 0 \\ \left|\xi_N^{(q)} - \xi_{N+1}^{(q)}\right| &\leq a_q \end{aligned}$$

This result is expanded to develop an experimental design optimization problem that focuses on reducing the uncertainty in the optimization variables by augmenting the designs specified by Problems (B.16) and (B.17) with the sensitivity of the optimization variables to the adjustable parameter estimates.

$$\max \quad trace\left[\left(\frac{dx}{d\beta}\right)\left(\sum_{i=1}^{r}\sum_{j=1}^{r}\sigma^{ij}\mathbf{X}_i'\mathbf{X}_j + \mathbf{V}_o^{-1}\right)\left(\frac{dx}{d\beta}\right)^T\right] \tag{B.18}$$

$$\begin{aligned} s.t. \quad f_i\left(y_{i,N+1}^{(v)}, \xi_{i,N+1}^{(q)}, \widehat{\theta}_p\right) &= 0 \\ \left|\xi_N^{(q)} - \xi_{N+1}^{(q)}\right| &\leq a_q \end{aligned}$$

By applying the optimization problems specified by Equations (B.16), (B.17) or (B.18) a difficulty that is realized is that the objective functions use values at the $N+1$ observation. Some of the observations will not be known at the $Nth$ stage. The values for $\widehat{\boldsymbol{\theta}}$ can be obtained from the nonlinear regression problem after the $Nth$ stage, and the values of the measured response variables $\left(y_{N+1}^{(v)}\right)$ will be controlled by the equality constraints defined by the process model. The solution of the optimization problem will provide a design point, contained within the trust region, that will produce high quality information to reduce decision variable uncertainty.

# Appendix C

# Case Study Data

## C.1 Williams-Otto Reactor Case Study (Section 2.4)

The nominal operating point in the case study for Section 2.4 is given in Table C.1.

Table C.1: Williams-Otto Reactor Nominal Operating Point

| Variable | Operating Point |
|---|---|
| FA | 0.4000 |
| $T_R$ | 0.6400 |
| AR | 0.1020 |
| BR | 0.4395 |
| CR | 0.0202 |
| ER | 0.2606 |
| PR | 0.1066 |
| GR | 0.0711 |

The measurement covariance matrix was found by generating a set of 1,000 process data points around the nominal value, shown in Table C.1, subject to a white noise distribution on the true plant values. Table 2.4 shows the applied standard deviations of the measurements, expressed as a percentage of the nominal case. The flow of reactant B and reactor temperature are considered to be in dependent of the compositions and have the following variances,

$$\begin{aligned} \sigma_{FB}^2 &= 0.1666 \times 10^{-4} \\ \sigma_{T_R}^2 &= 0.0956 \times 10^{-4} \end{aligned}$$

The covariance matrix for the compositions of the materials in the mixture exiting the reactor are shown below, with the columns and rows corresponding to the variables in the following order,

$$\left[ \; AR \quad BR \quad CR \quad ER \quad PR \quad GR \; \right]$$

$$
\mathbf{U} = 10^{-4} \times \begin{bmatrix} .0984 & .0020 & .0011 & -.0046 & -.0061 & -.0012 \\ .0020 & 1.6959 & -.0027 & -.0565 & -.0011 & .0054 \\ .0011 & -.0027 & .0036 & -.0016 & .0007 & .0003 \\ -.0046 & -.0565 & -.0016 & .5983 & -.0075 & .0015 \\ -.0061 & -.0011 & .0007 & -.0075 & .1079 & .0056 \\ -.0012 & .0054 & .0003 & .0015 & .0056 & .0491 \end{bmatrix}
$$

The initial measured data set used in the case study is shown in Table C.2.

Table C.2: Initial Measured Data Set

| Point | FA | $T_R$ | AR | BR | CR | ER | PR | GR |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.4141 | 0.6450 | 0.0914 | 0.4488 | 0.0174 | 0.2620 | 0.0997 | 0.0776 |
| 2 | 0.4138 | 0.6448 | 0.0938 | 0.4286 | 0.0180 | 0.2763 | 0.1045 | 0.0795 |
| 3 | 0.4056 | 0.6475 | 0.0935 | 0.4033 | 0.0183 | 0.2702 | 0.1029 | 0.0755 |
| 4 | 0.4111 | 0.6385 | 0.0985 | 0.4478 | 0.0194 | 0.2769 | 0.1042 | 0.0749 |
| 5 | 0.4175 | 0.6438 | 0.0924 | 0.4426 | 0.0174 | 0.2658 | 0.1122 | 0.0780 |
| 6 | 0.4005 | 0.6421 | 0.1006 | 0.4282 | 0.0202 | 0.2565 | 0.1085 | 0.0720 |
| 7 | 0.4249 | 0.6433 | 0.1009 | 0.4467 | 0.0183 | 0.2624 | 0.1076 | 0.0675 |
| 8 | 0.4108 | 0.6421 | 0.1028 | 0.4478 | 0.0176 | 0.2526 | 0.1078 | 0.0748 |
| 9 | 0.4025 | 0.6419 | 0.0965 | 0.4227 | 0.0189 | 0.2516 | 0.1091 | 0.0713 |
| 10 | 0.3997 | 0.6456 | 0.0967 | 0.4493 | 0.0186 | 0.2829 | 0.1019 | 0.0804 |

### C.1.1 Selecting Optimal Number of Historical Points

Table C.3 presents the parameter estimates and optimization results for tests #1 through #10.

Table C.3: Optimization Results for Selecting the Optimal Number of Historical Points

| Historical Data Points | Adjustable Parameter Estimates | | | Optimal Point Calculation | |
|---|---|---|---|---|---|
| | $A_1$ | $A_2$ | $A_3$ | FB | $T_R$ |
| 1 | 8.8603 | 13.3545 | 19.1436 | 0.3762 | 0.6554 |
| 2 | 9.6412 | 14.3841 | 19.4688 | 0.3824 | 0.6548 |
| 3 | 9.4510 | 14.7401 | 19.9013 | 0.3819 | 0.6551 |
| 4 | 9.4199 | 14.8713 | 19.5095 | 0.3841 | 0.6560 |
| 5 | 9.5157 | 14.7680 | 19.1659 | 0.3851 | 0.6561 |
| 6 | 9.6346 | 14.9456 | 19.3312 | 0.3853 | 0.6558 |
| 7 | 9.7821 | 15.1795 | 19.8104 | 0.3844 | 0.6550 |
| 8 | 9.8612 | 15.1337 | 19.4999 | 0.3855 | 0.6552 |
| 9 | 9.9219 | 15.1935 | 19.6767 | 0.3851 | 0.6549 |
| 10 | 9.9092 | 15.1288 | 19.9091 | 0.3839 | 0.6545 |

For tests #1 through #10 the following matrices were found,

**Test #1:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.9373 & 1.4328 & 2.4333 \\ 1.4328 & 2.6816 & 4.5483 \\ 2.4333 & 4.5483 & 9.4568 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0016 & -0.0047 & 0.0020 \\ -0.0022 & 0.0005 & -0.0008 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.7997 & -0.1057 \\ -0.1057 & 1.2437 \end{bmatrix}$$

**Test #2:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.5921 & 0.8992 & 1.3950 \\ 0.8992 & 1.6638 & 2.5599 \\ 1.3950 & 2.5599 & 4.9673 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0010 & -0.0046 & 0.0023 \\ -0.0021 & 0.0005 & -0.0008 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.6078 & -0.0193 \\ -0.0193 & 0.6459 \end{bmatrix}$$

**Test #3:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.3723 & 0.5933 & 0.9290 \\ 0.5933 & 1.1707 & 1.8190 \\ 0.9290 & 1.8190 & 3.5809 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0011 & -0.0045 & 0.0022 \\ -0.0021 & 0.0005 & -0.0007 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.4175 & -0.0287 \\ -0.0287 & 0.4326 \end{bmatrix}$$

**Test #4:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.2747 & 0.4430 & 0.6806 \\ 0.4430 & 0.8815 & 1.3412 \\ 0.6806 & 1.3412 & 2.5697 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0009 & -0.0046 & 0.0024 \\ -0.0022 & 0.0005 & -0.0007 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.3306 & -0.0214 \\ -0.0214 & 0.3265 \end{bmatrix}$$

**Test #5:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.2239 & 0.3540 & 0.5395 \\ 0.3540 & 0.6896 & 1.0432 \\ 0.5395 & 1.0432 & 1.9917 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0008 & -0.0047 & 0.0025 \\ -0.0021 & 0.0005 & -0.0008 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.2744 & -0.0187 \\ -0.0187 & 0.2624 \end{bmatrix}$$

**Test #6:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1915 & 0.3020 & 0.4583 \\ 0.3020 & 0.5876 & 0.8855 \\ 0.4583 & 0.8855 & 1.6776 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0008 & -0.0046 & 0.0025 \\ -0.0021 & 0.0005 & -0.0007 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.2274 & -0.0143 \\ -0.0143 & 0.2174 \end{bmatrix}$$

**Test #7:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1710 & 0.2741 & 0.4143 \\ 0.2700 & 0.5239 & 0.7995 \\ 0.4174 & 0.8047 & 1.5549 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0008 & -0.0045 & 0.0024 \\ -0.0021 & 0.0004 & -0.0007 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.1932 & -0.0175 \\ -0.0175 & 0.1894 \end{bmatrix}$$

**Test #8:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1514 & 0.2455 & 0.3766 \\ 0.2360 & 0.4534 & 0.7161 \\ 0.3594 & 0.6892 & 1.3421 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0008 & -0.0046 & 0.0024 \\ -0.0021 & 0.0004 & -0.0007 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.1893 & -0.0189 \\ -0.0189 & 0.1666 \end{bmatrix}$$

**Test #9:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1359 & 0.2112 & 0.3232 \\ 0.2112 & 0.4045 & 0.6179 \\ 0.3232 & 0.6179 & 1.2028 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0008 & -0.0045 & 0.0023 \\ -0.0020 & 0.0004 & -0.0007 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.1613 & -0.0158 \\ -0.0158 & 0.1468 \end{bmatrix}$$

**Test #10:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1215 & 0.1761 & 0.2732 \\ 0.1883 & 0.3600 & 0.5338 \\ 0.2927 & 0.5575 & 1.1080 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0009 & -0.0045 & 0.0023 \\ -0.0020 & 0.0004 & -0.0007 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.1422 & -0.0156 \\ -0.0156 & 0.1315 \end{bmatrix}$$

