uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Luis Felipe Dominguez Palomeque
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A.Sc. (Chemical Engineering)
GRADE / DEGREE

Department of Chemical Engineering
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Investigation of Robust Techniques for the Design of Chemical Processes under Uncertainty

TITRE DE LA THÈSE / TITLE OF THESIS

Dr. David McLean
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dr. Arturo Macchi

Dr. Jason Zhang

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Investigation of Robust Techniques for the Design of Chemical Processes under Uncertainty

by

Luis Felipe Dominguez Palomeque

A thesis submitted to the Faculty of Graduate and Postdoctoral Studies
in partial fulfilment of the requirements for the degree

Master of Applied Science

in the Department of Chemical Engineering
University of Ottawa
December, 2006

Your file Votre référence
ISBN: 978-0-494-49190-4
Our file Notre référence
ISBN: 978-0-494-49190-4

Statement of Contributions of Collaborators

I hereby declare that this thesis is the result of two years of research carried out by myself. The computer codes for the sampling techniques used in this work as well as the optimization formulations in GAMS® and MATLAB® were all made by me.

I personally wrote this thesis and my supervisor, Dr. David McLean, of the Department of Chemical Engineering at the University of Ottawa supervised my work and made editorial corrections.

Signature: Date: December, 14 2006

To my parents,
Luis Felipe and Marilú

Abstract

The design of a chemical process is a complex optimization problem. Process models are used to describe the interrelationship of the variables being analyzed and provide the means for representing the description of the problem in mathematical terms. Traditionally, deterministic optimization has been employed to obtain values of design and control variables which minimize capital and operating costs. Subsequently, over-design or safety factors are applied to design variables to account for the fact that plant operating parameters and input variables are uncertain. This design strategy has some limitations. Over-design factors unnecessarily increase capital costs and do not guarantee that the process will have feasible operation. Moreover, the deterministic optimization fails to explore the benefits of integrating operability issues, such as flexibility and robustness, at the design stage. This work examined integrated robust techniques for the design of chemical processes and employed a rigorous stochastic optimization approach that fully take into account the uncertainty of technical parameters and the inherent variability of input variables. In addition to capital and operating costs, quality costs were considered. The influence of uncertain parameters on the final cost was also studied and sensitivity analyses were performed to determine their impact. It was found that the optimal values of the process variables were affected by the uncertainty of the parameters. That is, they presented considerable sensitivity. The location of the optimal values of these variables was also influenced by the components of the overall cost. It is shown that capital and operating costs tend to increase when quality costs contribute in great extent to the overall cost. The overall cost was also influenced by the desired extent of robustness in quality variables.

The stochastic optimization formulation turns the objective function into a probabilistic representation, commonly in form of expectations. Contrary to the common approach, where quadrature or cubature formulas are used, this work employed sampling techniques to estimate the expected value of the objective function of the stochastic optimization problem. It was found that, at a modest computational effort, the Hammersley Sequence Sampling

(HSS) provides a more accurate estimate of the objective function than cubatures and Monte Carlo sampling techniques. Three quasi-Monte-Carlo sampling techniques, based on the sequences of Halton, Faure and Sobol, were also examined for their potential used in the solution of stochastic optimization problems. These sampling techniques were implemented in a computer code in MATLAB and tested in two engineering problems. It was found that the proposed sampling techniques, Halton Sequence Sampling (HalSS), Faure Sequence Sampling (FSS) and Sobol Sequence Sampling (SSS), can compete with Hammersley Sequence Sampling (HSS) in terms of their low required sample sizes and number of function evaluations to achieve sufficient accuracy.

Résumé

L'élaboration d'un processus chimique est un problème d'optimisation complexe. Les modèles d'un processus sont utilisés pour décrire l'interdépendance des variables analysées; ils permettent de présenter une description du problème en termes mathématiques. Traditionnellement, on emploie l'optimisation déterministe pour obtenir des valeurs de calcul et des variables de contrôle qui réduisent les coûts capitaux et opérationnels. Plus tard, les coefficients de sécurité ou de surdimensionnement sont appliqués aux variables du calcul pour justifier le fait que les paramètres d'exploitation de l'usine et les variables d'entrée sont incertains.Cette stratégie de surdimensionnement augmente inutilement les frais en capitaux et ne garantissez pas que le processus aura une opération rentable. D'ailleurs, l'optimisation déterministe n'explore pas les avantages d'intégration des questions d'opération, telles que la flexibilité et la robustesse, à l'étape de la conception. Ce travail a examiné des techniques robustes intégrées pour la conception des processus chimiques et utilise des techniques rigoureuses d'optimisation stochastique qui tiennent compte entièrement de l'incertitude des paramètres techniques et de la variabilité inhérente des variables entrées. En plus des frais capitaux et d'opération, coûts de les maintient de qualité ont été considérés. L'influence des paramètres incertains sur le coût final a été également étudiée et des analyses de sensibilité ont été exécutées pour déterminer leur impact. On a découvert que les valeurs optimales des variables du processus étaient influencées par l'incertitude des paramètres, c'est-à-dire qu'elles possédaient une sensibilité considérable. L'emplacement des valeurs optimales de ces variables était aussi influencé par les éléments du coût total. On voit que le capital et les coûts d'opération ont tendance à augmenter lorsque les coûts de qualité contribuent dans une grande mesure au coût total. Le degré de robustesse des variables de qualité a aussi influencé le coût de la conception.

Cette formulation stochastique transforme la fonction objectif en une représentation probabiliste, généralement sous la forme de valeurs probables. Contrairement à l'approche commune, dans laquelle les formules de quadrature ou celles de cubature sont utilisées, dans ce travail, des méthodes d'échantillonnage ont été employées pour évaluer la fonction

objectif du problème d'optimisation stochastique. On a découvert qu'avec un peu de travail informatique, l'échantillonnage par suite de Hammersley fournit des évaluations plus exactes de la fonction objectif que les méthodes d'échantillonnage de cubature et de Monte Carlo. Trois méthodes d'échantillonnage presque similiaires à celle de Monte Carlo, s'appuyant sur les suites de Halton, Faure et Sobol, ont aussi été examinées en raison de leur utilisation potentielle dans la résolution des problèmes d'optimisation stochastique. Ces méthodes d'échantillonnage ont été implantées dans un code machine dans MATLAB et mises à l'essai dans deux problèmes d'ingénierie. On a découvert que les méthodes d'échantillonnage proposées, l'échantillonnage de la suite de Halton, l'échantillonnage de la suite de Faure et l'échantillonnage de la suite de Sobol, peuvent se mesurer à l'échantillonnage de la suite de Hammersley comme peu d'évaluations de fonctions.

Acknowledgments

I would like to express my gratitude to my supervisor, Dr. McLean, for his support and guidance during my Masters studies. I am also grateful for introducing me to the area of experimental design and optimization under uncertainty and for giving me the opportunity to carry out my research on such an interesting topic.

I am grateful to my thesis committee, professors Arturo Macchi and Jason Zhang, for providing valuable suggestions and recommendations that could improve the content of this thesis. I also would like to thank to people at the Department of Chemical Engineering at the University of Ottawa, professors, peers and staff for their assistance and companionship during my stay here.

I am indebted to my grandparents Manuel Gonzales Salvador and Georgette de Gonzales for encouraging me to pursue graduate studies. I could not have achieved one of my most yearned goals without their help and support throughout the years of my life. Thank you for believing in me.

Most importantly thanks to my parents, Luis Felipe Dominguez Diaz and Maria de Lourdes Palomeque Diaz for their unconditional love and for their encouragement. I owe to them who I am.

Financial support from the Mexican Council of Science and Technology (CONACyT) through its scholarship program is greatly appreciated.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

The design of a chemical process/product is a complex process. It requires extensive research and a long period of development. Currently, the procedure used to design a chemical process is through the hierarchical approach (Douglas, 1998). In the hierarchical approach the initial design is made as simple as possible by initially solving basic design problems and then updating the design by adding successive amounts of detail. Throughout the process, design objectives and constraints are gradually modified.

This design strategy has, however, some limitations. Typically, consideration of process control is left as final design project, often resulting in a more difficult or impossible control system design. A cost-effective solution, which would provide a more accurate assessment of the overall economics, would be obtained if process control and process design were considered as an integrated task. Moreover, such an integrated strategy would minimize the effect of disturbances, like changes in a feedstock's composition, on the quality of products and provide a more flexible process. Recognizing that integrated design leads to a more efficient, easy-to-control and cost-effective chemical plant, researches have attempted to integrate operability issues such as controllability, flexibility and robustness.

Controllability is concerned with the stability of the dynamic response of processes. In an attempt to integrate control and design at the same stage researchers have studied the interactions between design and control variables (Blanco and Bandoni, 2003; Luyben and Floudas 1994; Bansal *et al.*, 2000). Flexibility is concerned with the problem of ensuring feasible operation of a plant once it has been constructed. As in the case of controllability, flexibility has been broadly studied and different measures of flexibility have been proposed (Swaney and Grossman, 1983; Chacon-Mondragon and Himmelblau, 1988; Straub and Grossmann, 1990; Pistikopoulos and Mazzuchi, 1990). Robustness is concerned with the ability of the process to tolerate uncontrollable variations in factors internal and external to the process. Special emphasis is given to the robustness of quality characteristics since designs must ensure that consistent quality is achieved to satisfy clients' demands. Quality has become an issue of growing importance for the processing industries due to market demands and economic competition. By addressing quality issues early in the development process, unforeseen technical, regulatory and economic consequences of design choices can be anticipated.

Within the field of quality engineering, robust quality studies have been performed and techniques to achieve robustness have been proposed. A key element in the development of robust quality methods has been the work of G. Taguchi who developed an approach for designing products and processes that are insensitive (i.e., robust) to uncertainty and variability. He defined his methodology as Robust Parameter Design. Unlike mechanical and electrical engineering, Taguchi's method has not been broadly applied to the chemical engineering field. For the design of chemical processes, a few studies have been performed to try to accommodate robustness. Straub and Grossman (1993) proposed an approach for maximizing the flexibility of a design and adapted Taguchi's ideas for minimizing quadratic losses. They argued that the incorporation of the Taguchi metric in their approach may not yield satisfactory results due to its inability to handle hard constraints (i.e., those that must be satisfied). Boudriga (1990) applied Monte Carlo simulation coupled with a nonlinear optimization routine to arrive at the design that minimized the variance of a quality response. Diwekar and Rubin (1994) adapted the robust parameter design of Taguchi in the design of a chemical process using a simulator rather than physical experiments. As in the study of Boudriga (1990), they focused on the minimization of variance rather than quality costs. Samsatli *et al.* (1998) proposed different robustness metrics capable of dealing with various types of constraints that may be imposed on a design or on the performance metric.

All the above approaches attempted to account for variability by optimizing a performance metric, such as variance excluding capital and operating costs. Only a few studies have been devoted to incorporating robustness in terms of product quality costs along with design and operating costs. Georgiadis and Pistikopoulos (1999) presented a unified framework for incorporating flexibility and robustness. They defined a net profit value

(NPV), the difference between the expected profit and quality costs, as the objective to be optimized. Bernardo *et al.* (2001) proposed a systematic design framework for process quality, embedding Taguchi's quality loss concept and robustness criteria within a stochastic optimization formulation, and explored the trade-offs between profitability and robustness.

An important element of the design of chemical processes is the accuracy and precision of the values of physical constants used in design equations. Uncertainty in the values of these parameters can significantly affect the design due to the sensitivity of equipment sizes and final costs to the parameter values. Precise and accurate data are needed to determine if a design is feasible and economically viable. Also, they play an important role in responding to changing conditions. Thermodynamic, transport and physical properties are examples of data that are essential to process design (Larsen, 1986). Accurate data tighten designs and reduce energy costs (Williams and Albright, 1976). Errors in activity coefficient can result in increasing the number of separation stages and thus increase the capital cost. The design process is a complex optimization problem where the objective is expressed in terms of cost or profit, and reliable data can make the difference between an attractive low-cost process and one which cannot be justified.

Despite this fact, the effect of the uncertainty on the final cost of a design has largely been overlooked. Certainly, there's a potential for reducing the overall cost by identifying those uncertain parameters which have an important effect on the overall cost of a process. A sensitivity analysis on the impact of uncertain parameters on the final cost can identify those parameters having greatest influence on costs and suggest where improved values are required. Robust design techniques also improve the ability of a process to tolerate different sources of uncertainty.

1.1 Objectives

The objectives of this thesis were:

1) To evaluate existing robust design techniques for quality improvement within a stochastic optimization framework for design of chemical processes;
2) To demonstrate an evaluate the integration of capital operating and quality costs within the design of chemical processes under uncertainty;
3) To study the impact of uncertain parameters on the final cost;
4) To compare current integration techniques and to devise a method of integration that further reduces the computational demand of stochastic optimization algorithms.

To meet the above objectives, steady-state simulations and optimizations of a reactor and a reactor-exchanger system were carried out using MATLAB® and the General Algebraic Modelling System (GAMS) (Brook *et al.*, 2005).

1.2 Thesis Outline

This thesis consists of six chapters. In Chapter Two a literature review and background for robust parameter design methodology and process design, via stochastic mathematical programming, are presented. Chapter Three, the first paper, focuses on the minimization of quality costs. Quality costs are modelled with a quadratic loss function as suggested by Taguchi (1986). Chapter Four, the second paper, deals with the integration of capital and operating costs with quality cost. Using the proposed design framework of Bernardo *et al.* (2001), the one-stage stochastic optimization approach is extended to study the effect of uncertainty on the overall cost of a chemical process. A difference from work of Bernardo et al. (2001), where cubatures were used, in this work, HSS is employed as a method of integration. Different uncertain parameters and their level of uncertainty are also studied. In Chapter Five, the third paper, three quasi-Monte-Carlo sampling techniques are proposed as methods of integration for the solution of stochastic optimization problems. Finally, conclusions, contributions and recommendations are summarized in Chapter Six.

References

Bansal, V., Ross, R., Perkins, J. D., and Pistikopoulos, E. N., “The interactions of design and control: double-effect distillation”, Journal of Process Control, Vol. 10, 2000b.

Bernardo, F. P., Pistikopoulos, E. N., and Saraiva, P. M., “Quality cost and robustness criteria in chemical process design optimization”, Computers and Chemical Engineering, Vol. 21, 27-40, 2001.

Blanco, A. M., and Bandoni, J. A., ”Interactions between process design and operability of chemical processes: an eigenvalue optimization approach”, Computers and Chemical Engineering, Vol. 27, 2003.

Boudriga, S., “Evaluation of parameter design methodologies for the design of chemical process units”, M.Sc. Thesis, University of Ottawa, 1990.

Chacon-Mondragon, O. L., and Himmelblau, D. M., “Integration of flexibility and control in process design”, Computers and Chemical Engineering, Vol. 20, 1995.

Chacon-Mondragon, O. L., and Himmelblau, D. M., “A new definition of flexibility for chemical process design”, Computers and Chemical Engineering, Vol. 12, No. 5, 1988.

Diwekar U. M., and Kalagnanam, R. J., “Efficient sampling technique for optimization under uncertainty”, AIChE J., Vol. 43, No. 2, 1997.

Diwekar, U. M., and Rubin, E. S., “Parameter design methodology for chemical processes using a simulator”, Industrial Engineering Chemistry Research, Vol. 33, 292-298, 1994.

Douglas, J.M., “Conceptual design of chemical processes”, McGraw-Hill, 1988.

Larsen, A. H., “Data quality for process design”, Fluid Phase Equilibria, Vol. 29, 47, 1986.

Lin, D. K., and Tu, W., "Dual response surface optimization", Journal of Quality Technology, Vol. 27, No.1, 1995.

Luyben, M.L., and Floudas, C. A., "Analyzing the interaction of design and control-1. A multiobjective framework and application to binary distillation synthesis", Computers and Chemical Engineering, Vol. 18, No. 10, 1994.

Pistikopoulos, E. N., and Mazzuchi, T. A., "A novel flexibility analysis approach for processes with stochastic parameters", Computers and Chemical Engineering, Vol.4, No. 9, 1990.

Prasad, B., "Concurrent engineering fundamentals", Prentice-Hall, Englewood Cliffs, NJ., 1996.

Rosenblatt, A., and Watson, G. F. , "Concurrent engineering" , IEEE spectrum, July 1991.

Straub, D. A., and Grossman, I. E., "Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties", Computers and Chemical Engineering, Vol. 14, No. 9, 967-985, 1990.

Swaney, R. E., and Grossman, I. E., "An index for operational flexibility in chemical process design", AIChE J., Vol. 31, No. 4, 1985.

Vassilis, S., Perkins, J. D., and Pistikopoulos, E. N., "Recent advances in optimization-based simultaneous process and control design", Computers and Chemical Engineering, Vol. 29, 2004.

Vining, G. G., and Myers, R. H., "Combining Taguchi and response surface philosophies: A dual response approach", Journal of Quality Technology, Vol. 22, No.1, 1990.

Taguchi, G., "Introduction to quality engineering", Asian productivity association, 1986.

Williams, C. C., and Albright, M. A., "Better data saves energy", Hydroc. Proc., Vol, 55, No. 5, 115, 1976.

Chapter 2

Background and Literature Review

The user's need is the driving force for the creation of a process or product, and the degree of satisfaction that will be achieved is interpreted by its quality. A step that leads to a high-quality product/process is the identification of the optimal design specification for its design/manufacturing. This process is called product/process development or engineering design.

Most nonconformities are due to design specification. Such is the case for the chemical industry where about 50% of the problems are due to development (Gryna *et al.*, 2006). This is common since both the product and the process for manufacturing it, are usually devised at the same stage. The proper design of a process is of paramount importance since its operation will impact the final cost of a product. An important factor in the final cost of a product is the quality cost. Quality cost or cost of poor quality involves other costs such as the cost incurred for reprocessing defective material, the cost incurred due to waste derived from poor performance of processes, or the implicit cost incurred for losing customers. The main categories involved in quality cost are shown in Figure 2.1. For a more detailed discussion on the definition of quality costs the reader is referred to chapter two of the book due to Gryna *et al.* (2006).

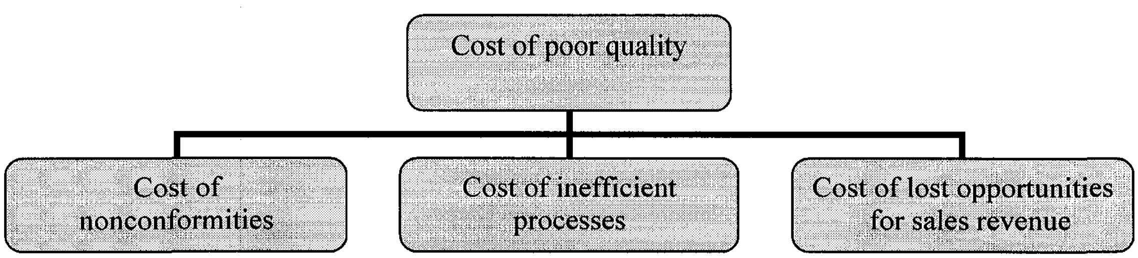

Figure 2.1 Classification of the cost of poor quality (Gryna *et al.*, 2006)

The cost of nonconformities may be attributed to the impossibility to repair a defective product (i.e., scrap), the reprocessing of defective or out of specification products (i.e., rework), or the reinspection of product that has been reprocessed. Among the factors affecting the cost of inefficient processes are the variability in process performance, the loss of capacity due to equipment failure, or losses due to recycling time. The cost of loss of opportunities is related to the potential loss of costumers because of poor quality or the potential loss of revenue due to the inability to satisfy clients' needs.

Traditionally, process/product development is done through the assignation of nominal values of design variables and the establishment of lower and upper specification limits of a quality characteristic with the belief that, by keeping it within those limits, quality costs are not incurred. Unfortunately, this approach does not eliminate quality costs. One of the first to realize this was G. Taguchi, who suggested in his book (Taguchi and Wu, 1980) that quality losses are incurred whenever the quality characteristic departs from its target value regardless of whether or not it falls within the specification limits. Taguchi proposed the use of a quadratic approximation to represent the economic losses due to variations in quality characteristics. The mathematical form for Taguchi's loss function is given by Equation 2.1 and its graphical representation by Figure 2.2,

$$L(y) = k(y-\tau)^2 \tag{2.1}$$

where L(y) represents the loss imparted to the product's user when the quality variable, y, deviates from its target value, τ, and k is a constant factor relating deviation from target to economic loss. The value of k can be determined if the actual loss is known for a particular value of y. For instance, for the case in which τ =5 and y = 10 the loss would be L = 50. Solving for k would give k = 2. Thus the actual loss function for this specific process would be $L = 2(y-\tau)^2$. Although this may not be reliable, it may be difficult to obtain data that provide complete and accurate coverage of quality costs required for building accurate models for quality losses. Nevertheless, the location of the minimum quality cost, using the quadratic loss function, will be independent of k and requires only knowledge of the quality characteristic to perform the optimization. On the other hand, knowledge of the value of k

will be critical if the quality loss is to be combined with other costs associated with process design such as capital and operating costs.

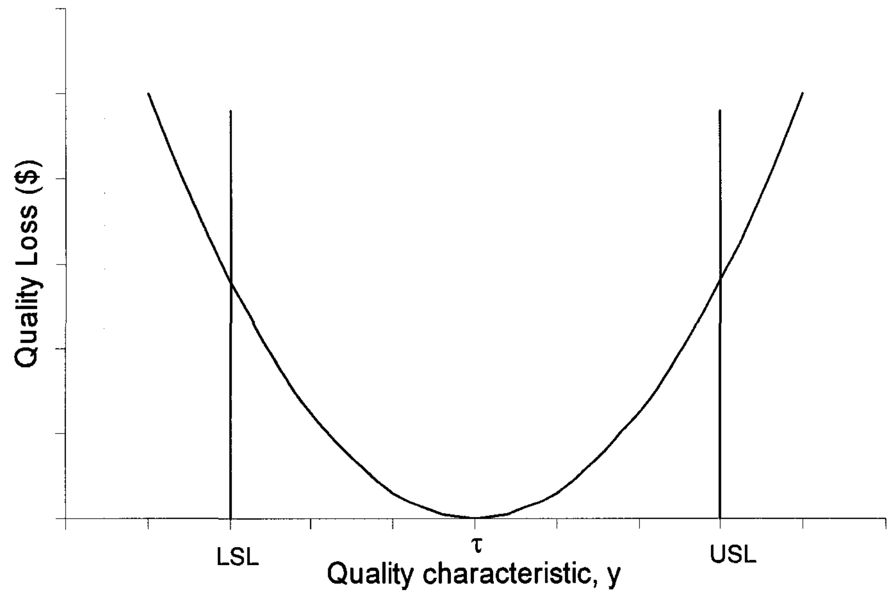

Figure 2.2 Representation of quality losses

Taguchi's loss function makes it clear why variability in the quality characteristic plays an important role in reducing quality costs. Along with the realization of this fact came the definition of "robustness". A product or process is said to be "robust" if not affected by variability in its inputs during operation. This implies that a process can be designed to minimize the impact of input variations on quality characteristics thereby leading to less quality loss. However, creating a robust design is not an easy task. It requires (1) the identification of the factors or variables that affect the performance of the quality characteristic and (2) the identification of the nominal values of those variables. This situation gets more complicated when there are design variables that affect both the mean and variance of the quality characteristic.

Based on experimentation and the construction of prototypes, Taguchi and Wu (1980) developed a methodology for determining the optimal values of the product and process variables that minimize variability while keeping the mean on target. Taguchi called this methodology Robust Parameter Design (RPD). Over the years, his methodology has been improved and his ideas have been implemented in different frameworks. New approaches were devised that incorporated other types of costs. These new approaches focussed not only on the minimization of quality losses but also on capital and operating costs. Of special interest for this work is the stochastic optimization approach since it does not involve physical experimentation.

Next the principles of RPD, as introduced by Taguchi and Wu (1980), and the mathematical foundations of the stochastic optimization approach for process design under uncertainty are presented.

2.1 Taguchi approach to design for quality improvement

Considering products and processes as systems, Taguchi (1980) classified the inputs that influence the quality of the systems as control variables and noise variables. Control variables were defined as those that are easily controlled and manipulated. Noise variables were defined as those that are difficult and/or costly to control and thus uncontrollable. Defining $z = [\mu_{z_1}\ \mu_{z_2} \ldots \mu_{z_n}]$ as the vector of control variables representing the mean values of the control variables and $\boldsymbol{\theta} = [\theta_1\, \theta_2 \ldots \theta_n]$ as the vector of noise variables, a quality characteristic, y, is represented as, $f(z,\boldsymbol{\theta})$. In order to identify the settings of control variables that make a system insensitive to variation, Taguchi proposed different criteria to be optimized. He proposed the use of performance statistics, called signal-to-noise ratios, to estimate the effect of noise variables on quality variables and to minimize the expected loss derived from variations. While quality characteristics are functions of z and $\boldsymbol{\theta}$, signal-to-noise ratios are functions of the z only. It is because of this assumed condition that proper adjustment of the control variables makes possible the minimization of the expected loss.

Signal-to-noise ratios were created so that when maximized, quality costs are minimized. They involve the parameters of the distribution of y (i.e., mean and variance) which are defined by

$$\mu_y(z) = E[y] \qquad (2.2)$$

and

$$\sigma_y^2(z) = \mathrm{E}[y - \mu_y(z)^2] \tag{2.3}$$

If the functional form for the performance characteristic is not well known, it needs to be estimated through experimentation or simulation. In order to maximize the appropriate signal-to-noise ratio, Taguchi employed orthogonal array designs for his experiments to find the optimal values of the control variables.

In engineering applications, three of the most commonly used signal-to-noise ratios are smaller-the-better, larger-the-better and target-is-best. They were derived from the definitions of the mean square error (MSE), the bias (deviation from target), and the square of the inverse of the coefficient of variation of y and are given by the Equations 2.4 to 2.6

$$\mathrm{MSE}\ (z) = \mathrm{E}[(y - \mu_y(z)^2] \tag{2.4}$$

$$\mathrm{B}\ (z) = \mu_y(z) - \tau \tag{2.5}$$

$$\xi\ (z) = \mu_y^2(z)/\sigma_y^2(z) \tag{2.6}$$

Kackar (1985) analyzed the Taguchi approach and explained the origin of signal-to-noise ratios; a summary is given here.

In the *smaller-the-better* case, the performance characteristic, y, is considered to be normally distributed with the target value $\tau = 0$. The loss function, L(y), increases as y increases from zero, and the expected loss in Equation 2.1 is proportional to

$$\mathrm{MSE}\ (z) = \mathrm{E}[(y - 0)^2] = \mathrm{E}[y^2] \tag{2.7}$$

Taguchi suggests maximizing the performance statistic

$$\mathrm{S/N} = -10\ \mathrm{Log}\ \mathrm{MSE}(z) \tag{2.8}$$

which minimizes the mean squared error, MSE(z), and can be estimated by

$$\mathrm{MSE}(z) = \sum_{i=1}^{n} y_i^2 / \mathrm{N_p} \tag{2.9}$$

where N_p is the number of samples taken from the experiment.

In the *larger-the-better* case, it is assumed that the quality characteristic, y, has a positive distribution and the target value is infinity. The performance statistic is formulated as the inverse of the MSE of the smaller-the better case, and the loss function $L(y)$ decreases as y increases from zero. The performance statistic is given by

$$\mathrm{S/N} = -10 \text{ Log } \mathrm{MSE}(z) \tag{2.10}$$

where the MSE can be estimated by

$$\mathrm{MSE}\ (z) = \mathrm{E}[(1/y)^2] = \sum_{i=1}^{n} (1/y_i^2)/\mathrm{N_p} \tag{2.11}$$

In the *target-is-best* case, it is assumed that the quality characteristic, y, has a specific target value, $\tau = \tau_0$, the loss function $L(y)$ increases as y deviates from τ_0 in either direction and $E[y] \neq \tau_0$. Accordingly, the expected loss in Equation 2.1 is exactly the same as in Equation 2.12

$$\mathrm{MSE}\ (z) = \mathrm{E}[(y-\tau_0)^2] \tag{2.12}$$

The target-is-best case is the performance measure most frequently used. Therefore, a more detailed derivation is given here. If we add and subtract $\mu_y(z)$, Equation 2.12 can be written as

$$\mathrm{MSE}(z) = \mathrm{E}[(y - \mu_y(z)) + (\mu_y(z) - \tau_0)]^2 \tag{2.13}$$

Expanding this expression one obtains

$$\mathrm{MSE}(z) = \mathrm{E}[(y-\mu_y(z))^2 + 2(y-\mu_y(z))(\mu_y(z)-\tau_0) + (\mu_y(z)-\tau_0)^2] \qquad (2.14)$$

Since, $\mathrm{E}[y-\mu_y(z)] = \mathrm{E}[y]-\mu_y(z) = 0$ the middle term becomes zero, thus

$$\mathrm{MSE}(z) = \mathrm{E}[y-\mu_y(z)]^2 + \mathrm{E}[\mu_y(z)-\tau_0]^2 \qquad (2.15)$$

Furthermore, by definition,

$$\mathrm{E}[y-\mu_y(z)]^2 = \sigma_y^2(z) \qquad (2.16)$$

and

$$\mathrm{E}[\mu_y(z)-\tau_0]^2 = (\mu_y(z)-\tau_0)^2 \qquad (2.17)$$

Substituting Equations 2.16 and 2.17 into 2.15, Equation 2.12 becomes

$$\begin{aligned} \mathrm{MSE}(\mathbf{z}) &= \sigma_y^2(z) + (\mu_y(z) - \tau_0)^2 \\ &= \sigma_y^2(z) + (\mathrm{B}(z))^2 \end{aligned} \qquad (2.18)$$

In engineering applications two situations are common. (1) The mean and variance of the quality characteristic are independent and (2) the mean and variance are dependent. Taguchi assumes the existence of adjustment variables that minimize the squared bias term. Thus, for the first situation, the bias term in Equation 2.18 is eliminated and the MSE reduces to the variance term only

$$\mathrm{MSE}(z) = \sigma_y^2(z) \qquad (2.19)$$

For the latter situation, the relationship between the mean and variance can be expressed by the coefficient of variation, $\nu(z) = \sigma_y(z)/\mu_y(z)$. Since the adjustment variables take action on the deviation of the mean's quality characteristic from its target, the remaining objective is to minimize the coefficient of variation, $\nu(z)$, or, using the term described the Equation 2.6, to maximize the square of the inverse of the coefficient of variation $\xi(\mathbf{z})$. This is commonly known as a 'two-step' procedure. For the maximization case, the MSE takes the form

$$\mathrm{MSE}(z) = \frac{\mu_y^2(z)}{\sigma_y^2(z)} \tag{2.20}$$

Taguchi suggests maximizing the performance statistic

$$\mathrm{S/N} = 10 \text{ Log } \mathrm{MSE}(z) \tag{2.21}$$

which maximizes the mean squared error, MSE(**z**), and can be estimated by

$$\mathrm{MSE}(z) = \bar{y}^2 / s^2 \tag{2.22}$$

or

$$\mathrm{MSE}(z) = (\bar{y}^2 / s^2) - (1 / \mathrm{N_p}) \tag{2.23}$$

This can be easily demonstrated by considering, $\mathrm{N_p}\bar{y}^2 - s^2$, the unbiased estimator of $\mathrm{N_p}[\mu_y^2(z)]$. Thus

$$\mu_y^2(z) = \frac{\mathrm{N_p}\bar{y}^2 - s^2}{\mathrm{N_p}} \tag{2.24}$$

and

$$\frac{\mu_y^2(z)}{\sigma_y^2(z)} = \frac{\mathrm{N_p}\bar{y}^2 - s^2}{\mathrm{N_p} s^2} = \frac{\bar{y}^2}{s^2} - \frac{1}{\mathrm{N_p}} \tag{2.25}$$

Equation 2.22 and 2.23 are equivalent since both maximize the MSE when the performance statistic, Equation 2.21, is maximized.

Taguchi's approach relies on the existence of adjustment variables in order to bring the mean quality characteristic on target. In addition, in his methodology, it is assumed that any change in the control variables, in order to minimize the variance, will have an insignificant or no effect on the mean of the quality characteristic. Certainly, the assumption of availability of adjustment variables in all engineering problems is one of the drawbacks of Taguchi's approach since such adjustment variables not always exists as pointed out by Myers *et al.* (1992). Leon *et al.* (1987) investigated the relationship between the adjustment variables and the 'two-step' procedure and showed that under certain models (e.g., quadratic loss function), signal-to-noise ratios do minimize quality losses. For the case in which different models hold, they proposed the use of a different performance statistic, called performance measurement independent of adjustment (PERMIA), to perform the 'two-step' optimization.

Most of the recent research has been devoted to creating alternative approaches to parameter design. Boudriga (1990) and Vining and Myers (1990) applied Response Surface Methodology (RSM) techniques (Myers, 1976). Vining and Myers (1990) defined an appropriate region of interest within their experimental design and used second-order response surface designs to obtain data from which they fitted both a primary response, namely $\sigma_y^2(z)$, and a secondary response, $\mu_y(z)$. They formulated the robust deign problem as that of minimizing $\sigma_y^2(z)$ subject to the constraint, $\mu_y(z)=\tau_0$, for the target is best case. Essentially, this dual response problem was posed as the determination of the set of conditions $\mathbf{z}^* \in$ R, which minimize $\sigma_y^2(z)$ subject to $\mu_y(z)=\tau_0$. However, the limitation of this approach is that the true optimum can be overlooked since there may be situations in which, by allowing some bias in the target value, better operating conditions can be obtained. For the larger- and smaller-the-better cases they proposed modelling $\mu_y(z)$ as the primary response subject to a maximum value for the variance.

The optimization procedure proposed by Vining and Myers (1990) was based on the optimization procedure used in RSM. Essentially, the operating region was defined by specifying, for instance, the radius of the hypersphere, ρ. Thus the additional constraint, $\mathbf{z}^T\mathbf{z}=\rho$, was imposed in the optimization. The solution of the optimization was obtained by fixing the Lagrange multiplayer of equality constraint defined by the secondary response, λ_1, and then varying the Lagrange multiplayer of the second equality constrained, λ_2, defined by the operating region, within the interval $-\pi \leq \lambda_2 \leq \pi$.

Del Castillo and Montgomery (1993) noted that the range values for the Lagrange multiplier of the equality constraints describing the operating region was too restrictive and that it may lead to suboptimal solutions. They proposed the addition of more constraints in form of inequalities to include not only a specific value but also a range of possible values such as upper and lower bounds for the secondary response. With the inclusion of inequality constraints, the experimental design could also include different region of interest such as cuboidal (e.g., factorial designs) or spherical (e.g., central composite design). Because the optimization problem involved linear and nonlinear inequality constraints, they proposed the used of standard nonlinear programming techniques, specifically, the generalized reduce gradient (GRG) algorithm to perform the minimization.

Tu and Lin (1995) proposed a minimization of the mean squared error (MSE), given by Equation 2.18, as an alternative to the dual response approach. They pointed out that the MSE approach is not only restricted to polynomial second-order models but also can handle more realistic nonlinear mechanistic models. They concluded that forcing the estimated secondary response to equal a certain value as a constraint could be misleading as the optimization procedure would exclude globally preferred values.

Copeland and Nelson (1996) suggested formulating the dual response problem as that of minimizing $\sigma_y^2(z)$ subject to a constraint that specifies how far one would be willing to allow $\mu_y(z)$ to deviate from its target value τ_0. They made use of the Nelder-Mead simplex procedure to perform the required minimization.

Ding *et al.* (2004) proposed the minimization of a weighted combination of the squared bias and variance making up the MSE as expressed in Equation 2.18. Their approach is useful when the relative importance between bias and variance is not clear.

All the previous works create empirical models for the mean and variance of the quality characteristic. Polynomial response surface models are fitted to data obtained form experiments. Since the design is replicated at each design point, estimates of the mean and variance for the quality characteristic can be obtained and optimization algorithms are applied to arrive at conditions that minimize the quality costs. Empirical models have been used because the functional relationship between the mean or variance of the quality characteristic and control variables is not known nor the distributions that are involved in minimizing the expected loss. Consequently, designers resort on experimental designs and data analysis to identify such functional relationship and, to determine the nominal values for the control variables that make the process insensitive to variations induced by noise inputs. For the design a new process, however, this represents a limitation; especially if a plant is not available for running experiments before the process is designed. Fortunately, in the field of chemical engineering, process models based on first principles are available. For instance, one can model the phenomenon taking place inside a reactor using the principles of reaction kinetics along with material and energy balances or model the separation taking place in a distillation column using principles of vapour-liquid equilibria and tray hydraulics, along with the corresponding material and energy balances in each tray.

Currently, the use of computer-aids for the design of the chemical process is widely spread and, process models have been implemented in the design and optimization of chemical processes using computer packages such as Aspen® or Hysys®. However, most of the design and optimization of chemical process are carried out in a deterministic environment leaving aside the uncertainties that may be present in some parameters or ignoring the inherent variability in some process variables. In an attempt to minimize variability transmitted by some input variables Boudriga (1990) made use of process models for the design of a continuous stirred-tank reactor, and employed Monte Carlo simulations coupled with a nonlinear optimization routine to arrive at the best design to minimize variability. Diwekar and Rubin (1991) simulated an integrated gasification power plant using Aspen® then evaluated its performance in the face of uncertain conditions during its operation. With the stochastic modeling capability built around the simulator they performed sensitivity analysis which allowed them to identify critical variables for the design of such process. Later the same authors proposed a stochastic optimization framework for implementing the robust parameter design method using the ASPEN® process simulator. They explored the effect of different sampling techniques to improve the precision of the results and concluded that Latin Hypercube Sampling provided a higher precision in addition to showing a more consistent behaviour than Monte Carlo Sampling (Diwekar and Rubin, 1994).

Although the above approaches provided the means for dealing with parameter uncertainty and process variability, they focussed on the minimization of variability or quality costs only, excluding capital and operating costs. The realization of a process design into a fully operating plant highly depends on the economics involved and on its attractiveness for the investors. For the economics of a prospective process to be properly evaluated, it is necessary that knowledge about capital and operating costs be obtained (Biegler *et al.*, 1997). Combination of capital and operating costs with other costs then results in complete estimation of the profitability of the process (Douglas, 1998). Unfortunately, the inclusion of quality costs and the explicit treatment of robustness (variability) of some variables and uncertainty in model parameters, at the initial stage of the design, are rarely considered. This is due to the fact that the explicit treatment of uncertainties is not easy. It requires of systematic methods and specialized algorithms to handle different sources of uncertainty (Samsatli *et al.*, 1998). Furthermore, the solution of problems derived from handling uncertainty is computationally demanding.

Approaches have been proposed to make the treatment of uncertainties and variability tractable. These are based on nonlinear mathematical programming; they differ in how the uncertainty is modeled and the way the feasibility constraints are satisfied. Although different approaches exist, this work will only be dealing with the two-stage stochastic programming approach (Pai and Hughes, 1987; Pistikopoulos and Ierapetritou, 1995). This modeling is suitable for process design problems in chemical engineering. It allows the distinction between design and control variables, and allows constraints, which are usually encountered at the design and operating stage of a process, to be explicitly treated within a stochastic formulation. In the next section, the general formulation for the design of chemical processes under uncertainty and the two-stage stochastic optimization are presented.

2.2 Optimization under uncertainty

The design of chemical processes under uncertainty relies on the assumption that after the plant has been built, design variables, $\mathbf{d}$, remain fixed and cannot be changed. On the other hand, control variables, $\mathbf{z}$, can be adjusted to account for variations in the parameters, $\boldsymbol{\theta}$. It should be pointed out that, although variables presenting variability are not parameters, throughout this work they will be treated as such since uncertain variables can take any values according to the realization of their distributions.

The vector of design variables, $\mathbf{d}$, generally consists of variables related to the sizing of the equipment (e.g., volume, area, height). Variables that belong to $\mathbf{d}$ are assumed to be fixed and do not change during the operation of a plant. A vector of control variables, $\mathbf{z}$ (e.g., temperature, flow, level, pressure), consists of variables that can be adjusted to compensate fluctuations in parameter values, allowing the process to operate within specification and minimal cost. A vector of state variables, $\mathbf{x}$ (e.g., temperature, conversion), consists of variables that are functions of the control variables, design variables, and parameters. The last term to be included in the model is $\boldsymbol{\theta}$, which is a vector that contains the uncertain parameters of the model.

A performance measure, $C(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta})$, by which the design is evaluated, is formulated and may represent a profit or return to be maximized or a cost to be minimized. Equality constraints, $\mathbf{h}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta})$, typically represent a system of equations for material and energy balances. Inequality constraints, $\mathbf{g}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta})$, typically represent the mathematical formulation of process specifications. Therefore, the problem of designing a chemical plant can be represented by the nonlinear program (NLP)

$$
\begin{aligned}
\min \quad & C(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) \\
\text{w.r.t.} \quad & \mathbf{d},\mathbf{z},\mathbf{x} \\
\text{s.t.} \quad & \mathbf{h}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) = 0 \\
& \mathbf{g}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) \le 0 \\
& \mathbf{d} \in D, \quad \mathbf{z} \in Z, \quad \mathbf{x} \in X, \quad \boldsymbol{\theta} \in \Theta
\end{aligned}
\tag{2.26}
$$

Three approaches have been proposed to solve optimization problems under uncertainty depending on how the uncertainty is modeled: (1) the deterministic approach (Grossman and Sargent, 1978; Halemane and Grossman, 1983), (2) the parametric approach (Acevedo *et al.*, 1996; Acevedo and Pistikopoulos, 1997; Pertsinidis *et al.*, 1998), (3) the probabilistic approach (Pistikopoulos and Ierapetritou, 1995; Acevedo and Pistikopoulos, 1998)

In the deterministic approach boundary values of the uncertain parameters, $\boldsymbol{\theta}$, are specified and the parameter space, Θ, is approximated by a set of N_p scenarios or periods. The uncertainty space is expressed as

$$
\Theta = \left\{ \boldsymbol{\theta} : \boldsymbol{\theta}_i^L \le \boldsymbol{\theta}_i \le \boldsymbol{\theta}_i^U, \ i = 1,2,..N_p \right\}
\tag{2.27}
$$

and the expected cost is approximated by a weighted cost function (Grossman and Sargent, 1978). With this discretization strategy, the expected cost takes the form

$$
\sum_{i=1}^{N_p} w_i C(\mathbf{d}, \mathbf{x}_i, \mathbf{z}_i, \boldsymbol{\theta}_i)
\tag{2.28}
$$

where the weights, w_i, correspond to discrete probabilities for the selected finite number of parameter points $\boldsymbol{\theta}_i \in \Theta$, $i=1,2, \ldots N_p$. This deterministic approach it is suitable when it is known that a plant operates with parameter value $\boldsymbol{\theta}_i$ in a specified period and the length of the period is proportional to w_i. Because the optimization is performed for the set of scenarios or periods it becomes a 'multiperiod optimization'. When information on confidence limits of an unspecified distribution is available, this approach can also be applied since confidence limits can be considered the specified bounds.