### C.1.2 Benefits of Experimental Design

For the second case study of Section 2.4, the results for the optimization problems of the first RTO cycle, using the initial data set (see Table C.2) are,

Table C.4: First RTO Cycle

| Optimization Scheme | Optimal Point Calculation | |
|---|---|---|
| | FA | TR |
| Economic | 0.3819 | 0.6551 |
| Experimental Design | 0.4242 | 0.6482 |

These points were implemented and provided the following measurements,

Table C.5: Implemented Optimization Result

| | FA | TR | AR | BR | CR | ER | PR | GR |
|---|---|---|---|---|---|---|---|---|
| Econ. | 0.3812 | 0.6563 | 0.0787 | 0.3918 | 0.0143 | 0.2781 | 0.1080 | 0.1109 |
| DOE | 0.4422 | 0.6497 | 0.0807 | 0.4588 | 0.0155 | 0.2495 | 0.1082 | 0.0796 |

Table C.6: Economic Optimization Results - Second RTO Cycle

| **First Optimization Routine** | **Adjustable Parameter Estimates** | | | **Optimal Point Calculation** | |
|---|---|---|---|---|---|
| | $A_1$ | $A_2$ | $A_3$ | FA | TR |
| Economic | 9.8181 | 14.2520 | 19.3681 | 0.3809 | 0.6530 |
| Experimental Design | 9.7146 | 14.2424 | 19.3931 | 0.3822 | 0.6546 |

Table C.6 presents the economic optimization results for the second RTO cycle with the augmented data sets,

For the final economic optimization problem the following matrices were found.

**Test #1**: Data set augmented with an economic optimization result.

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.4022 & 0.5862 & 0.9028 \\ 0.5862 & 1.0707 & 1.6779 \\ 0.9028 & 1.6779 & 3.6172 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0009 & -0.0046 & 0.0022 \\ -0.0020 & 0.0005 & -0.0008 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.5296 & -0.0659 \\ -0.0659 & 0.4354 \end{bmatrix}$$

**Test #2**: Data set augmented with an experimental design optimization result.

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.3959 & 0.1761 & 0.2732 \\ 0.5819 & 1.0693 & 0.5338 \\ 0.9184 & 1.6835 & 3.3109 \end{bmatrix}$$

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = \begin{bmatrix} 0.0014 & -0.0048 & 0.0023 \\ -0.0019 & 0.0004 & -0.0008 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-5} \times \begin{bmatrix} 0.3822 & -0.0299 \\ -0.0299 & 0.4220 \end{bmatrix}$$

## C.2 Williams-Otto Reactor Case Study (Section 3.3)

Two case studies were performed in Section 3.3. The first examined if the RTO system can be improved by adding an information generation component to reduce uncertainty in the decision variables. The second case study investigated if the improvement can also be applied to a situation where plant/model mismatch is present

The design cost criterion developed by Forbes and Marlin (1996), defined in Equation (3.11), is used to evaluate the performance of the RTO designs over 25 RTO intervals.

### C.2.1 Tuning Parameter Selection

The data presented in this section is used to determine if the RTO system incorporating information generation could improve the performance of the RTO system when structural plant/model mismatch is not present. This section will also determine which set of tuning parameters produces the best results.

#### Raw Data

The following summarizes the results of the ten runs performed for each of the test conditions as outlined in Table 3.2.

**Test #1**: $\alpha = 0.90, c_1 = 0.90, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 168.9739 | 6 | 138.7053 |
| 2 | 108.9626 | 7 | 144.7672 |
| 3 | 161.0817 | 8 | 122.9528 |
| 4 | 199.4699 | 9 | 122.1651 |
| 5 | 119.6551 | 10 | 97.4428 |

**Test #2**: $\alpha = 0.90, c_1 = 0.90, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 87.5640 | 6 | 111.9854 |
| 2 | 129.0496 | 7 | 79.9772 |
| 3 | 106.1350 | 8 | 89.8739 |
| 4 | 98.9029 | 9 | 86.5970 |
| 5 | 135.5707 | 10 | 125.7413 |

**Test #3**: $\alpha = 0.90, c_1 = 0.90, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 64.3635 | 6 | 42.1058 |
| 2 | 59.3382 | 7 | 48.1035 |
| 3 | 58.5313 | 8 | 99.7822 |
| 4 | 51.8252 | 9 | 100.3348 |
| 5 | 85.5607 | 10 | 78.1025 |

**Test #4**: $\alpha = 0.90, c_1 = 0.90, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 49.4408 | 6 | 104.9123 |
| 2 | 36.1893 | 7 | 79.2948 |
| 3 | 83.5946 | 8 | 65.5477 |
| 4 | 102.4943 | 9 | 109.8425 |
| 5 | 101.0841 | 10 | 64.4793 |

**Test #5**: $\alpha = 0.90, c_1 = 0.95, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 84.7164 | 6 | 76.5408 |
| 2 | 83.3318 | 7 | 125.9892 |
| 3 | 102.6479 | 8 | 129.0275 |
| 4 | 106.8161 | 9 | 131.3550 |
| 5 | 69.6106 | 10 | 84.9149 |

**Test #6**: $\alpha = 0.90, c_1 = 0.95, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 96.5524 | 6 | 73.3786 |
| 2 | 99.9522 | 7 | 106.5585 |
| 3 | 71.1664 | 8 | 91.8085 |
| 4 | 105.5905 | 9 | 97.6200 |
| 5 | 81.7921 | 10 | 90.8741 |

**Test #7**: $\alpha = 0.90, c_1 = 0.95, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 76.8567 | 6 | 140.9674 |
| 2 | 84.0154 | 7 | 83.3168 |
| 3 | 73.4299 | 8 | 109.6220 |
| 4 | 51.2654 | 9 | 101.7906 |
| 5 | 85.3139 | 10 | 54.6005 |

**Test #8**: $\alpha = 0.90, c_1 = 0.95, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 47.4791 | 6 | 51.2276 |
| 2 | 81.1773 | 7 | 45.0073 |
| 3 | 60.9576 | 8 | 69.3417 |
| 4 | 44.7973 | 9 | 56.7401 |
| 5 | 43.9572 | 10 | 51.0764 |

**Test #9**: $\alpha = 0.90, c_1 = 0.98, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 62.1334 | 6 | 74.5592 |
| 2 | 69.9219 | 7 | 97.8965 |
| 3 | 113.6927 | 8 | 82.9108 |
| 4 | 84.6803 | 9 | 93.9651 |
| 5 | 88.2086 | 10 | 117.9644 |

**Test #10**: $\alpha = 0.90, c_1 = 0.98, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 62.2818 | 6 | 77.9029 |
| 2 | 119.7443 | 7 | 110.6427 |
| 3 | 88.1447 | 8 | 69.3734 |
| 4 | 67.7189 | 9 | 108.7104 |
| 5 | 75.5310 | 10 | 46.5410 |

**Test #11**: $\alpha = 0.90, c_1 = 0.98, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 41.1222 | 6 | 50.3953 |
| 2 | 31.1025 | 7 | 73.5192 |
| 3 | 62.4170 | 8 | 58.2678 |
| 4 | 51.9452 | 9 | 82.5846 |
| 5 | 73.3096 | 10 | 49.7328 |

**Test #12**: $\alpha = 0.90, c_1 = 0.98, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 36.9032 | 6 | 54.8662 |
| 2 | 64.5065 | 7 | 38.6966 |
| 3 | 40.3181 | 8 | 74.4314 |
| 4 | 48.2629 | 9 | 44.0546 |
| 5 | 59.1141 | 10 | 61.7208 |

**Test #13**: $\alpha = 0.95, c_1 = 0.90, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 132.2776 | 6 | 76.1489 |
| 2 | 119.0740 | 7 | 103.2282 |
| 3 | 74.2833 | 8 | 135.4752 |
| 4 | 78.1025 | 9 | 90.9129 |
| 5 | 112.2124 | 10 | 30.4390 |

**Test #14**: $\alpha = 0.95, c_1 = 0.90, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 34.4847 | 6 | 73.5946 |
| 2 | 72.1879 | 7 | 126.4227 |
| 3 | 78.7290 | 8 | 100.1029 |
| 4 | 112.6247 | 9 | 87.5974 |
| 5 | 88.8239 | 10 | 127.9877 |

**Test #15**: $\alpha = 0.95, c_1 = 0.90, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 107.3330 | 6 | 99.4819 |
| 2 | 67.9797 | 7 | 68.3998 |
| 3 | 84.7118 | 8 | 64.6729 |
| 4 | 103.8840 | 9 | 64.0724 |
| 5 | 103.6230 | 10 | 49.6661 |

**Test #16**: $\alpha = 0.95, c_1 = 0.90, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 85.2506 | 6 | 80.6371 |
| 2 | 74.3260 | 7 | 42.0248 |
| 3 | 33.7690 | 8 | 40.4515 |
| 4 | 52.8478 | 9 | 58.2586 |
| 5 | 98.3108 | 10 | 45.9694 |

**Test #17**: $\alpha = 0.95, c_1 = 0.95, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 88.3308 | 6 | 79.2440 |
| 2 | 94.8196 | 7 | 91.8431 |
| 3 | 107.4474 | 8 | 80.0029 |
| 4 | 54.3981 | 9 | 75.9166 |
| 5 | 48.2388 | 10 | 117.0343 |

**Test #18**: $\alpha = 0.95, c_1 = 0.95, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 58.6462 | 6 | 37.8554 |
| 2 | 35.9932 | 7 | 50.2430 |
| 3 | 72.4180 | 8 | 107.0450 |
| 4 | 64.8992 | 9 | 96.1476 |
| 5 | 90.6438 | 10 | 92.1749 |

**Test #19**: $\alpha = 0.95, c_1 = 0.95, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 49.2729 | 6 | 70.9381 |
| 2 | 60.0744 | 7 | 82.7769 |
| 3 | 60.2586 | 8 | 55.3565 |
| 4 | 38.4515 | 9 | 66.4915 |
| 5 | 32.9305 | 10 | 64.7392 |