In the parametric approach, the uncertain parameters, $\boldsymbol{\theta}$, are given by ranges and a solution is obtained for a parametric profile of uncertain parameters. Accordingly, the solution is a function of the uncertainty. Among the advantages of this approach is the information it provides on the sensitivity of the system's flexibility to values of the design variables (Bansal *et al.*, 2002). A disadvantage though is that the formulation becomes complex requiring specialized algorithms which solution demands high the computational effort, especially when more than ten uncertain parameters are involved.

The probabilistic approach is the most frequently used in the literature of process design under uncertainty (Grossman *et al.*, 1983). Designers take into account detailed statistical properties of parameter variations. Uncertain parameters are modeled as random variables, following a probability density function (PDF) and are approximated by specifying the statistical moments that correspond to a specified distribution. For instance, the first two moments, mean (μ_θ) and variance(σ_θ^2), characterize the uncertainty of normally distributed parameters. The parameter space is expressed as

$$\Theta = \{\boldsymbol{\theta} : \boldsymbol{\theta}_i \in \mathrm{j}(\boldsymbol{\theta})\} \tag{2.29}$$

This expression states that any realization of the uncertainty in the parameters, $\boldsymbol{\theta}_i$, is an element of a joint probability distribution set, $\mathrm{j}(\boldsymbol{\theta})$.

In this work, the probabilistic approach has been used since (1) Information about the distribution of the parameters can be extracted from process data of previous plants or from statistical analysis of experimental data obtained at the laboratory scale, and the appropriate PDF depends on the information gathered, with the uniform PDF representing a maximum degree of ignorance (Rudd and Watson, 1968), (2) Information concerning the precision of parameters can obtained from the variance of random errors. (3) Experimental designs and fitted models can be used for determining the level of uncertainty about the true values of parameters, (4) Parameters are commonly reported in nominal values with corresponding standard deviation or sigma limits. Therefore, the probabilistic approach seems to be a more realistic approach to representing the uncertainty in the parameters.

2.2.1 Two-stage stochastic optimization

The common approach to solve design problems that involve uncertain parameters characterized by PDFs is through two-stage stochastic programming (Halemane and Grossman, 1983). This approach assumes a perfect control scheme and is based on selecting optimal values for the vector $\mathbf{d}$ at the design stage (also called 'first stage') while seeking feasibility/flexibility at the operating stage (or 'second stage').

In the two-stage stochastic programming approach, values for design variables, $\mathbf{d}$, are selected at the design stage and it is assumed that the plant will operate optimally no matter what the values the parameters $\boldsymbol{\theta}$ take. The objective at the operating stage is to determine the optimal values of the control variables, $\mathbf{z}$, for each realization of the parameters in Θ. That is, there will be an optimal and feasible value of $\mathbf{z}$ for each $\boldsymbol{\theta}\in\Theta$.

It is common in practice to eliminate the state variables, $\mathbf{x}$, by expressing them as implicit functions of the control variables. This is possible since the dimension of the state vector, $\mathbf{x}$, equals the number of equality constraints. This is represented as

$$\mathrm{h}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) = 0 \rightarrow \mathbf{x} = \mathbf{x}(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \tag{2.30}$$

$$\mathrm{g}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) = \mathrm{g}(\mathbf{d},\mathbf{z},\mathbf{x}(\mathbf{d},\mathbf{z},\boldsymbol{\theta}),\boldsymbol{\theta}) = \mathrm{f}(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \le 0 \tag{2.31}$$

The optimal operation of a plant that minimizes the cost is found by solving the NLP,

$$\begin{aligned} &\min && \mathrm{C}(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \\ &\text{w.r.t.} && \mathbf{z} \\ &\text{s.t.} && \mathbf{f}(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \le 0 \\ & && \mathbf{d}\in \mathrm{D}, \quad \mathbf{z}\in \mathrm{Z}, \quad \boldsymbol{\theta}\in\Theta \end{aligned} \tag{2.32}$$

The solution obtained at the first stage represents the cost function $\mathrm{C}^*(\mathbf{d},\boldsymbol{\theta})$ that is optimal during operation of the plant for the current fixed values of $\mathbf{d}$ and $\boldsymbol{\theta}$. When the optimization is performed for each realization, $\boldsymbol{\theta} \in \Theta$, the average cost of operation corresponds to the expected value of $\{C^*(\mathbf{d},\boldsymbol{\theta})\}$.

In order to ensure that a feasible solution of the optimization problem at the operating stage exists for each parameter realisation, $\boldsymbol{\theta} \in \Theta$, that is to say, there exist values of **z** that satisfy the constraints in Equation 2.32, proper values of the design vector, **d,** must be chosen. In addition, for the design to be optimal, these values give rise to a minimum expected cost, $E_{\theta \in \Theta}\{C^*(\mathbf{d},\boldsymbol{\theta})\}$, over the parameter space Θ. This is an important condition since, if wrong values for the vector **d** are chosen, it will be impossible to find values of the control variables, **z,** that satisfy the constraints in Equation 2.32.

The mathematical formulation proposed by Halemane and Grossman (1983) to solve the above problem is given by

$$\begin{aligned}
&\min \quad E_{\theta \in \Theta}\{C^*(\mathbf{d},\boldsymbol{\theta})\} \\
&\text{w.r.t.} \quad \mathbf{d} \\
&\text{s.t.} \quad \forall \boldsymbol{\theta} \in \Theta \ \left\{ \exists \mathbf{z} \left(\forall j \in J[f_j(\mathbf{d},\mathbf{z},\boldsymbol{\theta})] \leq 0 \right] \right) \right\} \\
&\qquad \mathbf{d} \in D, \ \ \mathbf{z} \in Z, \ \ \boldsymbol{\theta} \in \Theta
\end{aligned} \tag{2.33}$$

where j = 1, 2,...J is the index set for the components of vector **f**. The constraint in Equation 2.33 is defined as the feasibility constraint because the feasibility of operation of a plant over the region, Θ, is ensured if and only if this constraint is satisfied. This logical constraint states that for each parameter realization, $\boldsymbol{\theta} \in \Theta$, there must exist at least one realization of the control vector, **z**, such that all elements in the constraint function set (i.e., J) are non positive. In other words, regardless of the values of the parameters within the region Θ, the plant with design vector, **d**, has the flexibility to satisfy all the process and product constraints.

Having identified the critical points in the optimization formulation, the problem of optimal design under uncertainty can be formulated as the following two-stage infinite stochastic programming problem (Halemane and Grossmann, 1983),

$$
\begin{aligned}
&\min \quad E_{\theta\in\Theta}\left\{\min_{\mathbf{z}} \; C(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \middle| f(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \le 0\right\} \\
&\text{w.r.t.} \quad \mathbf{d} \\
&\text{s.t.} \quad \forall \boldsymbol{\theta} \in \Theta \; \left\{\exists \mathbf{z}\left(\forall j \in J[f_j(\mathbf{d},\mathbf{z},\boldsymbol{\theta})] \le 0]\right)\right\} \\
&\qquad \mathbf{d} \in D, \;\; \mathbf{z} \in Z, \;\; \boldsymbol{\theta} \in \Theta
\end{aligned}
\tag{2.34}
$$

Since there are an infinite number of possible values of $\boldsymbol{\theta}\in\Theta$ the overall solution is dependent on $\boldsymbol{\theta}$ and the number of possible values for decision variables (i.e., **d**, **z**) involved in the problem described in Equation 2.34 is also infinite.

Since the vector of uncertain parameters, $\boldsymbol{\theta}$, follows a probability density function j($\boldsymbol{\theta}$), the expected value of the cost function C(**d**,**z**,$\boldsymbol{\theta}$) can be represented by the integral (Pistikopoulos and Ierapetritou, 1995) .

$$
E_{\theta\in\Theta}\left\{ \min_{\mathbf{z}} C(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \middle| f(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \le 0\right\} = \int_{\theta\in\Theta} C^*(\mathbf{d},\boldsymbol{\theta}) \;\; j(\boldsymbol{\theta}) \;\; d\boldsymbol{\theta} \tag{2.35}
$$

where

$$
C^*(\mathbf{d},\boldsymbol{\theta}) = \min_{\mathbf{z}} \; \left\{C(\mathbf{d},\boldsymbol{\theta}) \middle| \mathbf{f}(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \le 0\right\} \tag{2.36}
$$

The problem formulated in Equation 2.34 can be recast as

$$
\begin{aligned}
&\min \quad \int_{\theta\in\Theta} C^*(\mathbf{d},\boldsymbol{\theta}) \;\; j(\boldsymbol{\theta}) \;\; d\boldsymbol{\theta} \\
&\text{w.r.t.} \quad \mathbf{d} \\
&\text{s.t.} \quad \forall \boldsymbol{\theta} \in \Theta \; \left\{\exists \mathbf{z}\left(\forall j \in J[f_j(\mathbf{d},\mathbf{z},\boldsymbol{\theta})] \le 0]\right)\right\} \\
&\qquad \mathbf{d} \in D, \;\; \mathbf{z} \in Z, \;\; \boldsymbol{\theta} \in \Theta
\end{aligned}
\tag{2.37}
$$

The direct solution of this problem is impractical, demanding a high degree of computational effort when the number of parameter and decision variables is large. Consequently, approaches have been proposed to avoid the direct solution of Equation 2.37.

Pistikopoulos and Ierapetritou (1995) proposed a decomposition approach to solve the above problem. The problem was formulated as one determining the design that maximizes an expected profit while simultaneously measuring the flexibility of a design. The maximization of an expected profit was evaluated inside an associated feasible region of the design vector $\mathbf{d}$ (i.e., $R_n(\mathbf{d}) = \boldsymbol{\theta} \big| \; \forall \boldsymbol{\theta} \in R \;\; \exists \mathbf{z} : f(\mathbf{d}, \mathbf{z}, \boldsymbol{\theta}) \le 0$, where n is the number of uncertain parameters).

Having defined the feasible region, Equation 2.37 can be re-written as

$$\begin{aligned} &\min \int_{R_n(d)} C^*(\mathbf{d}, \boldsymbol{\theta}) \;\; j(\boldsymbol{\theta}) \;\; d\boldsymbol{\theta} \\ &\text{w.r.t. } \mathbf{d} \\ &\qquad \mathbf{d} \in D \end{aligned} \tag{2.38}$$

Using the expectancy operator, the objective function in Equation 2.38 can be represented as

$$E_{R_n(d)} \left\{ \min_{\mathbf{z}} C(\mathbf{d}, \mathbf{z}, \boldsymbol{\theta}) \big| f(\mathbf{d}, \mathbf{z}, \boldsymbol{\theta}) \le 0 \right\} \tag{2.39}$$

According to Straub and Grossman (1990), flexibility of a design corresponds to the measure of the ability of a design to cope with uncertainties for a given discrete state. When the region of uncertainty is described by a joint probability distribution function the flexibility of the design is designated as the stochastic flexibility because it corresponds to the portion of the joint probability distribution that lies within the region of feasible operation of the design. That is to say, a n-dimensional integral of $j(\boldsymbol{\theta})$ over the region $R_n(\mathbf{d})$. The stochastic flexibility metric is represented as

$$SF = \int_{\theta \in R(d)} j(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{2.40}$$

where

$$R_n(\mathbf{d}) = \boldsymbol{\theta} \big| \;\; \forall \boldsymbol{\theta} \in R \;\;\; \exists \mathbf{z} : f(\mathbf{d}, \mathbf{z}, \boldsymbol{\theta}) \le 0$$

Figure 2.3 shows the stochastic flexibility of a design with two uncertain parameters corresponding to the shaded area of the constraints.

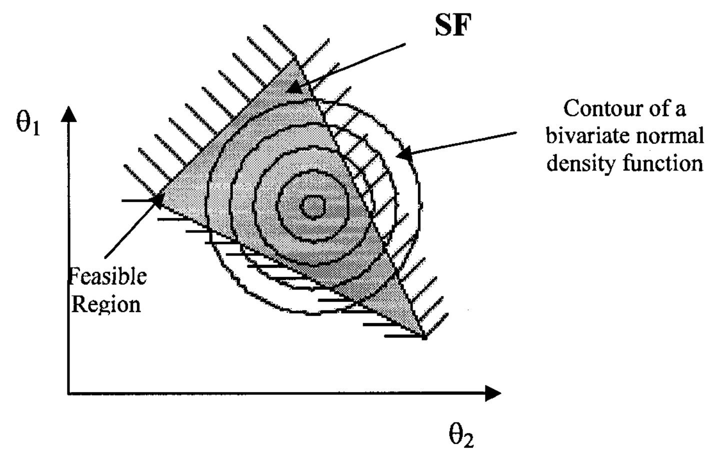

Figure 2.3 Stochastic flexibility of a design

The major difficulty of the decomposition approach arises from the fact that the feasible region is unknown and the integrand of the integral in Equation 2.38 is only implicitly defined when the optimization in the operating stage is performed. To circumvent these difficulties Pistikopoulos and Ierapetritou (1995) made use of the principles of the Generalized Benders Decomposition algorithm selecting the vector of design variables, **d**, as "complicating variables". Then the multiple integral for the expected profit evaluation was approximated through a Gaussian quadrature formula with unknown quadrature points. These unknown quadrature points were determined as a part of the optimization through the solution of a sequence of feasibility sub-problems. With this procedure, the boundary of the feasible region was properly defined. As a result, infeasibilities were avoided and the estimation of the expected profit/cost could effectively be performed.

A disadvantage of the decomposition approach proposed by Pistikopoulos and Ierapetritou (1995) is that the issue of partial feasibility (e.g. partial demand satisfaction, partially fulfilled market demands) was not addressed by their formulation. Permanent feasibility is the common assumption used for process design because it can be easily represented mathematically; however, from the process design point of view, permanent feasibility may not represent a realistic assumption (Pai and Hughes, 1987). Permanent feasibility is equivalent to defining an infinite cost for every pair of $(\mathbf{d},\boldsymbol{\theta})$ that yields no solution for Equation 2.38. This implies that an infinite penalty or loss is incurred whenever the plant becomes inoperable. Of course this is not always true. For instance, a chemical plant failing to meet product specifications or unable to satisfy market demands does not become inoperable leading to an infinite loss. Out-of-specification products can always be reprocessed or sold at a lesser price, capacity of the plant can be increased or damaged equipment can be replaced in time. Therefore, partial feasibility does need to be addressed in process design problems.

On the other hand, when strict safety considerations have to be followed in a design, permanent feasibility may also need to be incorporated. For instance, it may be necessary that for safety the reactor temperature does not surpass the upper limit established by the designer, or that the concentration of toxic component be strictly controlled within the limits established by environmental regulations. These two perspectives of feasibility have led to classifying the types of constraints to be incorporated in a design. Wellons and Reklaitis (1989), classified the types of constraints involved in a design into "hard" constraints, those that must always be satisfied, and "soft" constraints, those that can be violated for some realizations of the uncertain parameters.

Pai and Hughes (1987) suggested to assign a large but finite value to a lost function, $L(\mathbf{d},\boldsymbol{\theta})$ whenever the operating stage on Equation 2.38 yields an infeasible solution. Bernardo *et al.* (2001) proposed to treat quality constraints as "soft" and used the Taguchi approach to quality engineering to quantify the losses when quality characteristics deviate from their target values.

Bernardo *et al.* (2001) used the Taguchi's perspective to quality to differentiate between "hard" and "soft" constraints. Consider the quality variable, y, in the interval $y^L \leq y \leq y^U$ depicted in Figure 2.4; while the hard constraint perspective is binary, penalizing the objective function with an infinite value when y falls outside its boundaries or assigning a value of zero penalty when y falls inside the feasible region, the Taguchi's perspective, on the other hand, is continuous, penalizing whenever the quality value departs from its target value, τ, where the penalty is zero.

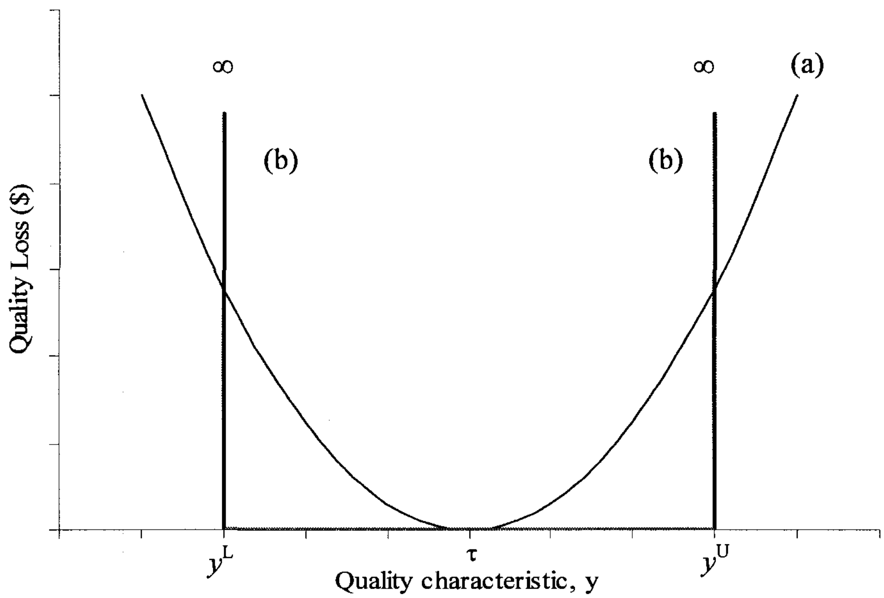

Figure 2.4 Quality loss models according to (a) Taguchi's loss perspective and (b) hard constraint perspective (Bernardo *et al.*, 2001)

To account for process robustness and quality cost, Bernardo *et al.* (2001) proposed an optimization formulation in which a penalty term, defined by the Taguchi loss function, was used to penalize the infeasibilities that may be encounter at the operating stage when quality constraints were not met. It was assumed that hard constraints were not sources of infeasibilities (design variables do not restrict the feasible region) and that for any realization of the uncertain parameters, optimal values for the control variables were available. Under these assumptions the integral in Equation 2.38 becomes total uncertainty region, Θ, so the objective function can be formulated using the expectancy operator, E_Θ, The process design under uncertainty involving quality is represented by the following stochastic optimization problem

Design stage:

$$\min \int_{\Theta} C^*(\mathbf{d},\boldsymbol{\theta}) \;\; j(\boldsymbol{\theta}) \;\; d\boldsymbol{\theta} = E_{\Theta}\{C^*(\mathbf{d},\boldsymbol{\theta})\}$$

$$\text{w.r.t.} \quad \mathbf{d}$$

$$\mathbf{d} \in D, \quad \boldsymbol{\theta} \in \Theta$$

Operating stage:

$$C^*(\mathbf{d},\boldsymbol{\theta}) = \min C(\mathbf{d},\mathbf{z},y_q,\boldsymbol{\theta}) + \sum_{q=1}^{Q} L_q(y_q,\tau_q) \tag{2.41}$$

$$\text{w.r.t.} \quad \mathbf{z}, y_q$$

$$\text{s.t.} \quad \mathbf{g}(\mathbf{d},\mathbf{z},\boldsymbol{\theta}) \le 0$$

$$\mathbf{z} \in Z, \quad y_q \in Y, \quad q = 1,2,...,Q$$

where Q is the total number of quality variables, L_q is the loss function and τ_q is the target value for quality variables q=1,2,.. Q.

Different approximation techniques have been proposed to estimate the expected value of the objective function in Equation 2.41, these defer in the integration techniques used. In the next section, these integration techniques are presented.

2.3 Integration techniques for stochastic optimization algorithms

Numerical integration is a study of the methods that seek a numerical value of an integral (Davis and Rabinowitz, 1975). It can be both simple and exceedingly difficult, so it requires from simplistic methods to more complicated ones. The difficulty lies in two factors. It may require of considerable computing time, sometimes leading to impractical ones, and it may require an in depth analysis for which special methods may not be currently available. For the solution of integrals involving joint PDF's, two Integration techniques can be stated: integration formulas and sampling techniques.

2.3.1 Integration formulas

Numerical integration by integration formulas is the method of approximate integration, wherein an integral is approximated by a linear (or nonlinear) combination of the values of the integrand (Davis and Rabinowitz, 1975). Common integration formulas used are Gaussian quadratures (Carnahan *et al.*, 1969) and Gaussian cubatures (Stroud, 1971; Engels, 1980).

2.3.1.1 Gaussian quadratures

The Integration through Gaussian quadratures seeks the best numerical estimate of the objective function by choosing optimal locations, $\mathbf{u}_i$, with their corresponding weighting coefficients, w_i, at which the function $f(\mathbf{u}_i)$ is evaluated. The locations (abscissa) are not evenly spaced but are chosen so that the sum of the weighted functional values in the integral yields an integral exactly when the function $f(\mathbf{u}_i)$ is a polynomial of degree 2n+1 or less (Carnahan *et al.*, 1969). Gaussian quadratures differ in the type of polynomial used for approximation and the type of integrand. For instance, Gauss-Legrendre, Gauss-Laguerre, Gauss-Chebyshev, and Gauss-Hermite are quadratures derived from the Legrendre, Laguerre, Chebyshev, and Hermite orthogonal polynomials.

A generalized formula of Gauss-type can be represented as

$$\int_a^b \int_a^b \int_a^b \ldots f(u)\varpi(u)du_1du_2\ldots du_n \doteq \sum_{i=1}^{N_p} w_i f(\mathbf{u}_i)\varpi(\mathbf{u}_i) \qquad (2.42)$$

where f is a function of the vector of independent variables $\mathbf{u}$, and ϖ is a weight function. These quadrature formulas have a total of N_p points $\mathbf{u}_i$ where w_i is the product of the weight factors for single dimensions. When uncertain parameters follow a normal PDF, denoted by $\boldsymbol{\theta}\sim N(\mu,\Sigma)$, the expected value of their function can be represented by the following integral.

$$E[f(\boldsymbol{\theta})] = \int_{-\infty}^{+\infty} f(\boldsymbol{\theta}) j(\boldsymbol{\theta}) d\boldsymbol{\theta} \tag{2.43}$$

where $j(\boldsymbol{\theta})$ is the joint probability density function of the uncertain parameters $\boldsymbol{\theta}$ with vector of means $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ described by

$$j(\boldsymbol{\theta}) = \frac{1}{2\pi^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2} (\boldsymbol{\theta} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta} - \boldsymbol{\mu}) \right] \tag{2.44}$$

One can envision the similarities between the Equation 2.42 and Equation 2.43 since one can consider $j(\boldsymbol{\theta})$ the weight function of $f(\boldsymbol{\theta})$.

The common Gauss formula used to estimate Equation 2.43 is the Gauss-Legrendre quadratue. However, because the expected value of a function of uncertain parameters corresponds to an indefinite integral, this quadrature cannot to be directly applied. Instead, the uncertainty region has to be truncated. This is commonly accomplished by using $\pm$ 3σ limits of the parameter distributions. Moreover, because an interval [-1, 1] in each integral dimension is used to integrate a n-dimensional cube, a suitable transformation of the form

$$\theta_m(u_m) = 0.5[\theta_m^U(1 + u_m) + \theta_m^L(1 - u_m)] \tag{2.45}$$

has to be applied in order to bound the uncertainty space as $\Theta = \{\boldsymbol{\theta}: \theta_m^L \leq \theta_m \leq \theta_m^U,\ m = 1,2,..\ n\}$.

Substituting this transformation in Equation 2.42 results in the Gaus-Legrendre formula for the calculation of the expected value of functions of normally distributed parameters

$$E[f(\boldsymbol{\theta})] \doteq \prod_{m=1}^{n} \frac{\theta_m^U - \theta_m^L}{2} \sum_{i=1}^{N_p} w_i \theta(\mathbf{u}_i) j[\theta(\mathbf{u}_i)] \tag{2.46}$$

Gauss-Legrendre quadratures have been successfully applied to problems involving uncertainties expressed as normal PDF's (Pistikopoulos and Ierapetritou, 1995; Straub and

Grossman, 1990) A disadvantage though, is that the number of function evaluations increases exponentially with the number of uncertain parameters. As a result, quadratures are usually applied to problems involving no more than two or three uncertain parameters. Moreover, due to the fact that an integral need to be definite for the Gauss-Legrendre quadrature to be applied, the distribution of the parameters has to be truncated, decreasing the accuracy on the estimates of the expected objective function.

2.3.1.2 Gaussian Cubatures

Gaussian cubatures (also called cubatures) are generalization formulas, derived from the principles of one-dimensional quadratures, for multidimensional integration (Bernardo *et al.*, 1999). Numerical integration through cubatures attempts to approximate a multidimensional definite integral by using information about the integrand only at a set of discrete points, $\mathbf{u}_i$, where the integrand is defined (Stroud, 1973).

Since cubatures are applied to integrals of higher dimension, a notation in Euclidian space, E_n, will be used to denote the region of integration, R_n, of the n-dimensional integral. A general formula for cubatures is of the form

$$\int_{R_n} f(\mathbf{u})\varpi(\mathbf{u})d\mathbf{u} \doteq \sum_{i=1}^{N_p} w_i f(\mathbf{u}_i) \qquad (2.47)$$

where again, f is a scalar function of the vector of independent variables $\mathbf{u}$, ϖ a weight function. Like quadratures, cubatures have N_p points $\mathbf{u}_i$, with corresponding weights w_i with each point $\mathbf{u}_i$ lying inside E_n.

Cubatures and quadratures differ in two aspects: accuracy and computational time used in function evaluations. Though cubatures require less number of function evaluations than quadratures, they not always reach the same level of accuracy achieved by quadratures. On the other hand, the number of function evaluations for quadratures increases exponentially with the integral dimension whereas, for cubatures, it increases polynomially.

An advantage of cubatures is that these do not need to be truncated in order to integrate uncertain parameters that follow a certain distribution. However, cubatures require of a transformation to map the integration region, R_n, to the uncertainty space, Θ (Bernardo *et al.*, 1999). This is done through a transformation of the form (Davis and Rabinowitz, 1980)

$$\theta = \phi(\mathbf{u}) \tag{2.48}$$

where $\mathbf{u} \in R_n$ and $\boldsymbol{\theta} \in \Theta$. With this transformation each $\mathbf{u}_i$ in R_n is mapped into $\boldsymbol{\theta}_i$, in Θ. The weight factor, w_i, is obtained by an additional transformation of the form

$$w_i^* = w_i \left|J(\mathbf{u}_i)\right| \quad i = 1, 2 \ldots N_p \tag{2.49}$$

where $J(\mathbf{u}_i)$ is the jacobian matrix of the function $\phi(\mathbf{u})$ evaluated at the point $\mathbf{u}_i$.

Based on the rules reported by Stroud (1971), Bernardo *et al.* (1999) constructed specialized cubatures (SC) suitable for estimated expected values of objective functions of uncertain parameters where uncertainty was described by a normal PDF. These cubatures made use of a weight function of the form

$$w_c(\mathbf{u}) = \exp(-\mathbf{u}^T\mathbf{u}) \tag{2.50}$$

to integrate the n-dimensional region in E_n. In order to integrate over the uncertainty space $\Theta=\{\boldsymbol{\theta}:\boldsymbol{\theta}\sim N(\mu,\Sigma)\}$, the following transformation was used

$$\theta(\mathbf{u}) = \boldsymbol{\mu} + \mathbf{I}_{\sqrt{2}}\Sigma^{1/2}\mathbf{u} \tag{2.51}$$

where $\mathbf{I}_{\sqrt{2}}$ is the diagonal matrix with all the diagonal elements equal to $\sqrt{2}$.

Applying this transformation to Equation 2.43, the expected value of the function is given by

$$E_\Theta[f(\boldsymbol{\theta})] = \frac{1}{\pi^{n/2}} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f[\theta(\mathbf{u})]\, w_c(\mathbf{u}) d\mathbf{u} \doteq \frac{1}{\pi^{n/2}} = \sum_{i=1}^{Np} B_i f(\theta(\mathbf{u}_i))] \tag{2.52}$$

Three types of SC's, proposed by Bernardo *et al.* (1999), were used throughout this work and are presented in the appendix of Chapter 4. These were constructed specifically to integrate normally distributed uncertainties and differ in the degrees of exactness. According to Stroud (1971) "A cubature formula has degree of exactness, *d*, if it is exact for all polynomials $u_1, u_2, \ldots u_n$ of degree $\leq d$ and there is at least one polynomial of degree $d + 1$ for which it is not exact".

Applying any the integration formula to Equation 2.41, the formulation the two-stage stochastic optimization becomes as a single-stage stochastic optimization of the form

$$
\begin{aligned}
&\min \sum_{i=1}^{N_P} w_i C(\mathbf{d},\mathbf{z}_i,y_{i,q},\tau_q,\boldsymbol{\theta}_i) \\
&\text{w.r.t.} \quad \mathbf{d},\mathbf{z}_i,y_{i,q} \\
&\text{s.t.} \\
&\qquad C(\mathbf{d},\mathbf{z}_i,y_{i,q},\tau,\boldsymbol{\theta}_i) = C(\mathbf{d},\mathbf{z}_i,y_{i,q},\boldsymbol{\theta}_i) + \sum_{k=q}^{Q} L_q(y_{i,q},\tau_q) \qquad (2.53) \\
&\qquad \mathbf{g}(\mathbf{d},\mathbf{z}_i,\boldsymbol{\theta}_i) \leq 0 \\
&\qquad \mathbf{d} \in D, \quad \boldsymbol{\theta}_i \in \Theta, \quad \mathbf{z}_i \in Z, \quad y_{i,q} \in Y \\
&\qquad i = 1,2, \ \ldots., \ N_p; \qquad q = 1,2, \ldots, Q
\end{aligned}
$$

2.3.2 Sampling techniques

Sampling techniques are part of the approximation methods for the calculation of integrals. They are based on statistical methods using random numbers. Among these methods, Monte Carlo and Quasi-Monte-Carlo (also called number-theoretic) can be mentioned. These two techniques are relevant for the approximate calculation of multi dimensional integrals, especially with large dimensions, for complicated regions and those cases where only a rough approximation is required (Engels, 1980).

2.3.2.1 Monte Carlo Method

The classical Monte Carlo method approximates a multiple integral of the form (Stroud, 1971).

$$I(f) \doteq \int .. \int f(u_m,..u_n)du_m,...du_n \tag{2.54}$$

For a specified number of N_p function evaluations, N_p points are chosen at random, uniformly distributed in the integration region R_n. The integral is then estimated by

$$I(f) \doteq \frac{V}{N_p}\sum_{i=1}^{Np} f(\mathbf{u}_i) \tag{2.55}$$

where $V \cong I(1)$ is the n-dimensional volume of R_n. If u_m are regarded as independent random variables then I(f) is a random variable with mean, $I(f)$ and standard deviation $\sigma_f / \sqrt{N_p}$, where

$$\sigma_f^2 = V[I(f^2)] - [I(f)]^2 \tag{2.56}$$

Therefore, the standard deviation is taken as the measure of error to be expected in the estimate in Equation 2.68.

Two distinct sampling techniques that fall into the category of the Monte Carlo methods are Monte Carlo Sampling (MCS) and Latin Hypercube Sampling (LHS). Next these two techniques are presented.

2.3.2.1.1 Monte Carlo Sampling (MCS)

MCS, as its name indicates, is based on the crude Monte Carlo method. It relies on the method for generating random real-valued numbers. MCS selects a random number that lies in an interval in an n-dimensional design space, $[u^L, u^U b]^n$, in which the sample site is an ordered n-tuple.

In contrast to quadratures, MCS does not suffer from the exponential time-increase in evaluating the function, *f*, when a high number of uncertain parameters is involved. However, its accuracy increases only as the square root of N_p, the number of sample points, increases. If the accuracy requirement is modest, or if the allowable computational time is large, then this sampling technique is highly recommended (Press *et al.*, 1992).

On the other hand, while MCS is simple to implement, the sample procedure will often leave large regions of the uncertainty space unexplored. This is the result of the random and independent nature of the samples generated by the random number generator (Guinta *et al.*, 2003). Because the effectiveness, in terms of accuracy, of a sampling technique to estimate an integral relies on the equidistribution properties of the sets of points at which the integrand values are computed and not on the randomness of the sampling procedure, other methods have been suggested. A stratified sampling procedure is a method that provides a more uniform sampling of the uncertainty space as compared to the crude MCS.

2.3.2.1.2 Latin Hypercube Sampling (LHS)

LHS is a stratified sampling technique that divides the distribution of uncertain parameters into strata then applies random sampling within each stratum. The stratification procedure is as follows. First, the sample space of each uncertain parameter is divided in blocks of equal probability, and then, N_p values are obtained by sampling separately each block rather than the entire distribution. Samples are taken at random from within each interval and then paired in a random manner with the N_p values of the second uncertain parameter. These N_p pairs are combined with N_p values of the third uncertain parameter and so forth to form N_p n-tuplets.

An advantage of LHS is that it guarantees all the sample values are representative of the entire range of the distribution. Figure 2.4 is an example of a two-dimensional case. The sample placement for LHS guarantees that only one stratum is selected in each row and column. For N_p=10 there are 10 partitions in both u_1 and u_2. This leads to a total of 100 strata in which each sample is randomly drawn. The points in Figure 3.2 represent 10 samples in this example, where each sample is randomly located in its stratum.

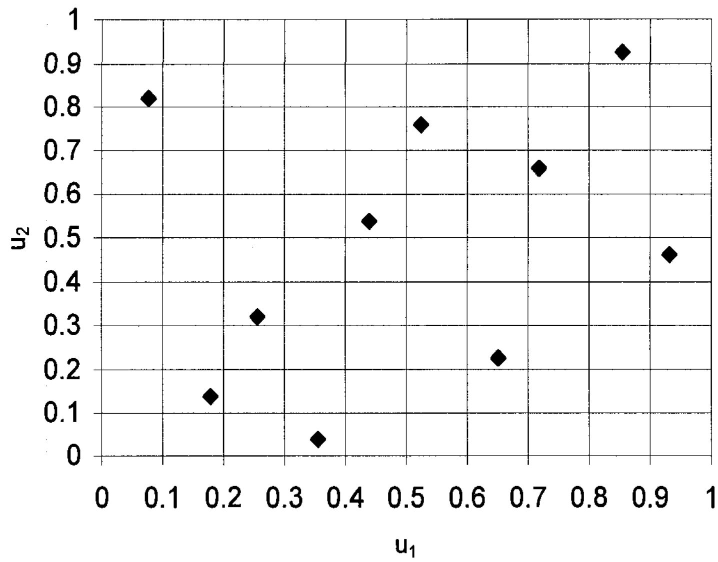

Figure 2.4 LHS with ten samples drawn from the joint PDF of two uncertain parameters u_1 and u_2

2.3.2.2 Quasi-Monte-Carlo Methods

The Quasi-Monte-Carlo method is a generalization of the crude Monte Carlo method. It alleviates the dependence on the number of samples in order to increase the accuracy in the estimation of integrals. In the Quasi-Monte-Carlo method the quantity $\sigma_f / \sqrt{N_p}$ does not decrease very rapidly as N_p increases (Stroud, 1971). In contrast, the error of approximation, $\varepsilon(f)$, in Equation 2.69 is of the form

$$I(f) \doteq \frac{V}{N_p} \sum_{i=1}^{Np} f(u_i) + \varepsilon(f) \qquad (2.57)$$

and is decreased by finding suitable points $(\mathbf{u}_1, \ldots \mathbf{u}_{Np})$ with uniform distribution.

Because methods for choosing the $\mathbf{u}_{Np}$ points are base on the theory of numbers, this approximation method is also called number-theoretic method. One method of choosing the $\mathbf{u}_{Np}$ points is through the discrepancy of the points $\mathbf{u}_i$, i=1,..., N_p. Discrepancy is related to the error in Equation 2.71 through the *Koksma-Hlawka* inequality (Hlawka, 1961). Discrepancy refers to a quantitative measure of how much the distribution of the points ($\mathbf{u}_1$,...$\mathbf{u}_{Np}$) deviates from the ideal uniform distribution. Different sampling techniques exist that are based on low-discrepancy sequence of the points ($\mathbf{u}_1$, ...$\mathbf{u}_{Np}$). In the next section, sampling techniques based on low-discrepancy sequences used in this work are presented.

2.3.2.2.1 Hammersley Sequence Sampling (HSS)

HSS is based on the Hammersley sequence (Hammersley, 1960). The Hammersley sequence is a generalization of the *van der Corput* (1935) sequence in base two for higher dimensions sine it makes use other bases of prime numbers. The sequence starts from zero and takes values in the interval [0,1).

The Hammersley sequence is constructed based on specific positive integers, p, written in radix-R notation as follows:

$$p \equiv p_m p_{m-1} \dots p_2 p_1 p_0 = p_0 + p_1 R + p_2 R^2 + \dots + p_m R^m \qquad (2.58)$$

where m = [$\log_R p$]=[(ln p)/(ln R)], and the square brackets denote the integral part. A unique fraction between 0 and 1, called the inverse radix number, can be constructed by reversing the order of the digits of n about the decimal point as follows:

$$\phi_R(p) = 0.p_0 p_1 p_2 \dots p_m = p_0 R^{-1} + p_1 R^{-2} + \dots + p_m R^{-m-1} \qquad (2.59)$$

The Hammersley points in n-dimensions is given by the following sequence

$$z_n(p)=\left(\frac{p}{N_p},\phi_{R_1}(p),\phi_{R_2}(p),\ldots,\phi_{R_{n-1}}(p)\right)\quad p=1,2,\ldots,N_p \tag{2.60}$$

where R_1, R_2, ..., R_{n-1} are the first n-1 prime numbers (2,3,5,7,11,13,17,....). With this algorithm N_p points are in the n-dimensional design space $[0,1]^n$.

2.3.2.2.2 Halton Sequence Sampling (HalSS)

HalSS is based on the Halton sequence (Halton, 1960). Like HSS, it is based on the generalization of van der Corput sequence for higher dimensions. It also makes use of different prime bases for each dimension of the sequence. However, a difference from HSS and HalSS uses the van der Corput sequence base two for the first dimension. Therefore, the formula in Equation 2.74 can be modified to express the Halton Sequence as

$$z_n(p)=\left(\phi_{R_1}(p),\phi_{R_2}(p),\ldots,\phi_{R_{n-1}}(p)\right)\quad p=1,2,\ldots,N_p \tag{2.61}$$

where R_1, R_2, ..., R_{n-1} are the first n-1 prime numbers (2,3,5,7,11,13,17,....). With this algorithm N_p points are in the n-dimensional design space $[0,1]^n$.

2.3.2.2.3Faure Sequence Sampling (FSS)

FSS is based on the Faure sequence (Faure, 1982) and is derived from linking low discrepancy theory and combinatorial theory for vector reordering. FSS differs from HSS and HalSS in that it only uses one base for the construction of sequences of all n-dimensions and that the elements of the sequence of each dimension are permuted. In addition, the base used in the Faure sequence is the smallest prime number that is larger than or equal to the number of n-dimensions. For instance if n=50, in the Halton sequence, the prime number used as base for the 50th dimension is 229 whereas in the Faure sequence, the base prime number for the 50th dimension is 53. Since the last base in Faure is lower than that in the Halton sequence, the computational burden to generate the samples is less in the Faure

Sequence Sampling. An advantage of the Faure sequence is that the reordering prevents correlation problems for sequential high-dimensions as occurred with Halton and Hammersley sequence.

2.3.2.2.4 Sobol Sequence Sampling (SSS)

SSS is based on the Sobol sequence (Sobol, 1972). Like the FSS, SSS has the same base for all dimensions and the element vectors are reordered within each dimension (Bratley and Fox, 1988). Contrary to FSS, SSS uses the van der Corput base 2 for all dimensions. That is, the sequence is generated such that the first 2^m terms of each dimension for m=0,1,2,... are a permutation of the corresponding terms of the van der Corput base 2. For a more detailed discussion of the construction of the Sobol sequence, the reader is referred to the works of Bratley and Fox, (1998) and Press and Teukolsky (1989). Bratley and Fox (1998) describe a Fortran implementation of a Sobol sequence generator and compare it in some detail to the generator suggested by Faure. Press and Teukolsky (1989) discuss the Sobol's sequence and computational methods for generating them.

Considerably work has focused on demonstrating which type of sequence performs better. However, a definitive answer is not available. While Paskov (1994) and Papageorgiou and Traub (1996) argue that low-discrepancy sequences are more efficient than random sequences, Bratley, Fox and Niederreiter (1992) conclude that rather, random sequences are more efficient than low-discrepancy sequences for dimension higher than twelve. In regards to low-discrepancy sequences, some low-discrepancy sequences may be more efficient than others; this will depend on the type of function being integrated and the type of uncertainties under consideration. Part of this work will be devoted to investigate which type of low-discrepancy sequence performs better in the integration of objective functions involving normally distributed uncertainties.