**Test #20**: $\alpha = 0.95, c_1 = 0.95, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 55.9796 | 6 | 30.9996 |
| 2 | 49.6339 | 7 | 29.9424 |
| 3 | 53.0745 | 8 | 29.6597 |
| 4 | 86.9884 | 9 | 38.0886 |
| 5 | 49.5277 | 10 | 59.1456 |

**Test #21**: $\alpha = 0.95, c_1 = 0.98, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 108.7272 | 6 | 76.1016 |
| 2 | 114.3510 | 7 | 83.0410 |
| 3 | 72.0505 | 8 | 97.8824 |
| 4 | 62.2935 | 9 | 69.2071 |
| 5 | 94.1982 | 10 | 59.7491 |

**Test #22**: $\alpha = 0.95, c_1 = 0.98, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 58.0631 | 6 | 56.9751 |
| 2 | 63.6741 | 7 | 90.7860 |
| 3 | 65.7920 | 8 | 81.5819 |
| 4 | 65.3398 | 9 | 63.6108 |
| 5 | 50.9128 | 10 | 76.4298 |

**Test #23**: $\alpha = 0.95, c_1 = 0.98, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 49.7182 | 6 | 43.5084 |
| 2 | 71.2392 | 7 | 71.3139 |
| 3 | 27.5545 | 8 | 65.2105 |
| 4 | 54.0717 | 9 | 61.5667 |
| 5 | 38.6724 | 10 | 79.3463 |

**Test #24**: $\alpha = 0.95, c_1 = 0.98, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 40.5651 | 6 | 47.1045 |
| 2 | 58.6305 | 7 | 48.9982 |
| 3 | 56.3008 | 8 | 59.9520 |
| 4 | 48.8264 | 9 | 46.4230 |
| 5 | 61.6517 | 10 | 55.5608 |

**Test #25**: $\alpha = 1.00, c_1 = 0.90, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 69.5440 | 6 | 52.0883 |
| 2 | 67.7424 | 7 | 131.5164 |
| 3 | 116.1176 | 8 | 57.7948 |
| 4 | 73.1151 | 9 | 31.0817 |
| 5 | 107.4577 | 10 | 83.5324 |

**Test #26**: $\alpha = 1.00, c_1 = 0.90, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 91.7534 | 6 | 101.6411 |
| 2 | 59.3221 | 7 | 60.7266 |
| 3 | 51.7479 | 8 | 68.5976 |
| 4 | 81.7442 | 9 | 66.2866 |
| 5 | 57.1239 | 10 | 56.7718 |

**Test #27**: $\alpha = 1.00, c_1 = 0.90, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 23.8771 | 6 | 26.6952 |
| 2 | 49.6608 | 7 | 52.7168 |
| 3 | 44.3460 | 8 | 42.1973 |
| 4 | 33.4714 | 9 | 35.6005 |
| 5 | 53.9971 | 10 | 35.6704 |

**Test #28**: $\alpha = 1.00, c_1 = 0.90, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 34.5701 | 6 | 31.2399 |
| 2 | 62.1698 | 7 | 59.0538 |
| 3 | 27.2904 | 8 | 52.0581 |
| 4 | 27.7631 | 9 | 29.2634 |
| 5 | 51.1363 | 10 | 62.2634 |

**Test #29**: $\alpha = 1.00, c_1 = 0.95, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 56.1095 | 6 | 72.0398 |
| 2 | 75.7058 | 7 | 94.8073 |
| 3 | 66.3432 | 8 | 76.2360 |
| 4 | 54.7580 | 9 | 63.4637 |
| 5 | 83.6947 | 10 | 73.5897 |

**Test #30**: $\alpha = 1.00, c_1 = 0.95, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 51.1804 | 6 | 38.0492 |
| 2 | 66.6875 | 7 | 55.9263 |
| 3 | 60.6956 | 8 | 79.8668 |
| 4 | 67.6280 | 9 | 85.4986 |
| 5 | 60.4899 | 10 | 59.6848 |

**Test #31**: $\alpha = 1.00, c_1 = 0.95, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 40.7527 | 6 | 70.6195 |
| 2 | 40.0294 | 7 | 45.2501 |
| 3 | 50.5438 | 8 | 25.9621 |
| 4 | 28.8752 | 9 | 31.3566 |
| 5 | 68.8668 | 10 | 51.0000 |

**Test #32**: $\alpha = 1.00, c_1 = 0.95, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 34.1643 | 6 | 38.2833 |
| 2 | 57.4373 | 7 | 56.1119 |
| 3 | 50.6741 | 8 | 42.9426 |
| 4 | 57.7549 | 9 | 16.8579 |
| 5 | 47.1314 | 10 | 36.9385 |

**Test #33**: $\alpha = 1.00, c_1 = 0.98, c_2 = 0.990$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 66.7232 | 6 | 102.3196 |
| 2 | 107.8724 | 7 | 88.6620 |
| 3 | 90.5158 | 8 | 82.7894 |
| 4 | 60.3525 | 9 | 69.0183 |
| 5 | 99.4410 | 10 | 60.5312 |

**Test #34**: $\alpha = 1.00, c_1 = 0.98, c_2 = 0.995$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 38.2605 | 6 | 33.6000 |
| 2 | 32.5347 | 7 | 70.6053 |
| 3 | 71.6893 | 8 | 33.3607 |
| 4 | 65.5382 | 9 | 65.3857 |
| 5 | 81.5723 | 10 | 43.1265 |

**Test #35**: $\alpha = 1.00, c_1 = 0.98, c_2 = 0.997$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 21.0812 | 6 | 38.3519 |
| 2 | 36.0147 | 7 | 40.5558 |
| 3 | 32.8397 | 8 | 49.5153 |
| 4 | 37.4645 | 9 | 49.2175 |
| 5 | 49.9815 | 10 | 47.9357 |

**Test #36**: $\alpha = 1.00, c_1 = 0.98, c_2 = 0.999$

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 39.5367 | 6 | 32.0875 |
| 2 | 36.3081 | 7 | 44.6890 |
| 3 | 34.8633 | 8 | 38.8428 |
| 4 | 24.7737 | 9 | 20.3499 |
| 5 | 21.6485 | 10 | 56.6957 |

The results for the RTO design that did not incorporate an information generation section in the command conditioning subsystem are:

**Base Case:**

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 59.5374 | 6 | 51.5857 |
| 2 | 86.1428 | 7 | 55.6214 |
| 3 | 54.7361 | 8 | 69.0702 |
| 4 | 64.9741 | 9 | 64.5537 |
| 5 | 55.1566 | 10 | 63.0957 |

### Hypothesis Tests

The rest of this section will illustrate the results of the calculations performed to determine if the RTO systems with DOE are significantly different from the base case RTO system. Please see *Appendix D* for a description of the statistical calculations. The following statistics were used,

$$
\begin{aligned}
F_{1-0.05,240,240} &= 0.8083 \\
F_{0.05,240,240} &= 1.2371 \\
F_{1-0.05,9,9} &= 0.3146 \\
F_{0.05,9,9} &= 3.1789 \\
t_{0.05,20} &= 1.725 \\
t_{0.05,18} &= 1.734 \\
t_{0.05,15} &= 1.753 \\
t_{0.05,14} &= 1.761 \\
t_{0.05,13} &= 1.771 \\
t_{0.05,12} &= 1.782 \\
t_{0.05,11} &= 1.796
\end{aligned}
$$

Table C.7 shows the results of the one-sided hypothesis tests, to determine whether the data used to find the sample mean for each test, may have the same variance as the data set produced by the base case. If the lower statistic is greater than 1, or if the upper statistic is less than 1, then it can be stated, with a significance level of 5%, that the variances are different.

Table C.7: Identifying Different Sample Variances Among RTO Systems

| Test # | Standard Deviation | Lower Stat. | Upper Stat. | Test # | Standard Deviation | Lower Stat. | Upper Stat. |
|---|---|---|---|---|---|---|---|
| 36 | 11.09 | 0.39 | | 26 | 16.67 | 0.87 | |
| 35 | 9.26 | | 0.98 | 18 | 25.19 | 1.99 | |
| 27 | 11.48 | 0.41 | | 29 | 12.22 | 0.47 | |
| 28 | 14.94 | 0.70 | | *25* | *31.07* | 3.03 | |
| 32 | 12.86 | 0.52 | | 4 | 25.34 | 2.01 | |
| 31 | 15.43 | 0.75 | | 15 | 20.94 | 1.38 | |
| 20 | 17.60 | 0.97 | | 10 | 23.71 | 1.76 | |
| 12 | 12.61 | 0.50 | | 33 | 17.73 | 0.99 | |
| *24* | *6.95* | | 0.55 | 17 | 21.34 | 1.43 | |
| 34 | 19.09 | 1.14 | | 21 | 19.19 | 1.15 | |
| 8 | 12.21 | 0.47 | | 7 | 26.40 | 2.19 | |
| 23 | 16.45 | 0.85 | | 9 | 17.96 | 1.01 | |
| 11 | 15.87 | 0.79 | | *14* | *28.26* | 2.50 | |
| 19 | 14.91 | 0.70 | | 6 | 12.46 | 0.49 | |
| 16 | 22.03 | 1.52 | | *13* | *32.03* | 3.22 | |
| **Base** | **10.01** | | | 5 | 23.01 | 1.66 | |
| 30 | 13.58 | 0.58 | | 2 | 19.80 | 1.23 | |
| 22 | 12.17 | 0.46 | | *1* | *30.94* | 3.00 | |
| 3 | 20.99 | 1.38 | | | | | |

For the tests where the variances were determined not to be equal, the following lists the estimates of the degrees of freedom found from solving Equation (D.10).

| v | Test # |
|---|---|
| 11 | 1, 13, 25 |
| 12 | 4, 7, 14, 18 |
| 13 | 5, 10, 16 |
| 14 | 2, 3, 15, 17 |
| 15 | 9, 21, 34 |
| 18 | 24 |
| 20 | 35 |

Table C.8 shows the results of the one-sided hypothesis test to determine which sets of tuning parameters produce a RTO system that may have a different sample mean from the base case. If the lower statistic is greater than 0, or if the upper statistic is less than 0, then it can be stated, with a 5% significance level, that the RTO system that incorporates information generation is different than the system that does not.