For sampling techniques, all sampling points possess the same weight, $w_i=1/N_p$. Therefore, we can express Equation 2.41 as

$$
\begin{aligned}
&\min \quad \frac{1}{N_p}\sum_{i=1}^{N_P} C(\mathbf{d},\mathbf{z}_i,y_{i,q},\tau_q,\boldsymbol{\theta}_i)\\
&\text{w.r.t.} \qquad \mathbf{d},\mathbf{z}_i,y_{i,q}\\
&\text{s.t.}\\
&\qquad C(\mathbf{d},\mathbf{z}_i,y_{i,q},\tau_q,\boldsymbol{\theta}_i) = C(\mathbf{d},\mathbf{z}_i,y_{i,q},\boldsymbol{\theta}_i) + \sum_{q=1}^{Q} L_q(y_{i,q},\tau_q) \qquad (2.62)\\
&\qquad \mathbf{g}(\mathbf{d},\mathbf{z}_i,\boldsymbol{\theta}_i) \le 0\\
&\qquad \mathbf{d}\in D, \quad \mathbf{z}_i \in Z, \quad y_{i,q}\in Y \quad \boldsymbol{\theta}_i \in \Theta,\\
&\qquad i = 1,2,\ \ldots,\ N_p; \qquad q = 1,2,\ldots,Q
\end{aligned}
$$

2.4 Summary

Quality costs are of paramount importance for the profitability of a process. Market demands are driven by economic competition. Therefore quality issues need to be addressed in order to ensure the client's satisfaction and product sells. By including quality related issues such as robustness in quality variables at the initial design of a process, additional costs such as the costs for out-of-specification product, rework or scrap can be anticipated. In this work, robust techniques are studied and implemented in process design algorithms using rigorous mathematical programming. The overall objective is to design chemical processes that are robust and economically optimal.

The design of a chemical process under considerable uncertainty (e.g., uncertain market demands, uncertain prices) and variability (e.g. intrinsic dynamic response of process system, variability in quality of raw material) is not an easy task. This may be the reason for which optimization under a deterministic paradigm prevailed in the past. Design under uncertainty requires of specialized algorithms that allows modeling the uncertainty using mathematical expressions. One practical approach is to apply statistical techniques. Information about the uncertainty of model parameters and variables can be extracted from statistical analysis of experimental and/or process data, so a probabilistic approach to modelling uncertainty seems to be quite suitable. With the advent of fast computers considerable advances in the development of specialized algorithm have been forthcoming making possible the incorporation of uncertainty in process modelling and design. However, these algorithms still require of considerable computational time to perform the optimization. This is due to the fact that the optimization procedure involves and implicit integration of an objective function. In this work the objective function is that which minimizes capital, operating and quality costs. Because current integration techniques still demand considerable computational effort, an additional objective of this work is to devise integration techniques that can reduce the number of functions evaluations demanded by optimization algorithms in addition to providing a measure of the uncertainty of the result.

Nomenclature

d	Vector of design variables
z	Vector of control variables
z	Vector of control variables as defined by Taguchi
x	Vector of state variables
y	Vector of quality variables
h	Vector of equality constraints
g, f	Vector of inequality constraints
u	Vector of independent variables
μ	Vector of uncertain parameters
f	Scalar function
C	Scalar cost function
y	Quality variable
k	Taguchi loss constant
N_p	Number of sampling points
Qv	Number of quality variables
n	Number of uncertain parameters
j	Joint probability density function distribution
E, E_Θ	Expectancy operator; Expectancy operator over uncertainty region
E_n	Euclidian space with dimension n

L	Scalar quality loss/penalty cost function
w_i	Weighting coefficient at sampling point i.
$R_n(\mathbf{d})$,	Feasible region
R_n	Integration region
$J(\mathbf{u}_i)$	Jacobian matrix of the function $\phi(\mathbf{u}_i)$
$\mathbf{I}_{\sqrt{2}}$	Diagonal matrix with all elements equal to $\sqrt{2}$
B_i,	Weighting coefficients for cubatures
d	Degree of exactness for cubatures
V	n-dimensional volume of R_n
p	Integer number
R	Number in radix notation
$Z_n(p)$	Hammersley/Halton sequence in n dimension
m	Integer part of $\log_R p$=[(ln p)/(ln R)]
$I(f)$	Integral

Abbreviations

S/N	Signal-to-noise ratio
SF	Stochastic flexibility
MSE(z)	Mean squared error
FS	Fully symmetric permutation

Greek letters

$\boldsymbol{\theta}_i$	Vector of uncertain parameters
Θ	Domain of the uncertain parameters
τ, τ_0	Target value for the quality variable
σ_y	Standard deviation of quality variable
σ_y^2	Variance of quality variable
μ_y	Mean of quality variable

μ_z	Mean value of control variable
Σ	Variance-covariance matrix of uncertain parameters
ξ	Square of inverse of coefficient of variation
υ	Coefficient of variation
ρ	Radius of the hyper sphere describing the operating region
λ	Lagrange multiplayer of equality constraints
π	Boundary of Lagrange multiplayer
ϖ	Weight function
ϕ	Transform function
ε	Error of integration
$\phi_R(p)$	Inverse radix of p

Set theory

$\forall$	For all
$\in$	Element of
$\exists$	There exist
D	Set of design variables
X	Set of state variables
Z	Set of control variables
Y	Set of quality variables
R	Domain of real numbers
J	Set of constraint

Optimization acronyms

min	Minimize
w.r.t	With respect to
s.t.	Subject to

Subscripts

i	Sample point i
j	Constraint function j
m	Uncertain parameter m
q	Quality variable q

Superscripts

^{U}	Upper bound
^{L}	Lower bound

References

Antonov, I. A., and Saleev, V. M., "An economic method of computing LP_T-sequences", USSSR Comput. Math. Phys., 9, 252-256, 1979.

Acevedo, J., and Pistikopoulos, E. N., "Stochastic optimization based algorithms for process synthesis under uncertainty", Computers and Chemical Engineering, Vol. 22, 647-671, 1998.

Acevedo, J., and Pistikopoulos, E. N., "Parametric MINLP algorithm for process synthesis problems under uncertainty", Industrial and Engineering Chemistry Research, Vol. 35, 147, 1996.

Acevedo, J., Pistikopoulos, E. N. "A multiparametric programming approach for linear process engineering problems under uncertainty", Industrial and Engineering Chemistry Research, Vol. 36, 717, 1997.

Bansal, V., Perkins, J. D., and. Pistikopoulos E. N., "Flexibility analysis and design using a parametric programming framework", AIChE Journal, Vol. 48, No. 12, 2002.

Bernardo, F., P., Pistikopoulos E. N. and Saraiva P. M., "Quality cost and robustness criteria in chemical process design optimization", Computers and Chemical Engineering, Vol. 21, 27-40, 2001.

Bernardo F., P., Pistikopoulos E., N. and Saraiva P. M., "Integration and computational issues in stochastic design and planning optimization problems", Industrial and Engineering Chemistry Research, Vol. 38, 3056, 1999.

Biegler, L. T., Grossmann, I. E., and, Westerberg, W. "Systematic methods of chemical process design", Prentice Hall, New Jersey, 1997.

Boudriga, S., "Evaluation of parameter design methodologies for the design of chemical process units", M.Sc. Thesis, University of Ottawa, 1990.

Bratley, P, Fox, B. L.., and Niederreiter, H., "Implementation and tests of low-discrepancy sequences", ACM Transactions on Mathematical Software", Vol. 2, No. 3, 195-213, 1992

Bratley, P., and Fox, B., "Algorithm 659: Implementing Sobol's quasirandom sequence generator", ACM Transactions on Mathematical Software, Vol. 14, No. 1, 88-100, 1988.

Carnahan, B., Luther, H. A., and Wilkes. J.O., "Applied numerical methods", Wiley, New York, 1969.

Copeland, K. A. F., and Nelson P. R., "Dual response optimization via direct function minimization", Journal of Quality Technology, Vol. 28, No. 3, 1996.

Davis, P. J., and Rabinowitz, P., "Methods of numerical integration", Academic Press, New York, 1975.

Del Castillo, E., and Montgomery D. C., "A nonlinear programming solution to the dual response problem", Journal of Quality Technology, Vol. 25, No. 3 1993.

Ding, R., Lin, D. J. J., and Wei, D., "Dual response surface optimization: A weighted MSE approach", Quality Engineering, Vol. 16, No.3, 377-385, 2004.

Diwekar, U. M., and Rubin, E. S., "Parameter design methodology for chemical processes using a simulator", Industrial Engineering Chemistry Research, 33, 292-298, 1994.

Diwekar, U. M., and Rubin, E. S., "Stochastic modeling of chemical processes", Computers and Chemical Engineering, Vol. 15, No. 2, 105-114, 1991.

Douglas, J.M., "Conceptual design of chemical processes", McGraw-Hill, 1988.

Engels, H., "Numerical quadrature and cubature", Academic Press: London, 1980.

Faure, H., "Discrepance de suites associees a un systeme de numeration (en dimension s)", Acta Arithmetica, Vol. 16, 337-351, 1982.

Galanti, S., and Jung A., "Low-discrepancy sequences: Monte Carlo simulation of option prices", Journal of Derivatives, 5, 1, 63, 1997.

Grossmann, I. E., and Sargent, R. W. H., "Optimum design of chemical plants with uncertain parameters", AIChE J., Vol. 37, 1978.

Grossmann, I. E., Halemane, K. P., and. Swaney, R. E., "Optimization strategies for flexible chemical processes", Computers and Chemical Engineering, Vol. 7, No. 4, 439-462, 1983.

Gryna F. M., Chua R. C., and DeFeo J. A., "Juran's quality planning and analysis", Fifth edition, McGraw Hill, New York, 2006.

Halemane, K. P., and Grossmann, I. E., "Optimal process design under uncertainty", AIChE J. Vol. 29 439, 1983.

Halton, J., "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals", Numerische Mathematik, Vol. 2, 84-90, 1960.

Hammersley, J., "Monte Carlo methods for solving multivariable problems", Proceedings of the New York Academy of Science, Vol. 86, 844-874, 1960.

Hlawka, E., "Functionen von beschranker variation in deter therie der gleichverteilung". Ann. Math. Pura Appl., No., 54, 1961.

Kackar, R N., "Off-line quality control, parameter design, and the Taguchi method", Journal of Quality Technology, Vol. 17, No.4, 1985.

Leon, R. V., Shoemaker, A. C., and Kackar, R. N. "Performance measures independent of adjustment", Technometrics, 29, 253, 1987.

Lin, D. K., and Tu W., "Dual response surface optimization", Journal of Quality Technology, Vol. 27, No.1, 1995.

Myers R. H., Khuri, A. I. and Vining, G. "Response surface alternatives to the Taguchi robust parameter design approach", The American Statistician, Vol. 46, No. 2, 1992.

Myers, R. H., "Response surface methodology", Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, Va., 1976.

Pertsinidis, A, Grossman, I. A., and McRae, G. J., "Parametric optimization of MIL programs and a framework for the parametric optimization of MINLPs", Computers and Chemical Engineering, Vol 22, supplement, S205, 1998.

Pai, C. C., and Hughes, R. R., "Strategies for formulating and solving two-stage problems for process design under uncertainty", Computers and Chemical Engineering, Vol. 11, 695-706, 1987.

Papageorgiou, A., and Traub, J. "Beating Monte Carlo", Risk, Vol. 9, No. 6, 63-65, 1996.

Paskov, S., "Computing high dimensional integrals with applications in finance", Technical Report CUCS-023-94, Department of Computer Science, Columbia University, 1994.

Pignatiello, J. J., and Ramberg, J. S., "Discussion (of Kackar, 1985)", Journal of Quality Technology, Vol. 17, No.4, 198-206, 1985.

Pistikopoulos, E. N., and Ierapetritou, M. G., "Novel approach for optimal process design under uncertainty", Computers and Chemical Engineering, Vol. 19, No. 10, 1089, 1995.

Press, W. H., Teukolsy, S. A., Vetterling W. T., and Flannery, B. P., "Numerical recipes in Fortran – The art of scientific computing", Second Edition, Cambridge University Press, 1992.

Rudd, D. F., and Watson, C., "Strategies of process engineering", John Wiley and Sons, New York, 1968.

Samsatli, N. J., Papageorigou, L. G., and Shah, N., "Robustness metrics for dynamic optimization models under parameter uncertainty", AIChE Journal, Vol. 44, No. 9, 1998.

Straub, D. A. and Grossman I. E., "Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties", Computers and Chemical Engineering, Vol. 14 No.9 pp. 967-985, 1990.

Stroud, A. H., "Approximate calculation of multiple integrals", Prentice Hall: London, 1980.

Sobol, I., and Levitan, YL, "The production of points uniformly distributed in a multidimensional cube", Preprint IPM Akad. Nauk SSSR, No. 40, Moscow, 1976.

Taguchi, G., and Wu, "Introduction to off-line quality control", Central Japan Quality Control Association (Available from American Supplier Institute, 3100 Detroit Industrial Expressway, Romulus, MI 48174, 1980

Taguchi, G. "Introduction to quality engineering", Asian productivity association, 1986.

van der Corput, J. "Verteilungsfunktionen I & II", Nederl. Akad. Wetensch. Proc., Vol. 38, 813-820, 1935.

Vining, G. G., and Myers R. H., "Combining Taguchi and response surface philosophies: A dual response approach". Journal of Quality Technology, Vol. 22, No.1, 1990.

Wellons, H.S. and Rekleitis, G. V., "The design of multiproduct batch plants under uncertainty with staged expansion", Computers and Chemical Engineering. Vol. 13, 115, 1996.

Chapter 3 (Paper 1)

Parameter Design through Stochastic Optimization: CSTR problem revisited

L. F. Dominguez and D. D. McLean

Dept. of Chemical Engineering

University of Ottawa

Abstract

This article provides a comparison of current formulations of Robust Parameter Design (RPD) within a stochastic optimization framework. A computer simulation is embedded in a nonlinear optimization procedure using stochastic descriptions of the noise variables. Based on the principles of RPD, minimization of quality cost by means of reducing variability in a quality characteristic is carried out. The potential advantages of the stochastic optimization approach are highlighted using the example of designing a robust continuous stirred tank reactor (CSTR). The performance of three sampling techniques was also examined for their potential use in the stochastic optimization algorithm. The sampling techniques included in this work were: Monte Carlo Sampling (MCS), Latin Hypercube Sampling (LHS) and Hammersley Sequence Sampling (HSS)

Key words: Robust parameter design; stochastic optimization; sampling techniques

3.1 Introduction

The continuous operation during the lifetime of a product or process highly depends on its quality. By the mid 1980's a Japanese quality expert, G. Taguchi, showed that the monetary losses due to poor quality could be reduced by minimizing the variability in quality characteristics and by keeping their means as close as possible to target values. Assuming that quality costs can be approximated by a quadratic function, he proposed the loss function

$$\mathrm{E}[\mathrm{L}] = k\,\mathrm{E}[y-\tau]^2 \tag{2.63}$$

to quantify quality losses. With this expression, the expected monetary losses, E[L], could be easily quantified by multiplying a quality constant, k by the expected squared deviations of the quality characteristic, y, from its target value, τ. The constant k depended on the type of process or product and on information about costs involved in its production.

It can be easily demonstrated (Ryan, 2000) that when the expected value of y, E[y], differs from its target, τ, Equation 3.1 takes the form of the mean squared error (MSE)

$$\mathrm{E}[\mathrm{L}] = (\mathrm{E}[y]-\tau)^2 + \sigma_y^2 \tag{2.64}$$

where the term $(\mathrm{E}[y]-\tau)^2$ is the square bias and σ_y^2 the variance of the quality characteristic.

Taguchi (1980) also proposed techniques of experimental design as well as techniques for the data analysis to improve the quality of products and processes. He called this methodology Robust Parameter Design (RPD). RPD is performed with the aim of identifying nominal settings of the variables that make a process more robust to the impact of input variations and thus reduce quality costs. In RPD, the variables affecting the process performance are classified as control variables, which can be controlled and whose nominal values can be specified by the designer, and noise variables which cannot be controlled, and whose values impart variation to the process during operation. In addition, the mean and variance of responses obtained from the experiments are combined into a single performance measure known as signal-to-noise ratio (SNR) which corresponds to the metric to be optimized.

RPD has been subject to extensive studies and many articles on this topic have appeared in the research literature since its appearance. Kackar (1985) introduced the concepts of off-line quality control and explained the basis of RPD. Despite the success of this methodology for quality improvement, criticisms arose due to the non-orthodox experimental designs and data analysis techniques employed (Nair, 1992). Many works have been published which address the different areas of RPD. Box (1988) showed that the SNR advocated by Taguchi is only valid when the variance of the quality characteristic is not a function of its mean and proposed the minimization of MSE of the quality characteristic. Previously, Leon *et al.* (1987) had introduced alternative performance metrics, Performance Measures Independent of Adjustment (PerMIA). They showed that PerMIA could be used even when the mean and variance are dependent.

In the area of experimental design, alternative approaches to the orthogonal arrays recommended by Taguchi have been proposed. Using Response Surface Methodology (RSM), Vining and Myers (1990) (referred to as VM) proposed a dual response approach. The variance of the quality characteristic was considered the primary response to be minimized subject to an acceptable value of the mean considered the secondary response. In addition, they suggested specifying the Lagrange multiplier, λ_1, in the range of $-\pi \leq \pi \leq \pi$ for the constraint that restricts the search area to a spherical region. Identifying that this range is too conservative and that it may lead to suboptimal solutions, Del Castillo and Montgomery (1993), proposed the use of the generalized reduced gradient algorithm (GRG) to perform the optimization. With use of GRG more constraints or response function could be considered in addition to allowing the designer to specify the radius of the spherical region of interest.

Lin and Tu (1995), denoted by LT, found that forcing the secondary response to a specified value may not lead to the minimum cost, and showed that, by allowing a little bias in the response, a substantial reduction in variability can be obtained and thus in quality costs. They proposed modeling the objective function using the MSE approach. Noting that there may be situations in which deviations from target are restricted, Copeland and Nelson (1996) modified this approach by adding a restriction to the squared bias term. Using the same concept, Kim and Lin (1998) employed a fuzzy optimization methodology and proposed a membership function to determine the degree to which the squared bias equals zero. More recently, Ding *et al.* (2004) proposed a weighted MSE approach to minimize cost. In this approach the components of the MSE were assigned fractional weights. Data-driven procedures to determine the weights were recommended.

The Taguchi approach and the other formulations are based on analysis of statistically designed physical experiments to determine the settings of design variables that optimize a quality performance measure or measures in order to produce relatively accurate estimates. These techniques require that the performance be estimated for each set of values, possibly replicated, according to the experimental design. The advantage is that the computational effort required to obtain such estimates of the response is not demanding. However, one has to confirm the validity of the predicted location of the solution. Therefore, there exists a potential for arriving to erroneous conclusions. Accordingly, Taguchi recommended performing additional runs to verify that the optimum has been attained.

Statistical techniques are a viable approach for approximating a response in which the relationship between input-output is unknown. For the design of chemical processes, however, physical experiments are not necessary. Process models exist that allows the design to be carried out. One can perform computer simulations to estimate the variability of a response when some information about the uncertainty in the noise variables is available.

Stochastic optimization deals with the inherent system noise (Sen, 2003). It also copes with nonlinear and high dimensional models. Stochastic optimization algorithms have become widely spread for their success in solving such problems. Multidimensional integration is present in stochastic optimization algorithms due to the need to calculate expectations or probabilities. Two distinct approaches are applied: integration formulas and sampling techniques. While integration formulas approximate the integral by finding optimal locations with corresponding weights to evaluate the integrand (Davis and Rabinowitz, 1975), sampling techniques approximate the integral by generating samples uniformly distributed over the integration region to evaluate the integrand (Press *et al.*, 1992).

Because the RPD problem consists of finding values of control variables that minimize the effect of noise, it makes it suitable for stochastic optimization formulations. This work presents a comparison of current formulations of RPD within a stochastic optimization framework in which computer simulations are combined with stochastic descriptions of the noise variables. Quality costs are minimized by means of reducing variability in quality characteristics. The potential advantages of the stochastic optimization approach are highlighted using an example drawn from the chemical engineering literature. In Section 3.2, different formulations for the RPD problem and their mathematical statements are presented. In Section 3.3 the stochastic approach for RPD is introduced. The design of a Continuous Stirred-Tank Reactor (CSTR), is presented in Section 3.4 to demonstrate how the main goals of RPD can be achieved with this formulation. In section 3.5 results and discussion are presented. Finally, in Section 3.6 conclusions are summarized.

3.2 The robust parameter design problem

Parameter design classifies the variables that influence the performance of a quality characteristic into two categories: *control* and *noise variables*. Control variables are those for which the designer can choose nominal values and which specify the process. Noise variables are those that impart variation to the quality characteristics during the process' life span or across different process units

Let $z = [\mu_{z_1}\, \mu_{z_2} \ldots \mu_{z_n}]$ be a vector of the nominal values of the n control variables and $\boldsymbol{\theta} = [\theta_1\, \theta_2 \ldots \theta_k]$ a vector of k noise variables. Also, let the quality characteristic, y, be a function of **z** and $\boldsymbol{\theta}$ such that $y = f(z, \boldsymbol{\theta})$. During process design, the designer specifies set of nominal values for the control variables, $z = [\mu_{z_1}\ \mu_{z_2} ... \mu_{z_n}]$, so that the quality characteristic is kept as close as possible to the most desirable or target value, τ (i.e, $f(\mu_{z_1}\ \mu_{z_2} ... \mu_{z_n}) = \tau$).

The variability of the noise variables is transmitted to a quality characteristic causing it to deviate from its nominal vale. Changes in the noise parameters are of a random nature, so if information about the randomness of these changes becomes available, the uncertainty of the noise variables can be characterized by specifying the parameter of their distribution. When the variability in the noise parameters is transmitted to the quality characteristic, it becomes a random variable and will follow a certain distribution. Accordingly, the probability distribution function of y, depends on three factors: the distribution of the noise parameters, $\boldsymbol{\theta}$, the distribution of the control variables, z, and the type of function representing y.

The designer may be interested in one of three design objectives regarding the quality characteristic of a process (a) Minimize the variance around a specific target value (b) Minimize its mean with a maximum allowable variance. (c) Maximize its mean with a maximum allowable variance. The first objective is the most common. Therefore, we shall focus on minimizing the variance of a quality characteristic around a target value.

In this study, the RPD problem was formulated using three approaches via stochastic optimization: the dual response of VM, the MSE of LT and the WMSE of Ding *et al.* (2004).

VM defined the RPD problem as that of finding the set of nominal values of the control variables, **z**, that minimize σ_y^2 subject to the constraint $\mu_y = \tau$. A mathematical statement of the problem is given by

$$\begin{aligned} &\min && \sigma_y^2 \\ &\text{w.r.t.} && z \\ &\text{s.t.} && \mu_y = \tau \\ & && z \in Z \end{aligned} \tag{2.65}$$

where μ_y is the mean of the quality characteristic, σ_y^2 is its variance, τ is the target value and Z is the set of acceptable nominal values for the control variables, z.

LT defined the RPD problem as that of minimizing the MSE. A mathematical representation of the problem is given by

$$\begin{array}{ll} \min & \text{MSE} \\ \text{w.r.t.} & z \\ \text{s.t.} & \text{MSE} = (\mu_y - \tau)^2 + \sigma_y^2 \\ & z \in Z \end{array} \tag{2.66}$$

Ding *et al.* (2004) proposed the minimization of a weighted version of the MSE of LT. The choice of the appropriate weight depended on the relative importance of the terms forming the MSE, the squared bias and the variance. The optimization scheme they proposed is given by

$$\begin{array}{ll} \min & \text{WMSE} \\ \text{w.r.t.} & z \\ \text{s.t.} & \text{WMSE} = \omega(\mu_y - \tau)^2 + (1-\omega)\sigma_y^2 \\ & z \in Z \qquad \omega \in [0,1] \end{array} \tag{2.67}$$

where ω is the weight used to perform the optimization. It should be noted that if ω=0, this formulation reduces to the criterion of VM without any constraint on μ_y. If ω=1, it reduces to solely keeping the mean on target without concern for σ_y^2. If ω=0.5, it reduces to the criterion of LT. As can be seen, all the above formulations obtain optimal settings for the design variables that minimize the quality cost assuming a Taguchi quadratic loss function.

Having defined some of the current optimization schemes for RPD, next the stochastic optimization approach is introduced.

3.3 Stochastic optimization approach to parameter design

Stochastic optimization is a branch of optimization which allows the treatments of uncertainties involved in model parameters or data (Ki-Joo and Diwekar, 2002). It is commonly called stochastic programming because the objective function and constraints are represented probabilistically (e.g., most likely values, expected values, variances or fractiles). Such an approach does not rely on an approximation of a response surface and is more useful when highly non-linear models are used in which the response is not very smooth (Diwekar and Rubin, 1994). A mathematical representation of the stochastic formulation is given by

$$\begin{aligned}
&\min P_1[C(z,\boldsymbol{\theta})] \\
&\text{w.r.t.} \quad z \\
&\text{s.t.} \\
&\quad P_2[\mathbf{h}(z,\boldsymbol{\theta})=0] \geq \alpha_1 \\
&\quad P_3[\mathbf{g}(z,\boldsymbol{\theta}) \leq 0] \geq \alpha_2 \\
&\quad z \in Z \;\; \boldsymbol{\theta} \in \Theta
\end{aligned} \tag{2.68}$$

where C is a scalar function, $\boldsymbol{\theta}$ is the vector of uncertain variables, z is the vector of control variables, P_1 is the probabilistic function to be optimized, and P_2 and P_3 are the probabilistic representation of the model equality and inequality constraints, **g** and **h**, respectively. Again, Z is the set of optimal nominal values for the control variables, z, and Θ is the set describing the uncertainty region of the noise variables. Finally, α_1 and α_2 are the minimum probabilities of satisfying the equality and inequality constraints under existing uncertainties (i.e., minimum stochastic flexibility, Straub and Grossman, 1990). If P_1 is considered to be the probability of the most likely value, the objective function is given by expected value of the function $C(z,\boldsymbol{\theta})$. On the other hand, if no probability for feasible operation (i.e, satisfaction of equality and inequality constraints) is specified, they must be treated explicitly. The stochastic formulation then reduces to

$$
\begin{aligned}
&\min \quad E[C(z,\boldsymbol{\theta})] \\
&\text{w.r.t.} \quad z \\
&\text{s.t.} \\
&\quad \mathbf{h}(z,\boldsymbol{\theta}) = 0 \\
&\quad \mathbf{g}(z,\boldsymbol{\theta}) \leq 0 \\
&\quad z \in Z \quad \boldsymbol{\theta} \in \Theta
\end{aligned}
\tag{2.69}
$$

where E represents the mathematical expectation with respect to the $\boldsymbol{\theta}$.

Since $C(z,\boldsymbol{\theta})$ is a function of the noise variables with probability density function $j(\boldsymbol{\theta})$, its expected value is given by the following integral

$$
E[C(z,\boldsymbol{\theta})] = \int_{-\infty}^{+\infty} C(z,\boldsymbol{\theta})\, j(\boldsymbol{\theta})\, d\boldsymbol{\theta} \tag{2.70}
$$

The main difficulty for solving Equation 3.7 is evaluating the function $C(z,\boldsymbol{\theta})$ and its expectation, $E[C(z,\boldsymbol{\theta})]$, when the uncertainty in the noise variables is propagated to the objective function.

One approach to solving this optimization problem is through Gaussian quadratures. The objective of Gaussian quadratures (Carnahan *et al.*, 1969) is to find the best numerical estimate of the integral by choosing optimal locations (abscissas), $\boldsymbol{\theta}_i$, with their corresponding weighting coefficients (weight factors), w_i, at which the function $f(\boldsymbol{\theta})$ is evaluated. Thus, with this discretization strategy, the expected value of the objective function in Equation 3.7 takes the form

$$
E[C(\mathbf{z},\boldsymbol{\theta})] = \sum_{i=1}^{N_p} w_i C_i(\mathbf{z}_i, \boldsymbol{\theta}_i) \tag{2.71}
$$

where N_p is the total number of location points used in the estimation.

A disadvantage of this integration technique is that the number of function evaluations increases exponentially with the number of uncertain parameters. For instance, to integrate numerically a function of two uncertain parameters using 10 quadrature points in each dimension would require $10^2 = 100$ points, and if five uncertain parameters were involved this would require 10^5 =10,000 points which would create an impractical computational effort in an optimization algorithm for which the integrations would be carried out many times before the optimum is found. In addition, since limits for integration are required, the distribution would need be truncated thus contributing to a further decrease in accuracy.

As a result, an approach using a sampling technique has been used. Sampling techniques generate a set of N_p samples at which and the function $C(z,\boldsymbol{\theta})$ is evaluated. Then the expected value of the function is approximated by calculating the sample mean of all evaluations as

$$E[C(z,\boldsymbol{\theta})] = \frac{1}{N_p}\sum_{i=1}^{N_p} C_i(z_i,\boldsymbol{\theta}_i) \tag{2.72}$$

If the probabilitistic function in consideration, P_1, is the variance, the objective function in Equation 3.7 takes the form

$$\sigma^2 = \int_{-\infty}^{+\infty} \left(C(z,\boldsymbol{\theta}) - E[C(z,\boldsymbol{\theta})]\right)^2 j(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{2.73}$$

which can be approximated by the sample variance of N_p values of $C_i(z,\boldsymbol{\theta})$ given by

$$\mathrm{Var}[C(z,\boldsymbol{\theta})] = \frac{\sum_{i=1}^{N_p}\left(C_i(z_i,\boldsymbol{\theta}_i) - E[C(z,\boldsymbol{\theta})]\right)^2}{N_p - 1} \tag{2.74}$$

Because for sampling techniques, the number of sampling points does not increase with the dimension of the problem, sampling techniques will be used in the stochastic optimization algorithm. The stochastic optimization procedure is shown in Figure 3.1. An initial design is selected by defining initial values of the control variables. A joint probability density function for the uncertain variables is assigned along with the vector of means and standard deviations. Discrete values for the quality variable, objective function and constraints, are obtained by sampling the distribution of the noise variables and evaluating them at the generated sampling points.

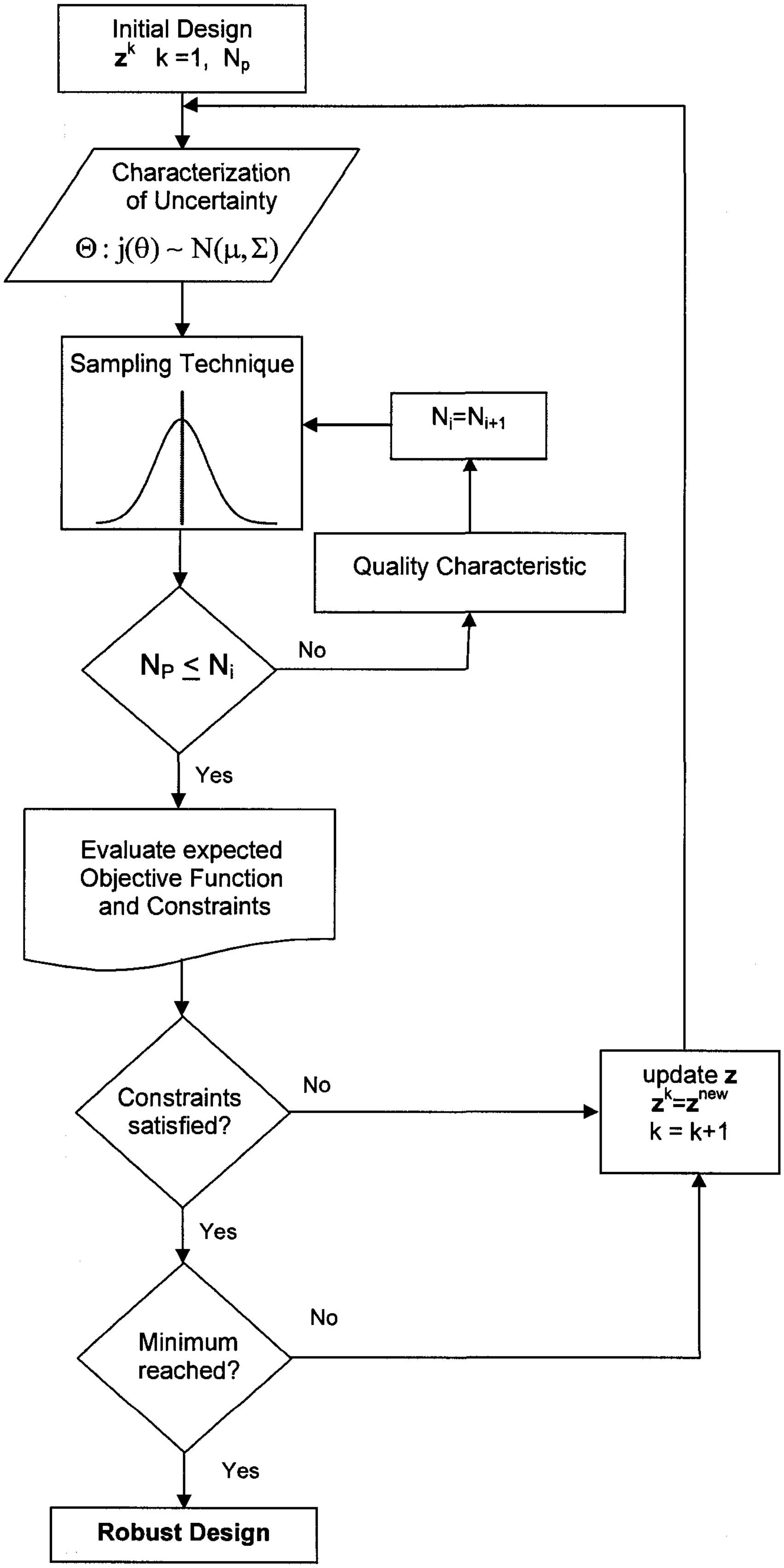

Figure 3.1 Stochastic optimization algorithm

Once discrete values have been obtained for all the samples, the probabilistic functions (e.g., mean or variance) of the quality variable, objective function and constraints are computed. If all the probabilistic constraints are satisfied, and the objective function cannot be further reduced, the stochastic optimization algorithm provides the optimal design. If not, the optimization algorithm provides a new set of values for the control variables and another iteration of the sampling objective function or evaluation procedure is undertaken.

It can be observed the computational effort involved in the stochastic optimization algorithm depend on the number of samples used since at each discrete value of the noise variables, obtained by sampling their probability distribution, the objective function is evaluated. Furthermore, the number of discrete values of the noise variables used in the optimization algorithm will depend on the desired accuracy for the estimation of the objective function. Therefore, the success of the stochastic optimization algorithm relies on determination of the optimal sampling technique and sample size that can provide a trade-off between accuracy and less computational effort and on the distribution of the sampled values of the noise variables over the sample space. Three sampling techniques have been studied in this work and are presented in the next section.

3.3.1 Sampling techniques used in the stochastic optimization algorithm

As explained in the previous section, during the optimization iteration, the model is run and the objective function is evaluated for each set of samples, hence making the computation intensive. Therefore a sampling technique that can provide a compromise between small sample size and accuracy of the estimate of the objective function is fundamental for increasing the efficiency of the stochastic optimization algorithm. In this study we examined the performance of three sampling techniques for the potential use in the stochastic optimization algorithm. Two Monte Carlo sampling techniques and one quasi-Monte-Carlo sampling technique were examined. The techniques were random Monte Carlo Sampling (MCS) and Latin Hypercube Sampling (LHS) for the former and Hammersley Sequence Sampling (HSS) for the latter.

3.3.1.1 Monte Carlo Sampling

MCS is one of the best known methods for numerical integration when the multidimensional integrals are involved. MCS evaluates the function at a pseudo-random sample points, and estimates its integral based on that random sample (Press *et al.*, 1992). In the MCS technique, a pseudo-random number generator is used to generate a uniformly distributed set of points in the interval [0,1] (Giunta *et al.*, 2003). This technique can readily be extended to an n-dimensional design space $[u_L,u_U]^n$ in which the sample site is an ordered n-tuple.

3.3.1.2 Latin Hypercube Sampling

Stratification is the process of selecting members of a population and grouping them. LHS (Mckay *et al.*, 1979) is part of the stratified sampling techniques that divide the distribution of noise variables into strata then applies random sampling within each stratum to select the values for the noise variables. In LHS, the sample space of each noise variable is divided in blocks of equal probability, and then, N_p values are obtained by sampling separately each block rather than the entire distribution. Samples are taken at random from within each interval and then paired in a random manner with the N_p values of the second noise variable. These N_p pairs are combined with N_p values of the third noise variable and so on to form N_p n-tuplets, where n is the number of noise variables. LHS guarantees that sample values are representative of the entire range of the distribution.

Figure 3.2 is an example of a two-dimensional case. The sample placement for LHS guarantees that only one stratum is selected in each row and column. For N_p=5 there are 5 partitions in both u_1 and u_2. This leads to a total of 25 strata in which each sample is randomly drawn. For all one-dimensional projections of the N_p samples and stratum, there will be one and only one sample in each stratum. The points in Figure 3.2 represent five samples sites in this example, where each sample is randomly located in its stratum.

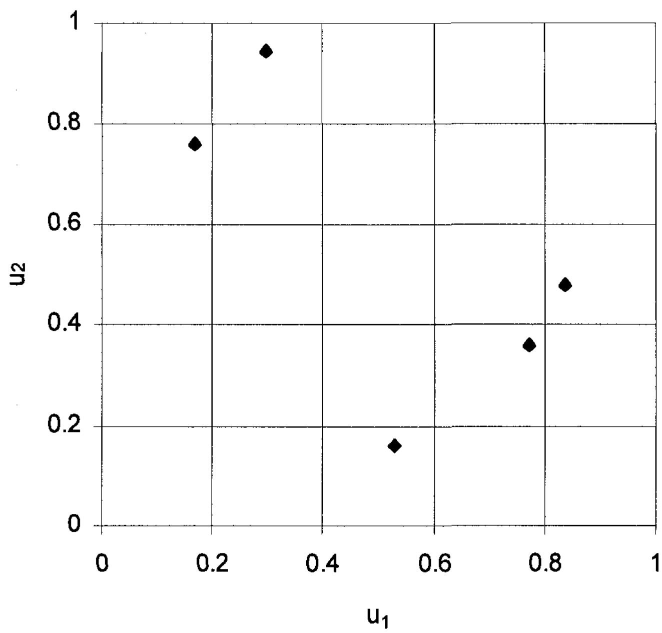

Figure 3.2 LHS with five samples drawn from the joint PDF of two variables u_1 and u_2.

3.3.1.3 Hammersley Sequence Sampling

HSS (Diwekar and Kalagnanam, 1997) is a sampling technique that uniformly samples a unit hypercube through the generation of a quasi-random sequence. The 'quasi-' prefix refers to the fact that the sequence is generated with a deterministic rather than a random generator. The samples generated with HSS possess low discrepancy. Discrepancy is a measure of uniformity for the distribution of the samples (Niederreiter, 1988). HSS make use of the Hammersley sequence (Hammersley, 1960) to generate the sample values for the noise variables. An algorithm for the generation of the Hammersley points can be found in Giunta *et al.* (2003) and in Diwekar and Kalagnanam (1997).

3.4 CSTR problem revisited

In order to introduce the stochastic optimization approach to parameter design, a common example found in the literature of chemical engineering design (Boudriga, 1990; Diwekar and Rubin, 1994; Kalagnanam and Diwekar, 1997) is used. The problem involved the design of a isothermal CSTR in which a first order series reaction $A \rightarrow B \rightarrow C$ takes place as shown in Figure 3.3.

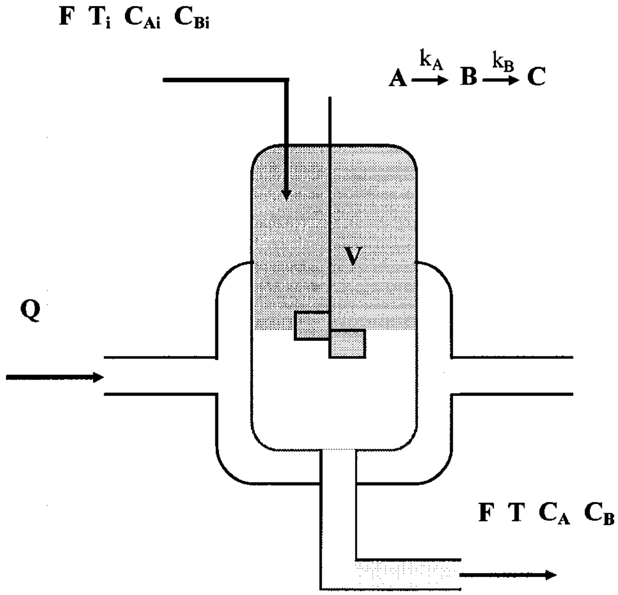

Figure 3.3 Temperature varying CSTR

The objective is to design a process with minimum quality costs. Quality costs depend on the quality characteristic of this process and are related, through the Taguchi loss function, to the variability of the quality characteristic and the deviation of its mean from the target value. The quality characteristic of this process is the rate of production of component B ($r_B V$) and the specified target is 60 moles/min. The rate of production of B, highly depends on the reaction temperature (as indicated by the Arrhenius expression). Therefore, the heat transferred to the reactor plays an important role for this process. Six input variables that transmit variability to the process performance were identified as the inlet concentration of component A, C_{Ai}, the inlet concentration of component B, C_{Bi}, the inlet temperature of the reactant, Ti, the heat stream added to the CSTR, Q, the bulk volume of the mixture in the reactor, V, and the volumetric flow rate, F.

$$Q = F\rho C_P (T - T_i) + V(r_A H_{rA} + r_B H_{rB}) \tag{2.75}$$

$$C_A = C_{Ai} / [1 + \tau k_A^0 \exp(-E_A / RT)] \tag{2.76}$$

$$C_B = [C_{Bi} + k_A^0 \exp(-E_A / RT) C_A] / 1 + \tau k_B^0 \exp(-E_B / RT) \tag{2.77}$$

$$-r_A = k_A^0 \exp(-E_A / RT) C_A \tag{2.78}$$

$$-r_B = k_B^0 \exp(-E_B / RT) C_B - k_A^0 \exp(-E_A / RT) C_A \tag{2.79}$$

$$\tau = V / F \tag{2.80}$$

$$r_B V = P_B \tag{2.81}$$

where C_A and C_B are the bulk concentrations of component A and B in the reactor outlet, $-r_A$ and $-r_B$, the rates of disappearance of A and B by reaction, and T is the bulk temperature of the mixture in the reactor. k_A^0, k_B^0 and E_A, E_B represent the pre-exponential Arrhenius constants and activation energies; H_{rA} and H_{rB} represent the molar heats of reactions that for this specific example are assumed to be independent of the temperature. Finally ρ and Cp are the densities and heat capacities of both streams. These parameters were assumed to be known exactly; therefore no uncertainty was involved. Their deterministic values are given in Table 3.1.