Table C.9 lists the sample variances and pooled variance of the profit around the expected optimum for the RTO systems described by the base case and test #36. It is assumed

Table C.8: Confidence Intervals to Identify Different RTO Systems

| Test # | Sample Mean | Lower Stat. | Upper Stat. | Test # | Sample Mean | Lower Stat. | Upper Stat. |
|---|---|---|---|---|---|---|---|
| 36 | 34.9795 | | -19.27 | 26 | 69.5715 | -3.54 | |
| 35 | 40.2960 | | -14.71 | 18 | 70.6066 | -7.12 | |
| 27 | 41.8233 | | -12.27 | 29 | 71.6748 | 0.56 | |
| 28 | 43.7139 | | -8.87 | 25 | 78.9990 | -1.99 | |
| 32 | 43.8296 | | -9.68 | 4 | 79.6880 | 1.89 | |
| 31 | 45.3256 | | -7.03 | 15 | 81.3832 | 6.01 | |
| 20 | 48.3040 | | -3.04 | 10 | 82.6591 | 5.80 | |
| 12 | 52.2874 | | -1.33 | 33 | 82.8225 | 9.21 | |
| 24 | 52.4013 | | -3.36 | 17 | 83.7276 | 8.15 | |
| 34 | 53.5673 | | 3.19 | 21 | 83.7602 | 9.32 | |
| 8 | 55.1762 | | 1.39 | 7 | 86.1179 | 7.76 | |
| 23 | 56.2202 | | 4.33 | 9 | 88.5933 | 14.75 | |
| 11 | 57.4396 | | 5.28 | 14 | 90.2553 | 10.91 | |
| 19 | 58.1290 | | 5.53 | 6 | 91.5293 | 20.32 | |
| 16 | 61.1846 | | 12.29 | 13 | 95.2154 | 13.72 | |
| **Base** | **62.4474** | | | 5 | 99.4950 | 22.99 | |
| 30 | 62.5707 | -9.13 | | 2 | 105.1397 | 30.34 | |
| 22 | 67.3165 | -3.77 | | 1 | 138.4176 | 57.50 | |
| 3 | 68.8048 | -6.60 | | | | | |

that the population variance among the ten runs for each RTO system are equal and can be pooled together according the following more general form of Equation (D.5),

$$s_p^2 = \frac{\sum_{i=1}^{N} (n_i - 1)\, s_i^2}{\left[\sum_{i=1}^{N} n_i\right] - N} \tag{C.1}$$

where $N$ represents the number of runs, and $n_i$ and $s_i^2$ are the number of RTO intervals and sample variance for run $i$.

Since $s_t^2 < s_b^2$ the one-sided hypothesis test defined in Problem (3.13), with a significance of 5%, was used to test the possible equality of the population variances. The computed value of the statistic is,

$$\frac{s_1^2}{s_2^2} F_{0.05,240,240} = 0.93$$

Since this value is less than 1, the null hypothesis is rejected and it can be stated that the variance is reduced with a DOE component added to the RTO system.

Table C.9: Variance Around Expected Optimum

| **Run** | **Variance** | |
|---|---|---|
| | **RTO with DOE** | **RTO without DOE** |
| | **($/s)** | **($/s)** |
| 1 | 4.37 | 5.16 |
| 2 | 2.68 | 11.39 |
| 3 | 3.79 | 11.28 |
| 4 | 2.52 | 10.27 |
| 5 | 3.67 | 2.09 |
| 6 | 3.04 | 7.26 |
| 7 | 17.39 | 4.20 |
| 8 | 7.36 | 5.45 |
| 9 | 0.87 | 5.71 |
| 10 | 6.39 | 6.67 |
| Pooled | 5.21 | 6.95 |

### C.2.2 Comparing RTO Systems When Plant/Model Mismatch is Present

Using the tuning parameters from Test #36,

$$\begin{aligned} \alpha &= 1.00 \\ c_1 &= 0.98 \\ c_2 &= 0.999 \end{aligned}$$

the performance of a RTO system which incorporates information generation was compared to one that does not for different levels of structural mismatch in the activation energy constants, which act as fixed parameters for this case study.

Similar to the first case study, the extended design cost criterion developed by Zhang and Forbes (2000), defined in Equation (3.11), is used to evaluate the performance of the RTO designs over 25 RTO intervals.

#### Raw Data

The following summarizes the results of the ten runs performed for each of the test conditions.

**1% Bias in the Fixed Parameters**

RTO with DOE:

| **Run #** | **Design Cost** | **Run #** | **Design Cost** |
|---|---|---|---|
| 1 | 40.5389 | 6 | 47.6782 |
| 2 | 30.0034 | 7 | 25.6798 |
| 3 | 31.6938 | 8 | 22.6565 |
| 4 | 42.5339 | 9 | 44.7728 |
| 5 | 31.6473 | 10 | 30.6044 |

RTO without DOE:

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 48.2840 | 6 | 40.2087 |
| 2 | 85.7633 | 7 | 49.1671 |
| 3 | 67.7771 | 8 | 53.9273 |
| 4 | 74.6062 | 9 | 49.6604 |
| 5 | 66.5513 | 10 | 70.4433 |

**3% Bias in the Fixed Parameters**

RTO with DOE:

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 89.8435 | 6 | 77.1870 |
| 2 | 76.7003 | 7 | 140.8297 |
| 3 | 99.5141 | 8 | 71.8534 |
| 4 | 88.6214 | 9 | 101.1337 |
| 5 | 86.0016 | 10 | 101.8449 |

RTO without DOE:

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 66.1463 | 6 | 100.8710 |
| 2 | 111.4954 | 7 | 109.4646 |
| 3 | 74.4643 | 8 | 38.9687 |
| 4 | 94.2750 | 9 | 85.7252 |
| 5 | 93.5043 | 10 | 79.4874 |

**5% Bias in the Fixed Parameters**

RTO with DOE:

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 162.6213 | 6 | 113.2262 |
| 2 | 135.7014 | 7 | 211.8640 |
| 3 | 184.7249 | 8 | 139.2375 |
| 4 | 148.5583 | 9 | 163.9797 |
| 5 | 151.8552 | 10 | 148.4595 |

RTO without DOE:

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 108.7392 | 6 | 100.2991 |
| 2 | 198.2947 | 7 | 107.0362 |
| 3 | 117.0234 | 8 | 180.2862 |
| 4 | 121.6399 | 9 | 137.3925 |
| 5 | 83.7631 | 10 | 153.5356 |

**10% Bias in the Fixed Parameters**

RTO with DOE:

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 456.7278 | 6 | 432.8457 |
| 2 | 419.7843 | 7 | 516.9571 |
| 3 | 396.5679 | 8 | 3999.6508 |
| 4 | 359.5052 | 9 | 402.6954 |
| 5 | 429.2966 | 10 | 445.2575 |

RTO without DOE:

| Run # | Design Cost | Run # | Design Cost |
|---|---|---|---|
| 1 | 300.2334 | 6 | 283.2105 |
| 2 | 415.9282 | 7 | 339.5513 |
| 3 | 336.8884 | 8 | 409.1073 |
| 4 | 376.0584 | 9 | 338.2261 |
| 5 | 307.8003 | 10 | 378.0159 |

### Hypothesis Tests

The rest of this section will illustrate the results of the calculations performed to determine if the RTO system using DOE still performs effectively in the presence of plant/model mismatch. Please see *Appendix D* for a description of the statistical calculations. The following statistics were used,

$$F_{1-0.05,240,240} = 0.8083$$
$$F_{0.05,240,240} = 1.2371$$
$$F_{1-0.05,9,9} = 0.3146$$
$$F_{0.05,9,9} = 3.1789$$
$$t_{0.05,18} = 1.734$$

Table C.10 shows the results of one-sided hypothesis tests, at a 5% significance level, to determine if incorporating DOE into the RTO system produces a significant economic benefit, if plant/model mismatch is present. The column labelled standard deviation refers to the variation in the data set used to evaluate the sample standard deviation of the extended design cost measures for a specific RTO system.

The results show that the sample means of the two RTO systems were found to be different for all tests, except the 3% bias case. However, the RTO system that incorporates information generation is only able to realize more profit for the test where there is a 1% bias in the fixed parameters.

Table C.11 lists the pooled variances of the profit around the expected optimum for each of the bias case studies, and the statistics used to determine their significance. One-sided hypothesis tests, at a significance level of 5% (see Problems (3.12)and (3.13)), were used to determine the significance of the variances. Similar to the previous set of case studies, it is assumed that the population variance among the ten runs for each RTO system are equal and can be pooled together according to Equation (D.5).

Table C.10: Effects of Bias on RTO Performance

| Bias (%) | DOE | Sample Mean | Lower Stat. | Upper Stat. | Standard Deviation | Lower Stat. | Upper Stat. |
|---|---|---|---|---|---|---|---|
| 1 | Yes | 34.7809 | | -16.67 | 8.49 | | 1.10 |
| | No | 60.6389 | | | 14.44 | | |
| 3 | Yes | 93.3772 | -10.47 | | 19.76 | | 2.54 |
| | No | 87.5840 | | | 22.11 | | |
| 5 | Yes | 156.0228 | 0.20 | | 27.30 | | 1.77 |
| | No | 130.8010 | | | 36.56 | | |
| 10 | Yes | 425.9288 | 43.41 | | 42.44 | | 2.80 |
| | No | 348.5020 | | | 45.24 | | |

Table C.11: Hypothesis Tests for Variance Results: Mismatch Case Study

| Bias (%) | DOE | Variance ($/s) | Lower Stat. | Upper Stat. |
|---|---|---|---|---|
| 1 | Yes | 4.472 | | 0.836 |
| | No | 6.620 | | |
| 3 | Yes | 8.872 | 1.039 | |
| | No | 6.903 | | |
| 5 | Yes | 14.654 | | 0.996 |
| | No | 18.205 | | |
| 10 | Yes | 78.036 | 3.192 | |
| | No | 19.763 | | |

Table C.11 shows the variances of the RTO systems are significantly different for all cases, with the DOE improving the RTO performance for the 1% and 5% bias tests.

# Appendix D

# Statistical Significance Tests

The statistical tests used to determine if the sample means of the different RTO systems are significant are cited from Montgomery and Runger (1994).