Table 3.1 Deterministic parameters and their values for the CSTR problem

Parameter	Values	Units
k_A^0	8.4×10^5	min^{-1}
k_B^0	7.6×10^5	min^{-1}
E_A	3.64×10^4	J/mol
E_B	3.46×10^4	J/mol
H_{rA}	-2.12×10^4	J/mol
H_{rB}	-6.36×10^4	J/mol
R	8,14	J/mol °K
C_p	3.2×10^2	J/kg °K
ρ	1180	kg/m^3

3.4.1 Degrees of freedom analysis

This system consists of 7 equations (Equations 3.13 through 3.19), 9 parameters (see Table 3.1) and 13 variables (C_{Ai}, C_{Bi}, T_i, V, F, Q, T, C_A, C_B, r_A, r_B, τ and P_B). Therefore there are 6 degrees of freedom associated with the system which can be used to perform the optimization.

If information on operating conditions is known and assumptions on the design of the reactor system are made, the 6 available variables can be defined. For this particular case, information about concentrations of the components, temperature and flow rate is available since the inputs of reactor system come from an upstream process. In addition, nominal heat input of 1710 kJ/min and a nominal mixture volume of 0.0391 m^3 were assumed. Nominal operation conditions of the upstream with allowable ranges are given in Table 3.2.

Table 3.2 Nominal values of input variables for the CSTR system

Input variable	Units	Values	Limits
C_{Ai}	mol/m^3	3118	1000 – 5000
C_{Bi}	mol/m^3	342	100 – 500
T_i	°K	300	290 – 330
F	m^3/min	0.0781	0.012 – 0.17
Q	kJ/min	1710	1.126e6 –2.394e6
V	m^3	0.0391	0.01 – 0.09

Having the input values properly defined, next the robust design is carried out.

3.4.2 Robust design

Variations in input variables (C_{Ai}, C_{Bi}, T_i, V, F, and Q) transmitted to the output of the process (i.e., production rate) were assumed to be independently normally distributed about their nominal values, μ,, with a standard deviation, σ, equal to 10 % of the nominal value (i.e. $\sigma = 0.1\mu$) and are given in Table 3.3. The variation transmitted to the output by each noise input was characterized by a Gaussian probability distribution function. The output distribution for the quality variable was characterized by calculating the sample mean and variance of the R_{Bi} values for the specified number of sampled values of the inputs.

Table 3.3 Noise variables with corresponding variations

	Noise variables	Units	M	σ
C_{Ai}	Inlet concentration of component A	mol/m^3	3118	311.8
C_{Bi}	Inlet concentration of component B	mol/m^3	342	34.2
T_i	Inlet temperature of the reactant	°K	300	30
V	Bulk volume of mixture in CSTR	m^3	0.0391	0.00391
F	Volumetric flow rate	m^3/min	0.0781	0.00781
Q	Heat supplied to the CSTR	kJ/min	1710	171

The variance of the output distribution, $Var[P_B(z,\boldsymbol{\theta})]$, was described by

$$Var[P_B(z,\boldsymbol{\theta})] = \int_{-\infty}^{+\infty} (P_B(z,\boldsymbol{\theta}) - E[P_B(z,\boldsymbol{\theta})])^2 j(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{2.82}$$

where E[R_B] represents the probabilistic functional mean, and can be defined as

$$E[P_B(z,\boldsymbol{\theta})] = \int_{-\infty}^{+\infty} P_B(z,\boldsymbol{\theta}) j(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{2.83}$$

where z is the vector of control variables $z = [\mu_{CAi}\ \mu_{CBi}\ \mu_{Ti}\ \mu_V\ \mu_F\ \mu_Q]$, $\boldsymbol{\theta}$, the vector of noise variables $\boldsymbol{\theta}$=$[C_{Ai}\ C_{Bi}\ T_i\ F\ V\ Q]^T$. $j(\boldsymbol{\theta})$ represents the joint probability density function of the noise variables. The expected value of the production rate, E[R_B], and its variance, $Var[P_B(z,\boldsymbol{\theta})]$, can be approximated using any of the sampling techniques discussed in the previous section as follows

$$E[R_B(z,\boldsymbol{\theta})] = \frac{1}{N}\sum_{i=1}^{N} P_{Bi}(z_i,\boldsymbol{\theta}_i) \tag{2.84}$$

and

$$Var[P_B(z,\boldsymbol{\theta})] = \frac{\sum_{i=1}^{N_p}\left(P_{Bi}(z_i,\boldsymbol{\theta}_i) - E[P_B(z,\boldsymbol{\theta})]\right)^2}{N_p - 1} \tag{2.85}$$

where $P_{Bi}(z_i,\boldsymbol{\theta}_i)$ is the value of R_B for the i^{th} sample. The number of samples to be used depends upon the accuracy required for estimating the mean and variance of the output distribution.

Due to the nonlinearity of the model representing this process, as described in Equation 3.12 through 3.17, a nonlinear optimization procedure was employed to perform the stochastic optimization. The solver used to perform the robust design for this process was *fmincon* from the Optimization toolbox of MATLAB®

3.5 Results and Discussion

This section presents the results obtained before and after the robust design.

3.5.1 Characterization of output distribution

The probability distribution of our model output was characterized by sampling the noise variables and computing the values for the quality characteristic at each set of sample values given the initial operating conditions. In this study, the first and second moments, (i.e., mean and variance) were chosen to characterize the output probability distribution and histograms were also created. Associating moments of a population to those of a sample drawn from such population is the most straightforward approach (Marion and Henrion, 1990).

It is known that the precision of an estimator increases as larger amounts of data are gathered and that the true value of an estimator can only be found for an infinite sample size. In practice one can only take a finite number of samples, and consequently parameters estimated with a large sample size can be considered "true" estimates. Therefore, estimates drawn from sample size of 10,000 were considered representative of the true output population.

Figure 3.4 shows the distribution of the quality characteristic for 10,000 samples using the operating conditions shown in Table 3.2. This distribution is bimodal, with modes falling between 90 and 110 and between 0 and 10 moles/min and widely spread covering a range of values between 0 and 180 moles/min. This broad range of values is indicative of significant sensitivity of the quality characteristic to relatively small changes (< 10%) in one or more of the noise variables

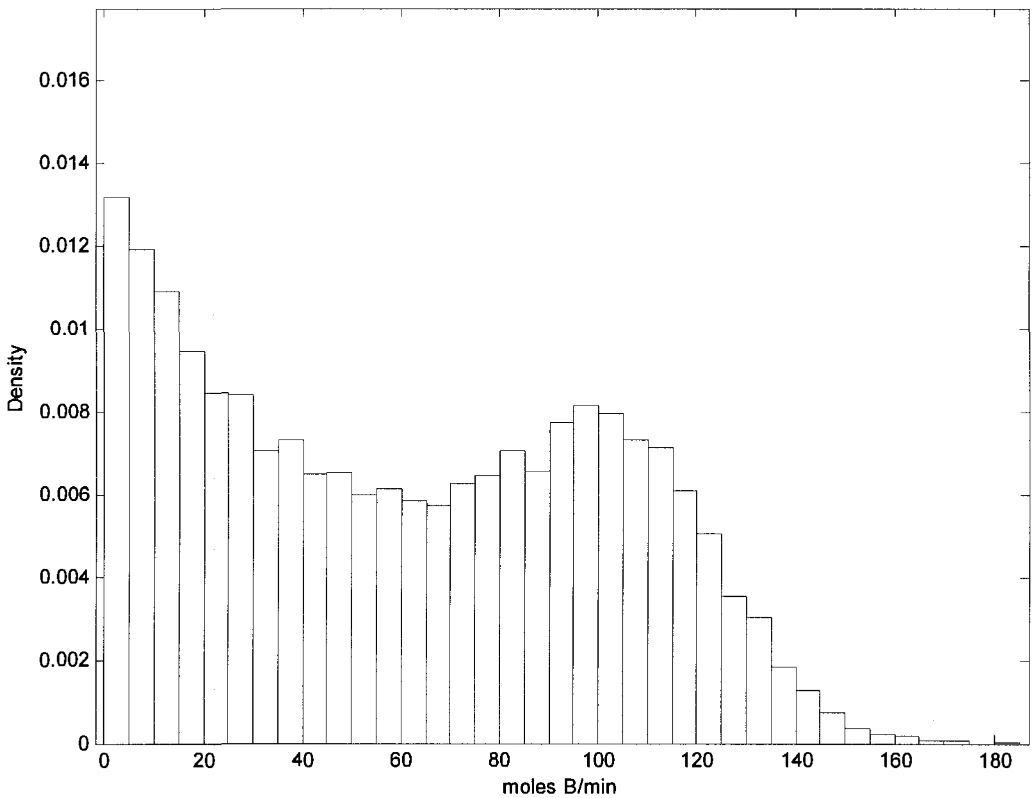

Figure 3.4 Output distribution under nominal conditions before robust design

Sensitivity analysis was performed in order to identify the noise variables that contributed the most to the widely spread distribution of the quality characteristic. Figure 3.5 shows the distribution of the quality characteristic when noise was added to one input variable at a time (i.e, each input variable was subjected to noise while the rest were held constant at their nominal values). Nominal values for C_{Ai}, C_{Bi}, V and Q produced a normal distribution for the quality characteristic. In contrast, the nominal value of Ti induced most of the variability to quality characteristic and explained its bimodality. Similarly, the nominal value of F produced a distribution highly skewed towards high values. The output distribution was sensitive to values in the inlet temperature, Ti, and the flow rate, F.

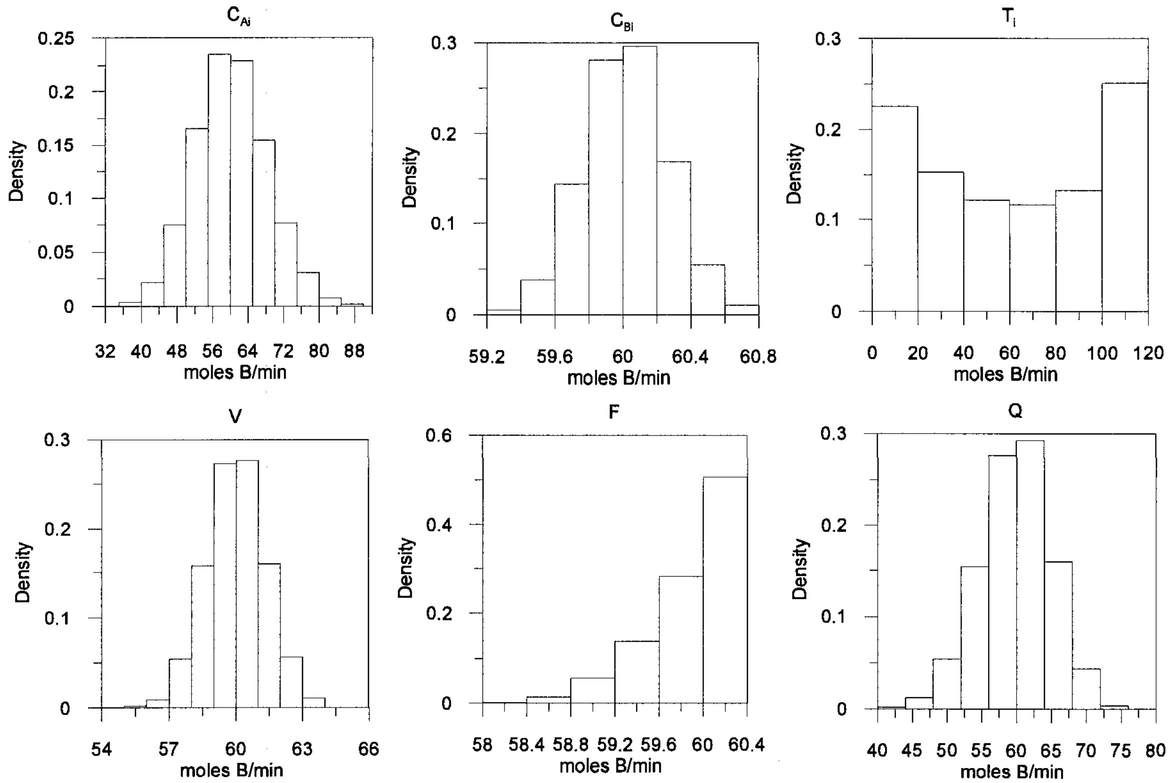

Figure 3.5 Distribution of quality variable when noise is added to one input variable one at a time.

In order reduce the computation time of the stochastic optimization algorithm while achieving the desired accuracy an optimal sample size was sought. The estimation of the parameters (mean and variance) of the output distribution, using the nominal design conditions, was performed for a sequence of sample sizes ranging from 50 to 5000 in steps of 50. The desired accuracy of the estimation was chosen to be $\pm$ 0.5% of the true values of the parameters. The true values of the mean and variance were considered those obtained for a sample size of 10,000 and were determined to be 60.44 moles/min and 1630 moles^2/min^2, respectively. This study was carried out for MCS, LHS and HSS.

Figure 3.6 shows the estimation of the mean for an increasing number of samples using the three sampling techniques. The estimates obtained using MCS (dotted line) displayed high variability and failed to achieve the desired accuracy. Diwekar and Kalagnanam (1997) have shown that MCS produces different estimates for different sample sizes. The random number generator used by MCS uses a different seed (i.e. initial value) to generate the samples; as a result the uniformity property is seriously affected. Moreover, the dimensionality of the problem contributes to distorting the uniformity of the samples generated thus resulting in a poor estimate of the distribution parameters of the output distribution.

Figure 3.6 also shows the estimates of the mean using LHS (dash line). LHS displayed much less variability than MCS. Although after a sample size of 250 the estimates of the sample mean predominately remain within the specified accuracy, it was only after 3000 samples that the estimates fully remained within $\pm$ 0.5% of the true mean. Figure 3.6 also shows the estimates using HSS (solid line). HSS displays the least variability and requires only 500 samples to achieve the required accuracy.

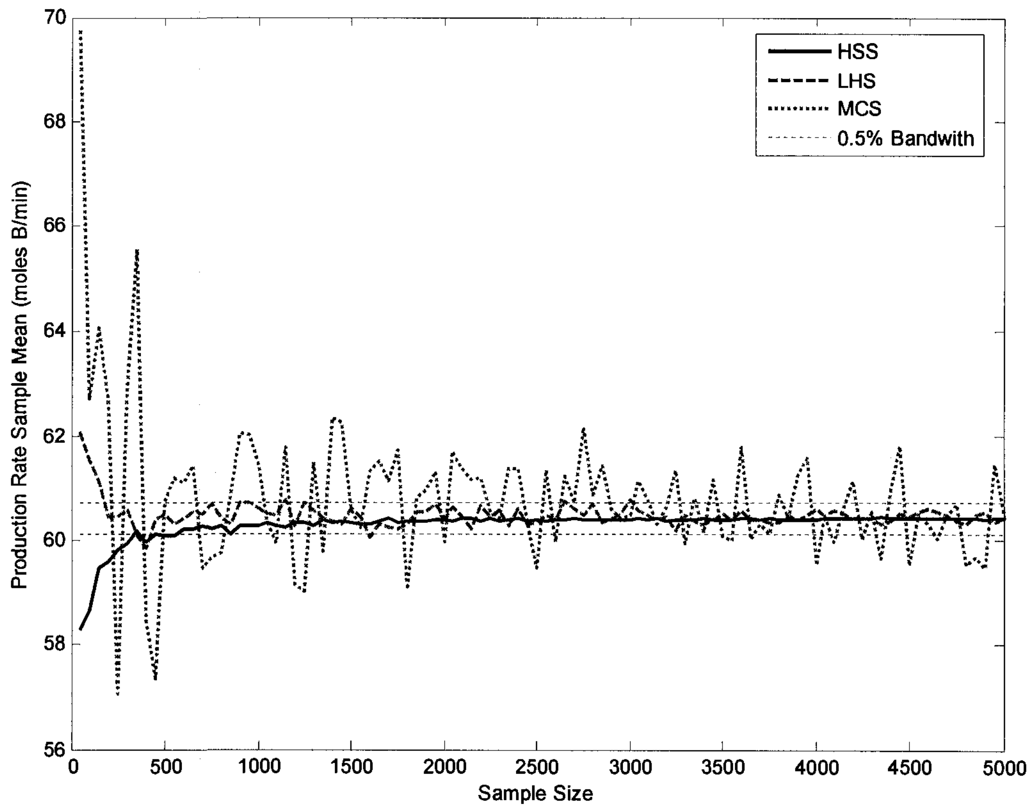

Figure 3.6 Estimation of the production rate sample mean as a function of sample size with three sampling techniques (HSS – Hammersley, LHS – Latin Hypercube, MCS – Monte Carlo)

Figure 3.7 displays the effect of sample size on the accuracy of the estimated variance of the production rate of B for the three sampling techniques. Again, estimates using MCS displayed the highest variability and failed to provide the required accuracy after a sample size of 5000. Similarly, estimates obtained using LHS, though presenting less variability than MCS, did not comply with the specified accuracy. Even with a sample size of 10,000 the required accuracy was not achieved for MCS or LHS. In contrast, estimates obtained with HSS displayed less variability in addition to providing the required accuracy with a sample size of 550. As a result, a sample size of 550 was used for solving the stochastic optimization problem.

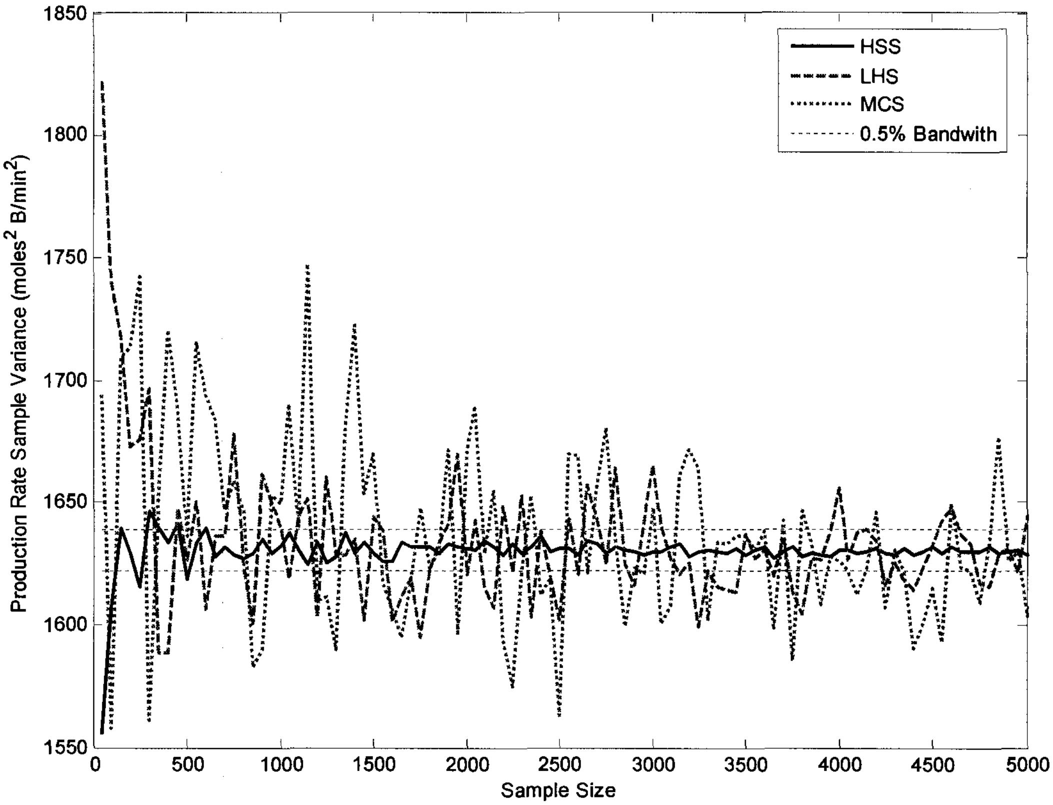

Figure 3.7 Estimation of the output variance as a function of sample size with three sampling techniques

3.5.2 Formulation of the objective function

Three formulations of the robust design via stochastic optimization for the CSTR were investigated in this work and are presented in this section.

3.5.2.1 Dual response approach

VM suggested minimizing the quality costs by minimizing the variance of the quality characteristic subject to the constraint that the mean value of the quality characteristic was on target. The mathematical representation of this problem using a stochastic optimization approach is given by

$$
\begin{aligned}
&\min\ k\ Var[P_B(z,\boldsymbol{\theta})] \\
&\text{w.r.t.}\quad z \\
&\text{s.t.} \\
&\quad \text{Equation 3.22} \\
&\quad \text{Equation 3.23} \\
&\quad E[P_B(z,\boldsymbol{\theta})]-60=0 \\
&\quad z=[\mu_{CAi}\ \mu_{CBi}\ \mu_{Ti}\ \mu_V\ \mu_F\ \mu_Q] \\
&\quad \boldsymbol{\theta}=[C_{Ai}\ C_{Bi}\ T_i\ V\ F\ Q] \\
&\quad 1000 \le C_{Ai} \le 5000 \\
&\quad 100 \le C_{Bi} \le 500 \\
&\quad 210 \le T_i \le 390 \\
&\quad 0.01 \le V \le 0.09 \\
&\quad 0.012 \le F \le 0.17 \\
&\quad 1126 \le Q \le 2394 \\
&\quad z\in Z;\ \boldsymbol{\theta}\in\Theta
\end{aligned}
\tag{2.86}
$$

where k is the quality loss constant proposed by Taguchi (1986).

For the solution of this stochastic optimization problem a k value of 20 $\$\cdot mol^{-2}\cdot hr^{2}$ was used. The minimum cost obtained by solving Equation 3.24 was \$7138 with nominal values for the input variables C_{Ai} = 3873 mol/m^3, C_{Bi} = 145 mol/m^3, T_i = 305 K, V = 0.07 m^3, F = 0.039 m^3/hr and Q = 1,913 kJ/min. The minimum variance obtained was 357 moles^2/min^2.

Figure 3.6 shows the distribution of the quality characteristic after performing the stochastic optimization. The distribution of the quality characteristic has become more nearly normal with mean value centered around the target value of 60 moles/min. More importantly, the variance of the quality characteristic has been reduced from 1630 to 357 moles^2/min^2.

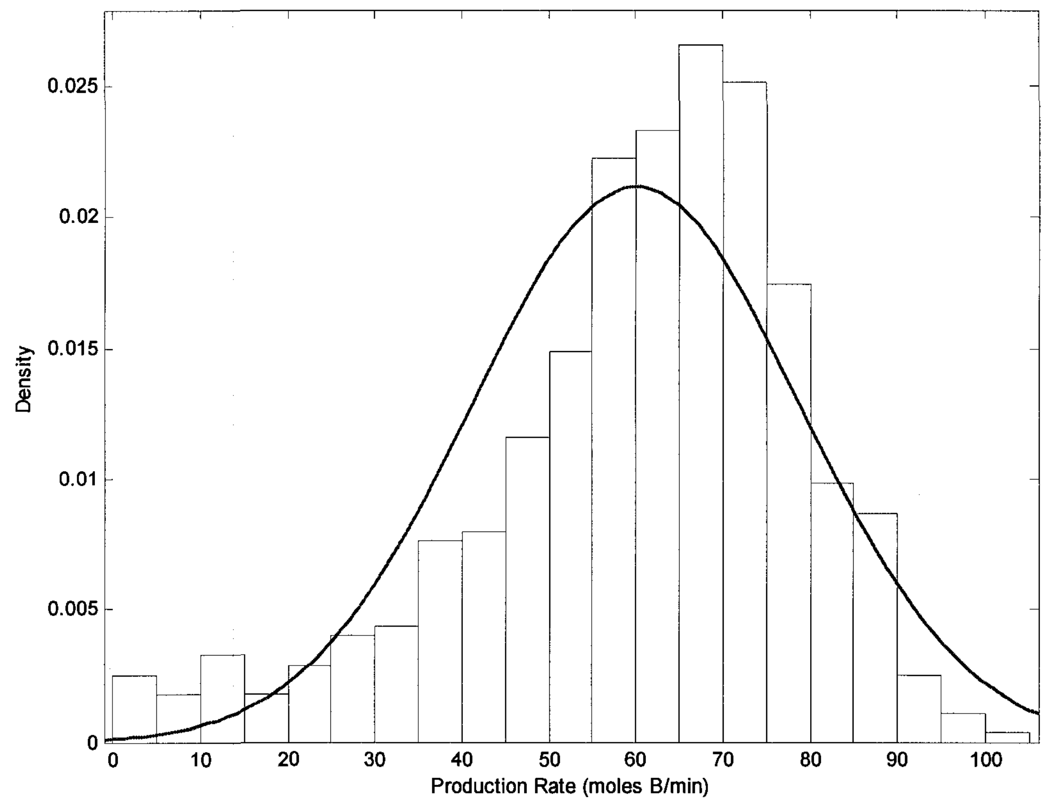

Figure 3.8 Distribution of the quality characteristic after optimization (DR approach)

Due to the non-convexity of the nonlinear objective function and constraints, global optimality cannot be assured. To gain greater confidence that the global optimum had been found, optimizations were performed using five sets of starting values based on HSS chosen from the feasible region for the decision variables. The design is shown in Table 3.4

Table 3.4 Computer design of experiment generated with HSS

Condition	C_{Ai}	C_{Bi}	T_i	V	F	Q
2	2600	200	316.66	0.042	0.062	1357
3	3400	400	294.44	0.058	0.084	1472
4	4200	150	307.77	0.074	0.105	1587
5	1160	350	321.11	0.013	0.127	1702

The same optimum was found for all sets of starting values providing strong confidence that the global optimum had been reached.

3.5.2.2 MSE approach

The second formulation of the stochastic optimization problem was based on the proposed approach of LT. The objective function in the stochastic optimization problem was modeled as the MSE. The stochastic formulation version of their problem is given by

$$
\begin{aligned}
&\min \ MSE[\mathrm{P_B}(z,\boldsymbol{\theta})] \\
&\text{w.r.t.} \quad z \\
&\text{s.t.} \\
&\quad MSE[\mathrm{P_B}(z,\boldsymbol{\theta})] = k\left(\sigma^2_{\mathrm{P_B}} + (\mathrm{E}[\mathrm{P_B}(z,\boldsymbol{\theta})]-60)^2\right) \\
&\quad \text{Equation 3.22} \\
&\quad \text{Equation 3.23} \\
&\quad z = [\mu_{\mathrm{CAi}}\ \mu_{\mathrm{CBi}}\ \mu_{\mathrm{Ti}}\ \mu_{\mathrm{V}}\ \mu_{\mathrm{F}}\ \mu_{\mathrm{Q}}] \\
&\quad \boldsymbol{\theta} = [\mathrm{C_{Ai}}\ \mathrm{C_{Bi}}\ \mathrm{T_i}\ \mathrm{V}\ \mathrm{F}\ \mathrm{Q}] \\
&\quad 1000 \le \mathrm{C_{Ai}} \le 5000 \\
&\quad 100 \le \mathrm{C_{Bi}} \le 500 \\
&\quad 210 \le \mathrm{T_i} \le 390 \\
&\quad 0.01 \le \mathrm{V} \le 0.09 \\
&\quad 0.012 \le \mathrm{F} \le 0.17 \\
&\quad 1126 \le \mathrm{Q} \le 2394 \\
&\quad z \in Z;\ \boldsymbol{\theta} \in \Theta
\end{aligned}
\qquad (2.87)
$$

The minimum cost obtained by solving Equation 3.25 was \$6092 with nominal values for the input variables C_{Ai} = 3874 mol/m^3, C_{Bi} = 145 mol/m^3, T_i = 305 K, V = 0.07 m^3, F = 0.035 m^3/min and Q = 1,913 kJ/min. The mean value of P_B was 54 moles/min indicating an offset from the target of 6 moles/min while the variance of P_B was 265 moles^2/min^2.

Figure 3.9 shows the distribution of the quality characteristic after performing the stochastic optimization. The quality characteristic has again become normally distributed. As in the previous formulation, the variance of the quality characteristic has been decreased to

265 moles^2/min^2. This reduction in variability corresponds to a 74% from the value using the Dual Response design. This result demonstrates clearly that relaxation of the "mean on target" constraint can allow the process to operate under conditions where the quality characteristic is less sensitive to input variation resulting in a reduction of the square bias and a further reduction in the variability of the quality characteristic.

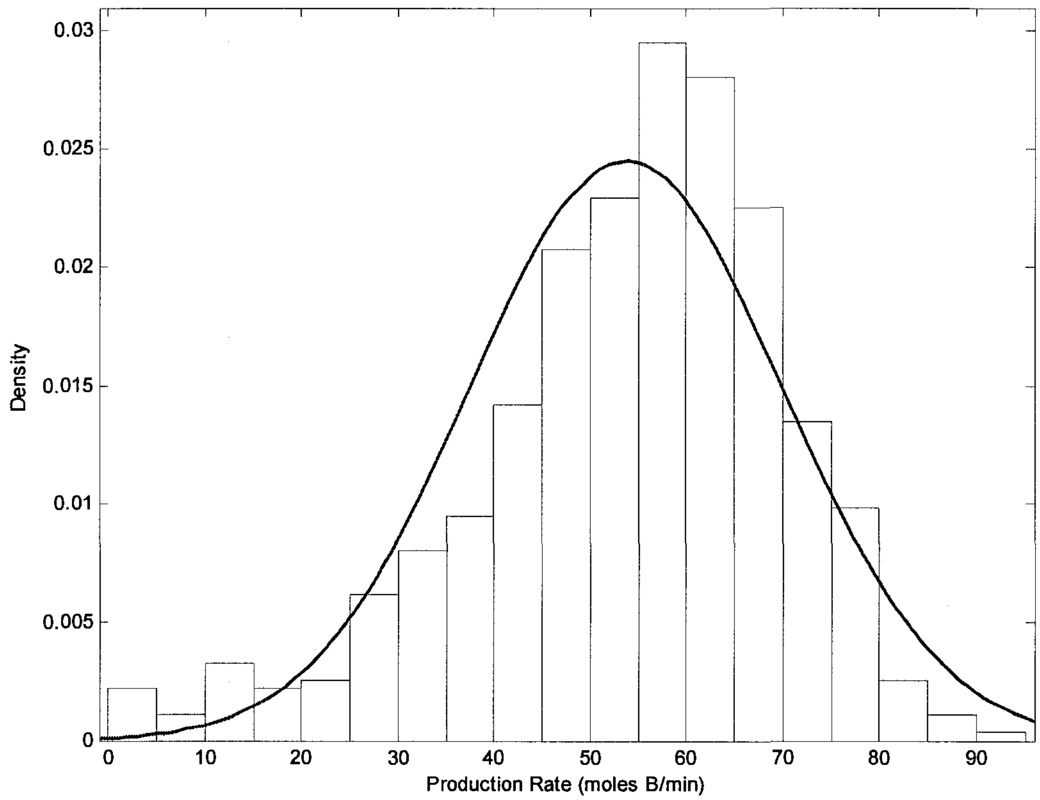

Figure 3.9 Distribution of the quality characteristic after optimization (MSE approach)

3.5.2.3 Weighted MSE approach

Finally, the objective function in the stochastic optimization problem was formulated with the proposed weighted MSE of Ding *et al.* (2004). The mathematical formulation for this stochastic optimization problem is given by

$$
\begin{aligned}
&\min\ WMSE[\mathrm{P_B}(z,\boldsymbol{\theta})] \\
&\text{w.r.t.}\quad z \\
&\text{s.t.} \\
&\quad WMSE[\mathrm{P_B}(\mathbf{z},\boldsymbol{\theta})] = k\left((1-\omega)\sigma^2_{\mathrm{P_B}} + \omega(\mathrm{E}[\mathrm{P_B}(z,\boldsymbol{\theta})]-60)^2\right) \\
&\quad \text{Equation 3.22} \\
&\quad \text{Equation 3.22} \\
&\quad z = [\mu_{CAi}\ \mu_{CBi}\ \mu_{Ti}\ \mu_{V}\ \mu_{F}\ \mu_{Q}] \\
&\quad \boldsymbol{\theta} = [\mathrm{C_{Ai}}\ \mathrm{C_{Bi}}\ \mathrm{T_i}\ \mathrm{V}\ \mathrm{F}\ \mathrm{Q}] \\
&\quad 1000 \le \mathrm{C_{Ai}} \le 5000 \\
&\quad 100 \le \mathrm{C_{Bi}} \le 500 \\
&\quad 210 \le \mathrm{T_i} \le 390 \\
&\quad 0.01 \le \mathrm{V} \le 0.09 \\
&\quad 0.012 \le \mathrm{F} \le 0.17 \\
&\quad 1126 \le \mathrm{Q} \le 2394 \\
&\quad z \in Z;\ \boldsymbol{\theta} \in \Theta;\ \omega \in [0,1]
\end{aligned} \tag{2.88}
$$

where ω is the appropriate weight according to the relative importance of both terms in the objective function.

Weights were applied to the objective function and the optimization was performed. Figure 3.10 shows how the quality costs, mean, variance and bias term of the objective function varies with increasing values of ω. The first plot shows that the minimum cost ($6092) of is found for a specific weight of 0.5. This is actually the result obtained with the same procedure proposed by LT. This is expected since the MSE is a special case of the WMSE approach. This plot also shows the maximum cost ($7138) which is the actual result obtained with the LT approach. The second plot shows the mean's quality characteristic as a function of ω. A mean value closer to the target value is obtained as the weight is increased. Similar trend is obtained in the third plot, the variance versus ω; as the weight is increased, a larger variance is obtained. Finally, the last plot shows the bias term versus ω. This plot shows the tendency of the bias term to decrease as the weight is increased.

These trends can be explained by the optimization procedure. As the weight is decreased, the variance term in the objective function becomes more important and is minimized whereas the bias, due to a low weight, is minimized to a lesser extent. Because the bias term in the objective function is not minimized, the cost is increased. On the other hand, when the weight is increased, the bias term in the objective function becomes more important and is minimized causing the mean to move closer to its target. In contrast, the variance is increased due to a lower weight and therefore, causes the cost to rise.

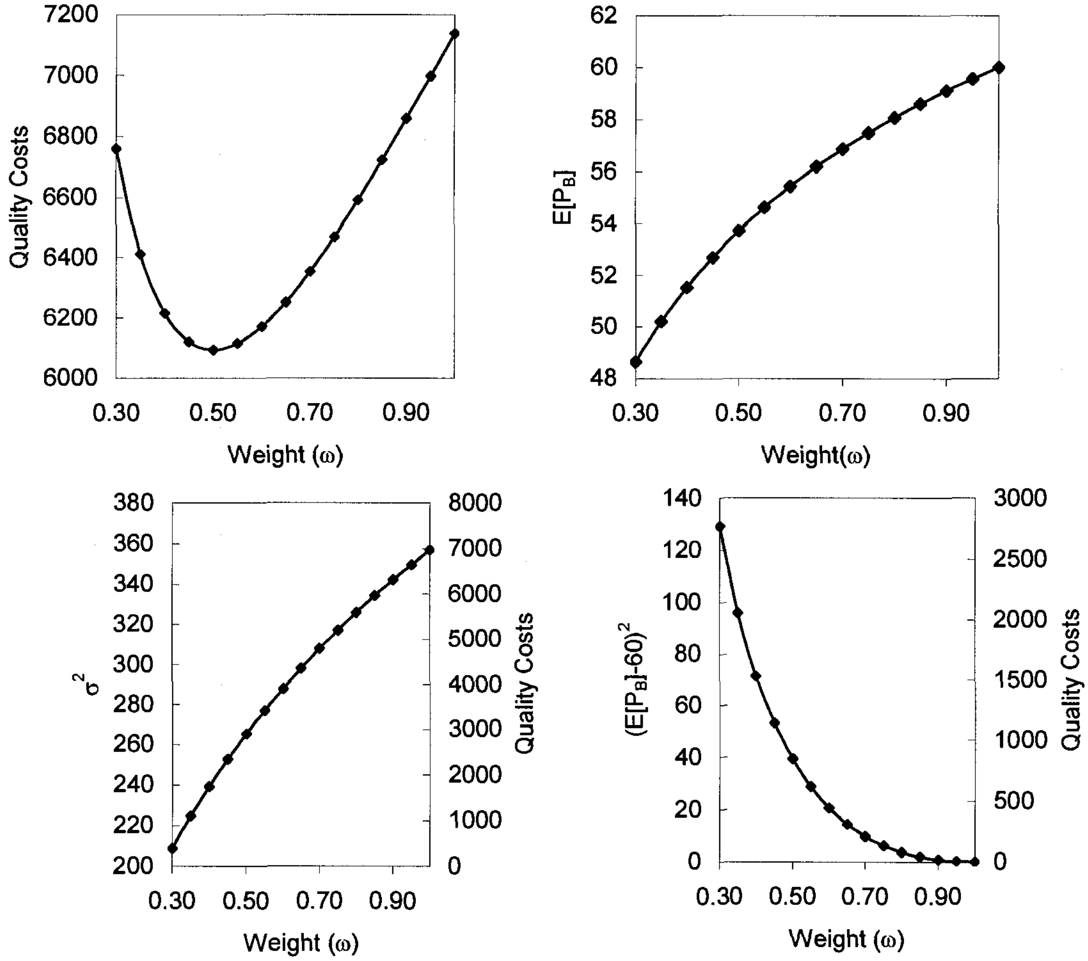

Figure 3.10 Effect of weight in the objective loss function (WMSE approach)

Overall, the three formulations provided optimal solutions that minimized the quality costs when variability is transmitted from the noise inputs to the quality variable. While with the MSE approach of LT provided the lowest cost ($6092), it also showed the largest deviation from target (μ = 54 mol/min). In contrast, the approach of VM provided a solution that was exactly on target with a minimum cost ($7138). On the other hand, the approach of

Ding *et al.* (2004) provided the flexibility to perform the optimization depending on the relative importance on both terms in the objective function. If the relative importance of both terms is known a priori, then the optimization can be performed directly using Equation 3.25. Otherwise it is necessary to plot optimal solutions for several weights and choose the weight that matches the robust constraints. For instance, if a deviation from target of $\pm$ 1 mol/min is allowed, using the efficient curve the appropriate weight can be chosen. For this particular case a weight of 0.9 provided the desired deviation from target (59 moles/min) with minimum variability (345 moles^2/min^2) and minimum cost ($6915).

Table 3.5 shows a comparison in the results obtained by previous authors and the results obtained in this study. The first column shows the results of Kalagnanam and Diwekar (1997). They focused on variance minimization and not on quality cost. Simulations were performed using their reported optimal values for control variables then quality costs were calculated. Their design ensures a low variability in the quality variable. However, it also provides the largest deviation from the target value. As a result the quality cost obtained with their design is very high. Columns two to four show the results obtained with the three approaches used in this study. They provided similar solutions with differences in the values of one of the control variables. The flow rate, F, was identified as the adjustment parameter that drives the mean of the quality variable close to its target.

Table 3.5 Comparison of results

	Kalagnanam and Diwekar (1997)	**This work**		
		MSE	**Weighted MSE ω =0.9**	**Dual Response**
μ_{CAi} **(mol/lt)**	3119.8	3874	3873	3873
μ_{CBi} **(mol/lt)**	342.24	145	145	145
μ_{Ti} **(K)**	350	305	305	305
μ_V **(m^3)**	0.05	0.07	0.07	0.07
μ_F **(m^3/min)**	0.043	0.035	0.038	0.039
μ_Q **(kJ/min)**	5,000	1913	1913	1913
μ_{RB}	37.69	54	59	60
σ^2_{RB}	338.28	265	345	357
σ_{RB}	18.39	16.27	18.57	18.89
Quality cost $	16,720	6,092	6,915	7,138

A sensitivity analysis was also performed after robust design in order to determine the impact of the optimal nominal values of the input variables on the variability of the quality variable. Simulations were performed considering a 10% noise of each input variable, one at a time, while keeping the rest of the variables constant at their optimal nominal values. Table 3.6 shows the contribution of each input variable on the variability of the quality variable before and after performing the robust design. The inlet temperature, $T_{i,}$, again presented the greatest contribution to the variability of the quality variable and therefore on the quality costs. This was followed by the mixture volume, V, and the inlet concentration of A, C_{Ai}.

Table 3.6 Contribution of the noise variables to variability in the quality variable

Variables	Initial	Final
C_{Ai}	4.0 %	14.00 %
C_{Bi}	0.003 %	0.0034 %
T_i	91.5 %	67.18 %
V	3.64 %	17.36 %
F	0.73 %	1.14 %
Q	0.12 %	0.40 %

The previous analysis shows that the inlet temperature (T_i) deserves special consideration in the design of this process followed by the inlet concentration of component A (C_{Ai}) and the mixture volume (V) to a lesser extent. The proposed change in the nominal inlet temperature (T_i) to a lower value close to its allowable minimum is due to the fact that this variable contributes mainly to the variability in the process as shown in the sensitivity analysis in Table 3.6. The robust design suggests operating the process with a higher nominal concentration of component A and with a lower nominal concentration of component B. Another required change is in the nominal mixture volume. An increase in the nominal volume provides the means to minimize the variability transmitted by the concentrations.

On the other hand, basing all these changes on minimizing just the quality cost by means of minimizing the variance of a quality variable will have an economic impact on the overall cost of the process (i.e., capital and operating costs). An economic function involving capital and operating costs as well as quality cost may be a more suitable choice since interactions between design, control, state and quality variables would be taken into account to achieve the optimum. An optimal methodology that takes account of capital, operating and quality costs at the design stage of a chemical process is under investigation and the benefits for this integration will be reported in a future article. Meanwhile, in this work, only quality costs were addressed by our stochastic optimization formulation.

In regards to the stochastic optimization approach for robust design some points deserve discussion. The applicability of the stochastic approach relies heavily on the information about the uncertainty of the noise variables. For the reactor problem it was assumed that information about the uncertainty associated with the noise variables was available in the form of probability density functions with known means and variances. In some situations, such information about the noise variables may not be available and represents a limitation for the application of the stochastic approach. In regards to the computational effort involved in the stochastic optimization algorithm, the computational time was reduced greatly due to an efficient sampling technique. HSS proved to be the most efficient technique since the estimate the objective function was obtained with the required accuracy with the least samples and thus fewest function evaluations.

3.6 Conclusions

This work has presented a comparison of the current formulations of the RPD problem using a stochastic optimization framework. The uncertainty in the noise variables was characterized by specifying the joint probability density function (PDF) of the noise variables as well as its parameters. Because the major computational challenge is the multidimensional integration involved to calculate the expectation or probability of the objective function, three sampling techniques were examined for the potential use in the stochastic optimization algorithm. The HSS proved to be the most efficient technique in terms of the fewest number of function evaluations involved. The application of the stochastic optimization algorithm to robust design was illustrated using a chemical engineering example, the design of a Continuous Stirred Tank Reactor (CSTR) and the potential limitations were discussed.

Acknowledgements

Financial support from the National Council of Science and Technology (CONACyT) of Mexico is greatly acknowledged.

Nomenclature

$\boldsymbol{Z}$	Vector of control variables as defined by Taguchi
$\mathbf{h}$	Vector of equality constraints
$\mathbf{g}$	Vector of inequality constraints
y	Quality variable
k	Taguchi loss constant
f	Scalar function
SN	Signal-to-noise ratio
MSE	Mean squared of error
N_p	Number of sampling points
j	Joint probability density function distribution (PDF)
E	Expectancy operator
L	Quality loss
C	Scalar function
P	Probabilistic function

Greek letters

$\boldsymbol{\theta}$	Vector of uncertain parameters
Θ	Domain of the uncertain parameters
τ	Target value
σ_y	Standard deviation of quality variable
μ_y	Mean of quality variable
μ	Mean vector of noise variables
σ	Standard deviation of noise variables

ω	Weight
λ	Lagrange multiplier

Set theory

$\in$	Element of
Z:	Set of control variables

Optimization acronyms

min	Minimize
w.r.t	With respect to
s.t.	Subject to

References

Boudriga, S., "Evaluation of 'Parameter Design' methodologies for the design of chemical Process Units", M.Sc. Thesis, University of Ottawa, 1990.

Box, G.E.P., "Signal-to-noise, performance criteria, and transformations", Technometrics, Vol. 30, 1-17, 1988.

Carnahan, B., Luther, H. A., and Wilkes. J.O., "Applied numerical methods", Wiley, New York, 1969.

Copeland, K. A. F., and Nelson, P. R., "Dual response optimization via direct function minimization", Journal of Quality Technology, Vol.28, No. 3, 1996.

Davis, P. J., and Rabinowitz, P., "Methods of numerical integration", Academic Press, New York, 1975.

Del Castillo, E., and Montgomery, D. C., "A nonlinear programming solution to the dual response problem", Journal of quality Technology, Vol. 25, No. 3 1993.

Ding, R., Lin, D. K. L., and Wei, D., "Dual response surface optimization: A weighted MSE approach", Quality Engineering, Vol. 16, No. 3, 377-385, 2004.

Diwekar, U. M., "Novel sampling approach to combinatorial optimization under uncertainty", Computation Optimization and Applications, 24, 335-371, 2003.

Diwekar, U. M., and Kalagnanam, R. J., "Efficient sampling technique for optimization under uncertainty", AIChE Journal, vol. 43, No. 2, 1997

Diwekar, U. M. ,. and Rubin, E. S. "Parameter design methodology for chemical processes using a Simulator", Industrial Engineering Chemistry Research, 33, 292-298, 1994.

Giunta, A. A., Wojtkiewicz Jr., S. F., and Eldred, M. S., "Overview of modern design of experiments methods for computational simulations", American Institute of aeronautics and Astronautics, AIAA2003-0649, 2003.

Hammersley, J. M., "Monte carlo methods for solving multivariate problems", Ann. N.Y. Acad. Sci., 86, 844, 1960.

Kackar, R N., "Off-line quality control, parameter design, and the Taguchi method", Journal of Quality Technology, Vol. 17, No.4, 1985.

Kalagnanam, J. R., Diwekar, U. M., "An efficient sampling technique for off-line quality control", Technometrics, V. 3, No. 3, 308-319, 1997.

Ki-Joo K., and Diwekar, U.M., "Hammersley stochastic annealing: efficiency improvement for combinatorial optimization under uncertainty", IIE Transactions, Vol. 34, 761-777, 2002

Leon, R. V., Shoemaker, A. C., and Kacker, R. N., "Performance measures independent of adjustment", Technometrics, Vol. 29, 253, 1987.

Lin, D. K., and Tu, W., "Dual response surface optimization", Journal of Quality Technology, Vol. 27, No.1, 1995.

McKay, M.D., Conover, W. J., and Beckman, R. J., "Ac comparison of three methods for selective values of input variables in the analysis of output from a computer code", Technometics, Vol. 21, 239-245, 1979.

Morgan, G, and Henrion, M., "Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis", Cambridge University press: New York 1990.

Nair, V. N. "Taguchi's parameter design: A panel of discussion", Technometrics, Vol. 34, 127-161, 1992.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P, "Numerical recipes in Fortran – The art of scientific computing", Second Edition, Cambridge University press, 1992.

Niederreiter, H., "Low-discrepancy and low-dispersion sequences", Journal of Number Theory, Vol. 30, 51–70, 1988.

Ryan, Thomas P. "Statistical methods for quality improvement", Second Edition, Wiley series in Statistics, New York, 2000.

Sen S., "Stochastic programming: computational issues and challenges", Encyclopedia of OR/MS, S. Gass and C. Harris (eds.), 2003.

Straub, D. A. and Grossman I. E., "Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties", Computers and Chemical Engineering, Vol. 14 No.9 pp. 967-985, 1990.

Taguchi, G. and Wu, Y, "Introduction to off-line quality control", Central Japan Quality Control Association, (American Supplier Institute), 1980.

Vining, G. G., and Myers, R. H., "Combining Taguchi and response surface philosophies: A dual response approach", Journal of Quality Technology, Vol. 22, No.1, 1990.

Chapter 4 (Paper 2)

Effect of Uncertainties on the Robust Design of Chemical Processes

Luis F. Dominguez and D. D. McLean

Dept. of Chemical Engineering

University of Ottawa

Abstract

Recently, the problem of designing robust chemical processes has been formulated as a single-stage stochastic optimization in which capital, operating and quality cost are taken into account. Based on the concept of Taguchi's quality loss, quality costs are treated via penalization of the objective function when quality constraints are not met. The stochastic optimization algorithm requires that the expected value of the objective function be calculated. This is usually approximated using quadrature and cubature formulas. In this work, a comparison of using Hammersley sampling to the application of cubatures for the calculation of the expected value of a function showed that the Hammersley sampling had a favourable computational demand and provided a measure of uncertainty for the expected value of the objective function. Both a penalty approach and a restriction approach on robustness metrics were used within the stochastic optimization formulation to study the effect of uncertainty/variability of parameters/variables on the overall cost of a chemical process. The optimal robust design of a reactor heat exchanger system was used to illustrate our results.

Key words: single-stage stochastic optimization, Robust design, quadrature, cubature, Hammersley sampling.

4.1 Introduction

The common approach to design chemical plants is to model the process though a series of mathematical equations representing the relationships between input and output variables. Then optimization is applied to identify optimal values for the design variables that reduce an overall cost or maximize a profit.

At the design stage, however, there are often technical or commercial parameters whose values are uncertain. The uncertainty in the parameter values may arise from a paucity of data gathered during experiments at a pilot or bench scale or from lack of information about the market for a product. In addition, the values of some process variables may be expected to change during operation of the process due to their inherent variability. In order to overcome these difficulties in the modelling, the designer assigns nominal values to the parameters and assumes that they will remain fixed during operation. The solution obtained after the optimization is performed is subsequently modified by applying overdesign or safety factors, based on heuristics or past experience, to variables involving equipment sizes in order to ensure flexible, safe operation of the process. This approach may be arguable since the designer may rely more on intuition and/or past experience rather than on the solution obtained based on his previous analysis. Moreover, this unnecessary overdesign increases capital and operating costs. On the other hand, under-design may take place when the applied overdesign factors are not enough to ensure that violation of the process constraints are not incurred or that unsafe operating conditions do not take place. It is therefore necessary to determine if a plant will operate under uncertainty of parameters and variability in input variables. In this regard, flexibility analysis provides the means for quantitatively determining if a process is able to cope with variation and uncertainty while satisfying all the constraints.

A chemical plant will maintain its operation as long as the company meets the specifications and complies with the satisfaction of its clients' needs. Therefore the continuous operation of a plant depends highly on the quality of the particular product it produces or on the quality of the process itself. Since satisfaction of clients' need is something that has to be constantly met, quality has to be characterized. This is usually done by specifying a target value for a quality characteristic and keeping track of not only its mean, but also its variance. Despite the difficulty of assigning a dollar value to quality costs, designers have started to approach quality related issues when designing chemical processes.

Taguchi (1986) and others (Robinson *et al.*, 2003; Ryan, 2000; Montgomery, 2001) have described methodologies that minimize quality loss; they consist mostly of (i) identifying and reducing input variability and (ii) changing the settings of control variables to minimize the transmission of input variability to the quality characteristic while keeping its mean on target. This has been done by running statistically designed experiments on prototypes and building empirical models to approximate the response to be optimized. These techniques are commonly called robust parameter design. However, direct applicability of such methodologies to the design of chemical processes under uncertainty is limited. In the case of chemical plants, this would require the plant or pilot plant to exist for an optimization to be performed. Moreover, this approach has neglected the impact of minimizing quality loss on capital and operating costs.

Approaches that address process design under uncertainty and focus on economic capital and operating costs have also appeared in the literature. For the design of chemical plants, different approaches have been proposed that take into account the design and operating stages of a plant as well as the uncertainty in model parameters and noisy process inputs (Grossman and Sargent, 1978; Halemane and Grossmann, 1982; Pistikopoulos, 1995; Ierapetritou *et al.*, 1996). These techniques share a common objective of optimizing an economic performance metric (e.g., design and operating costs or profitability) directly or indirectly.

When uncertainties are characterized by probability density functions (PDFs), the design problem is modeled as a two-stage stochastic programming problem (Dantzig, 1955; Pistikopoulos and Ierapetritou, 1995; Acevedo and Pistikopoulos, 1998). A disadvantage of this approach is that all the constraints involved are of the "hard" type (i.e., those that must be satisfied for all realizations of the uncertain parameters) and thus excluding potential benefits of relaxing some types of constraints that are not necessarily of the "hard" type. Another disadvantage is that this formulation leads to questionable implications. The two-stage stochastic approach assumes that the process will become inoperable if a solution for the optimization problem is not obtained (i.e., the problem is infeasible). This is, of course, not always true, especially when it comes to quality constraints. Failing to meet product specifications does not necessarily lead to the inoperability of a process. Out-of-specification products could always be reprocessed or sold at lesser price. It may lead to a more expensive cost of operation but not to inoperability.

Recently, Bernardo *et al.* (2001) explored this idea based on Taguchi's concept of quality loss given by $L=k(y-\tau)^2$; where L is the monetary loss, k is the Taguchi loss constant and y is the quality variable with corresponding target value, τ. Quality constraints were relaxed by penalizing the infeasibilities using the Taguchi loss function. With this penalization strategy, quality costs along with extents of robustness (e.g., variance, coefficient of variation) in quality characteristics were taken into account in a single-stage stochastic optimization framework. In their work, the proportional quality loss parameter, k, was assumed to be constant and their problem was solved for different values. In reality, the value of the quality loss constant is, however, not known exactly. Although the value of k can be estimated if the actual loss is known for one value of the quality characteristic (Ryan, 2000), it may difficult to obtain with enough accuracy for a range of values of the quality characteristic. Therefore, it is proposed that treating k as uncertain parameter would be a more realistic approach. Bernardo *et al.* (2001) also employed a cubature technique to compute the multiple integrals involved in their stochastic optimization problem. However, the use of such integration methods does not provide any information on the relative accuracy or precision obtained with their solution.

It was the objective of this work to extend the previous work of Bernardo *et al.* (2001) to more effectively deal with uncertainty in the quality loss constant and other parameters. Accordingly, the effect of the number of uncertain parameters involved on the solution of the optimization problem was studied by solving it for a different number of uncertain parameters. Another objective was to evaluate the use of a sampling technique rather than cubature for the calculation of integrals in the optimization algorithm. Hammersley Sequence Sampling (HSS) was chosen since it had been shown (Diwekar and Kalagnanam, 1997) to provide sufficiently accurate estimates of the mean of objective functions with modest computational effort.

In order to introduce the single-stage stochastic optimization formulation, in Section 4.2 the mathematical foundations of the stochastic optimization approach within a two-stage formulation is presented. In Section 4.3 the single-stage stochastic formulation for process quality proposed by Bernardo *et al.* (2001) is briefly present. Next, in Section 4.4, a design example chosen from the chemical engineering literature is employed to demonstrate and evaluate our approach. Finally, conclusions are summarized in Section 4.5

4.2 Two-stage stochastic optimization

The two-stage stochastic optimization formulation assumes a perfect control scheme and is based on selecting optimal values for the design vector **d** (here-and-now decisions) at the design stage (i.e., 'first stage') while seeking feasibility/flexibility by finding optimal values for the control vector **z** (wait-and-see decisions) at the operating stage (i.e., 'second stage').

Values for design variables, **d**, are selected at the design stage and remain fixed at the operating stage. The objective at the operating stage is to determine the optimal values of the control variables, **z**, for each realization of the uncertain parameters $\boldsymbol{\theta}$ falling in the feasible region $R_n(\mathbf{d})$:

$$R_n(\mathbf{d}) = \left\{ \boldsymbol{\theta} \in \Theta \,\middle|\, \exists(\mathbf{z}, \mathbf{x}) : h(\mathbf{d}, \mathbf{z}, \mathbf{x}, \boldsymbol{\theta}) = 0 \wedge g(\mathbf{d}, \mathbf{z}, \mathbf{x}, \boldsymbol{\theta}) \leq 0 \right\} \tag{2.89}$$

where n represents the number of uncertain parameters, **d**, **z**, and **x**, represent the vectors of design, control, and state variables; $\boldsymbol{\theta}$ corresponds to the vector of uncertain parameters. **h** and **g** represent the vectors of equality and inequality constraints, respectively.

Having defined the feasible region, $R_n(\mathbf{d})$, the mathematical statement of the two-stage stochastic optimization can be formally introduced

$$
\begin{aligned}
&\text{Design stage:} \\
&\qquad\qquad \min \quad E_{R_n(\mathbf{d})}\{C^*(\mathbf{d},\boldsymbol{\theta})\} \\
&\qquad\qquad \text{w.r.t.} \quad \mathbf{d} \\
&\qquad\qquad \mathbf{d} \in D, \quad \boldsymbol{\theta} \in \Theta \\
&\text{Operating stage:} \\
&C^*(\mathbf{d},\boldsymbol{\theta}) = \min C(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) \\
&\qquad\qquad \text{w.r.t.} \quad \mathbf{z},\mathbf{x} \\
&\qquad\qquad \text{s.t.} \quad \mathbf{h}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) = 0 \\
&\qquad\qquad\qquad \mathbf{g}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) \le 0 \\
&\qquad\qquad\qquad \mathbf{z} \in Z, \quad \mathbf{x} \in X
\end{aligned}
\tag{2.90}
$$

where C is the operating cost and C* is the minimum cost obtained at the operating stage. Because the parameter uncertainties are characterized by a joint PDF, the expected value of the objective function $C^*(\mathbf{d},\boldsymbol{\theta})$ is represented by the n-dimensional integral over the optimization problem at the design stage:

$$E_{R_n(\mathbf{d})}\{C^*(\mathbf{d},\boldsymbol{\theta})\} = \int_{R_n(\mathbf{d})} C^*(\mathbf{d},\boldsymbol{\theta}) \;\; j(\boldsymbol{\theta}) \;\; d\boldsymbol{\theta} \tag{2.91}$$

It has been shown (Ierapetritou and Pistikopoulos, 1995) that under this formulation, the integration region is a function of the design variables. The two-stage stochastic optimization relies on the fact that optimal values for the design variables, **d**, for which the inequality constraint, **g**, is satisfied exist. Pai and Hughes (1987) pointed out that there may be situations in which, for some values of the design vector, **d**, it may be impossible to find optimal values of the control vector, **z**, that satisfy all the constraints, thus making the design problem infeasible. This implies that the plant will become inoperable due to the lack of proper values for the control variables, **z**. Mathematically, this is equivalent to penalizing the objective function with an infinite value for every pair of $\boldsymbol{\theta}$ and **d** that yields an infeasible solution.

In real situations this may not be true. The operation of a plant may be subjected to many types of constraints. While, for one, satisfaction needs to be ensured (i.e., hard constraints), there are others for which possible violation does not necessarily lead to inoperability of a plant (i.e., soft constraints) (Wellons and Rekleitis, 1989). Accordingly, a plant of current design variables, **d**, may be able to operate even if some realizations of the parameter space lie outside of the feasible region. For instance, out-of-specification products can be sold at a lesser price, design requirements can be changed or equipment can be replaced so that the plant still operates.

This was, in fact, the basis of the proposed approach of Bernardo *et al.* (2001). Assuming that all the constraints leading to infeasibilities at the design stage are quality related constraints they proposed to treat quality constraints as soft constraints and to penalize the economic objective function with a quadratic function. Based on Taguchi's concept of quality losses, the penalty function was employed to represent the quality losses incurred when quality constraints were violated. The stochastic formulation for process quality is briefly described in the next section.

4.3 Stochastic optimization for process quality

In order to account for process quality two approaches can be stated: a penalty and a penalty plus explicit restriction on robustness metrics approach.

4.3.1 Penalty approach

The penalty approach can be formulated by defining a scalar function which penalizes the objective function when quality characteristics deviate from their targets. For quality variables, y_q, and target values, τ_q, q=1,2,..., Q, quadratic loss functions (Taguchi, 1980) can be implemented. The objective function of Equation 4.2 at the design stage can be modified as

$$C(\mathbf{d},\mathbf{z},\mathbf{x},y_q,\tau_q,\boldsymbol{\theta}) = C(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) + \sum_{q=1}^{Q} k_q (y_q - \tau_q)^2 \qquad (2.92)$$

where k_q is the quality constant associated with the loss incurred when a particular value of the quality characteristic, y_q, differs from its target value, τ_q. The quality constant k_q, may take different values depending upon the type of function used. Table 4.1 shows the possible values of k_q for four different formulations of the penalty function.

Table 4.1 Penalty terms based on Taguchi loss function

Taguchi S/N ratio	**k value**
Nominal-the-best Symmetric	Same k_q for all values of y_q
Nominal-the-best Asymmetric	$k_q = k_1$ if $y_q < \tau_q$ $k_q = k_2$ if $y_q \geq \tau_q$
Larger-the-better	$k_q = k_1$ if $y_q < \tau_q$ $k_q = 0$ if $y_q \geq \tau_q$
Smaller-the-better	$k_q = 0$ if $y_q < \tau_q$ $k_q = k_2$ if $y_q \geq \tau_q$

By treating soft quality constraints with this penalization strategy, the integration region R_n in Equation 4.2 becomes the entire uncertainty region, Θ. Thus the expected value of the objective function at the design stage takes the form

$$E_\Theta \left\{ C(\mathbf{d},\mathbf{z},\mathbf{x},y_q,\tau_q,\boldsymbol{\theta}) \right\} = \int_\Theta C(\mathbf{d},\mathbf{z},\mathbf{x},y_q,\tau_q,\boldsymbol{\theta})\ \ j(\boldsymbol{\theta})\ \ d\boldsymbol{\theta} \qquad (2.93)$$

Unlike Bernardo *et al.* (2001), in this work the above formulation was simplified by approximating the integration formula involved in the expectation operation of Equation 4.5 using the HSS. This sampling technique generates a grid of N_p points, $\boldsymbol{\theta}_i$, uniformly distributed over the integration region. Accordingly, with this discretization of the uncertainty region, the expected value of the objective function takes the form

$$E_{\Theta}\left\{C(\mathbf{d},\mathbf{z},\mathbf{x},y_q,\tau_q,\boldsymbol{\theta})\right\} \cong \frac{1}{N_p}\sum_{i=1}^{Np} C(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,y_{i,q},\tau_{i,q},\boldsymbol{\theta}_i) \tag{2.94}$$

Furthermore, substituting Equation 4.6 into the overall expected cost in Equation 4.2 leads to reformulation of the two-stage stochastic optimization as a single-stage stochastic optimization

$$\begin{aligned}
&\min \quad \frac{1}{N_p}\sum_{i=1}^{N_p} C(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,y_{i,q},\tau_q,\boldsymbol{\theta}_i) \\
&\text{w.r.t.} \quad \mathbf{d},\mathbf{z}_i,\mathbf{x}_i,y_{i,q} \\
&\text{s.t.} \quad C(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,y_{i,q},\tau_q,\boldsymbol{\theta}_i) = C(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,\boldsymbol{\theta}_i) + \sum_{q=1}^{Q} k_q (y_{i,q}-\tau_q)^2 \\
&\qquad h(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,\boldsymbol{\theta}_i) = 0 \\
&\qquad g(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,\boldsymbol{\theta}_i) \le 0 \\
&\qquad \mathbf{d}\in D, \quad \boldsymbol{\theta}_i \in \Theta, \quad \mathbf{z}_i \in Z, \quad \mathbf{x}_i \in X, \quad y_{i,q} \in Y \\
&\qquad i = 1,\ 2,\ ....,\ N_p; \quad q = 1,\ 2,...,Q
\end{aligned} \tag{2.95}$$

Equation 4.7 accounts for the quality costs incurred by penalizing the objective function when quality characteristics, y_q, deviate from their targets, τ_q.

An advantage of the single-stage stochastic formulation is that the same structure of the deterministic optimization formulation is conserved. Therefore, if the deterministic optimization version of the problem is convex, the stochastic formulation defined in Equation 4.7 is also going to be convex (Bernardo *et al.*, 2001; Georgiadis and Pistikopoulos, 1998). Accordingly, the stochastic optimization problem can be solved with any standard non-linear programming algorithms such as the generalized reduced gradient (GRG) method or the successive quadratic programming (SQP) algorithm (Edgar *et al.*, 2001). On the other hand, the penalty approach does not constrain the distribution of quality characteristics, thereby allowing them to have any distributional shape, depending on the location of the optimal overall cost. This approach can be easily modified to include robustness metrics as specified by the designer.

4.3.2 Penalty plus restriction on robustness metrics approach

If the designer is interested in attaining specific values of the parameters of the quality characteristic distributions, such as maximum mean or minimum standard deviation, then explicit restrictions on robustness metrics must be specified. The first two statistical moments, the mean and variance, characterize the location and variability in a quality characteristic. This robustness metrics are defined as

$$\mu_{y_q} = E_{\Theta}(y_q) \tag{2.96}$$

$$\sigma^2_{y_q} = E_{\Theta}\left\{(y_q - \mu_{y_q})^2\right\} \tag{2.97}$$

It can be easily demonstrated (Mogan and Henrion, 1990) that

$$\hat{\mu}_{y_q} = \frac{1}{N_p}\sum_{i=1}^{N_p} y_{i,q} \tag{2.98}$$

$$\hat{\sigma}^2_{y_q} = \frac{1}{N_p - 1}\sum_{i=1}^{N_p}(y_{i,q} - \mu_{y_q})^2 \tag{2.99}$$

These can be incorporated to a single-stage stochastic optimization formulation as

$$\begin{aligned}
\min \quad & \frac{1}{N_p}\sum_{i=1}^{N_p} C_q(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,y_{i,q},\tau_q,\boldsymbol{\theta}_i) \\
\text{w.r.t.} \quad & \mathbf{d},\mathbf{z}_i,\mathbf{x}_i,y_{i,q} \\
\text{s.t.} \quad & C_q(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,y_{i,q},\tau_q,\boldsymbol{\theta}_i) = C(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,\boldsymbol{\theta}_i) + \sum_{q=1}^{Q} k_q (y_{i,q} - \tau_q)^2 \\
& h(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,\boldsymbol{\theta}_i) = 0 \\
& \hat{\mu}_{y_q} = \frac{1}{N_p}\sum_{i=1}^{N_p} y_{i,q} \\
& g(\mathbf{d},\mathbf{z}_i,\mathbf{x}_i,\boldsymbol{\theta}_i) \le 0 \\
& \frac{1}{N_p - 1}\sum_{i=1}^{Np}(y_{i,q} - \hat{\mu}_{y_q})^2 \le \gamma_q \\
& \mathbf{d} \in D, \quad \boldsymbol{\theta}_i \in \Theta, \quad \mathbf{z}_i \in Z, \quad \mathbf{x}_i \in X, \quad y_{i,q} \in Y \\
& i = 1,\ 2,\ ....,\ N_p; \quad q = 1,\ 2, ..., Q
\end{aligned} \tag{2.100}$$

where γ_q is an upper bound for the variance of the quality variables q=1,2,.. Q.

In the next section, both approaches will be applied to the optimal design of a reactor heat exchanger system. Also, a comparison of the results obtained with cubatures and with the sampling technique used will be provided. The impact of the uncertainty/variability in internal parameters/variables in the overall design cost will be also demonstrated.

4.4 Design example

In order to demonstrate this approach, a test problem first introduced by Chacon-Mondragon and Himmelblau (1996), and later used by Bernardo *et al.* (2001) was employed (see Figure 4.1). A first order chemical reaction, $A \rightarrow B$, takes place in a CSTR at temperature T_1. Reactant A enters the reactor at temperature, T_0, with initial concentration, C_{A0}, and molar flow rate, F_0. The temperature inside the reactor is maintained by recycling reacting mixture with flow F_1 through a counter-current heat exchanger which reduces the reactor content temperature from T_1 to T_2. Cooling water enters the heat exchanger at flow rate, F_{w1} and temperature T_{w1}. The product is removed at flow rate, F_0, and temperature, T_1. In addition, to reduce the number of equations, it is assumed that no reaction takes place inside the exchanger or along the pipe transporting the reaction mixture.

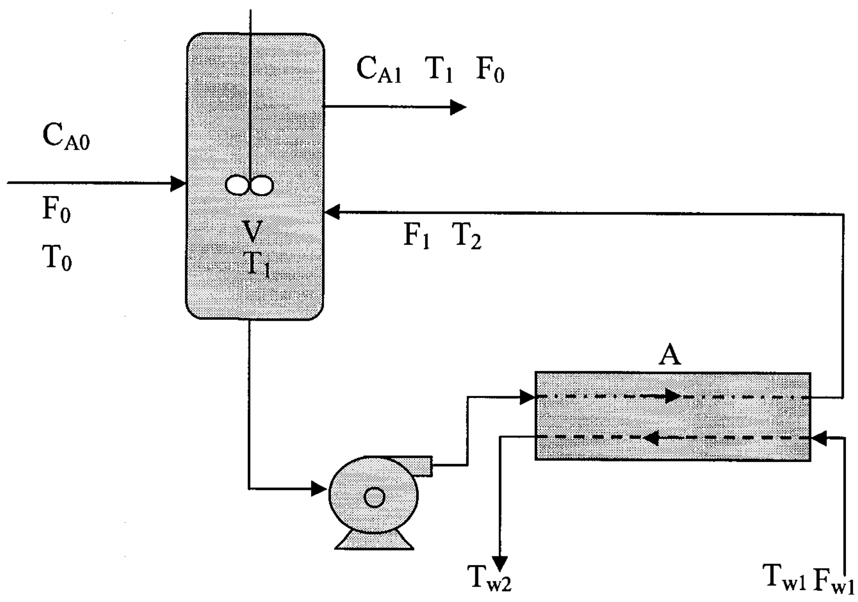

Figure 4.1 Reactor and heat exchanger system

The corresponding material and energy balances for this reactor are given by Equations 4.13 to 4.18.

$$x_A = (C_{A_0} - C_A)/C_{A_0} \tag{2.101}$$

$$F_0 x_A - k_R \exp(-E/RT_1) C_{A0}(1 - x_A)V = 0 \tag{2.102}$$

$$(-\Delta H) F_0 x_A + F_0 C_p (T_0 - T_1) - F_1 C_p (T_1 - T_2) = 0 \tag{2.103}$$

$$F_1 C_p (T_1 - T_2) = AU\Delta T_{lm} \tag{2.104}$$

$$\Delta T_{lm} = (T_1 - T_{w_2}) - (T_2 - T_{w_1}) / \ln[(T_1 - T_{w_2})/(T_2 - T_{w_1})] \tag{2.105}$$

$$F_1 C_p (T_1 - T_2) = F_w C_{p_w} (T_{w_2} - T_{w_1}) \tag{2.106}$$

where k_R, E, ΔH are the Arrhenius pre-exponential constant, activation energy and heat of reaction, respectively. Heat capacities of reactant and cooling water are given by C_p and C_{pw}, and U is the overall heat transfer coefficient in the heat exchanger. Equations 4.13 to 4.18 define the equality constraints for the system. The optimal design is considered to be that which minimizes the overall cost of this process. The objective function is partitioned into two terms: capital costs and operating costs, leading to

$$\begin{aligned} &\text{Capital Costs} = c_1 V^{0.7} + c_2 A^{0.6} \\ &\text{Operating Costs} = c_3 F_W + c_4 F_1 \\ &\text{Overall Cost} = \text{Capital Costs} + \text{Operating Costs} \end{aligned} \tag{2.107}$$

Coefficients $c_1$and c_2 in Equation 4.19 relate the capital cost to the sizes of the equipment, in this case, the CSTR reactor volume, V, and the heat transfer area of the heat exchanger, A. Coefficients c_3 and c_4 relate the operating costs to the use of the cooling fluid and pumping associated with the recycling of the reactor contents in the cooling loop, respectively. It was required that neither the reactor temperature, T_1, nor the recycling temperature, T_2, exceeded 389 K, and both had to be at least 311 K. For the heat exchanger, a minimum approach temperature difference between hot and cold streams of 11.1 K was specified. Also, constraints were imposed to ensure that the exit temperature of the hot stream was lower than its temperature at the inlet. Conversely, the exit temperature for the cold fluid had to be greater than the temperature at the inlet so another constraint was imposed. Finally, it was required that the cooling water be between 294K and 323K. Mathematically these constraints are represented by Equations 4.20 to 4.26

$$311 \le T_1 \le 389 \tag{2.108}$$

$$311 \le T_2 \le 389 \tag{2.109}$$

$$294 \le T_{w_2} \le 323 \tag{2.110}$$

$$T_1 - T_{w_2} \ge 11.1 \tag{2.111}$$

$$T_2 - T_{w_1} \ge 11.1 \tag{2.112}$$

$$T_{w_2} - T_{w_1} \ge 0 \tag{2.113}$$

$$T_1 - T_2 \ge 0 \tag{2.114}$$

The quality cost of this process was considered to be a function of the reactant concentration in the product. Accordingly, conversion was considered the quality variable, and the desired target value was set to 0.9. It was assumed that a higher conversion (i.e., higher purity of B) had no impact on the final cost of the product. In contrast, a decrease in conversion (i.e., lower purity) had a quadratic effect on the quality cost. Therefore, an asymmetric Taguchi loss function was used to represent the quality constraint and quality costs associated with violating such constraints

$$\begin{aligned} &\mathrm{Qcost} = k(x_A - 0.9)^2 \\ &k = \begin{cases} 6.4\mathrm{x}10^6 & x_A \ge 0.9 \\ 0 & x_A < 0.9 \end{cases} \end{aligned} \tag{2.115}$$

The nominal parameter values for this model were identical to those used by Bernardo *et al.* (2001) and are given in Table 4.2

Table 4.2 Deterministic Parameters for the RHE model

c_1	Cost per reactor volume	691.2 $\$\cdot yr^{-1}\cdot m^{-3}$
c_2	Cost per heat exchange area	873.6 $\$\cdot yr^{-1}\cdot m^{-2}$
c_3	Cost of coolant fluid	1.76 $\$\cdot yr^{-1}\cdot kg^{-1}\cdot s$
c_4	Cost of pumping	7.056 $\$\cdot yr^{-1}\cdot kmol^{-1}\cdot hr$
E/R	Activation energy/gas constant ratio	555.6 K
ΔH	Molar heat of reaction	-23260 $kJ\cdot kmol^{-1}$
C_p	Heat capacity of the reactant	167.4 $kJ\cdot kmol^{-1}\cdot K^{-1}$
C_{pw}	Heat capacity of Cooling water	4.184 $kJ\cdot kg^{-1}\cdot K^{-1}$

It was also assumed that the initial concentration, C_{A0}, the inlet flow rate, F_0, and the temperature of the feed, T_0, will fluctuate due to constant instability of an upstream process. It is also known that the temperature of the cooling water, T_w, will fluctuate over time. Variability in T_0 and T_w was characterized by a normal distribution, $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu}$ is a vector of means and $\boldsymbol{\Sigma}$ is the covariance matrix. Uncertainties about the values of the pre-exponential constant in the Arrhenius expression and the heat transfer coefficient were also present; these were also described by a normal distribution. Therefore, the vector of uncertain parameters, $\boldsymbol{\theta}$, contained four variables and two internal parameters, which presented sources of process variability/uncertainty. Nominal values of these terms are given with their corresponding standard deviations, σ, in Table 4.3

Table 4.3 Uncertain parameters for the RHE model

	Parameter	*Units*	μ	σ
F_0	Flow rate of the feed	kmol/hr	45.36	2.93
T_0	Temperature of the feed	K	333	4.31
Tw_1	Cooling water inlet temperature	K	293	3.79
k_R	Arrhenius rate constant	hr^{-1}	12	0.77
U	Overall heat transfer coefficient	$kJ.m^{-2}.h.\ K$	1635	105.82
C_{A0}	Initial concentration at the feed	$kmol.m^{-3}$	32.04	2.07

An increase in the feed temperature, T_0, is accompanied by an increase in the inlet temperature of the cooling water, F_w. Therefore, a positive correlation between T_0 and T_w of 0.7 was imposed. The covariance matrix, Σ, of all uncertain parameters was

$$\boldsymbol{\Sigma} = \begin{bmatrix} 8.58 & 0 & 0 & 0 & 0 & 0 \\ 0 & 18.57 & 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 14.36 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.59 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11193 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4.28 \end{bmatrix} \tag{2.116}$$

4.4.1 Degrees of freedom analysis

This system consists of 6 equations (Equations 4.13 through 4.18), 1 objective function (Equation 4.19), 8 deterministic parameters (c_1, c_2, c_3, c_4, E/R, ΔH, C_p, and C_{pw},), 6 uncertain parameters (F_0, T_0, Tw_1, k_R, U and C_{A0}) and 10 variables (A, V, F_1, F_w, X_A, T_1, T_2, T_{w2}, ΔT_{lm}, C_A,). The degrees of freedom analysis indicates that there are 4 free variables which can be used in the optimization. These can be further classified as design variables, involving equipment sizes, $\mathbf{d} = [A, \quad V]$, and control variables, $\mathbf{z} = [F_1, \quad F_w]$. The problem involves 6 equations that when solved determine value of 6 state variables. Conversion is the variable that describes the quality of the process. Therefore, $y = X_A$ and the remaining variables form the state vector $\mathbf{x} = [T_1, \quad T_2, \quad T_{w2}, \quad \Delta T_{lm}, \quad C_A]$.

4.4.2 Deterministic optimization

In order to demonstrate the traditional approach to designing chemical plants, a deterministic optimization was performed using the nominal values of the parameters. In addition, the quality constraint was reformulated as a "hard" constraint. The results are shown in Table 4.4.

Table 4.4 Values of the decision variables and cost for the deterministic case

A (m^2)	5.3
V (m^3)	4.4
F_1 (kmol/hr)	75.2
F_w (kg/s)	4177
Xa	0.9
Capital Costs ($/year)	4,347
Operating Costs ($/year)	7,883
Quality Costs ($/year)	0
Overall cost ($/year)	12,230

The deterministic approach provides a solution with an exchange area of 5.34 m^2 and a reactor volume of 4.42 m^3. It also avoids consideration of quality costs since the quality constraint is imposed as a hard constraint. On the other hand, the solution suggests operating the reactor at the maximum allowable temperature in order to maximize the conversion. Clearly, this deterministic optimization assumes that the parameter values will remain constant during the entire time of operation of this process. In order to prevent situations that may compromise its operation, designers apply over-design factors to the design variables to account for unexpected changes in the parameter values. This action results in a higher expected cost when designing and operating this process. For instance, applying a 15% over-design factor to the design variables and optimizing with respect to the operating variables, the total expected cost would result in an increase of 3.69 % of the original expected cost shown in Table 4.5.

Table 4.5 Expected cost when a 15% over-design factor has been applied

A (m^2)	6.09
V (m^3)	5.06
F_l (kmol/hr)	60.81
F_w (kg/s)	4272
Xa	0.91
Capital Costs ($/year)	4,733
Operating Costs ($/year)	7,949
Quality Costs ($/year)	0
Overall cost ($/year)	12,682

Again, no quality costs were considered since the quality constraint is satisfied. However, this design strategy does not guarantee that the overall cost of this process will be $12,682. The variability in input variables and uncertainty in model parameters will be transmitted to the operating costs causing a change in the location of the optimum. It is therefore, necessary to asses the effect of variability and uncertainty on the overall cost of the process. By considering the uncertainty and variability in the process design formulation, the problem of designing the process becomes a stochastic optimization problem. In the next section, the results obtained by solving a single-stage stochastic optimization problem are presented.

4.5 Stochastic optimization results

This section presents the results of the stochastic optimization formulation. Using a single-stage stochastic formulation, both the penalty approach and the penalty-plus-explicit-restriction-of-statistical-metrics are used to assess the effect of uncertainties in the process design.

4.5.1 Penalty Approach

Solution of the optimization problem using the penalty approach made use of the GAMS software (Brook *et al.*, 2005) using CONOPT3 as the specified solver. The quality constraint was represented by the asymmetric loss function as indicated in Equation 4.27. The first step in this endeavour was to determine the appropriate sample size to obtain an accurate estimate of the overall expected cost for this process. The parameter space was discretized using sample sizes between 50 and 1000 using HSS and a single-stage stochastic optimization was performed for each set of samples. A quality constant of 6.4 x 10^6 \$ was used. The estimation of the overall expected cost was also performed using specialized cubature formulas (Bernardo *et al.*, 1999) which are presented in Appendix A.

Figure 4.2 shows the evolution of the approximation of the expected cost as samples of larger sizes are used. Using sample sizes of 200 or more one can obtain an estimate of the overall expected cost within a 1% error band of the true mean.

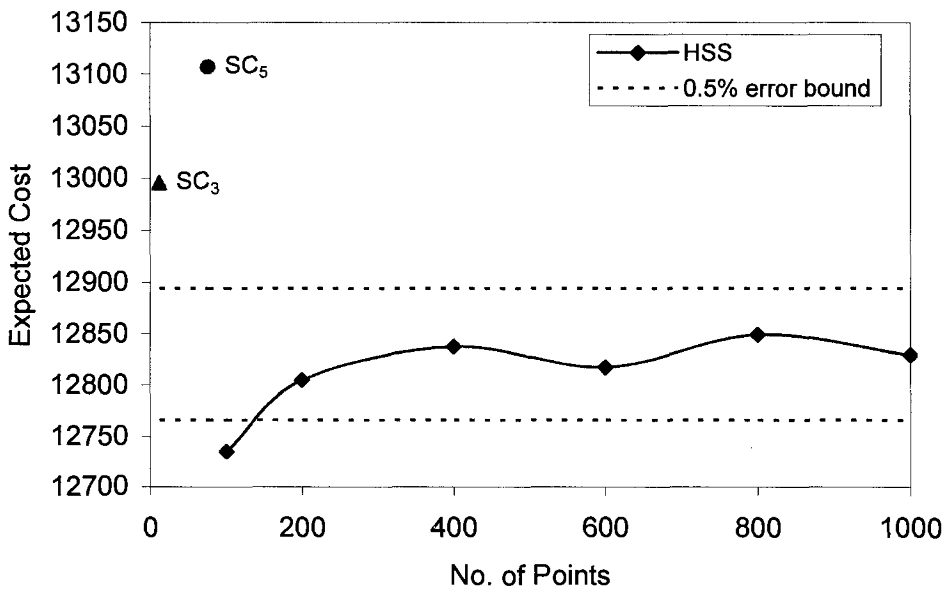

Figure 4.2 Estimation of expected cost using HSS and cubatures

The true mean was considered to be that obtained with a large sample size (i.e., 1000). Figure 4.2 also shows that values of the overall expected cost obtained using cubatures of degree 3 and 5 fell outside the specified error band.

The resulting process designs, using both HSS and cubatures, are presented in Table 4.6. Due to the non-convexity of the problem, the optimization algorithm could yield local minima. Therefore, the optimization was performed for different starting points. Although this tactic may not guarantee a global optimum, it is a simple and practical technique that increases the probability of obtaining the global solution. For this design problem, it was found that the use of HSS lead to the same solution obtained when different starting points were used. In contrast, cubatures led to different local minima. This implies that the use of cubatures may increase the non-convexity of the problem which is a disadvantage. Accordingly, for the case of cubatures, only the smallest solution (considered to be the global optimum) is reported.

Expected values of the overall, operating and quality costs are shown in Table 4.6 as well. Estimates of the overall expected cost obtained with HSS approached the true value as more points were used. Optimal values for the design variables changed very little, except for the case when 46 points were used and a relative lower exchange area was obtained (5.94 m^2). Estimates of the expected operating costs as well as their standard deviations approach the true values, as obtained with a very large sample size, as more points were used. Only 200 points with a computational time of 30 CPU seconds were necessary to achieve the required accuracy in the estimation of the overall expected cost. As more sample points were used the computation time increased. This was expected since as more points are used, more function evaluations are carried out.