When presented with two sets of randomly sampled data, the first step is to determine if the sample variances may be equal. The statistic used in the two-sided hypothesis test, defined in Problems (3.12) and (3.13), has the following form,

$$\frac{s_1^2}{s_2^2} F_{\frac{\alpha}{2}, n_2-1, n_1-1} \leq \frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} F_{1-\frac{\alpha}{2}, n_2-1, n_1-1} \tag{D.1}$$

where $\frac{\sigma_1^2}{\sigma_2^2} = 1$ if $\sigma_1^2 = \sigma_2^2$, and $s_1^2$ and $s_2^2$ are the sample variances of the random samples of sizes $n_1$ and $n_2$, respectively, from the two independent normal populations with unknown variances $\sigma_1^2$ and $\sigma_2^2$. $F_{\frac{\alpha}{2}, n_2-1, n_1-1}$ and $F_{1-\frac{\alpha}{2}, n_2-1, n_1-1}$ are the upper and lower $\frac{\alpha}{2}$ percentage points of the $F$ distribution with $n_2 - 1$ and $n_1 - 1$ degrees of freedom. If either statistic includes the number 1 in its range, it cannot be claimed that the standard deviations of the two processes are different with a $\alpha$% significance level. Please note that,

$$F_{1-\frac{\alpha}{2}, n_2-1, n_1-1} = \frac{1}{F_{\frac{\alpha}{2}, n_2-1, n_1-1}} \tag{D.2}$$

For one-sided hypothesis tests,

For Problem (3.12):

$$\frac{\sigma_1^2}{\sigma_2^2} \leq \frac{s_1^2}{s_2^2} F_{1-\alpha, n_2-1, n_1-1} \tag{D.3}$$

For Problem (3.13):

$$\frac{s_1^2}{s_2^2} F_{\alpha, n_2-1, n_1-1} \leq \frac{\sigma_1^2}{\sigma_2^2} \tag{D.4}$$

If the statistic in Inequality (3.12) is greater than one, or less than one in Inequality (3.13), than the population variances of the two data sets are considered different with an $\alpha$% significance level. One-sided hypothesis tests are useful if the experimenter is only concerned with one side of the distribution.

If the sample variances are considered equal with an $\alpha\%$ level of confidence, than the difference in the sample means can be found using the following formulation. For small sample means (*i.e.*, $n_1$ and $n_2$ do not exceed 30), the assumption is made that the data populations are normally distributed, and the statistics used in the hypothesis tests can be based on the $t$ distribution (Montgomery and Runger, 1994).

Since the sample variances are estimates of the common variance, a pooled estimator for $\sigma^2$ can be determined,

$$s_p^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2} \tag{D.5}$$

The statistic for the two-sided hypothesis test has the form,

$$\bar{x}_1 - \bar{x}_2 - t_{\frac{\alpha}{2}, n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{\frac{\alpha}{2}, n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \tag{D.6}$$

where $\mu_1 - \mu_2 = 0$ if $\sigma_1^2 = \sigma_2^2$, and $\bar{x}_1$ and $\bar{x}_2$ are the means of the random samples from two independent normal populations with unknown but equal variances and means $\mu_1$ and $\mu_2$, respectively. If either statistic includes zero in its range, then it cannot be claimed that the means of the two processes are different with an $\alpha\%$ level of confidence.

For one-sided hypothesis tests,

For Problem (3.14):

$$\bar{x}_1 - \bar{x}_2 - t_{\alpha, n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \le \mu_1 - \mu_2 \tag{D.7}$$

For Problem (3.15):

$$\mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{\alpha, n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \tag{D.8}$$

If the statistic in Inequality (3.14) is less than zero, or greater than zero in Inequality (3.13), than the population means of the two data sets are considered different with an $\alpha\%$ level of confidence.

If the sample variances are not considered equal then the statistic has the form,

$$\bar{x}_1 - \bar{x}_2 - t_{\frac{\alpha}{2}, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{\frac{\alpha}{2}, v} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \tag{D.9}$$

where the degrees of freedom are found from the following,

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1+1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2+1}} - 2 \tag{D.10}$$

# Appendix E

# Case Study Data For Chapter 4

This case study presents the data from the Williams-Otto plant (Williams and Otto, 1960) case study in Chapter 4.

The nominal operating point for this case study is given in Table E.1.

Table E.1: Williams-Otto Plant Nominal Operating Point

| Variable | Operating Point |
|---|---|
| FA | 0.350 |
| TR | 0.630 |
| FL | 0.400 |

Table E.2 displays the variances for the independent variables, and is followed by the covariance matrices for the compositions in streams $FR$ and $FS$, respectively.

Table E.2: Independent Variable Variances

| Variable | Variance |
|---|---|
| FB | $.1269 \times 10^{-4}$ |
| TR | $.1040 \times 10^{-4}$ |
| FL | $.1593 \times 10^{-4}$ |
| YR | $.5692 \times 10^{-7}$ |
| ps | $.7601 \times 10^{-8}$ |

The covariance matrix for compositions in stream $FR$, corresponding to the following order,

$$[\, AR \;\; BR \;\; CR \;\; ER \;\; PR \;\; GR \,]$$

$$\mathbf{U}_{FR} = 10^{-4} \times \begin{bmatrix} .1547 & .0085 & -.0006 & .0084 & -.0014 & .0004 \\ .0085 & 1.6248 & -.0002 & -.0424 & .0085 & .0036 \\ -.0006 & -.0002 & .0073 & .0013 & .0013 & 0 \\ .0084 & -.0424 & .0013 & .7432 & .0038 & -.0004 \\ -.0014 & .0085 & .0013 & .0038 & .0592 & .0004 \\ .0004 & .0036 & 0 & -.0004 & .0004 & .0073 \end{bmatrix}$$

The covariance matrix for compositions in stream $FS$, corresponding to the following order,

$$[\ AS\ \ BS\ \ CS\ \ ES\ \ PS\ ]$$

$$\mathbf{U}_{FS} = 10^{-4} \times \begin{bmatrix} .1867 & .0302 & -.0002 & .0006 & .0010 \\ .0302 & 1.8462 & -.0018 & -.0035 & -.0052 \\ -.0002 & -.0018 & .0085 & -.0034 & .0001 \\ .0006 & -.0035 & -.0034 & .9189 & -.0039 \\ .0010 & -.0052 & .0001 & -.0039 & .0107 \end{bmatrix}$$

The initial measured data set used in the case study is shown in Table E.3.

Table E.3: Williams-Otto Plant Initial Measured Data Set

| Point | FA | TR | FL | AR | BR | CR | ER | PR |
|---|---|---|---|---|---|---|---|---|
| 1 | .3624 | .6335 | .4319 | .1243 | .4278 | .0275 | .3142 | .0827 |
| 2 | .3737 | .6266 | .4299 | .1257 | .4313 | .0279 | .2694 | .0813 |
| 3 | .3751 | .6291 | .3978 | .1303 | .4544 | .0291 | .2888 | .0834 |
| 4 | .3613 | .6328 | .4147 | .1317 | .4350 | .0281 | .3097 | .0771 |
| 5 | .3782 | .6281 | .4060 | .1202 | .4530 | .0273 | .2672 | .0849 |
| 6 | .3721 | .6349 | .4074 | .1190 | .4801 | .0259 | .3087 | .0818 |
| 7 | .3483 | .6332 | .4266 | .1291 | .4038 | .0286 | .3085 | .0783 |
| 8 | .3550 | .6357 | .4287 | .1318 | .4255 | .0274 | .3072 | .0824 |
| 9 | .3783 | .6389 | .4023 | .1269 | .4443 | .0259 | .2944 | .0826 |
| 10 | .3661 | .6301 | .4187 | .1174 | .4595 | .0277 | .2940 | .0840 |

| Point | GR | AS | BS | CS | ES | PS | YR | ps |
|---|---|---|---|---|---|---|---|---|
| 1 | .0288 | .1336 | .4674 | .0303 | .3317 | .0377 | .0131 | .0048 |
| 2 | .0276 | .1335 | .4890 | .0255 | .3171 | .0384 | .0198 | .0072 |
| 3 | .0272 | .1417 | .4753 | .0317 | .3239 | .0374 | .0138 | .0048 |
| 4 | .0273 | .1376 | .4876 | .0298 | .3286 | .0342 | .0127 | .0045 |
| 5 | .0273 | .1325 | .5177 | .0299 | .3142 | .0346 | .0163 | .0060 |
| 6 | .0296 | .1344 | .4503 | .0288 | .3164 | .0383 | .0198 | .0071 |
| 7 | .0288 | .1392 | .4747 | .0302 | .3499 | .0342 | .0114 | .0041 |
| 8 | .0303 | .1323 | .4780 | .0308 | .3210 | .0389 | .0140 | .0049 |
| 9 | .0290 | .1300 | .4785 | .0277 | .3088 | .0400 | .0223 | .0078 |
| 10 | .0293 | .1367 | .4586 | .0298 | .3163 | .0395 | .0179 | .0064 |

## E.1 Selecting Optimal Number of Historical Points

Table E.4 presents the parameter estimates and optimization results for tests #1 through #10.

For tests #1 through #10 the following matrices were found,

$$\frac{\partial \mathbf{x}}{\partial \boldsymbol{\beta}} = 10^{-2} \times \begin{bmatrix} -0.6399 & -0.9962 & 0.7509 & -0.0082 \\ -0.0415 & -0.1233 & -0.0064 & -0.0009 \\ 1.5103 & 1.0833 & -1.1542 & 0.0201 \end{bmatrix}$$

Table E.4: Optimization Results for Selecting the Optimal Number of Historical Points

| **Historical Data Points** | **Adjustable Parameter Estimates** | | | |
|---|---|---|---|---|
| | $A_1$ | $A_2$ | $A_3$ | $\alpha$ |
| 1 | 10.2992 | 14.4715 | 19.9861 | 2.8475 |
| 2 | 8.6029 | 13.3954 | 19.9807 | 2.8790 |
| 3 | 8.8597 | 13.4044 | 19.9480 | 2.8865 |
| 4 | 9.3040 | 13.9659 | 20.0010 | 2.8731 |
| 5 | 8.9696 | 13.4312 | 19.9710 | 2.8685 |
| 6 | 9.1380 | 13.5956 | 19.9772 | 2.8407 |
| 7 | 9.1975 | 13.8733 | 20.0066 | 2.8303 |
| 8 | 9.3030 | 13.8765 | 19.9900 | 2.8514 |
| 9 | 9.4870 | 14.3288 | 19.9998 | 2.8348 |
| 10 | 9.6158 | 14.4449 | 19.9985 | 2.8309 |

**Test #1:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 1.8462 & 2.5709 & 4.5548 & 0 \\ 2.5709 & 4.4456 & 7.1882 & 0 \\ 4.5548 & 7.1882 & 15.2178 & 0 \\ 0 & 0 & 0 & 0.0030 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.1895 & 0.0088 & -0.2783 \\ 0.0088 & 0.0111 & -0.0048 \\ -0.2783 & -0.0048 & 0.4263 \end{bmatrix}$$

**Test #2:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.6704 & 1.0422 & 2.2504 & 0 \\ 1.0422 & 1.9996 & 4.1504 & 0 \\ 2.2504 & 4.1504 & 10.9319 & 0 \\ 0 & 0 & 0 & 0.0012 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.1380 & -0.0083 & -0.2206 \\ -0.0083 & 0.0050 & 0.0177 \\ -0.2206 & 0.0177 & 0.3614 \end{bmatrix}$$

**Test #3:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.4634 & 0.6921 & 1.4524 & 0 \\ 0.6921 & 1.2698 & 2.5329 & 0 \\ 1.4524 & 2.5329 & 6.4168 & 0 \\ 0 & 0 & 0 & 0.0010 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0765 & -0.0037 & -0.1207 \\ -0.0037 & 0.0032 & 0.0086 \\ -0.1207 & 0.0086 & 0.1957 \end{bmatrix}$$

**Test #4:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.3874 & 0.5792 & 1.1237 & 0 \\ 0.5792 & 1.0608 & 1.9464 & 0 \\ 1.1237 & 1.9464 & 4.5261 & 0 \\ 0 & 0 & 0 & 0.0009 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0510 & -0.0010 & -0.0789 \\ -0.0010 & 0.0027 & 0.0039 \\ -0.0789 & 0.0039 & 0.1266 \end{bmatrix}$$

**Test #5:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.2858 & 0.4210 & 0.8664 & 0 \\ 0.4210 & 0.7641 & 1.4828 & 0 \\ 0.8664 & 1.4828 & 3.7196 & 0 \\ 0 & 0 & 0 & 0.0006 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0458 & -0.0017 & -0.0716 \\ -0.0017 & 0.0019 & 0.0043 \\ -0.0716 & 0.0043 & 0.1150 \end{bmatrix}$$

**Test #6:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.2505 & 0.3668 & 0.7414 & 0 \\ 0.3668 & 0.6622 & 1.2575 & 0 \\ 0.7414 & 1.2575 & 3.0943 & 0 \\ 0 & 0 & 0 & 0.0005 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0378 & -0.0011 & -0.0587 \\ -0.0011 & 0.0017 & 0.0032 \\ -0.0587 & 0.0032 & 0.0940 \end{bmatrix}$$

**Test #7:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.2148 & 0.3222 & 0.6335 & 0 \\ 0.3222 & 0.5948 & 1.1013 & 0 \\ 0.6335 & 1.1013 & 2.6399 & 0 \\ 0 & 0 & 0 & 0.0005 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0321 & -0.0007 & -0.0497 \\ -0.0007 & 0.0015 & 0.0025 \\ -0.0497 & 0.0025 & 0.0795 \end{bmatrix}$$

**Test #8:**

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1937 & 0.2847 & 0.5586 & 0 \\ 0.2847 & 0.5176 & 0.9500 & 0 \\ 0.5586 & 0.9500 & 2.2637 & 0 \\ 0 & 0 & 0 & 0.0005 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0274 & -0.0005 & -0.0422 \\ -0.0005 & 0.0013 & 0.0019 \\ -0.0422 & 0.0019 & 0.0672 \end{bmatrix}$$

**Test #9**:

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1819 & 0.2760 & 0.5146 & 0 \\ 0.2760 & 0.5133 & 0.8973 & 0 \\ 0.5146 & 0.8973 & 2.0277 & 0 \\ 0 & 0 & 0 & 0.0004 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0242 & 0.0001 & -0.0369 \\ 0.0001 & 0.0013 & 0.0012 \\ -0.0369 & 0.0012 & 0.0583 \end{bmatrix}$$

**Test #10**:

$$\mathbf{Q}_\beta = \begin{bmatrix} 0.1677 & 0.2527 & 0.4646 & 0 \\ 0.2527 & 0.4661 & 0.8026 & 0 \\ 0.4646 & 0.8026 & 1.7875 & 0 \\ 0 & 0 & 0 & 0.0004 \end{bmatrix}$$

$$\mathbf{Q}_x = 10^{-3} \times \begin{bmatrix} 0.0214 & 0.0002 & -0.0324 \\ 0.0002 & 0.0012 & 0.0008 \\ -0.0324 & 0.0008 & 0.0510 \end{bmatrix}$$

## E.2 Structural Mismatch Absent

This section displays the results for the case study comparing the performance of two RTO systems with no structural mismatch, implemented for the Williams-Otto plant.

### E.2.1 Raw Data

The optimization data concerning the case study for the Williams-Otto plant without structural mismatch, is provided in Tables E.5, E.6 and E.7. Table E.5 lists the extended design cost measures for both systems, and Tables E.6 and E.7 display the profit profiles of the ten runs for the RTO systems with and without DOE, respectively. Note that the italic terms in Table E.7 indicate the first term in that run, which was assumed to have reached the region around the expected optimum. This value was selected by viewing a plot of the profit profile and estimating the expected optimum.

### E.2.2 Hypothesis Tests

This section will discuss the results of the calculations performed to determine if the RTO system using DOE out-performs the original RTO system for the Williams-Otto plant.

Table E.5: Extended Design Costs for RTO - No Mismatch

| Run # | Extended Design Cost RTO without DOE (% Return) | RTO with DOE (% Return) |
|---|---|---|
| 1 | 105.0888 | 67.8204 |
| 2 | 88.6466 | 79.3848 |
| 3 | 84.8074 | 74.6597 |
| 4 | 62.9347 | 71.0308 |
| 5 | 56.3707 | 65.7256 |
| 6 | 94.6863 | 61.9127 |
| 7 | 90.6579 | 87.9312 |
| 8 | 83.2549 | 69.8291 |
| 9 | 99.482 | 73.0171 |
| 10 | 135.6754 | 76.8301 |

Table E.6: Profit Profiles for RTO With DOE - No Mismatch

| RTO | % Return for Run # | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Intervals | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 62.25 | 62.40 | 62.11 | 62.11 | 62.40 | 62.11 | 57.46 | 62.11 | 62.11 | 62.11 |
| 2 | 62.41 | 60.05 | 61.33 | 63.06 | 63.17 | 62.88 | 60.80 | 63.17 | 63.14 | 62.98 |
| 3 | 63.25 | 61.93 | 60.23 | 57.33 | 63.35 | 63.02 | 61.73 | 53.79 | 58.79 | 61.21 |
| 4 | 62.83 | 62.60 | 61.56 | 63.48 | 64.19 | 63.39 | 62.53 | 62.80 | 64.33 | 62.20 |
| 5 | 60.69 | 61.35 | 63.28 | 63.78 | 59.61 | 63.36 | 63.43 | 62.79 | 63.85 | 63.85 |
| 6 | 60.22 | 62.81 | 62.12 | 63.17 | 58.98 | 59.80 | 60.91 | 60.09 | 60.95 | 63.26 |
| 7 | 63.67 | 62.32 | 62.71 | 63.52 | 62.20 | 61.18 | 59.80 | 63.87 | 59.40 | 62.79 |
| 8 | 64.11 | 63.56 | 61.76 | 60.14 | 59.31 | 62.56 | 59.52 | 63.55 | 59.42 | 62.28 |
| 9 | 59.48 | 61.62 | 63.58 | 61.16 | 59.50 | 62.30 | 59.49 | 59.59 | 59.57 | 60.15 |
| 10 | 59.98 | 59.81 | 63.60 | 60.65 | 63.42 | 63.81 | 59.30 | 59.75 | 59.91 | 58.66 |
| 11 | 61.51 | 58.97 | 62.19 | 60.92 | 63.42 | 62.33 | 62.82 | 59.72 | 63.36 | 60.49 |
| 12 | 59.70 | 58.98 | 59.34 | 63.51 | 63.73 | 61.16 | 58.20 | 60.71 | 64.24 | 59.71 |
| 13 | 60.35 | 59.92 | 59.32 | 62.58 | 64.07 | 63.27 | 61.64 | 59.55 | 59.80 | 58.90 |
| 14 | 59.88 | 63.06 | 64.30 | 64.22 | 59.74 | 63.34 | 59.93 | 59.98 | 59.90 | 60.83 |
| 15 | 60.35 | 59.66 | 62.13 | 60.09 | 59.74 | 60.17 | 61.48 | 63.82 | 59.87 | 59.58 |
| 16 | 62.81 | 60.16 | 59.78 | 60.02 | 61.17 | 60.03 | 58.15 | 62.62 | 59.71 | 59.59 |
| 17 | 62.72 | 60.18 | 58.99 | 64.30 | 59.67 | 59.77 | 57.63 | 62.63 | 59.62 | 62.93 |
| 18 | 63.49 | 59.85 | 58.77 | 64.02 | 59.71 | 63.59 | 61.39 | 62.98 | 59.90 | 61.90 |
| 19 | 63.53 | 63.80 | 59.43 | 60.18 | 60.11 | 63.52 | 61.96 | 63.39 | 63.86 | 61.65 |
| 20 | 60.20 | 63.97 | 55.35 | 63.03 | 63.54 | 62.83 | 60.99 | 63.34 | 61.93 | 63.80 |
| 21 | 59.38 | 63.43 | 61.96 | 57.61 | 61.98 | 63.19 | 62.80 | 59.37 | 63.61 | 63.94 |
| 22 | 59.50 | 59.86 | 62.18 | 61.16 | 60.99 | 59.84 | 61.96 | 59.54 | 63.56 | 59.41 |
| 23 | 63.96 | 59.49 | 61.84 | 60.48 | 62.90 | 59.78 | 63.52 | 62.16 | 63.31 | 59.89 |
| 24 | 60.10 | 59.65 | 63.00 | 56.69 | 63.07 | 59.76 | 62.96 | 63.93 | 61.88 | 59.86 |
| 25 | 64.20 | 59.61 | 62.92 | 60.19 | 62.74 | 59.50 | 60.10 | 63.36 | 59.41 | 59.64 |