Table 4.6 Optimal values of the design variables

	Np	E[Cost][a] [$/year]	σ[Cost] [$/year]	UB[b] [$/year]	LB[b] [$/year]	A[c] [m^2]	V[c] [m^3]	μ(X_A)	σ(X_A)	Capital Cost [$/year]	E[OpCost] [$/year]	σ[OpCost] [$/year]	E[QCost] [$/year]	σ[QCost] [$/year	CPU[d] (s)
HSS[e]	46	12650	1561	12894	12766	5.94	4.74	0.90	0.010	4598	7864	1434	188	534	2.95
	100	12735	1595	12894	12766	6.13	4.69	0.90	0.010	4630	7942	1549	163	414	3.00
	200	12805	1693	12894	12766	6.08	4.70	0.90	0.010	4623	7986	1567	196	594	29.59
	250	12801	1707	12894	12766	6.08	4.67	0.90	0.010	4615	7999	1595	187	544	45.46
	600	12818	1784	12894	12766	6.09	4.66	0.90	0.010	4612	8023	1665	183	539	556.70
	800	12849	1812	12894	12766	6.17	4.66	0.90	0.010	4633	8026	1698	190	594	975.97
	1000	12830	1772	12894	12766	6.09	4.66	0.90	0.010	4612	8031	1657	187	571	995.63
SC_3	12	12996	--	12894	12766	5.87	4.69	0.91	0.010	4567	8272	--	158	--	0.76
SC_5	76	13107	--	12894	12766	6.46	4.68	0.90	0.013	4710	8670	--	385	--	9.47

a. The objective tolerance used by CONOPT was the default value ($3e^{-13}$). The feasibility tolerance (i.e., maximum constraint violation) was set to $1e^{-6}$.
b. Upper and lower bounds constructed considering a relative error of 0.5% around the HSS solution with 1000 points
c. The specified tolerances for the design variables were set to $1e^{-3}$.
d. CPU is the time in seconds that has been used by the optimizer to arrive at the optimum.
e. The optimization procedure using the HSS led to the same optimum for all starting values. For the optimization procedure using cubatures, only the lowest minimum is reported.

The last two rows of Table 4.6 show the results obtained with SC of degrees 3 and 5. Unlike the use of the HSS approach, estimation of the uncertainty in the overall cost was not possible since, to our knowledge, a weighted formula capable of estimation of standard deviations is not available. The same applies to the estimation of the standard deviations for the operating and quality costs. This may represent a limitation, for instance, in planning problems where robustness in operating costs has to be ensured and quantified (Ahmed and Sahidinis, 1998). Despite their over-estimation in the overall expected cost, SCs provided optimal values for the design variables which were close to those obtained with HSS using 180 points.

Optimal values for the mean and standard deviation of the quality characteristic converged to the same values as those obtained with HSS. On the other hand, one can envision the benefits of using SCs since they demand a low computational time (0.76 and 9.47 CPU seconds). This was expected since these formulas are specifically constructed to integrate functions with the least possible number of points.

Table 4.7 shows the optimal values of the control variables. Using HSS, more accurate estimates of the expected values were obtained as more points were used. The same applies to estimates of standard deviations. On the other hand, SCs overestimated the expected values and standard deviations of the control variables.

Table 4.7 Optimal values of the control variables

	Np	***$E[F_w]$* [kg/s]**	***$Sdev[F_w]$* [kg/s]**	***$E[F]$* [kmol/hr]**	***$sdev[F]$* [kmol/hr]**
HSS	46	4207	739	65.11	22.35
	100	4263	807	62.37	22.00
	200	4283	816	63.53	22.31
	250	4290	829	63.69	23.20
	600	4303	865	63.72	23.73
	800	4311	884	62.26	23.60
	1000	4307	861	63.83	23.70
SC_3	12	4406	---	73.36	---
SC_5	76	4655	---	67.50	---

From the above discussion we can conclude that although specialized cubatures require lower computational time, they tend to over estimate the expected value of objective function. The same happens for the estimation of the expected values and standard deviations of the control variables. This is the major limitation since we are interested in obtaining estimates with sufficient accuracy. In contrast, HSS, although at a cost of a modest computational effort, provides more accurate estimates of the objective function compared to that obtained with a large sample size and control variables. Since we are working with models at steady state, the computational time is not crucial. Therefore, HSS seem to be the best choice for numerical integration.

Distributions of the control variables, the reactant flow rate to the heat exchanger, F, and the cooling water flowrate, F_w, are shown in Figure 4.3 These plots represent distributions of the optimal values of the control variables for which the effect of uncertainty in the parameters and variability in process variables is minimized. The distributions are somewhat skewed towards lower values.

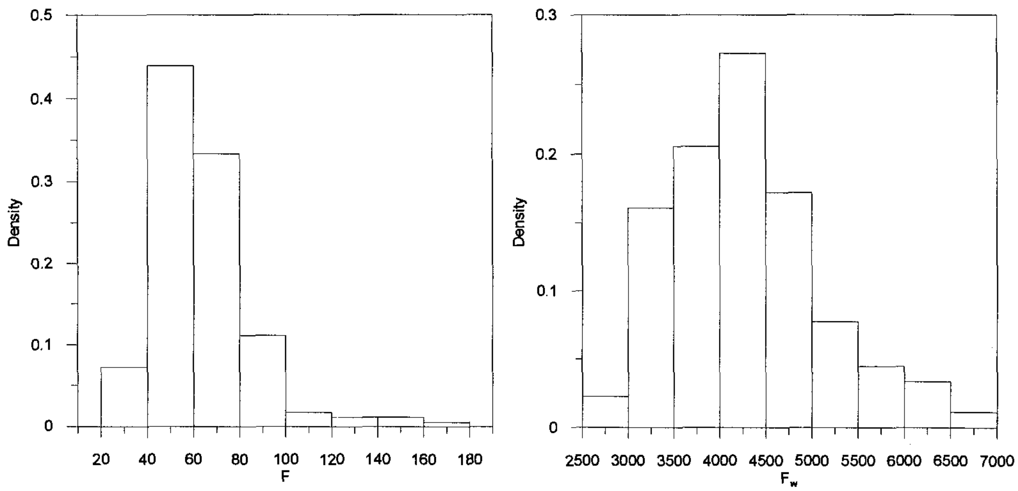

Figure 4.3 Distributions of optimal values of control variables F and F_w.

As shown in Table 4.7 the estimated optimal nominal value of the reactant flow rate was 63.53 kmol/hr with an standard deviation of 22.31 kmol/hr. The estimated coefficient of variation ($\hat{\sigma}/\hat{\mu}$) was 0.35. On the other hand, the resulting distribution of the cooling water, F_w, was less skewed with values tending to have relatively less variability (coefficient of variation = 0.19). The estimated optimal mean value was 4283 kg/sec with standard deviation of 816 kg/sec.

Due to the fact that the Taguchi loss constant, k, is difficult to obtain, especially for new products or processes, the design was carried out for three different values of k: 5.3×10^4, 1.5×10^5, 3.1×10^7. Distributions of the quality variable, x_A, for each value of k are shown in Figure 4.4.

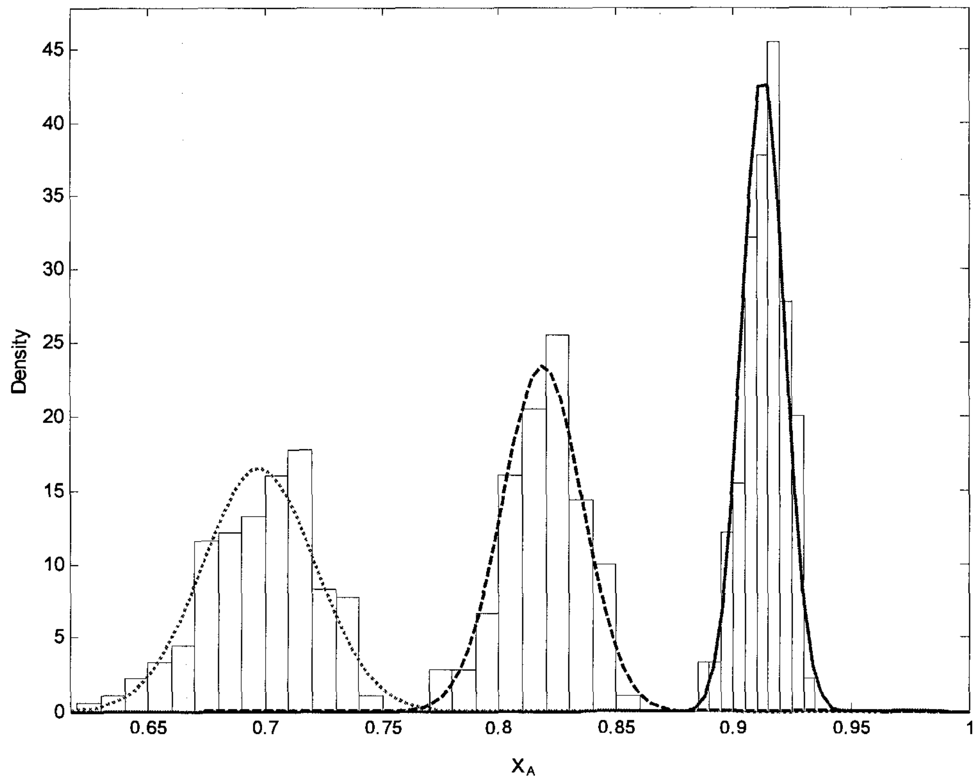

Figure 4.4 Distribution of the quality characteristic for different k values (solid line, $k = 3.1 \times 10^7$ \$; dash line, $k = 1.5 \times 10^5$ \$, dotted line, $k = 5.3 \times 10^4$ \$)

As expected, as the value of k increased, and consequently the quality loss, the optimal operation of the process moves to a region of higher conversion (>0.9) where the objective function is not penalized (see Table 4.8). The distribution of the quality characteristic was narrower when a high quality loss constant was involved. This was also expected since one of the terms of the Taguchi loss function is the variance of the quality characteristic and therefore minimizing the loss function, resulted in minimization of variance.

On the other hand, as k decreased the quality costs increased (see Table 4.8) and the optimal operation lies in regions of lower conversion. The variability of the quality characteristic increases but this has almost no impact on the overall cost due to low values of k. Figure 4.4 shows that the mean and variance are inversely related; as the variance is decreased the mean is increased.

Table 4.8 provides the expected costs for these scenarios along with corresponding optimal values of the design variables. For the case in which the value of the loss constant, k, is high, a high overall cost is expected. This is the result of two factors. First, there is a high capital cost due to larger equipment sizes. The reactor volume must be larger to provide high enough residence time to achieve conversion levels to meet the quality constraint. Due to larger capacity in the reactor, a large amount of reactant is converted into products generating a considerable amount of heat. As a result, cooling requirements increase and these are met by increasing the heat exchanger area. Second, there is also a high operating cost associated with handling large flow rates. With higher k values, most product meets specification. This is expected since as more cost elements are included in k (e.g., costs for downstream separation, costs for reprocessing low specification product), penalties for violating the quality constraint become higher thus increasing the overall cost. On the other hand, when the loss constant, k, takes lower values, the optimal values of the design and control variables are lower thus reducing the capital and operating costs and in consequence, the overall expected cost. This would imply that, if no additional costs, such as those previously mentioned, were included in k, the quality constraint would be violated increasing the quality costs. However, its impact on the overall cost would be compensated by lower capital and operating costs.

Table 4.8 Optimal values for three different *k* values

k	**5.3×10^4**	**1.5×10^5**	**3.1×10^7**
E[Cost] ($/year)	9603	11199	13067
Capital Cost	2673	3546	4806
E[Op Cost]	4722	6620	8126
E[Q Cost]	2208	1036	134
A (m^2)	3.69	5.11	6.18
V (m^3)	1.15	2.25	5.23
E[F_w](kg/s)	2535	4363	4358
E[F](kmol/hr)	37	52	65
μ_{XA}	0.7	0.82	0.91
σ_{XA}	0.024	0.017	0.009

This analysis has shown how the mean value of *k* affects the design. As the quality cost becomes an important part of the design, as indicated by the value of *k*, the penalty approach guarantees that the design does not violate quality constraints and ensures that quality costs are kept minimal. On the other hand, it does not provide any insight in how the uncertainty in *k* affects the design. In the next section a sensitivity analysis that studies the level of uncertainty in *k* is presented. The quality constant is considered an additional uncertain parameter following a normal distribution.

4.5.1.1 Sensitivity analysis

A single-stage stochastic optimization was performed considering the loss constant, *k*, as an uncertain parameter. As a result the parameter vector contained seven uncertain parameters with their corresponding errors. Error levels were defined to represent low, intermediate and high magnitudes of uncertainty. These were 0.32%, 16% and 30% for the three nominal values of *k* used previously. Table 4.9 to Table 4.11 show the results for these scenarios. 2σ limits are reported to provide a sense of the precision attained on the estimates of the expected costs. For the case when the loss constant, *k*, was small (Table 4.9), the overall expected cost decreased as the level of error increased. This decrease was, however, insignificant. The precision of the estimates of the expected costs as well as the optimal values of the design variables displayed no significant change.

Table 4.9 Error levels studied for $\mu_k = 5.3\text{x}10^4$

σ x 10^4	**0.00**	**0.17**	**0.85**	**1.59**
E(cost) ($/year)	9603 ($\pm$ 2434)	9601 ($\pm$ 2438)	9598 ($\pm$ 2562)	9587 ($\pm$ 2836)
Design Cost	2673	2672	2669	2665
E[Op Cost]	4722 ($\pm$ 2424)	4720 ($\pm$ 2420)	4711($\pm$ 2422)	4702 ($\pm$ 2424)
E[Q Cost]	2208 ($\pm$ 1064)	2209 ($\pm$1080)	2215 ($\pm$ 1308)	2221 ($\pm$ 1766)
A (m^2)	3.69	3.69	3.68	3.68
V (m^3)	1.15	1.15	1.14	1.14
μ_{XA}	0.70	0.70	0.70	0.70
σ_{XA}	0.024	0.024	0.024	0.024

For the case in which the loss constant took intermediate values (Table 4.10), the same trend seemed to follow. As the error level in k increased, the overall expected cost decreased very little with no significant change in the precision of the estimates of the expected costs. The last scenario (Table 4.11) contrasts the two previous ones. As the error level in k increased, the overall expected cost increased. This increase was, again, not significant. As occurred for the above two scenarios, optimal values of the design variables as well as the precision of the estimates of the expected costs remained almost unchanged. The insensitivity of the quality cost to the error level in k can be attributed to the asymmetric nature of the penalty function since only realizations of the quality variable that fall below the target value penalize the objective function.

Table 4.10 Error levels studied for $\mu_k = 1.5\text{x}10^5$

σ x 10^5	**0.00**	**0.016**	**0.079**	**0.15**
E(cost) ($/year)	11199 ($\pm$ 2866)	11199 ($\pm$ 2870)	11196 ($\pm$ 2906)	11193 ($\pm$ 2976)
Design Cost	3546	3543	3541	3539
E[Op Cost]	6620 ($\pm$ 2752)	6619 ($\pm$ 2752)	6616 ($\pm$ 2752)	6613 ($\pm$ 2750)
E[Q Cost]	1036 ($\pm$ 878)	1037($\pm$ 884)	1039 ($\pm$ 964)	1041 ($\pm$ 1134)
A (m^2)	5.11	5.11	5.11	5.10
V (m^3)	2.25	2.25	2.24	2.24
μ_{XA}	0.82	0.82	0.82	0.82
σ_{XA}	0.017	0.017	0.017	0.017

Table 4.11 Error levels studied for $\mu_k = 3.1\text{x}10^7$

σ x 10^7	**0.00**	**0.10**	**0.50**	**0.93**
E(cost) ($/year)	13067 ($\pm$ 3484)	13068 ($\pm$ 3490)	13073 ($\pm$ 3510)	13078 ($\pm$ 3538)
Design Cost	4806	4807	4809	4812
E[Op Cost]	8126 ($\pm$ 3178)	8127 ($\pm$ 3178)	8128 ($\pm$ 3178)	8130 ($\pm$ 3178)
E[Q Cost]	134 ($\pm$1410)	134 ($\pm$ 1412)	135 ($\pm$ 1430)	135 ($\pm$ 1458)
A (m^2)	6.18	6.18	6.18	6.19
V (m^3)	5.23	5.23	5.24	5.25
μ_{XA}	0.91	0.91	0.91	0.91
σ_{XA}	0.009	0.009	0.009	0.009

Identification of the critical uncertain parameters that contribute the most to the overall expected cost was performed. Interaction effects were neglected. Each uncertain parameter was eliminated from the parameter vector one at a time and another stochastic optimization was carried out. Figure 4.5 the results obtained.

The uncertain parameter that contributes the most in reducing the overall expected cost is the cooling water inlet temperature, T_{wI}. By eliminating its variability and keeping it on target the overall expected cost was reduced from $12805 to $12625 per year. This corresponds to a 1.4% reduction of the original expected cost. The second uncertain parameter that was identified is the inlet concentration, C_{A0}. Keeping its mean value on target (i.e., 32.04 kmol/m^3) and eliminating its variability led to an overall expected cost of $12677 per year. Also, the feed temperature and feed flow rate, T_0, and F_0 reduced the overall expected cost. Keeping them on target values led to $12694 and $12695 per year respectively. Considering the hypothetical situation of investing research effort on finding the true value of the Arrhenius pre-exponential constant would also lead to lower expected cost ($12694 per year). Also, exact knowledge of the overall heat transfer coefficient would reduce the overall expected cost from $12805 to $12730 per year.

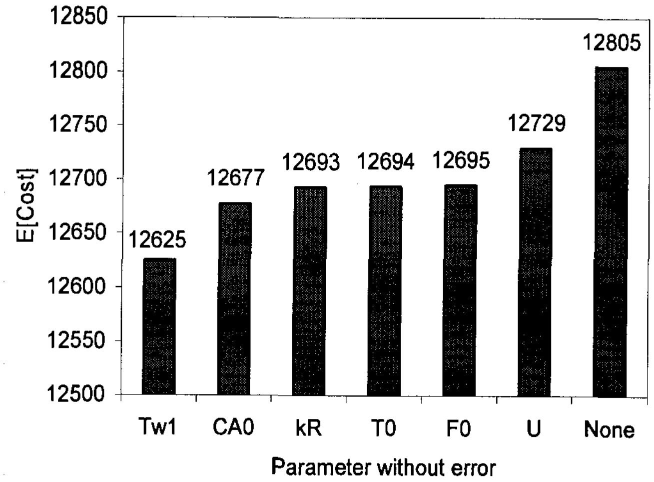

Figure 4.5 Overall expected costs after eliminating the uncertainty in each parameter

The effect of the number of uncertain parameters on the overall expected cost was also studied. Uncertain parameters that contributed the most to reducing the overall cost were eliminated sequentially and stochastic optimizations were performed. Figure 4.6 shows the effect of eliminating the uncertainty in those parameters. As the number of uncertain parameters decreased, the overall expected cost also decreased. This decrease in the overall expected cost can be explained with the help of Table 4.12

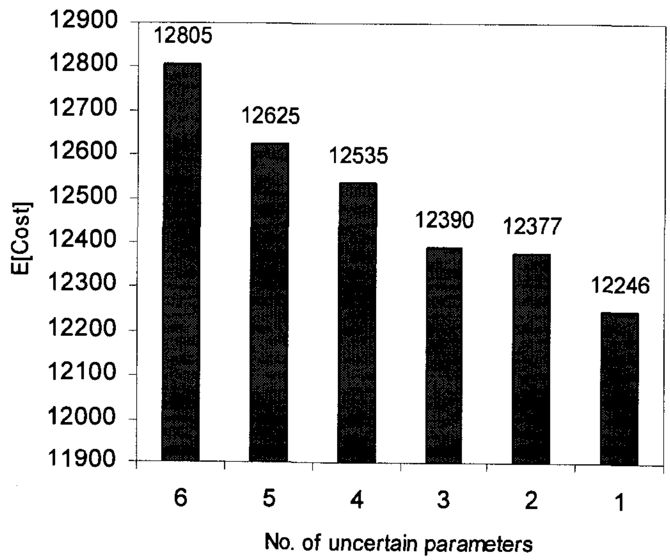

Figure 4.6 Effect of the number of uncertain parameters on the overall expected cost

Quality costs decreased because the variability in the quality characteristic was decreased. As the variabilities in the cooling water, Tw_1, initial concentration of the feed, C_{A0}, and feed temperature, T_0, were removed, the variability in the quality characteristic decreased further. This decrease in variability of the inputs translated into smaller sizes of equipment thus reducing the capital cost. This was expected since higher volumes and areas provide a damping effect on the variability of the conversion. Therefore, by having less variability in the system, smaller sizes of equipment were required. The elimination of variability also resulted in lower flow rates for the cooling water and recycling mixtures which decreased the operating costs. Lower flow rates were used to compensate the input variation.

Table 4.12 Effect of the number of uncertain parameters on the optimal values for the design and control variables.

	No. of uncertain parameters					
	6	5	4	3	2	1
E[Cost]	12805	12625	12535	12390	12377	12246
Capital Cost	4623	4580	4543	4495	4448	4395
E[Op Cost]	7986	7864	7839	7780	7813	7807
E[Q Cost]	196	181	153	115	116	44
A	6.08	5.99	5.98	5.93	5.75	5.67
V	4.70	4.64	4.53	4.41	4.41	4.30
μ_{X_A}	0.904	0.903	0.902	0.900	0.900	0.897
σ_{X_A}	0.010	0.009	0.008	0.006	0.006	0.005
μ_F	64	64	64	64	66	67
σ_F	22	21	20	19	15	12
μ_{F_w}	4283	4212	4199	4166	4173	4168
σ_{F_w}	816	445	454	365	243	154

In summary, it has been demonstrated that efforts to improve the precision of the loss constant, *k*, will not significantly change values of the overall cost. When the loss constant bears low and intermediate values, the overall cost appears to decrease as the error level in *k* increases. This decrease is, however, not significant. On the other hand, when the loss constant attains high values and as the level of error in *k* increases, slightly higher expected costs are obtained.

On the other hand, eliminating the variability in key uncertain variables does lead to decreasing the overall expected cost. By eliminating the variability in T_{w1}, a decrease in the overall cost of 1.4 % is expected. Moreover, a further decrease can be obtained by eliminating the variability in the variables, T_0, F_0 and C_{A0}, and the uncertainty in the pre-exponential rate constant and the heat transfer coefficient. This has the impact of reducing the equipment sizes thus decreasing capital costs. It also reduces the values for the control variables and minimizes operating costs.

Concerning the formulations used for the stochastic optimization problem, the penalty approach through the relaxation of the quality constraints avoided infeasibilities, typically faced in the two-stage approach. No difficulties were encountered in finding optimal solutions for the design problem which facilitated the performance of the sensitivity analysis on the level of uncertainty of the parameters. The penalty approach should be employed when the distribution type of the quality characteristic and when degree of variation are of no interest for the designer. When specific distribution properties of the quality characteristic are part of the design objectives, a restriction approach is more appropriate.

4.5.2 Penalty approach plus explicit restriction on robustness metrics

The optimization problem of restricting the variability in the quality characteristic was also formulated using Equation 4.12 and a quality constant of 6.4×10^6 \$. The extent of robustness was established by defining a maximum degree of variability for the quality characteristic. Three cases were considered: standard deviations lower than or equal to 0.009, 0.006 and 0.004.

Expected values of the objective with corresponding optimal solutions are shown in Table 4.11. Note that the overall expected cost in the first scenario was the lowest of the three cases, though there were losses in quality incurred (\$30). Quality losses were incurred because a portion of the distribution lied in a region lower than the target value (see Figure 4.6). Low capital costs were obtained because of relatively small sizes of equipment (i.e., 6.11 m^2 for heat exchanger area and 5.2m^3 for reactor volume). Low operating costs resulted from low expected flow rates in the control variables.

For the case in which a lower variability was required (i.e., $\sigma_{XA} \leq 0.006$), the overall expected cost increased. The optimal solution was found where design variables attained higher values (i.e., 6.35 m^2 for heat exchanger area and 8.19 m^3 for reactor volume). Operating costs also increased due to the increase in expected flow rates (4567 kg/s of cooling and 75 kmol/hr of recycling reactant as shown in Table 4.11). Finally, there was no quality cost involved since the region of operation of this process lied outside the region of penalization as shown in Figure 4.7.

When the highest robustness was required (i.e., $\sigma_{XA} \leq 0.004$) the overall expected cost increased further. Again no quality losses were involved due to satisfaction of the quality constraint. Although there was no quality costs, capital costs increased the overall expected cost dramatically. This was the result of a very large equipment size obtained (i.e., 6.56 m^2 for heat exchanger area and 12.76 m^3 for reactor volume). Operating costs also contributed to the overall expected costs. This was the result of higher expected values of the control variables.

Table 4.13 Optimal solutions for different levels of robustness

Robustness Criteria	$\sigma_{X_A} \leq$ **0.009**	$\sigma_{X_A} \leq$ **0.006**	$\sigma_{X_A} \leq$ **0.004**
E[Cost]) ($/year)	12984	14226	15714
Design Cost	4780	5659	6808
E[Op Cost]	8175	8567	8906
E[Q Cost]	30	0	0
A (m^2)	6.11	6.35	6.56
V (m^3)	5.20	8.19	12.76
E[Fw] (kg/s)	4361	4567	4743
E[F] (kmol/hr)	71	75	79
μ_{X_A}	0.91	0.94	0.96
σ_{X_A}	0.009	0.006	0.004
$P(X_{Ai} \leq 0.9)$	0.13	0	0

The distribution of the quality characteristic for different degrees of robustness is shown in Figure 4.7. As can be observed, the lower the variability required on the conversion, the higher its mean becomes. Conversely, as the variability on the conversion is relaxed, a lower mean is obtained. As discussed previously, the mean and variance are dependent on each other, so constraining one will cause an inverse effect on the other.

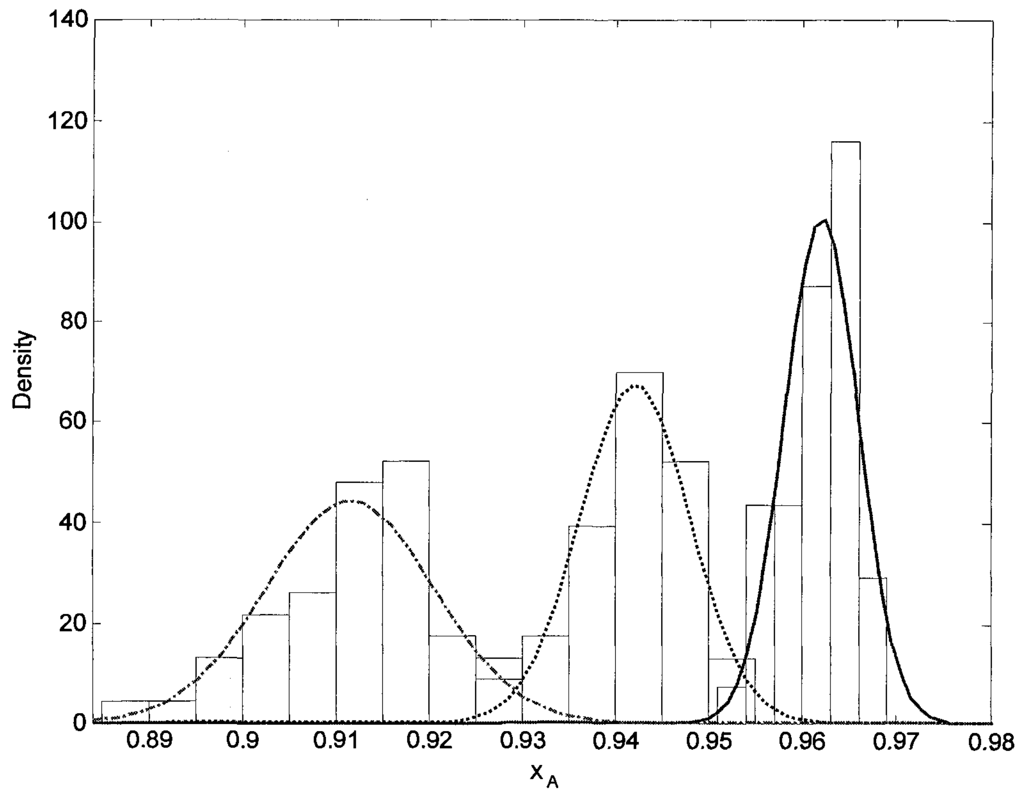

Figure 4.7 Distribution of the quality variable for different extent of robustness (dashed-dot line, $\sigma \leq 0.009$; dotted line, $\sigma \leq 0.006$; solid line, $\sigma \leq 0.004$).

4.5.2.1 Sensitivity analysis

As done for the penalty approach, a sensitivity analysis was performed to determine which parameter could reduce the overall expected cost by eliminating its uncertainty/variability. Each uncertain parameter was eliminated from the parameter vector one at a time and the optimization was performed for the case in which a minimum standard deviation in the quality characteristic of 0.004 was required. Figure 4.8 shows the effect of eliminating the uncertainty in each parameter.

By eliminating the uncertainty in the Arrhenius pre-exponential constant, k_R, the overall expected cost was reduced from \$15714 to \$14534 per year. This corresponded to a 7.5% reduction of the original expected cost. Also, eliminating the variability of the feed flow rate, F_{A0}, would lead to an overall expected cost of \$14727 per year. If variability in the cooling water inlet temperature, T_{wq}, and the inlet concentration of the feed, C_{A0}, were removed, this could reduce the overall expected cost to \$14814 and \$5088 pear year, respectively.

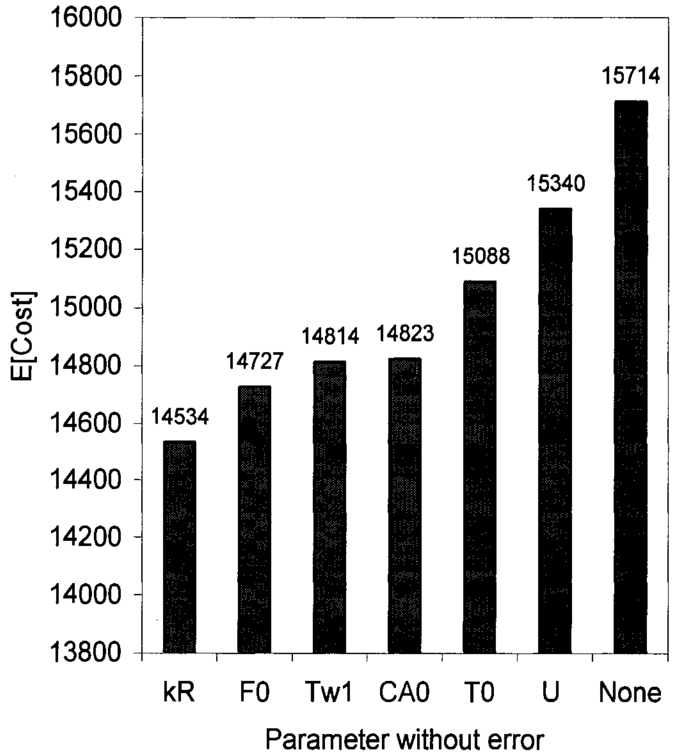

Figure 4.8 Overall expected costs after eliminating the uncertainty in one parameter for $\sigma \leq 0.004$ case.

Compared to the sensitivity analysis for the penalty approach, different uncertain parameters had larger effects on the overall costs. For this specific problem, the uncertain parameter that increases the overall cost is the cooling water inlet temperature whereas it is the Arrhenius pre-exponential constant when the objective is a minimum overall cost with a restricted variability in the quality characteristic.

As shown, one can incorporate the penalty approach, explicit restrictions on robustness metrics or both approaches simultaneously depending on the type of problem under consideration. If a designer wants a design that leads to a quality variable with a specified maximum standard deviation and at the same time to obtain estimates of the costs for violating quality constraints, he can achieve his objective by using the penalty approach with an explicit restriction on variance. On the other hand, there may be cases in which values for the design and operating variables that satisfy the robustness constraint imposed may not exist. This would represent a limitation since an infeasible solution could not be obtained. In those cases only the formulation using the penalty approach should be used.

4.6 Conclusions

The robust design of a chemical process was carried out using a one-stage stochastic optimization approach. Two formulations, a penalty approach and an explicit restriction on the robustness of the quality characteristic, were used to identify the uncertain parameter/variable that lead to process improvement. The potential advantages and disadvantages of the two single-stage formulations were also discussed. The overall expected cost involving seven uncertain parameters characterized by normal distributions was estimated by Hammersley Sequence Sampling (HSS) and specialized cubatures of degree 3 and 5. It was shown that, at a modest computational cost, HSS provided more accurate estimates than SC. It was also shown that reducing the variability in key variables and reducing the uncertainty of key parameters results in a lower overall expected cost. Finally, it was found that the uncertainty in the quality constant does not have a significant impact on the overall expected cost.

Acknowledgments

L. D. gratefully acknowledges support from the Mexican Government through the Consejo Nacional de Ciencia y Tecnología, CONACyT.

Appendix: Specialized Cubatures

Bernardo *et al.*, (1999) proposed different types of quadratures and cubatures to integrate the uncertainty region characterized by normally distributed parameters. Based on the decision rules presented in their work, cubatures that satisfy the following conditions were chosen:

1. All discrete points must lie within the parameter space.
2. All weights have to be positive.

Only two formulas satisfied the above criteria: the specialized cubatures degree 3 and 5 (Stroud, 1971).

To integrate the n-dimensional parameter space, these formulas make use of a weight function of the form

$$w_c(u) = \exp(-u^T u) \tag{A.1}$$

and the transformation

$$\theta(u) = \mu + I_{\sqrt{2}} \Sigma^{1/2} u \tag{A.2}$$

where $I_{\sqrt{2}}$ represents a diagonal matrix with all the diagonal elements equal to $\sqrt{2}$. The transformation is done to map the region R_n into the uncertainty space $\Theta = \{\theta : \theta \sim N(\mu, \Sigma)\}$

Applying this transformation to Equation 4.5, the expected value of the objective function is given by

$$E_\Theta(f) = \frac{1}{\pi^{n/2}} \int_{-\infty}^{+\infty} \ldots \int_{\infty}^{+\infty} f[\theta(u)] w_c(u) du \approx \frac{1}{\pi^{n/2}} = \sum_{i=1}^{Np} B_i f(\theta(u_i))] \tag{A.3}$$

A.1 Specialized cubature of degree 3 (SC_3)

For the SC_3, the grid of points, u_i, are constructed by a fully symmetric permutation of the elements and signs of a set of points

$$(r, 0, \ldots, 0)_{FS} \qquad B_0 \tag{A.4}$$

The parameter, r, and weights, B_0, are functions of the n-dimensional volume $V = \pi^{n/2}$.

$$r^2 = \frac{n}{2} \tag{A.5}$$

$$B_0 = \frac{1}{2n} V \tag{A.6}$$

This formula has N_p=2n points.

A.2 Specialized cubature of degree 5 (SC_5)

For the SC_5 the grid of points, u_i, are constructed the same way as in the SC_3, with the difference of the introduction of an additional parameter, s. Finally, the grid of points u_i are given by

$$(r, 0, \ldots, 0)_{FS} \qquad B_0 \tag{A.7}$$

$$(s, s, \ldots, s)_{FS} \qquad B_1 \tag{A.8}$$

The parameters r and s are given by

$$r^2 = \frac{n+2}{4} \tag{A.9}$$

$$s^2 = \frac{n+2}{2(n-2)} \tag{A.10}$$

with weights B_0 an B_1

$$B_0 = \frac{4}{(n+2)^2} V \tag{A.11}$$

$$B_1 = \frac{(n-2)^2}{2^n (n+2)^2} V \tag{A.12}$$

This formula is only defined for a number of uncertain parameters more than or equal to 3 and has a total number of points given by $N_p = 2^n + 2n$.

Nomenclature

d	Vector of design variables
z	Vector of control variables
x	Vector of state variables
y	Vector of quality variables
h	Vector of equality constraints
g	Vector of inequality constraints
k	Taguchi loss constant
SN	Signal to noise ratio
N_p	Number of sampling points
F_0	Flow rate of the feed
T_0	Temperature of the feed
Tw_1	Cooling water inlet temperature
k_R	Arrhenius rate constant
U	Overall heat transfer coefficient
C_{A0}	Initial concentration at the feed
c_1	Cost per reactor volume
c_2	Cost per heat exchange area
c_3	Cost of coolant fluid
c_4	Cost of pumping
E/R	Activation energy/gas constant ratio
ΔH	Molar heat of reaction
C_p	Heat capacity of the reactant
C_{pw}	Heat capacity of cooling water
V	Dimensional volume of integration
r	Cubature parameter
s	Cubature parameter
B_0	Cubature weight
B_1	Cubature weight 2

Q	Quality variable

Greek letters

$\boldsymbol{\theta}_i$	Vector of uncertain parameters
Θ	Domain of the uncertain parameters
τ	Target value
σ	Standard deviation of individual uncertain parameters
μ	Vector of means of uncertain parameters
Σ	Variance-covariance matrix of uncertain parameters
γ	Robustness metric

Set theory

$\forall$	For all
$\in$	Element of
$\exists$	There exist
D	Set of design variables
X	Set of state variables
Z	Set of control variables
Y	Set of quality variables

Optimization acronyms

min	Minimize
w.r.t	With respect to
s.t.	Subject to

Subscripts

i	Sample point i

j	Constraint function j
q	Quality variable q

References

Acevedo, J., and Pistikopoulos, E. N., "Stochastic optimization based algorithms for process synthesis under uncertainty", Computers and Chemical Engineering, Vol. 22, 647-671, 1998.

Ahmed, S., and Sahinidis, N. V., "Robust process planning under uncertainty", Industrial and Engineering Chemistry Research, Vol. 37, 1883, 1998.

Bernardo, F., P., Pistikopoulos E. N. and Saraiva P. M., "Quality cost and robustness criteria in chemical process design optimization", Computers and Chemical Engineering, Vol. 21, 27-40, 2001.

Bernardo F., P., Pistikopoulos E., N. and Saraiva P. M., "Integration and computational issues in stochastic design and planning optimization problems", Industrial and Engineering Chemistry Research, Vol. 38, 3056, 1999.

Brook, A., Kendrick, D., Meeraus, A. and Raman, R., "*GAMS* — a user's guide", Washington, DC, 2005.

Chacon-Mondragon, O., L., and Himelblau, D. M. "Integration of flexibility and control in process design", Computers and Chemical Engineering, Vol. 20, 447, 1996

Dantzing, G. B., "Linear programming under uncertainty", Management Science, Vol. 1, 197-20, 1955.

Diwekar, U. M., and Kalagnanam, R. J., "Efficient sampling technique for optimization under uncertainty", AIChE J., vol. 43, No. 2, 1997

Edgar, T. F., Himmelblau, D. M., and Lasdon, L. S., "Optimization of chemical processes", McGraw Hill, New York 2001.

Georgiadis, M. C., and Pistikopoulos, E. N., "An integrated framework for robust and flexible process systems", Ind. Eng. Chem. Res., Vol. 38, 133-143, 1999.

Grossmann, I. E. and Sargent, R. W. H., "Optimum design of chemical plants with uncertain parameters", AIChE J. Vol. 37, 1978.

Halemane, K. P. and Grossmann I. E., "Optimal process design under uncertainty", AIChE J., Vol. 29, 439, 1983.

Montgomery, D. C., "Introduction to statistical quality control", Forth edition, John Wiley & Sons, Inc, 2001.

Morgan, G, and Henrion, M. "Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis". Cambridge University press, New York 1990.

Pai, C. C. D and Hughes R. R., "Strategies for formulating and solving two-stage problems for process design under uncertainty", Computers and Chemical Engineering, Vol. 11, No. 6 pp. 695-706, 1987.

Pistikopoulos, E. N., "Uncertainty in process design and operations", Computers and Chemical Engineering, Vol. 19, Supplement, s553-s563, 1995.

Pistikopoulos, E. N., Ierapetritou, M. G. "Novel approach for optimal process design under uncertainty", Computers and Chemical Engineering Vol. 19, No. 10, 1089, 1995.

Robinson, T. J., Borrow, C. M., and Myers, R., "Robust parameter design: a review", Quality and Reliability Engineering international, Vol. 20, 81-101, 2004.

Ryan, Thomas P. “Statistical methods for quality improvement”, Second Edition, Wiley series in Statistics, 2000.

Stroud, A. H., “Approximate calculation of multiple integrals, Prentice Hall, London 1971.

Straub, D. A. and Grossman I. E., “Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties”, Computers and Chemical Engineering, Vol. 14 No.9 pp. 967-985, 1990.

Taguchi, G., “Introduction to quality engineering”, Asian productivity association, 1986.

Wellons, H.S. and Rekleitis, G. V., “The design of multiproduct batch plants under uncertainty with staged expansion”, Computers and Chemical Engineering. Vol. 13, 115, 1996.

Chapter 5 (Paper 3)

Evaluation of Quasi-Monte-Carlo Sampling Techniques for Stochastic Optimization

L. F. Dominguez and D. D. McLean

Dept. of Chemical Engineering

University of Ottawa

Abstract

This paper compares the performance of several quasi-Monte-Carlo sampling techniques for the solution of process design problems under uncertainty through stochastic optimization. The objective function of this type of problems is considered to be an average performance metric which is usually an overall expected cost to be minimized or an overall expected profit to be maximized. Their solution requires the evaluation of an objective function over the uncertainty region described by the probability density function of the uncertain parameters. Because integration of the objective function over the uncertainty region represents the major computational task in the algorithm, an efficient integration technique that can reduce the number of function evaluations without sacrificing the accuracy in the solution is necessary. In this work four quasi-Monte-Carlo sampling techniques, based on the sequences of Hammersley, Halton, Faure and Sobol, are evaluated through the solution of two engineering process design examples.