Please see *Appendix D* for a description of the statistical calculations. The following statis-

Table E.7: Profit Profiles for RTO Without DOE - No Mismatch

| **RTO** | **% Return for Run #** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Intvl.** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| 1 | 53.69 | 53.69 | 53.69 | 53.69 | 53.69 | 53.69 | 53.69 | 53.69 | 53.69 | 53.69 |
| 2 | 53.69 | *56.47* | 53.69 | *62.69* | *57.42* | 53.04 | *58.83* | *61.27* | *57.55* | 53.69 |
| 3 | 53.69 | 60.35 | *60.82* | 62.69 | 60.40 | 53.04 | 58.83 | 61.27 | 62.02 | 53.69 |
| 4 | *59.82* | 61.83 | 59.54 | 62.69 | 60.40 | 53.04 | 64.32 | 61.27 | 62.02 | 53.69 |
| 5 | 63.00 | 61.83 | 59.54 | 62.69 | 64.09 | 53.04 | 63.13 | 60.79 | 60.47 | *61.92* |
| 6 | 61.28 | 61.83 | 64.34 | 64.33 | 63.71 | *61.15* | 63.13 | 60.25 | 61.33 | 62.24 |
| 7 | 61.28 | 61.83 | 64.34 | 64.33 | 63.71 | 63.05 | 60.43 | 63.80 | 61.33 | 62.24 |
| 8 | 61.28 | 61.83 | 64.33 | 64.33 | 63.71 | 63.05 | 60.43 | 63.80 | 61.33 | 58.13 |
| 9 | 61.73 | 61.83 | 64.33 | 64.33 | 62.30 | 63.05 | 44.25 | 63.80 | 62.37 | 58.13 |
| 10 | 61.73 | 61.93 | 59.79 | 62.82 | 63.87 | 63.05 | 59.66 | 58.94 | 62.37 | 61.34 |
| 11 | 61.73 | 60.69 | 59.79 | 62.66 | 62.42 | 63.05 | 59.66 | 58.94 | 62.79 | 61.34 |
| 12 | 61.73 | 60.69 | 64.19 | 62.66 | 62.42 | 63.05 | 62.29 | 61.89 | 44.68 | 61.34 |
| 13 | 61.73 | 63.34 | 61.47 | 62.66 | 62.42 | 63.05 | 62.29 | 60.89 | 44.68 | 61.76 |
| 14 | 61.73 | 63.34 | 61.47 | 62.66 | 56.87 | 63.05 | 62.24 | 60.89 | 56.95 | 61.76 |
| 15 | 62.16 | 59.19 | 61.47 | 63.23 | 62.84 | 64.00 | 64.24 | 60.72 | 56.95 | 61.76 |
| 16 | 62.16 | 63.48 | 62.06 | 59.43 | 62.84 | 64.00 | 64.24 | 58.06 | 61.57 | 50.87 |
| 17 | 59.73 | 63.48 | 62.06 | 57.99 | 64.14 | 57.57 | 64.24 | 63.49 | 64.31 | 57.94 |
| 18 | 59.73 | 62.70 | 59.81 | 57.99 | 64.14 | 57.57 | 62.70 | 63.49 | 64.31 | 56.92 |
| 19 | 59.73 | 62.32 | 59.81 | 57.99 | 64.14 | 62.81 | 63.83 | 63.49 | 64.31 | 60.22 |
| 20 | 59.73 | 62.01 | 59.81 | 63.90 | 64.14 | 62.81 | 62.42 | 63.49 | 64.31 | 60.22 |
| 21 | 59.72 | 56.43 | 59.81 | 63.90 | 60.30 | 62.81 | 62.42 | 61.57 | 64.31 | 60.22 |
| 22 | 59.72 | 59.68 | 64.32 | 63.90 | 60.35 | 62.81 | 62.49 | 58.38 | 64.31 | 58.25 |
| 23 | 64.27 | 59.68 | 62.90 | 60.65 | 63.92 | 62.81 | 57.84 | 58.38 | 63.67 | 60.46 |
| 24 | 58.84 | 58.02 | 62.90 | 60.65 | 63.92 | 63.26 | 60.71 | 61.32 | 63.67 | 60.46 |
| 25 | 59.44 | 61.28 | 62.90 | 60.65 | 63.92 | 61.88 | 57.46 | 61.32 | 63.67 | 60.46 |

tics were used,

$$
\begin{aligned}
F_{1-0.05,240,220} &= 0.8040 \\
F_{0.05,240,220} &= 1.2438 \\
F_{1-0.05,9,9} &= 0.3146 \\
F_{0.05,9,9} &= 3.1789 \\
t_{0.05,11} &= 1.7959
\end{aligned}
$$

Table E.8 lists the sample variances of the % Return around the expected optimum for all the runs performed, and the pooled variance of the set of ten runs for each RTO system. Similar to the case study in Chapter 3, it is assumed that the population variance among the ten runs for each RTO system are equal and can be pooled together according to Equation (D.5).

Since $s_t^2 < s_b^2$ the one-sided hypothesis test defined in Problem (3.13), with a significance of 5%, was used to test the possible equality of the population variances. The computed

Table E.8: Variance Around Expected Optimum

| Run | Variance | |
|---|---|---|
| | **RTO with DOE** (% Return) | **RTO without DOE** (% Return) |
| 1 | 2.97 | 1.86 |
| 2 | 2.79 | 3.25 |
| 3 | 4.05 | 3.28 |
| 4 | 4.81 | 4.64 |
| 5 | 3.20 | 3.34 |
| 6 | 2.44 | 10.36 |
| 7 | 3.24 | 18.53 |
| 8 | 5.39 | 3.77 |
| 9 | 3.73 | 31.39 |
| 10 | 2.83 | 8.24 |
| Pooled | 3.55 | 8.94 |

value of the statistic is,

$$\frac{s_1^2}{s_2^2} F_{0.05,240,220} = 0.49$$

Since this value is less than 1, the null hypothesis is rejected and it can be stated that the variance is reduced with a DOE component added to the RTO system.

The next set of tests evaluates whether the RTO system with DOE, produces a significantly lower extended design cost measure. It must first be determined if the variance of the two data sets may be considered equal to determine the statistic that should be used to compare the means. Table E.5 lists the extended design costs (EDC) over ten runs for each RTO system. Table E.9 presents the results of the one-sided hypothesis test with 5% significance.

Table E.9: Sample Standard Deviation of EDC

| | **Standard Deviation of EDC** | **Upper Stat.** |
|---|---|---|
| RTO with DOE | 7.43 | 0.3623 |
| RTO without DOE | 22.02 | |

Since the upper statistic is less than one, it can be stated using a one-sided hypothesis test of 5% significance, that the variances of the two EDC data sets are different. Therefore the statistic defined in Inequality (D.9) must be used to evaluate the means of the data sets.

With the variances of the two data sets considered not equal, Equation (D.10) is used to solve for the degrees of freedom to be used in the hypothesis test,

$$v = 11.48$$

Since $\bar{x}_t < \bar{x}_b$ the one-sided hypothesis test defined in Problem (3.15), with a significance of 5%, was used to test the possible equality of the population means. The computed value of the statistic is,

$$\bar{x}_1 - \bar{x}_2 + t_{\alpha,v}\sqrt{\frac{s_t^2}{n_1} + \frac{s_b^2}{n_2}} = -4.1$$

Since this value is less than 0, the null hypothesis is rejected and it can be stated that the extended design cost is reduced with a DOE component added to the RTO system.

## E.3 Parametric Mismatch

This section displays the results for the case study comparing the performance of two RTO systems when parametric mismatch exists in the activation energy constants, implemented for the Williams-Otto plant.

### E.3.1 Raw Data

The optimization data concerning the case study for the Williams-Otto plant with parametric mismatch, is provided in Tables E.10, E.11 and E.12. Table E.10 lists the extended design cost measures for both systems, and Tables E.11 and E.12 display the profit profiles of the ten runs for the RTO systems with and without DOE, respectively. Please note that it is assumed that the RTO system has approached the expected value of the optimum by the first interval, in all the recorded runs.

Table E.10: Extended Design Costs for RTO: Mismatch Case Study

| Run # | Extended Design Cost | |
|---|---|---|
| | RTO without DOE (% Return) | RTO with DOE (% Return) |
| 1 | 167.2492 | 117.5420 |
| 2 | 147.5965 | 123.9808 |
| 3 | 184.0046 | 121.4065 |
| 4 | 154.2578 | 131.2441 |
| 5 | 119.3818 | 106.8801 |
| 6 | 154.5645 | 114.8462 |
| 7 | 106.6910 | 114.7603 |
| 8 | 137.3302 | 126.7799 |
| 9 | 174.4018 | 109.6003 |
| 10 | 129.1460 | 120.3422 |

### E.3.2 Hypothesis Tests

This section will discuss the results of the calculations performed to compare the performances of two RTO systems for the Williams-Otto plant, when parametric mismatch exists.

Table E.11: Profit Profiles for RTO With DOE: Mismatch Case Study

| RTO | % Return for Run # | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Intervals | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 59.18 | 59.18 | 59.18 | 59.18 | 59.18 | 59.08 | 59.08 | 59.18 | 59.18 | 59.18 |
| 2 | 56.16 | 60.93 | 60.50 | 59.78 | 60.75 | 60.32 | 60.83 | 59.93 | 59.76 | 56.77 |
| 3 | 58.70 | 59.15 | 56.92 | 59.34 | 61.22 | 59.32 | 58.58 | 52.97 | 59.54 | 56.23 |
| 4 | 52.25 | 60.33 | 58.48 | 58.18 | 59.86 | 59.54 | 60.14 | 57.39 | 59.86 | 58.39 |
| 5 | 57.16 | 57.82 | 59.66 | 61.19 | 60.47 | 58.06 | 62.07 | 60.35 | 60.11 | 59.90 |
| 6 | 59.75 | 59.20 | 60.31 | 60.80 | 59.82 | 58.88 | 60.92 | 58.95 | 59.76 | 57.95 |
| 7 | 61.28 | 59.40 | 60.79 | 61.77 | 61.80 | 63.50 | 60.11 | 59.94 | 62.91 | 60.23 |
| 8 | 61.46 | 60.16 | 57.48 | 57.66 | 61.79 | 59.54 | 59.43 | 60.03 | 59.67 | 59.96 |
| 9 | 60.06 | 59.31 | 58.94 | 59.13 | 60.84 | 58.87 | 60.04 | 60.79 | 60.80 | 61.60 |
| 10 | 59.36 | 58.11 | 59.46 | 59.17 | 60.38 | 59.90 | 58.72 | 59.51 | 57.31 | 60.73 |
| 11 | 60.30 | 59.56 | 59.57 | 61.23 | 61.28 | 60.32 | 58.88 | 59.64 | 59.05 | 60.22 |
| 12 | 59.47 | 59.32 | 59.21 | 60.06 | 59.91 | 60.32 | 58.89 | 58.99 | 59.01 | 60.67 |
| 13 | 58.99 | 58.08 | 58.61 | 61.08 | 60.29 | 60.32 | 59.05 | 59.17 | 58.98 | 61.09 |
| 14 | 60.12 | 58.08 | 60.55 | 59.62 | 59.14 | 60.50 | 59.80 | 59.06 | 60.55 | 60.77 |
| 15 | 59.77 | 59.04 | 61.08 | 58.89 | 59.77 | 59.55 | 61.12 | 58.91 | 63.29 | 59.47 |
| 16 | 61.05 | 58.90 | 60.66 | 59.39 | 60.47 | 59.11 | 57.91 | 59.31 | 61.19 | 60.23 |
| 17 | 61.05 | 59.06 | 59.96 | 59.58 | 60.38 | 59.16 | 59.37 | 62.07 | 61.31 | 59.70 |
| 18 | 60.50 | 58.91 | 60.06 | 59.87 | 59.89 | 60.19 | 58.23 | 59.71 | 61.50 | 61.20 |
| 19 | 58.53 | 58.08 | 60.90 | 60.72 | 59.42 | 59.12 | 59.10 | 58.68 | 59.71 | 58.75 |
| 20 | 58.88 | 59.10 | 59.16 | 59.77 | 60.96 | 59.68 | 59.59 | 58.49 | 60.26 | 58.70 |
| 21 | 61.25 | 63.43 | 58.21 | 58.62 | 58.86 | 60.00 | 59.96 | 60.42 | 58.59 | 59.61 |
| 22 | 61.24 | 61.59 | 58.21 | 57.77 | 58.61 | 60.15 | 60.39 | 59.99 | 58.88 | 59.58 |
| 23 | 60.41 | 59.95 | 58.82 | 55.33 | 58.61 | 60.85 | 59.45 | 59.89 | 59.81 | 58.58 |
| 24 | 60.41 | 58.81 | 61.36 | 55.07 | 58.80 | 58.38 | 61.67 | 59.13 | 58.97 | 59.22 |
| 25 | 63.10 | 58.96 | 58.95 | 53.98 | 59.04 | 58.91 | 61.33 | 59.17 | 58.82 | 59.34 |

Please see *Appendix D* for a description of the statistical calculations. The following statistics were used,

$$\begin{aligned} F_{1-0.05,240,240} &= 0.8083 \\ F_{0.05,240,240} &= 1.2371 \\ F_{1-0.05,9,9} &= 0.3146 \\ F_{0.05,9,9} &= 3.1789 \\ t_{0.05,11} &= 1.7959 \end{aligned}$$

Table E.13 lists the sample variances of the % Return around the expected optimum for all the runs performed, and the pooled variance of the set of ten runs for each RTO system. Similar to the previous case study, it is assumed that the population variance among the ten runs for each RTO system are equal and can be pooled together according to Equation (??).

Since $s_t^2 < s_b^2$ the one-sided hypothesis test defined in Problem (3.13), with a significance of 5%, was used to test the possible equality of the population variances. The computed

Table E.12: Profit Profiles for RTO Without DOE: Mismatch Case Study

| RTO Intervals | % Return for Run # | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 |
| 2 | 53.72 | 58.61 | 49.90 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 | 62.65 |
| 3 | 59.54 | 57.52 | 61.00 | 52.18 | 62.23 | 56.67 | 62.65 | 62.65 | 62.65 | 62.65 |
| 4 | 58.11 | 57.52 | 62.42 | 61.83 | 60.17 | 49.79 | 62.65 | 62.65 | 49.90 | 62.65 |
| 5 | 59.44 | 57.52 | 62.42 | 58.85 | 60.17 | 53.42 | 57.17 | 62.65 | 61.00 | 52.18 |
| 6 | 59.44 | 62.87 | 61.06 | 58.85 | 63.15 | 53.42 | 62.17 | 62.65 | 62.42 | 61.83 |
| 7 | 59.44 | 62.87 | 49.03 | 63.29 | 63.15 | 53.42 | 62.17 | 60.24 | 62.42 | 58.85 |
| 8 | 59.44 | 59.73 | 49.03 | 63.29 | 63.15 | 53.42 | 60.44 | 56.41 | 61.06 | 58.85 |
| 9 | 59.44 | 53.34 | 49.03 | 63.29 | 61.06 | 62.22 | 60.44 | 56.41 | 49.03 | 63.29 |
| 10 | 59.44 | 53.34 | 51.39 | 62.41 | 59.85 | 58.51 | 60.44 | 56.41 | 49.03 | 63.29 |
| 11 | 59.44 | 60.39 | 51.39 | 59.50 | 59.85 | 53.59 | 60.44 | 56.41 | 49.03 | 63.29 |
| 12 | 57.83 | 60.39 | 51.39 | 59.50 | 55.25 | 62.63 | 59.41 | 56.41 | 51.39 | 62.41 |
| 13 | 63.55 | 60.39 | 61.82 | 47.73 | 55.25 | 57.90 | 57.38 | 62.13 | 51.39 | 59.50 |
| 14 | 63.55 | 60.39 | 59.31 | 59.67 | 55.25 | 53.23 | 60.20 | 62.13 | 51.39 | 59.50 |
| 15 | 52.19 | 60.39 | 63.01 | 45.96 | 62.41 | 55.99 | 60.20 | 62.13 | 61.82 | 47.73 |
| 16 | 52.19 | 60.39 | 56.37 | 60.08 | 59.77 | 63.69 | 60.20 | 62.13 | 59.31 | 59.67 |
| 17 | 52.19 | 60.39 | 57.02 | 60.08 | 57.10 | 63.69 | 60.20 | 59.81 | 63.01 | 45.96 |
| 18 | 55.69 | 58.49 | 58.98 | 60.08 | 57.10 | 63.69 | 56.37 | 59.81 | 56.37 | 60.08 |
| 19 | 55.69 | 52.74 | 58.98 | 60.08 | 57.10 | 63.69 | 57.79 | 59.81 | 57.02 | 60.08 |
| 20 | 52.41 | 57.62 | 58.98 | 60.08 | 57.10 | 63.69 | 58.46 | 60.64 | 58.98 | 60.08 |
| 21 | 62.26 | 55.20 | 57.85 | 55.04 | 59.19 | 52.13 | 58.46 | 52.86 | 58.98 | 60.08 |
| 22 | 62.26 | 51.31 | 57.85 | 55.04 | 61.31 | 59.51 | 57.83 | 52.86 | 58.98 | 60.08 |
| 23 | 52.65 | 56.41 | 57.85 | 61.86 | 61.31 | 59.51 | 63.08 | 52.86 | 57.85 | 55.04 |
| 24 | 52.65 | 60.18 | 57.85 | 46.10 | 56.36 | 63.63 | 58.88 | 52.86 | 57.85 | 55.04 |
| 25 | 55.98 | 60.18 | 57.85 | 54.09 | 56.43 | 51.13 | 58.88 | 52.86 | 57.85 | 61.86 |

Table E.13: Variance Around Expected Optimum: Mismatch Case Study

| Run | Variance | |
|---|---|---|
| | RTO with DOE (% Return) | RTO without DOE (% Return) |
| 1 | 4.49 | 14.84 |
| 2 | 1.50 | 10.18 |
| 3 | 1.32 | 22.45 |
| 4 | 3.76 | 27.69 |
| 5 | 0.91 | 7.60 |
| 6 | 1.10 | 23.38 |
| 7 | 1.25 | 3.76 |
| 8 | 2.50 | 14.44 |
| 9 | 1.78 | 24.92 |
| 10 | 1.68 | 21.55 |
| Pooled | 2.03 | 17.08 |

value of the statistic is,

$$\frac{s_1^2}{s_2^2} F_{0.05,240,240} = 0.15$$

Since this value is less than 1, the null hypothesis is rejected and it can be stated that the variance is reduced with a DOE component added to the RTO system.

The next group of tests compares the mean of the extended design cost measure for the two RTO systems. Similar to the last case study, it must first be determined if the variance of the two data sets may be considered equal to determine the statistic that should be used to compare the means. Table E.10 lists the EDC over ten runs for each RTO system. Table E.14 presents the results of the one-sided hypothesis test with 5% significance.

Table E.14: Sample Standard Deviation of EDC: Mismatch Case Study

| | **Standard Deviation of EDC** | **Upper Stat.** |
|---|---|---|
| RTO with DOE | 7.56 | 0.30 |
| RTO without DOE | 24.58 | |

Since the upper statistic is less than one, it can be stated using a one-sided hypothesis test of 5% significance, that the variances of the two EDC data sets are different. Therefore the statistic defined in Inequality (D.9) must be used to evaluate the means of the data sets.

With the variances of the two data sets considered not equal, Equation (D.10) is used to solve for the degrees of freedom to be used in the hypothesis test,

$$v = 11.06$$

Since $\bar{x}_t < \bar{x}_b$ the one-sided hypothesis test defined in Problem (3.15), with a significance of 5%, was used to test the possible equality of the population means. The computed value of the statistic is,

$$\bar{x}_1 - \bar{x}_2 + t_{\alpha,v}\sqrt{\frac{s_t^2}{n_1} + \frac{s_b^2}{n_2}} = -14.1$$

Since this value is less than 0, the null hypothesis is rejected and it can be stated that the extended design cost is reduced with a DOE component added to the RTO system.