Key words: stochastic optimization, probability density function, quasi-Monte-Carlo sampling

5.1 Introduction

Considerable research effort has been focussed on devising more effective integration techniques that can reduce the computational time involved in estimating the expected value of a function, f, of uncertain parameters, $\boldsymbol{\theta}$, with required accuracy. When the uncertainty in such parameters is described by probability density functions (PDF), j, the expectancy operation takes the form of a multidimensional integral given by

$$\mathrm{E}[f(\boldsymbol{\theta})] = \int_{\Theta} f(\boldsymbol{\theta}) j(\boldsymbol{\theta}) \mathrm{d}\boldsymbol{\theta} \tag{5.1}$$

Equation 5.1 is commonly encountered in process design and planning optimization problems under uncertainty where the objective function is an average performance metric, usually a cost function to be minimized or a profit function to be maximized (Pistikopolos, 1995; Acevedo and Pistikopoulos, 1998). The most common, current approach to solve the above problems is through a two-stage formulation. The two stage approach differentiates design and operating variables. Design variables are assumed to be fixed and remain so during the operating phase of the plant whereas the control variables are optimally adjusted to compensate for the uncertainty and variability of model parameters and input variables. The two-stage stochastic formulation, for a minimization case, is given by

$$\begin{array}{l}
\text{Design stage:} \\
\qquad \min \ \ \mathrm{E}\{\mathrm{C}^*(\mathbf{d},\boldsymbol{\theta})\} \\
\qquad \text{w.r.t.} \ \ \mathbf{d} \\
\qquad \forall \boldsymbol{\theta} \in \Theta \ \ \left\{\exists \mathbf{z}\left(\forall \mathrm{j} \in \mathrm{J}[\mathrm{g}_{\mathrm{j}}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta})] \leq 0]\right)\right\} \\
\qquad \mathbf{d} \in \mathrm{D}, \quad \boldsymbol{\theta} \in \Theta \\
\text{Operating stage:} \\
\mathrm{C}^*(\mathbf{d},\boldsymbol{\theta}) = \min \mathrm{C}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) \\
\qquad \text{w.r.t.} \ \ \mathbf{z},\mathbf{x} \\
\qquad \text{s.t.} \ \ \ \mathbf{h}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) = 0 \\
\qquad \qquad \mathbf{g}(\mathbf{d},\mathbf{z},\mathbf{x},\boldsymbol{\theta}) \leq 0 \\
\qquad \qquad \mathbf{z} \in \mathrm{Z}, \quad \mathbf{x} \in \mathrm{X}
\end{array} \tag{5.2}$$

where $\mathbf{d}$, $\mathbf{z}$, and $\mathbf{x}$ represent the vectors of design, control, and state variables with corresponding sets D, Z and X. $\boldsymbol{\theta}$ corresponds to the vector of uncertain parameters, and $\mathbf{h}$ and $\mathbf{g}$ the vector of equality and inequality constraints. C and C* are the operating and minimum costs at the operating stage; J and Θ are the sets of the constraint function and characterization of the uncertainty.

The formulation in Equation 5.2 is hard to solve since the feasible region at the operating stage is implicitly defined wherein it involves the evaluation of the feasible region in both design and operating stages (Pistikopoulos and Ierapetritou, 1995). The above problem can be further simplified with a relaxation strategy that allows the integration region to correspond to the entire uncertainty region described by the PDF of the uncertain parameters (Bernardo *et al.*, 2001). Thus, the simplified problem can be written as

$$
\begin{aligned}
&\min \quad \int_{\boldsymbol{\theta}\in\Theta} C^*(\mathbf{d},\mathbf{x},\mathbf{z},\boldsymbol{\theta}) \;\; j(\boldsymbol{\theta}) \;\; d\boldsymbol{\theta} \\
&\text{w.r.t.} \qquad \mathbf{d},\mathbf{x},\mathbf{z} \\
&\text{s.t.} \\
&\qquad\qquad \mathbf{h}(\mathbf{d},\mathbf{x},\mathbf{z},\theta) = 0 \\
&\qquad\qquad \mathbf{g}(\mathbf{d},\mathbf{x},\mathbf{z},\theta) \le 0 \\
&\qquad\qquad \mathbf{d}\in \mathrm{D}, \quad \mathbf{x}\in \mathrm{X}, \quad \mathbf{z}\in \mathrm{Z}, \quad \theta\in\Theta
\end{aligned}
\tag{5.3}
$$

The similarity of the integrals in Equation 5.3 and Equation 5.1 is clear since $\mathrm{E}_{\Theta}\left[C(\mathbf{d},\mathbf{x},\mathbf{z},\boldsymbol{\theta})\right] = \int_{\boldsymbol{\theta}\in\Theta} C(\mathbf{d},\mathbf{x},\mathbf{z},\boldsymbol{\theta}) \;\; j(\boldsymbol{\theta}) \;\; d\boldsymbol{\theta}$.

One approach to solving the optimization problem in Equation 5.3 is to use Gaussian quadratures to evaluate the integral. The objective of Gaussian quadratures (Carnahan *et al.*, 1969) is to find the best numerical estimate of the integral by choosing optimal locations (abscissas), $\boldsymbol{\theta}_i$, with their corresponding weighting coefficients (weight factors), w_i, at which the function $f(\boldsymbol{\theta})$ is evaluated. Thus, with this discretization strategy, the formulation of Equation 5.3 takes the form

$$
\begin{aligned}
&\min \sum_{i=1}^{N_p} w_i C_q(\mathbf{d}, \mathbf{z}_i, \mathbf{x}_i, \boldsymbol{\theta}_i) \\
&\text{w.r.t.} \quad \mathbf{d}, \mathbf{z}_i, \mathbf{x}_i \\
&\text{s.t.} \\
&\qquad h(\mathbf{d}, \mathbf{z}_i, \mathbf{x}_i, \boldsymbol{\theta}_i) = 0 \\
&\qquad g(\mathbf{d}, \mathbf{z}_i, \mathbf{x}_i, \boldsymbol{\theta}_i) \leq 0 \\
&\qquad \mathbf{d} \in D, \quad \boldsymbol{\theta}_i \in \Theta, \quad \mathbf{z}_i \in Z, \quad \mathbf{x}_i \in X, \quad \mathbf{y}_i \in Y \\
&\qquad i = 1,\ 2,\ \ldots.,\ N_p
\end{aligned}
\tag{5.4}
$$

where C_q is the cost function that includes the penalization for infeasibilities due to soft constraint violations.

Gaussian quadratures were successfully applied to many problems involving uncertainties expressed as PDF's (Pistikopoulos and Ierapetritou, 1995; Straub and Grossman, 1990) Nevertheless, due to the exponential increase of number of points with an increasing number of uncertain parameters, quadratures are usually applied to problems involving no more than two or three uncertain parameters. To improve the usefulness of Gaussian formulas, cubatures were proposed to alleviate this limitation. Bernardo *et al.* (1999) presented two types of formulas, the Specialized Product Gauss (SPG) and the Specialized Cubature (SC) formulas, particularly constructed to integrate normally distributed uncertainties.

Another approach used to evaluate such integrals is through sampling techniques. When this integration strategy is used Equation 5.3 takes the form.

$$
\begin{aligned}
&\min \frac{1}{N_p} \sum_{i=1}^{N_p} C_q(\mathbf{d}, z_i, x_i, \theta_i) \\
&\text{w.r.t.} \quad \mathbf{d}, z_i, x_i \\
&\text{s.t.} \\
&\qquad h(\mathbf{d}, z_i, x_i, \theta_i) = 0 \\
&\qquad g(\mathbf{d}, z_i, x_i, \theta_i) \leq 0 \\
&\qquad \mathbf{d} \in D, \quad \theta_i \in \Theta, \quad z_i \in Z, \quad x_i \in X, \quad y_i \in Y \\
&\qquad i = 1,\ 2,\ \ldots.,\ N_p
\end{aligned}
\tag{5.5}
$$

where $\boldsymbol{\theta}_i$, i=1,2,... Np represents realizations of the uncertainty and $\mathbf{z}_i$, $\mathbf{x}_i$, the optimal values of the control and state values for each realization.

Among the commonly used sampling techniques are Monte Carlo methods. In contrast to Gaussian quadratures, Monte Carlo methods do not suffer from the exponential time-increase in evaluating the function, $f(\boldsymbol{\theta})$, when a higher number of uncertain parameters is involved. Two sampling techniques that fall within the category of Monte Carlo methods are Monte Carlo Sampling (MCS) and Latin Hypercube Sampling (LHS). MCS and LHS have been implemented in stochastic algorithms used to solve problems of robust parameter design and optimization under uncertainty (Diwekar and Rubin, 1994; Diwekar and Kalagnanam, 1997). However, these sampling techniques were inefficient since they require a large number of sampling points to estimate the expected value of a function with enough accuracy. To improve the efficiency of sampling techniques, Diwekar and Kalagnanam (1997) devised a new sampling technique based on Quasi-Monte-Carlo methods. They called this new sampling technique Hammersley Sequence Sampling (HSS) because it uses the Hammersley quasi-random sequence to generate, through inversion of the joint probability distribution, sampling points of the uncertain parameters of interest. In their work it was shown that HSS is 10 times faster than LHS and 100 times faster MCS.

Since the development of HSS, no other sampling technique that tries to alleviate the computational effort involved in stochastic optimization algorithms has been proposed. There are other sequences besides the Hammersley sequence that can be used for numerical integration such as Halton (1960), Faure (1982), Sobol (1967). Studies were performed to refine the above sequences so that they could be successfully applied in the estimation of integrals found in finance (Lemieux and L'Ecuyer, 2001). Despite the success of these sequences for integrating financial functions, no studies have been performed to explore their use to approximate the integrals involved in stochastic optimization problems encountered in process design.

In this work, three sampling techniques are proposed based on the sequences previously mentioned. Their performance on the estimation of functions derived from multivariate normal distributions is examined. Also, the equidistribution property of the points generated for uniform distributions, which is an important element for the performance (Marion and Henrion, 1980) of a sampling technique, is assessed by calculating their discrepancy for various dimensions and number of points. Based on the rank correlation method proposed by Iman and Conover (1982), correlation structures on the sequence of points are implemented in order to obtain accurate estimates of functions involving normal correlated parameters. The work is organized as follow. Without embarking on a detailed discussion on the construction of the sequences used in the sampling techniques, general overview for their design is given in Section 5.2. In section 5.3 the concept of discrepancy as a measure of uniformity is discussed and values for the discrepancy are obtained as a function of number of points and dimensions. Section 5.4 presents a procedure of transformation to sample normally distributed uncertainties. Also, a method to induce correlation among the parameters is discussed. In Section 5.5 the results from a large numerical experiment involving various functions, distributions and number of parameters and designed to determine the rate of convergence to a specified accuracy for the proposed quasi-Monte-Carlo sampling techniques are presented. Section 5.6 presents two engineering problems, the design of two reaction systems in the presence of uncertain model parameters and variability, to demonstrate the effectiveness of the proposed sampling techniques for process design. Finally, in Section 5.7 conclusions are summarized.

5.2 Quasi-Monte-Carlo methods for multivariate integration

Quasi-Monte-Carlo methods differ from regular Monte Carlo methods in that the former are based on low-discrepancy sequences and the latter are based on sequences of pseudorandom numbers. Discrepancy is a quantitative measure of the uniformity property of a sampling technique and represents the deviation of a sequence of points from a uniform distribution.

A term commonly employed in determining the effectiveness of a technique to perform multivariate integration is the concept of *average case complexity*. The average case complexity of multivariate integration and function approximation refers to the average cost (defined as the number of sample function evaluations) of computing an approximate solution (Wozniakowski, 1991; Novak, 1988). Another definition is the minimal expected cost among algorithms whose expected errors do not exceed ε (i.e., the average case error) (Wasilkowski, 1996). It has been shown (Wozniakowski, 1991) that the average case complexity of quasi-Monte-Carlo methods is bounded by $O(1/\varepsilon^2)$ (i.e. the number of points needed to keep the average error within ε is proportional to $1/\varepsilon^2$) and that the average case error, ε, is related to the discrepancy of the sequence used by a sampling technique.

Some of the best known low-discrepancy sequences are those by Hammersley (1960), Halton (1960), Faure (1982), Sobol (1967). The building block for the construction of these sequences is the *van der Corput* (1935) sequence in base 2, although it can include sequences using another base of prime numbers. This sequence starts from zero and takes values in the interval (0,1).

The van der Corput sequence, for an integer p of base b, is generated by a three step procedure:

2. The decimal-base number p is expanded in the base b.
3. The number in base b is reflected and converted into a number in the interval (0,1) by reflecting it about the decimal point.
4. The reflected number is again expanded in base b.

For example, the procedure to generate the 4th number of the van der Courput sequence (p=1,2,...4) is as follow:

Step 1. Expand the integer in base 2 (i.e., convert 4 to the base 2 number system). Because "100" is the base 2 representation of 4, the expansion can be represented as

$$4 = 1\text{x}2^2 + 0\text{x}2^1 + 0\text{x}2^0$$

Step 2. Reflect "100" (i.e., reverse the digits as "001") and shift the binary decimal point all the way to the left. Thus,

$$100 \rightarrow 001 \rightarrow 0.001$$

Step 3. Expand the binary fraction "0.001" in the base 2 as

$$0.001 = 0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3} = 1/8$$

This procedure can be repeated for all p=1,2,...N_p. Table 5.1 summarizes the three-step procedure and shows the first 5 van der Courput numbers.

Table 5.1 Three-step procedure for the generation of the van der Corput sequence

N	Base 2	Reflected base-2 No.	Expanded Reflected base-2 No.	van der Corput Sequence
1	1	1	1×2^{-1}	1/2
2	10	01	$0 \times 2^{-1} + 1 \times 2^{-2}$	1/4
3	11	11	$1 \times 2^{-1} + 1 \times 2^{-2}$	3/4
4	100	001	$0 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3}$	1/8
5	101	101	$1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3}$	5/8

Having defined the building block, next the low-discrepancy sequences used by the Quasi-Monte-Carlo sampling techniques proposed in this work are presented.

5.2.1 Hammersley Sequence

Hammersley sequence is part of the quasi-Monte-Carlo methods for the computation of integrals. The average case complexity of Hammersley Sequence is $(1/\varepsilon)(\log(1/\varepsilon))^{(n-1)/2}$ where n is the integral dimension (Wozniakowski, 1991; Papageoriou and Wasilkowski, 1990). In contrast to Monte Carlo method in which the error bound is probabilistic (i.e. the error of convergence, ε, is of the order of $O(N_p^{-1/2})$), the Hammersley sequence has a deterministic error bound of $N_p^{-1}(\log N_p)^{n-1}$.

The algorithm for generating the Hammersley sequence makes use of the radix-R notation of an integer. That is, a specific integer, p, in radix-R notation can be represented as

$$p \equiv p_m p_{m-1} \cdots p_2 p_1 p_0 = p_0 + p_1 R + p_2 R^2 + \ldots + p_m R^m \quad (5.6)$$

where m = [$\log_R p$]=[(ln p)/(ln R)], and the brackets denote the integral part. For example, in the familiar base-10 (i.e., radix-10) number system, the integer 256 has p_0=6, p_1=5, p_2=2 with R=10 and m=2. A unique fraction between 0 and 1, called the inverse radix number, can be constructed by reversing the order of the digits of p about the decimal point as follows:

$$\phi_R(p) = 0.p_0p_1p_2 \ldots p_m = p_0R^{-1} + p_1R^{-2} + \ldots + p_mR^{-m-1} \tag{5.7}$$

The Hammersley points in n-dimensions is given by the following sequence

$$z_n(p) = \left(\frac{p}{N_p}, \phi_{R_1}(p), \phi_{R_2}(p), \ldots, \phi_{R_{k-1}}(p) \right) \tag{5.8}$$

where p=,1,2,... N_p-1; and de values of R_1, R_2, ..., R_{n-1} are the first n-1 prime numbers (2,3,5,7,11,13,17...).

As an illustration, consider the 10th point of the Hammersley sequence for N_p=10 and n=5. Therefore, p=10 with p_0=0 and p_1=1. The integer p can be expressed in radix-R notation as

p	Radix-2	Radix-3	Radix-5	Radix-7
10	1010	101	100	13

Reversing the order of the digits about the decimal points leads to

p	Radix-2	Radix-3	Radix-5	Radix-7
10	0.0101	0.101	0.001	0.31

Expanding the radix-R numbers one obtains

Radix-2	Radix-3	Radix-5	Radix-7
$0x2^{-1} + 1x2^{-2} + 0x2^{-3} + 1\,x2^{-4}$	$1x3^{-1}+0x3^{-2}+1x3^{-3}$	$0x5^{-1}+0x5^{-2}+1x5^{-3}$	$3x7^{-1}+1x7^{-2}$

which leads to the 10th Hammersley point

$p/10$	$\phi_2(10)$	$\phi_3(10)$	$\phi_5(10)$	$\phi_7(10)$
10/10	5/16	10/27	1/125	22/49

The same procedure can be applied to any integer, p, of the Hammersley sequence. Table 5.2 shows the first 10 Hammersley points for dimensions 1 to 5 for N_p=10 generated by the Hammersley sequence algorithm.

Table 5.2 Hammersley points in $[0,1]^5$ produced by the Hammersley sequence algorithm

	Dimension				
N_i	1	2	3	4	5
1	1/10	1/2	1/3	1/5	1/7
2	1/5	1/4	2/3	2/5	2/7
3	3/10	3/4	1/9	3/5	3/7
4	2/5	1/8	4/9	4/5	4/7
5	1/2	5/8	7/9	1/25	5/7
6	3/5	3/8	2/9	6/25	6/7
7	7/10	7/8	5/9	11/25	1/49
8	4/5	1/16	8/9	16/25	8/49
9	910	9/16	1/27	21/25	15/49
10	10/10	5/6	10/27	4/5	22/49

5.2.2 Halton Sequence

The Haltom sequence like the Hammersley sequence is based on the van der Corput sequence. It also makes use of different prime bases for each dimension of the sequence. The difference from the Hammersley sequence is its use of the van der Corput sequence for base two for the first dimension. Therefore, the formula in Equation 5.8 can be modified to express the Halton Sequence as

$$z_n(p) = \left(\phi_{R_1}(p), \phi_{R_2}(p), \ldots, \phi_{R_{n-1}}(p)\right) \tag{5.9}$$

where R_1, R_2, ..., R_{n-1} are the first n-1 prime numbers. For a more detail description of the Halton sequence, the reader is referred to the works due to Halton (1960) and Kocis and Whiten (1997).

5.2.3 Faure Sequence

The Faure sequence was derived from linking low discrepancy theory and combinatorial theory for vector reordering. The Faure sequence differs from the Hammersley and Halton sequences in that it only uses one base for the construction of sequences of all n-dimensions and that the elements of the sequence of each dimension are permuted. In addition, the base used in the Faure sequence is the smallest prime number that is larger than or equal to the number of n-dimensions. For instance if n=50, in the Halton sequence, the prime number used as base for the 50^{th} dimension is 229 whereas in the Faure sequence, the base prime number for the 50^{th} dimension is 53.

Like the previously mentioned sequences, the Faure sequence makes use of Equation 5.6, and Equation 5.7. However, there is a combinatorial rearrangement of the elements of the sequence. The generation of the Faure sequence is rather complicated and it is best to check the papers of Faure (1982) and Fox (1986) for a more complete description of its construction.

5.2.4 Sobol Sequence

The Sobol, like the Faure sequence, has the same base for all dimensions and the element vectors are reordered within each dimension. However, contrary to Faure, the Sobol sequence uses the van der Corput base 2 for all dimensions. A detailed description of the Sobol sequence is out of the scope of this paper. The reader is referred to the works of Bratley and Fox, (1998) and Press and Teukolsky (1989) for a detailed discussion of its construction. Bratley and Fox (1998) describe a Fortran implementation of a Sobol sequence generator and compare it in some detail to the generator suggested by Faure. Press and

Teukolsky (1989) discuss the Sobol's sequence and computational methods for generating them.

5.3 Discrepancy in quasi-Monte-Carlo sampling techniques

It has been shown (Morokof and Caflisch, 1994) that sampling techniques with sequences presenting better equidistribution properties lead to smaller errors in integration. Figure 5.1 shows the equidistribution property of 100 points unit from a uniform distribution in a 2-dimensional unit space generated by Monte Carlo Sampling techniques. MCS places the samples in a random manner and leaves some parameter space empty. This randomization leads to poor uniformity. On the other hand, LHS shows a more uniform sampling scheme. LHS divides the parameter space into strata and then randomly places the samples within each stratum. This ensures a better coverage of the parameter space. Nevertheless, some of the parameter space still remains empty thus distorting its uniformity.

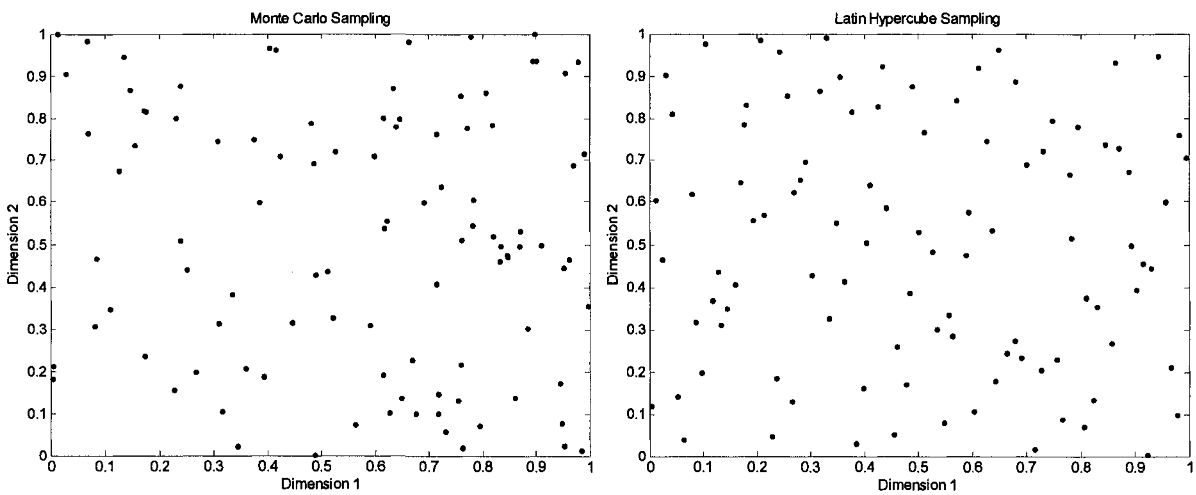

Figure 5.1 Distribution property of Monte Carlo sampling techniques

Similar plots, generated with two Monte Carlo techniques, are shown in Figure 5.2. It can be observed that all four sampling techniques present better uniformity than either MCS or LHS. When comparing the quasi-Monte-Carlo sequences, it can be observed that HSS present better uniformity than HalSS, FSS, and SSS. This is obvious since HalSS, FSS and SSS present some relatively large empty spaces among their samples.

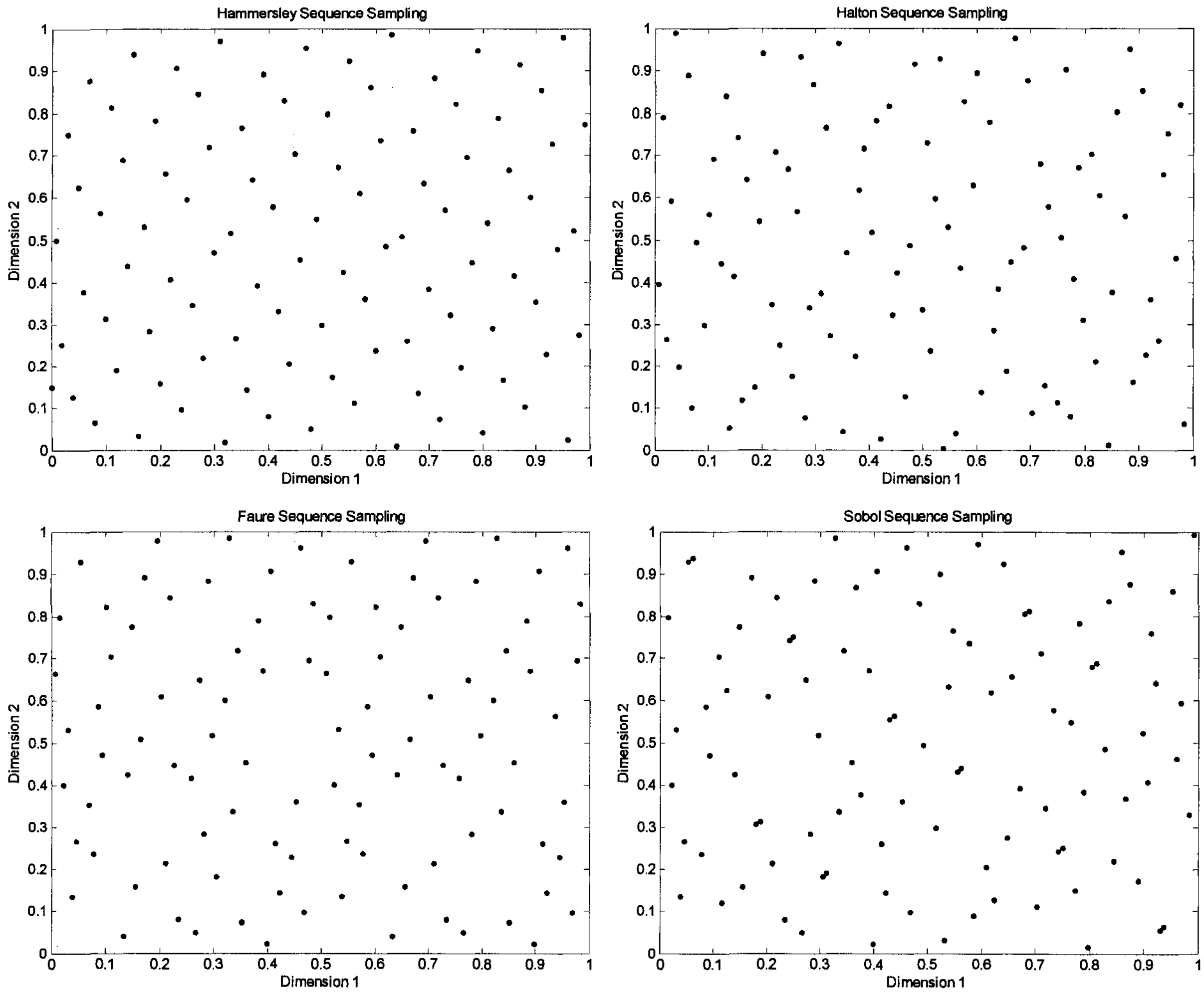

Figure 5.2 Distribution property of quasi-Monte-Carlo sampling techniques

Figure 5.2 gives us a qualitative picture of the uniformity of the sequences used by the sampling techniques. However, it is well known that the uniformity properties of quasi-random sequences start to degrade as the dimension increases. Figure 5.3 shows worst cases of uniformity for different quasi-Monte-Carlo techniques obtained in a two dimensional projection of 1000 points for a six-dimensional case. The poorest uniformity of Hammersley points is present in the projection of the first and sixth dimension. Its uniformity is in a way distorted due to an apparent correlation among the points. For the HalSS, the worst uniformity is present in the projection onto the first and fifth dimension. It can be observed that there are points that tend to agglomerate leaving some empty spaces. This projection appears to be considerably more uniform than that of the HSS. Also, there seems to be a negative correlation among the points since a weak inclined vertical pattern is observed. For the FSS case, the projection onto the first and fifth dimension shows the poorest uniformity. A correlation among the points seems to be apparent inducing an inclined pair-wise vertical pattern. On the other hand, for the SSS case, the worst uniformity is revealed in the projection onto the first and second dimension. The Sobol points fall in the unit square with a distinct pattern leaving some areas with many points inside while leaving others empty. This seems to be an intrinsic characteristic of the Sobol sequence. Press and Teukolsy (1989) explain this fact and show how the Sobol's sequence generator fills out the unit square.

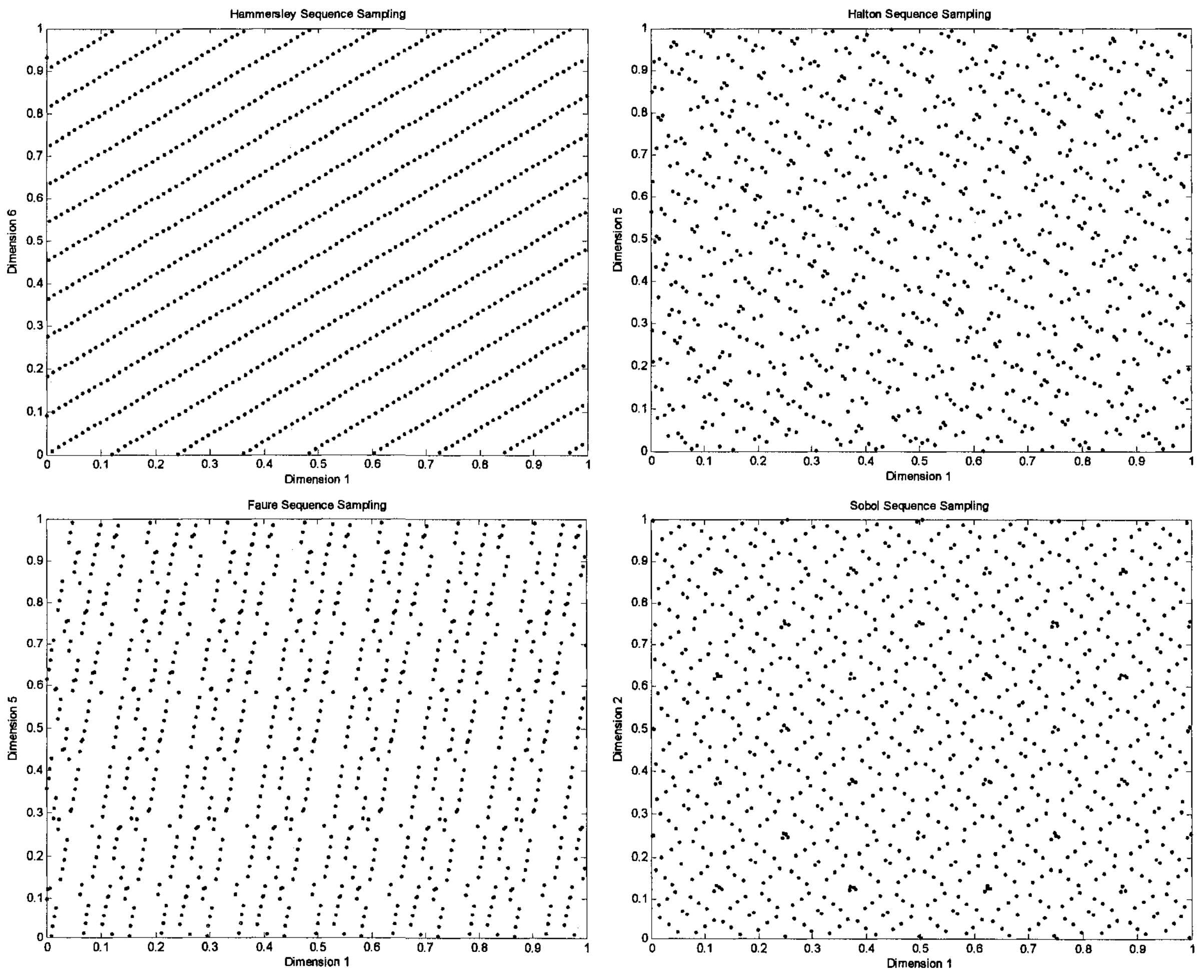

Figure 5.3 Projections of 1000 points with works case of uniformity in a six dimensional case for different quasi-Monte-Carlo sampling techniques.

The qualitative evaluation of the uniformity properties of these projections was possible because the projection onto two dimensions was visually examined one at a time. Generalization of these plots for more than three dimensions, however, is impossible. Therefore, it is necessary to define a measure of uniformity in order to assert conclusions on the uniformity of a sequence used by a sampling technique.

A sequence may have poor two-dimensional projections. Nevertheless, it may still be fairly uniform with respect to the total n-dimensional hypercube. Good uniformity is crucial for the estimation of multidimensional integrals. Sequences with better uniformity lead to smaller errors in integration. This is the result of their low discrepancy (Markoff and Cafish, 1994). The relationship between the discrepancy and the integration error has been demonstrated through the *Koksma-Hlawka* inequality (Hlawka, 1961). This inequality states that the integration error of a sampling technique with low-discrepancy sequence is bounded by the product of the function and the discrepancy of the sequence. One commonly used measure of discrepancy is the star-discrepancy (Niederreiter, 1992). The computation of the star discrepancy is not an easy task, especially for more than two dimensions. It involves a series of computations and optimizations to find the optimal arrangement that reduces the discrepancy among the sampling points (Dobkin and Eppstein, 1994). Fortunately, algorithms that allow calculating the star discrepancy of a sequence of points are available (Thiémard, 2001).

In this work, the computation of the star discrepancy for the proposed techniques is presented as a means of comparison and as a method for predicting their performance. The discrepancy values of the quasi-Monte-Carlo techniques corresponding to the two-dimensional case in Figure 5.2 are shown in Table 5.1. For these computations a relative error of 0.001 was used.

From Table 5.1 it is clear that HSS presents the lowest discrepancy among its points for a two dimensional case. In order to gain an insight in how the discrepancy is affected by the dimension involved and the number of points used, the discrepancies using 100 points for different dimensions was calculated.

Table 5.3 Star discrepancy values of quasi-Monte-Carlo sampling techniques for the two-dimensional case displayed in Figure 5.2

Sampling Technique	$D_N^*(P)$
HSS	0.03788
HalSS	0.05023
FSS	0.04818
SSS	0.03923

Figure 5.4 shows how the discrepancy increases with increasing the dimension. This is expected since, as more dimensions are added, more empty spaces are generated. It can be observed that FSS presents the highest discrepancy whereas the HSS, HalSS and SSS remain close to each other between dimension 3 and 4. On the other hand, after the 6th dimension it is clear that the discrepancies start to separate, being the SSS with the lowest discrepancy.

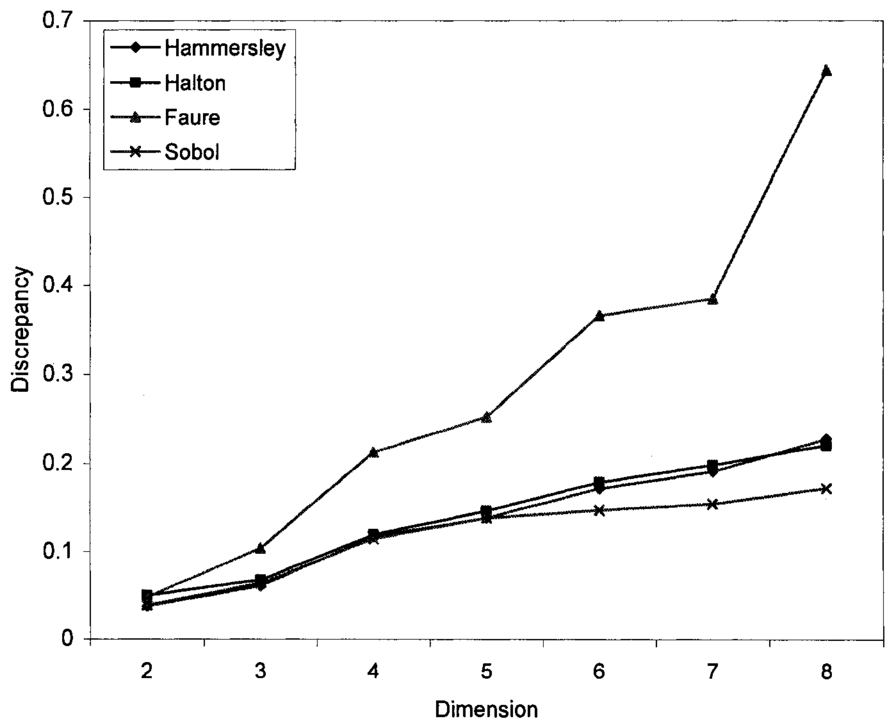

Figure 5.4 Discrepancies of Quasi-Monte-Carlo sampling techniques as a function of dimension.

Figure 5.4 shows how the discrepancy decreases with the number of points used in a simulation for a particular six-dimensional case. It can be observed that with a small number of points used FSS and SSS present the highest discrepancies. However, as the number of points increases the discrepancy of the SSS decreases sharply being the lowest of the four at the highest dimension. The SSS, HSS and HSS discrepancies remain close to each other as the number of points increases whereas the discrepancy in the FSS sequence continues to be the highest.

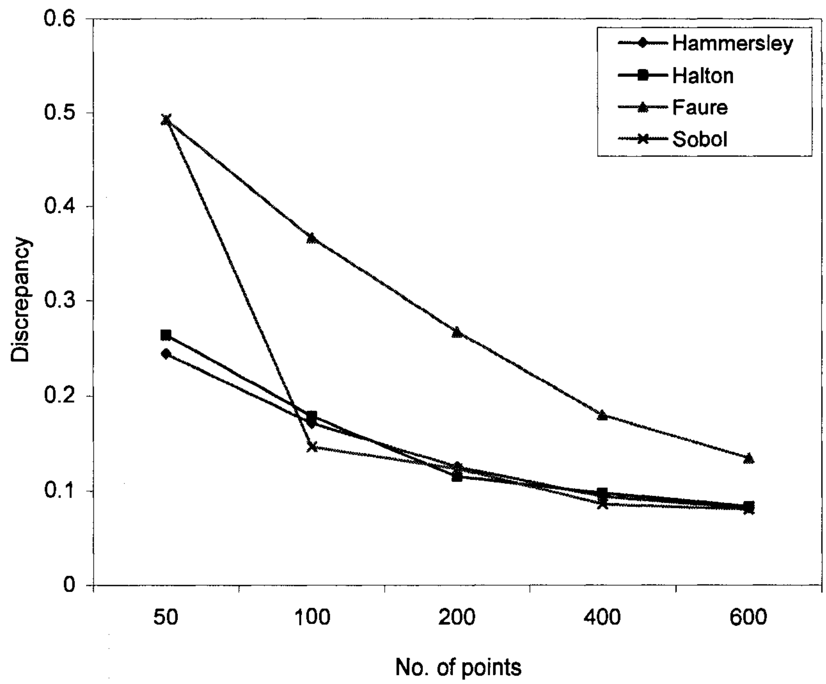

Figure 5.5 Discrepancies of Quasi-Monte-Carlo sampling techniques as a function of number of points.

In summary, the discrepancy of the quasi-Monte-Carlo sequences increases as the number of dimension increases. This increase, however, can be alleviated by increasing the number of points. It has been shown that the discrepancies among the quasi-Monte-Carlo sequences are very close to each other except for Faure which presents a noticeable higher discrepancy.

5.4 Sampling normally distributed uncertainties

So far, only uniform distributions have been considered. It is often the case that the parameters under consideration follow different distributions. In such a case, the sampling procedure requires of a transformation to map the sample value form a uniform distribution to that of the distribution type of the parameters. There are two methods which can provide samples of different distributions. The transformation method for the generation of samples derived from Exponential and Normal distributions and the rejection method for the generation of samples derived from Gamma, Poisson and Binomial distributions (Press *et al.*, 1992). Since this work is devoted to solving problems involving normally distributed uncertainties, only the transformation method will be presented.

The transformation method is based on the fundamental transformation law of probabilities. Let u be a sample drawn from a standard uniform distribution with probability density function, $\mathrm{j}(u)$, and z a sample drawn from a standard normal distribution with probability density function, j(z). Their normalized cumulative density functions, $\mathrm{J}(u)$ and $\mathrm{J}(z)$, can be expressed as

$$\int_{-\infty}^{+\infty} \mathrm{j}(u)\mathrm{d}u = \int_{-\infty}^{+\infty} \mathrm{j}(z)\mathrm{d}z = 1 \qquad (5.10)$$

The sample value, z, can be obtained from the current value of u, by solving the differential equation

$$\frac{\mathrm{d}u}{\mathrm{d}z} = \mathrm{j}(z) \qquad (5.11)$$

which solution is just $u = \mathrm{J}(z)$, where $\mathrm{J}(z)$ is the indefinite integral of $\mathrm{j}(z)$. The transformation which turns a sample uniformly distributed into one normally distributed with probability density function $\mathrm{j}(z)$ is therefore

$$z = \mathrm{J}^{-1}(u) \qquad (5.12)$$

where J^{-1} is the inverse function of J. The inverse function of the integral of j(z) is easy to compute and is available in many computer packages such as MATLAB. The procedure of transformation is shown in Figure 5.6.

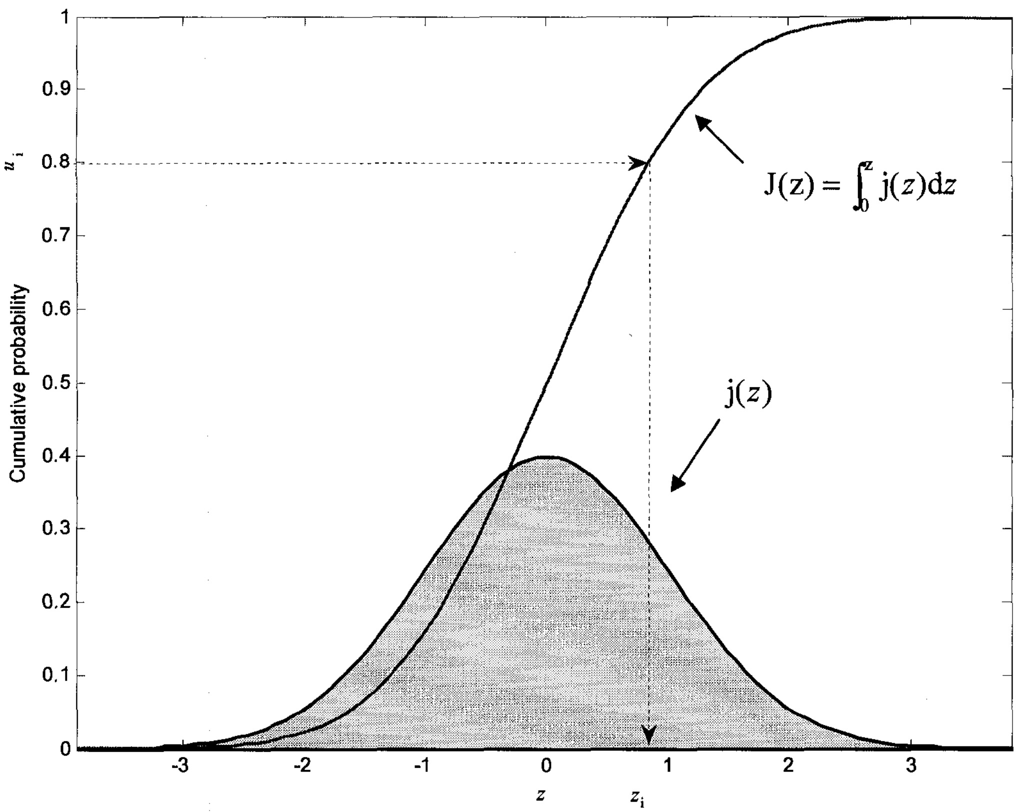

Figure 5.6 Procedure of transformation.

Since z is a sample generated from a standard normal distribution N(0,1), an additional transformation is performed in order to generate a sample from a distribution with mean, μ_θ, and standard deviation, σ_θ. This transformation is given by

$$\theta = \mu_\theta + \sigma_\theta z \qquad (5.13)$$

This procedure can be easy generalized to handle multivariate normally distributed uncertainties. For normal distributed parameters with mean vector, $\boldsymbol{\mu}$, and diagonal variance matrix, $\mathbf{V}$, the transformation is given by

$$\boldsymbol{\theta} = \boldsymbol{\mu} + \mathbf{V}^{1/2} z \tag{5.14}$$

5.3.1 Treatment of correlation among normally distributed uncertain parameters

While some uncertain parameters vary independently in a model, there are others between which dependencies may exist. Although for some process models the assumption of independence in their parameters is not detrimental, for others such as constrained process models, the independence among the uncertain variables is of paramount importance since the feasible region may be bounded by the joint distribution of the uncertain parameters. The relationship among uncertain parameters is commonly described by their correlation. Therefore, it is important to incorporate any correlation structure into the samples drawn from the joint distribution of the uncertain parameters.

There are different ways for incorporating dependencies. A rank correlation approach (Iman and Conover, 1982) was chosen to impose correlation structures on the samples generated with the proposed techniques because it can be used with all types of distribution functions, it is simple, and can be applied to any sampling scheme. This method consists of restricting the way the parameters are paired, based on the rank correlation of some target value, using the Cholesky decomposition of the correlation matrix. Consider the matrix $\mathbf{\Phi}$ of independent uncertain parameters with correlation matrix Γ, and $\mathbf{C}$ the desired correlation matrix. Matrix $\mathbf{C}$ can be decomposed as $\mathbf{C} = \mathbf{P} \times \mathbf{P}^T$ (Cholesky factorization) where $\mathbf{P}$ is the lower triangular matrix. Multiplying the matrix $\mathbf{\Phi}$ by $\mathbf{P}^T$ yields parameters with correlation matrix $\mathbf{C}$. The method can be summarized as follows:

- Generate matrix $\mathbf{\Phi}$ using any quasi-Monte-Carlo sampling technique of n parameters and sample size N_p, and calculate Γ, the correlation matrix of $\mathbf{\Phi}$.
- Calculate the $\mathbf{P}$ lower triangular matrix of the target correlation matrix $\mathbf{C}$ using Cholesky factorization $\mathbf{C} = \mathbf{P} \times \mathbf{P}^T$ and also $\mathbf{Q}$, the lower triangular matrix of Γ.
- Solve to obtain matrix $\mathbf{S}$ such that $\mathbf{S} \times \Gamma \times \mathbf{S} = \mathbf{C}$, which is calculated from $\mathbf{S} = \mathbf{P} \times \mathbf{Q}^{-1}$.

- Calculate the target correlation matrix $\mathbf{R}^* = \mathbf{\Phi} \times \mathbf{S}^T$, which has correlation matrix equal to **C.**
- Rearrange the values of each parameter in $\mathbf{\Phi}$ so they have the same rank as the target matrix $\mathbf{R}^*$

To illustrate the implementation of correlation in the proposed sampling techniques, let us consider two uncertain parameters described by a joint normal PDF with mean

$$\boldsymbol{\mu} = [6, 6]^T$$

and variance-covariance matrix

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

Contour plots of two uncertain parameters are shown in Figure 5.7 along with the sampling points generated with the proposed techniques. Although, all the sampling techniques tend concentrate their points in the center for the distribution, their uniformity is somewhat perturbed by the procedure of transformation. The question of whether or not the uniformity in samples is affected by the transformation, again, can only be answered by a measure of uniformity in spaces other than a uniform hypercube. This, however, is not possible since, to the knowledge of the authors, such a measure is not available. Accordingly, to determine if the uniformity after transformation seriously affect the effectiveness of the proposed sampling techniques in evaluating functions of normally distributed parameters, a series of numerical experiments, using various type of functions and distributions, were performed. In the next section the results obtained from this numerical experiment are summarized.

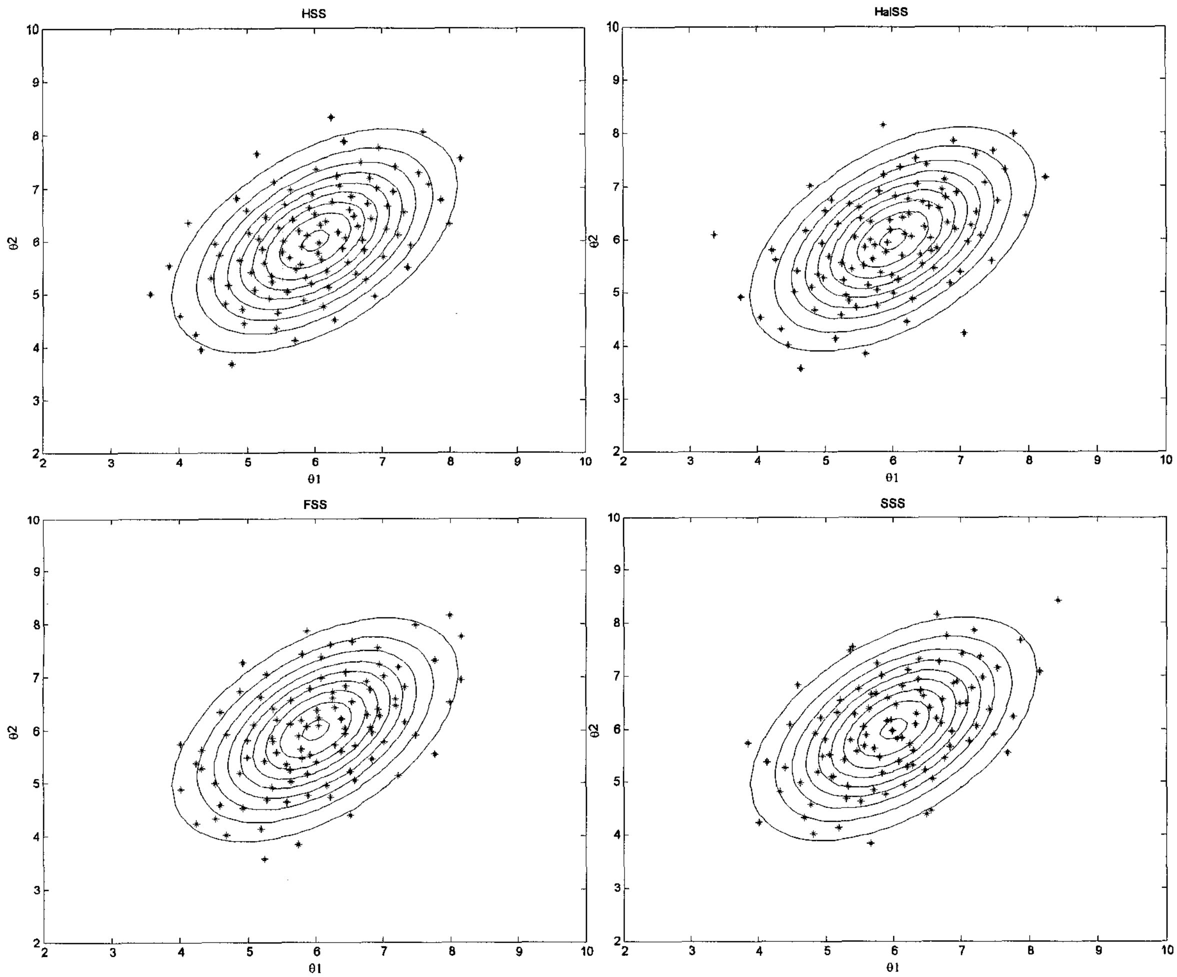

Figure 5.7 Sampling points with correlation imposed

5.4 Numerical experiments

This section provides a comparison of the performance of the proposed sampling techniques for evaluating the expected value of various types of functions. This is the same type of experiment carried out by Kalagnanam and Diwekar (1997) to demonstrate the effectiveness of HSS when compared to two Monte Carlo sampling techniques, MCS and LHS. This comparison is performed by generating samples from a joint probability distribution function of the parameters with the proposed sampling techniques, evaluating the expected value of different functions and then measuring the number of samples required to converge to within an error of 2% of the "true mean". The "true" mean was determined to be that obtained with a very large sample size (e.g., 50,000 samples) using HSS. This is graphically presented by plotting the calculated value of the mean as a function of number of samples used in the calculation.

For this numerical experiment, functions that are analytically simple and for which the "true" mean can be estimated exactly were chosen. It will be shown that the proposed sampling techniques converge to the true means.

Functions: The functions used in this numerical experiment were of the following types:

Type 1: Linear additive: $f(\theta) = \sum_{i=1}^{n} \theta_i$

Type 2: Multiplicative: $f(\theta) = \prod_{i=1}^{n} \theta_i$

Type 3: Quadratic: $f(\theta) = \sum_{i=1}^{n} \theta_i^2$

Type 4: Exponential: $f(\theta) = \sum_{i=1}^{n} \theta_i \exp(\theta_i)$

Type 5: Logarithmic: $f(\theta) = \sum_{i=1}^{n} \log(\theta_i)$

Distributions: Two types of symmetric distributions were used to characterize the uncertainties of the parameters, θ_i: Normal and Uniform.

Number of parameters: The numbers of parameters used were: 2, 6 and 10.

Sampling techniques: Four sampling techniques were compared: Hammersley Sequence Sampling (HSS), Halton Sequence Sampling (HalSS), Faure Sequence Sampling (FSS) and Sobol Sequence Sampling (SSS).

5.4.1 Results

This numerical experiment consisted of 5 functions, 2 distribution types, 3 number of parameters and 4 sampling techniques. This led to 120 data sets –a very large amount of information. For the sake of space and clarity, only the main findings of these numerical experiments will be reported.

Figure 5.8 plots the estimates of the mean for the linear function using 2 uncertain parameters that are uncorrelated. A uniform distribution (0,1) was used for both parameters. All the sampling techniques converged to within 2% of the true man. For SSS and FSS it only required 100 samples to obtain accurate estimates of the mean (as determined by the specified error band) whereas for HSS and HalSS it required 200 and 300 samples, respectively.

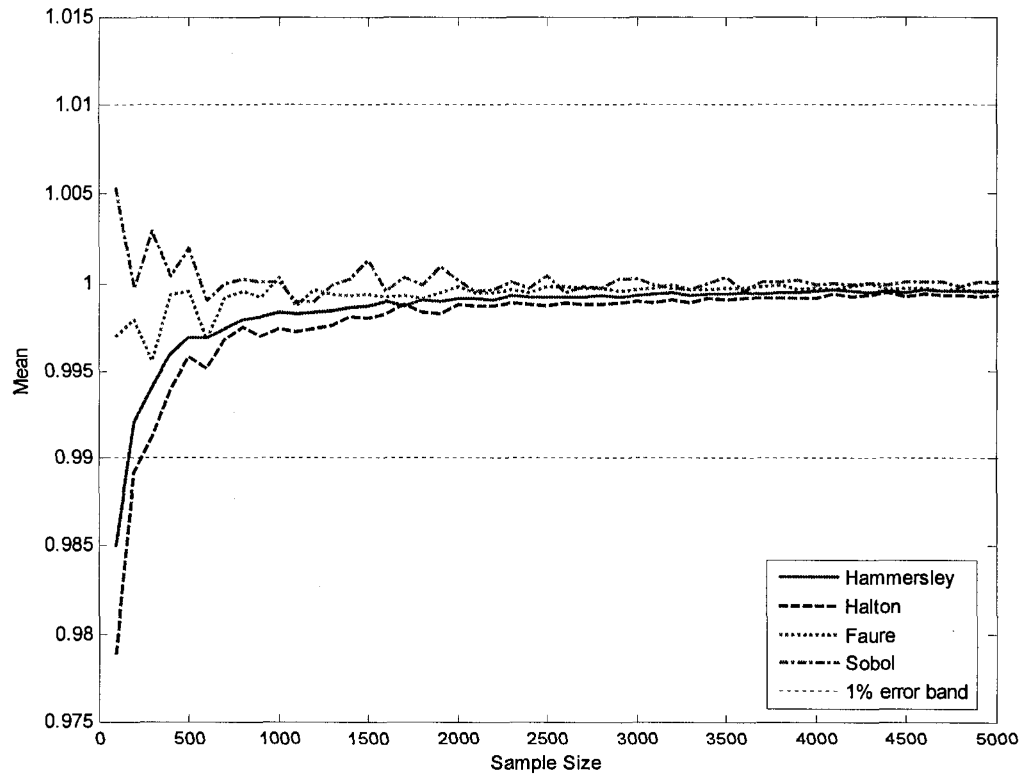

Figure 5.8 Estimates of the mean for the linear function as a function of sample size – Uniform distribution of two uncertain parameters.

Figure 5.9 shows the same type of plot but for two uncorrelated normally distributed parameters with vector μ = $[0.5\ 0.5]^T$ and variance matrix $\Sigma = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}$. The convergence rate was slower than the case when a uniform distribution was considered. For SSS, it took 500 samples to converge to the specified error bandwidth whereas it took 550 samples for FSS. For HSS and HalSS, the convergence rate was slower requiring 1800 and 2000 samples, respectively.

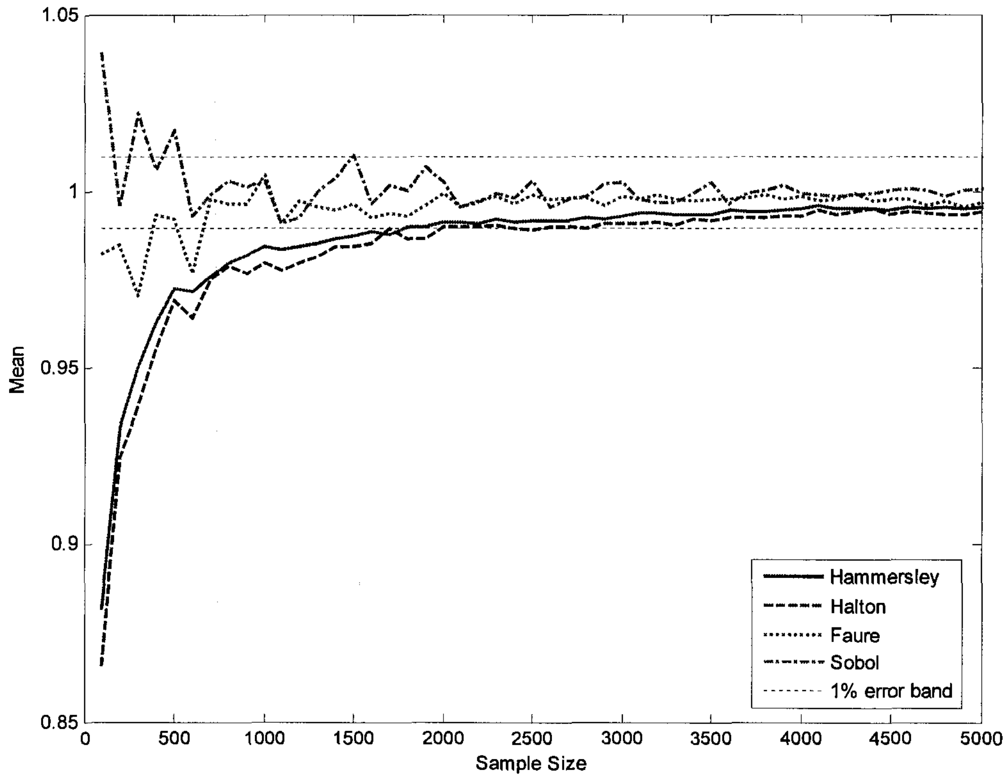

Figure 5.9 Estimates of the mean for the linear function as a function of sample size –Normal distribution of two uncertain parameters.

Figure 5.10 shows the convergence rates of the quasi-Monte-Carlo sampling techniques obtained considering uniform distributions only. SSS presented the fastest rates of convergence for the number of parameters and function types studied in this numerical experiment. The second technique showing fast rate of convergence was FSS. FSS seemed to work quite well for the 2-dimensional case, requiring low number of samples except from the logarithmic function. The next showing fast convergence rates were HSS and HalSS. Interestingly, the multiplicative function required the most number of samples for all the sampling techniques for the 6- and 10-dimensional cases. This may be the result of the interactions among their discrepancy, dimensionality of the problem and this type of function.

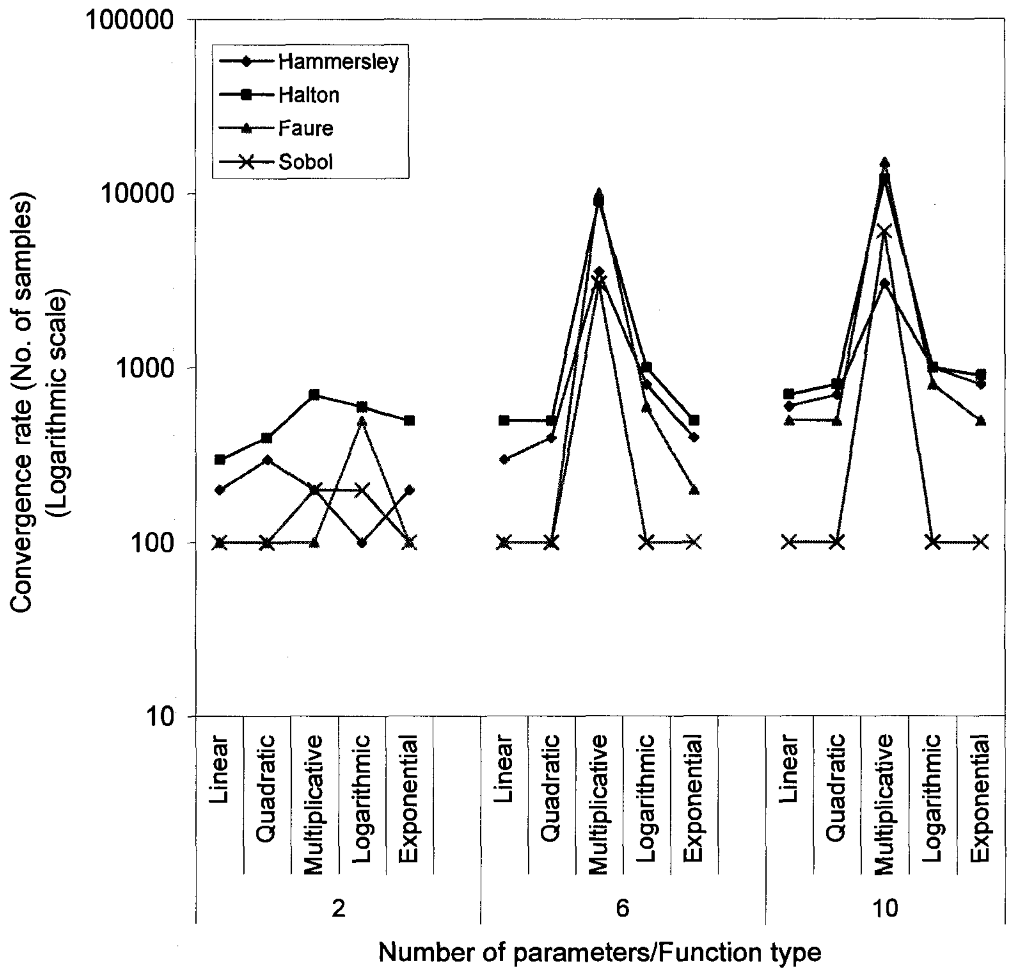

Figure 5.10 Convergence rate of quasi-Monte-Carlo sampling techniques as a function of number of uncertain parameters and function type –Uniform distribution.

Similarly, Figure 5.11 shows the convergence rates of the quasi-Monte-Carlo sampling techniques obtained considering normal distributions only. The sampling techniques required more samples to achieve the specified accuracy thus decreasing their convergence rate. This may be the result of the additional nonlinearity introduced by the normal distribution. It can be observed that the rate of convergence of all quasi-Monte-Carlo sampling techniques were very similar, though SSS seemed to have a slightly faster convergence rate than the rest of the techniques. For the 2-dimensional case, the convergence rate of all the sampling techniques remained close to each other. For the 6-dimensional case, the multiplicative function required the most number of samples to converge to the required accuracy only for SS and HSS whereas HalSS and FSS maintained a faster convergence rate. For the 10-dimensional case, again all the quasi-Monte-Carlo sampling techniques presented similar convergence rates with SSS having the fastest and FSS the slowest convergence rates, respectively.

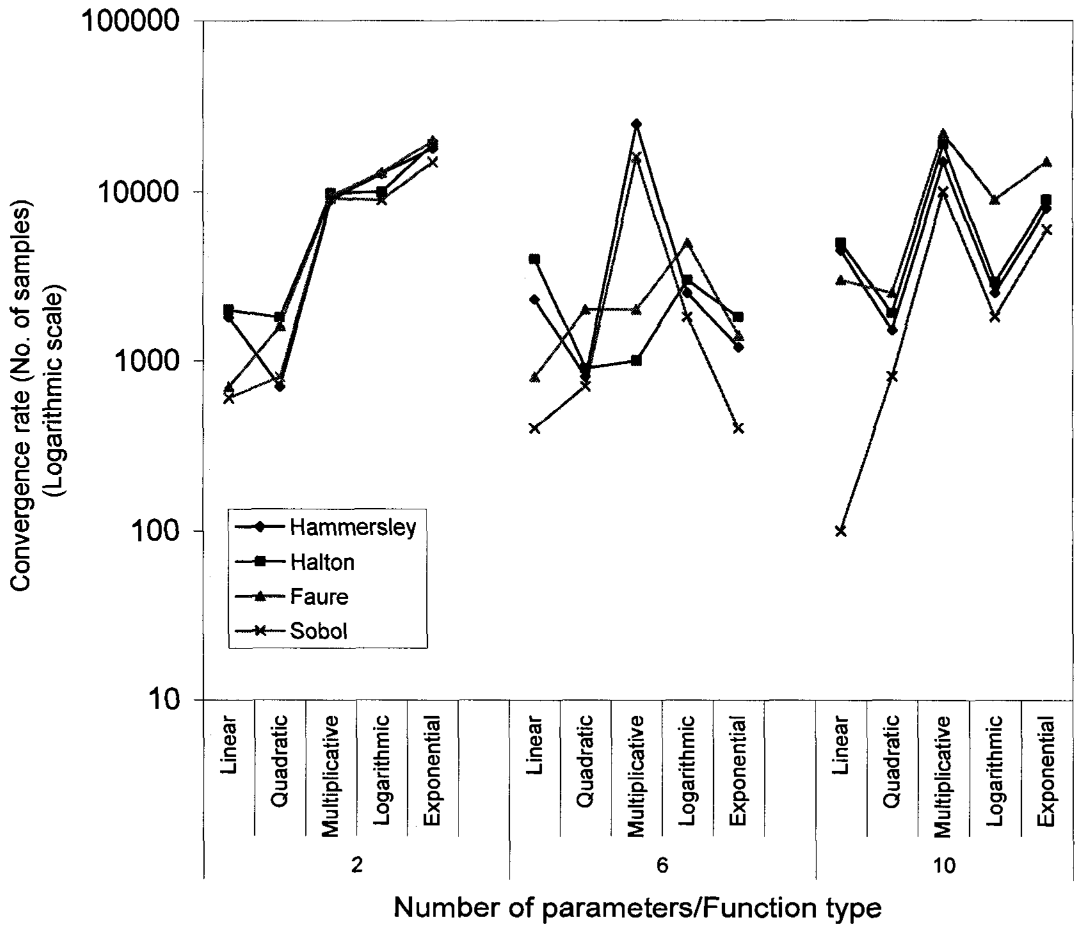

Figure 5.11 Convergence rate of quasi-Monte-Carlo sampling techniques as a function of number of uncertain parameters and function type –Normal distribution.

Overall, the proposed sampling techniques showed convergence rates similar to that of HSS with some of them presenting even faster convergence rates. These results have demonstrated that a potential for reducing the computational effort involved in the solution of stochastic optimization problems can be obtained with the used of the proposed quasi-Monte-Carlo sampling techniques. They showed fast convergence rates for the specified accuracy requirement established for this numerical experiment.

Having estimated the convergence rates of the proposed quasi-Monte-Carlo sampling techniques for various function and distribution types, next, two engineering examples, one of them presenting correlated parameters in their model, will be used to evaluate their effectiveness on the solution of stochastic optimization problems involved in process design.

5.5 Case studies

In this section the usefulness of the proposed sampling techniques is evaluated when applied to stochastic optimization. Two process design examples, subjected to uncertainty in model parameters and variability, are used and solved through single-stage stochastic optimization.

Reactor-heat exchanger model. The first problem considered is the problem of designing a reactor-heat-exchanger system. The reactor is assumed to be an ideal CSTR where a single first order chemical reaction A $\rightarrow$ B takes place (Bernardo, 2001; Chacon-Mondragon and Himmelblau, 1996; Dominguez and McLean, 2006). The objective is to determine the optimal reactor volume, V, and heat exchanger area, A, which ensures 90% conversion of reactant A, and a minimum overall annual cost consisting of capital, operating and quality costs. The mathematical model for this system is shown in Table 5.4

Table 5.4 Reactor-heat-exchanger model

Objective function ($/year) :

$$\text{Overall Cost} = \text{Capital Cost} + \text{Operating Costs} + \text{Quality Costs}$$

$$\text{Capital Cost} = c_1V^{0.7} + c_2A^{0.6}$$

$$\text{Operating Costs} = c_3F_W + c_4F_1$$

$$\text{Quality Costs} = k(x_A - 0.9)^2$$

Material balance in reactor:

$$F_0x_A - Vk_R\exp(E/RT_1)C_{A0}(1-x_A)V = 0 \quad x_A = (C_{A_0} - C_A)/C_{A_0}$$

Energy balance in reactor:

$$(-\Delta H)F_0x_A + F_0C_p(T_0 - T_1) - F_1C_p(T_1 - T_2) = 0$$

Heat exchanger design equation:

$$F_1C_p(T_1 - T_2) = AU\Delta T_{lm}, \ \Delta T_{lm} = ((T_1 - T_{w_2}) - (T_2 - T_{w_1}))/\ln[(T_1 - T_{w_2})/(T_2 - T_{w_1})]$$

Heat exchanger energy balance:

$$F_1C_p(T_1 - T_2) = F_wC_{p_w}(T_{w_2} - T_{w_1})$$

Temperature Bounds (K):

$$311 \le T_1 \le 389, \ 311 \le T_2 \le 389, \ 294 \le T_{w_2} \le 323$$

Heat exchanger operation constraints:

$$T_1 - T_{w_2} \ge 11.1, \ T_1 - T_2 \ge 0, \ T_{w_2} - T_{w_1} \ge 0, \ T_2 - T_{w_1} \ge 11.1$$

Quality constraint

$$x_A \ge 0.9$$

[a] Reactor volume, V, and the area of the heat exchanger, A, are given in m^3 and m^2, respectively. The flowrate from the reactor to the heat exchanger, F_1, is given in $kmol\cdot hr^{-1}$ and the flow rate of the cooling fluid, F_w, is given in $kg\cdot s^{-1}$. The reactor temperature, T_1, the reactant temperature exiting the heat exchanger, T_2, and the cooling fluid temperature exiting the heat exchanger, Tw_2, are given in *K*. Conversion is x_A,

The mathematical model contains six uncertain parameters described by a normal distribution with mean vector, μ, and variance-covariance matrix, **Σ**. Mean values along with standard deviations of the uncertain parameters are given in Table 5.5.

Table 5.5 Uncertain parameters of reactor-heat-exchanger model

	Parameter	Units	μ	σ
F_0	Flow rate of the feed	$kmol\cdot hr^{-1}$	45.36	2.93
T_0	Temperature of the feed	K	333	4.31
Tw_1	Cooling water inlet temperature	K	293	3.79
k_R	Arrhenius rate constant	hr^{-1}	12	0.77
U	Overall heat transfer coefficient	$kJ.m^{-2}\cdot h^{-1}\cdot K^{-1}$	1635	105.82
C_{A0}	Initial concentration at the feed	$kmol\cdot m^{-3}$	32.04	2.07

The feed temperature, T_0, and the cooling water inlet temperature, T_{W1}, are assumed to have a positive correlation of 0.7. Therefore, the variance-covariance matrix is given by

$$\Sigma = \begin{bmatrix} 8.58 & 0 & 0 & 0 & 0 & 0 \\ 0 & 18.57 & 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 14.36 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.59 & 0 & 0 \\ 0 & 0 & 0 & 0 & 11193 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4.28 \end{bmatrix} \tag{5.15}$$

Finally, known parameter values for the above model are shown in Table 5.6.

Table 5.6 Deterministic parameters for the reactor-heat-exchanger model

c_1	Cost per reactor volume	691.2 $\$\cdot yr^{-1}\cdot m^{-3}$
c_2	Cost per heat exchange area	873.6 $\$\cdot yr^{-1}\cdot m^{-2}$
c_3	Cost of coolant fluid	1.76 $\$\cdot yr^{-1}\cdot kg^{-1}\cdot s$
c_4	Cost of pumping	7.056 $\$\cdot yr^{-1}\cdot kmol^{-1}\cdot hr$
E/R	Activation energy/gas constant ratio	555.6 K
ΔH	Molar heat of reaction	-23260 $kJ\cdot kmol^{-1}$
C_p	Heat capacity of the reactant	167.4 $kJ\cdot kmol\cdot K^{-1}$
C_{pw}	Heat capacity of Cooling water	4.184 $kJ\cdot kg^{-1}\cdot K^{-1}$
k	Taguchi loss constant	$6.4x10^6$ \$

The problem was solved using the formulation described in Equation 5.5 and using GAMS/CONOPT (Brook *et al.*, 2005) as the optimization routine. Figure 5.12 shows the estimated values for the overall expected cost obtained with the proposed quasi-Monte-Carlo Sampling techniques. As can be seen, SSS was the first technique to converge within the required accuracy. The accuracy band shown was constructed considering a relative error of 0.5% around the HSS solution with 1000 samples. Only 100 points were necessary to estimate the overall expected cost with the required accuracy using SSS, the overall expected cost being \$12,834 per year. HSS required only slightly more; it required 125 samples with an overall expected cost of \$12,820. HalSS, with 175 samples, provided an overall expected cost of \$12766 per year. Finally, FSS took the most number of points to provide an accurate solution. With 200 samples, FSS provided an overall expected cost of \$12851 per year.

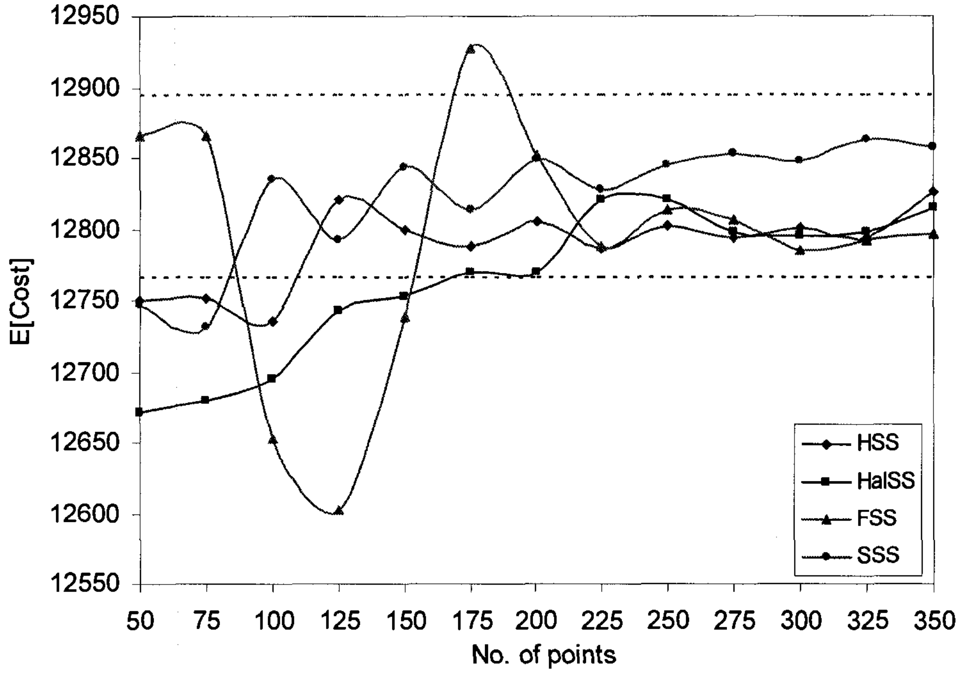

Figure 5.12 Estimations of the overall expected cost obtained with quasi-Monte-Carlo sampling techniques for the reactor-heat-exchanger problem.

Continuous stirred-tank reactor model. The second system used was a variant of the reactor design of a non-isothermal CSTR presented in Boudriga, (1990), and used by Kalagnanam and Diwekar (1997), Bernardo and Saravia, (1998) and Dominguez and McLean (2006). In the reactor two first-order reactions in series of the form A→B→C take place. The objective was to design a process with a minimal overall cost, consisting of capital, operating and quality costs, while maintaining a nominal rate of production of species B, P_B=60 moles/min, in order to facilitate further downstream separation. The cost of separation is directly related to the variability of the quality characteristic (P_B) through a Taguchi loss function (see Table 5.7). The capital cost is the annually amortized reactor investment which is expressed as a function of the reactor volume, V. The operating costs are associated with the cost of handling the flowrate, F, and the supplied heat, Q. The mathematical model involves an energy balance around the reactor and two material balances for components A and B, and is presented in Table 5.7.

Table 5.7 Non-isothermal CSTR model

Objective function ($/year) :

$$\text{Overall Cost} = \text{Capital Cost} + \text{Utility Costs} + \text{Pumping Costs} + \text{Quality Costs}$$

$$\text{Capital Cost} = 4.18(900/280)(937.7/3)(V/\pi)^{0.6227}$$

$$\text{Utility Costs} = (900/786)(7896 - 6327q + 4.764x10^4 q^2 - 1022x10^4 q^4),\ q = Q/1.71x10^6$$

$$\text{Pumping Cost} = (900/834)(38.60/3)(264.2F^{0.8050})$$

$$\text{Quality Costs} = k[(\mu_{PB} - 60)^2 + \sigma_{PB}^2]$$

Material balance in reactor:

$$C_A = C_{Ai}/(1 + k_A^0 e^{-E_A/RT}\tau),\ C_B = (C_{Bi} + k_A^0 e^{-E_A/RT} C_A)/(1 + k_B^0 e^{-E_B/RT}\tau),\ \tau = V/F,\ PB = r_B V$$

Energy balance in reactor:

$$Q = F\rho C_P(T - T_i) + V(r_A H_{RA} + r_B H_{RB})$$

Kinetics:

$$-r_A = k_A^0 e^{-E_A/RT} C_A,\ -r_B = k_B^0 e^{-E_B/RT} C_B - k_A^0 e^{-E_A/RT} C_A$$

Operating constraints:

$$0.012 \le F \le 0.090$$

$$1.6x10^4 \le Q \le 1.6x10^8$$

$$290 \le T \le 310$$

Quality constraint:

$$(\mu_{RB} - 60)^2 \le 1$$

* The reactor volume, V, is given in m^3. The flow rate, F, and the heat supplied to the reactor, Q, are given in $m^3 \cdot min$ and $J \cdot min^{-1}$, respectively. Exiting concentrations of component A and B, C_A and C_B, are expressed in $mol \cdot m^{-3}$ and the reaction rates, r_A and r_B, in $mol \cdot m^{-3} \cdot min^{-1}$.

Deterministic parameters are shown in Table 5.8.

Table 5.8 Deterministic parameters for the CSTR model

Parameter	Values	Units	Description
R	8,14	$J \cdot mol^{-1} \cdot K^{-1}$	Ideal gas constant
C_p	$3.2x10^2$	$J \cdot kg^{-1} \cdot K^{-1}$	System specific heat
k	5	$\$ \cdot min^2 \cdot mol^{-2}$	Taguchi loss constant

Variations in the inlet temperature and inlet concentrations were assumed to be described by a normal distribution with mean, μ, and standard deviation, σ. In addition, 7 parameters in the model, (i.e., ρ, k_A^0, k_B^0, H_{RA}, H_{RB}, E_A, E_B), were assumed to be uncertain. Nominal values with corresponding standard deviations for the uncertain parameters as well as for the variables subjected to variation are shown in Table 5.9.

Table 5.9 Uncertain parameters for the CSTR model

	Parameter	*Units*	μ	σ
C_{Ai}	Inlet concentration of A	$mol{\cdot}m^{-3}$	3118	62.36
C_{Bi}	Inlet concentration of B	$mol{\cdot}m^{-3}$	342	6.84
T_i	Inlet temperature	°K	300	6
ρ	System density	$kg{\cdot}m^{-3}$	1180	23.6
k_A^0	Pre-exponential constant of A	min^{-1}	$8.4\text{x}10^5$	16800
k_B^0	Pre-exponential constant of B	min^{-1}	$7.6\text{x}10^5$	1520
H_{RA}	Molar heat of reaction of A	$J{\cdot}mol^{-1}$	$-2.12\text{x}10^4$	-424
H_{RB}	Molar heat of reaction of B	$J{\cdot}mol^{-1}$	$-6.36\text{x}10^4$	-1272
E_A	Activation energy of A	$J{\cdot}mol^{-1}$	$3.64 \text{ x } 10^4$	728
E_B	Activation energy of B	$J{\cdot}mol^{-1}$	$3.46 \text{ x } 10^4$	692

Using the formulation described in Equation 5.5, the system was modeled in GAMS (Brook *et al.*, 2005) with CONOPT2 as specified solver. The overall expected cost was obtained using the proposed quasi-Monte-Carlo techniques for an increasing number of points. Figure 5.13 shows the results obtained. Again, the performance of the proposed sampling techniques was evaluated by estimating the minimum number of points needed to converge within $\pm$ 0.5% of the true overall expected cost. The first techniques to converge to an accurate solution were SSS and HSS, both with 200 points. SSS provided an overall expected cost of $10,920 per year whereas HSS provided $10,939 per year. These two techniques were followed by HalSS and FSS both requiring 250 samples. The estimated overall expected costs obtained were $10,913 and 10,993 per year, respectively. It should be noted that, despite the fact that HalSS required 250 samples to converge to the required accuracy, it remained close to the lower bound after 150 making its accuracy acceptable after less strict accuracy requirements. FSS presented highly oscillatory behaviour finally converging with 250 points.

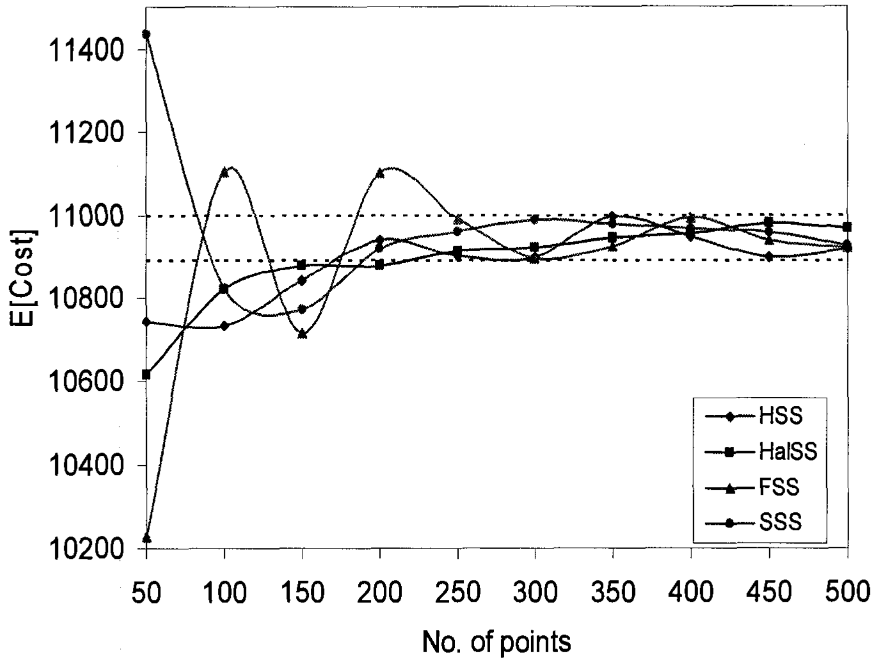

Figure 5.13 Estimations of the overall expected Cost obtained with quasi-Monte-Carlo sampling techniques for the CSTR problem.

5.6 Conclusions

In this work, the quasi-Monte-Carlo sequences of Halton, Faure and Sobol were examined for their potential use as sampling techniques. They were compared with an effective sampling technique currently used, HSS. Their uniformity was assessed by calculating their discrepancies. It was shown that Faure presents the highest discrepancy of the three. However, this did not seriously affect its effectiveness in the estimation of the expected values of various functions studied in this work. Interactions among the discrepancy, the dimensionality, number of samples and type of function being integrated influenced the estimation of expected values.

Uniform samples generated by the quasi-Monte-Carlo sequences were transformed to represent samples derived from normally distributed parameters. It was shown that despite the fact that the procedure of transformation somewhat distorted the uniformity of the proposed sampling techniques, this was not detrimental for their performance, though it slowed their rate of convergence.

The performance of the proposed techniques was also evaluated using two engineering design problems. The rate of converge to the specified accuracy was determined by solving the optimization problem for different number of points. It was found that SSS can perform equal or even better than HSS.

Nomenclature

$\mathbf{d}$	Vector of design variables
$\mathbf{z}$	Vector of control variables
$\mathbf{x}$	Vector of state variables
$\mathbf{h}$	Vector of equality constraints
$\mathbf{g}$	Vector of inequality constraints
$\mathbf{C}$	Desired correlation matrix of $\mathbf{\Phi}$
$\mathbf{P}$	Lower triangular matrix of $\mathbf{C}$
$\mathbf{Q}$	Lower triangular matrix of Γ
$\mathbf{R}^*$	Target correlation matrix
$\mathbf{S}$	Matrix obtained in the operation $\mathbf{P} \times \mathbf{Q}^{-1}$.
$\mathbf{V}$	Variance Matrix of uncertain parameters
k	Taguchi loss constant
N_p	Number of sampling points
w_i	weighting coefficients
f	Scalar function
j	Probability density function
J	Probability distribution function
p	Decimal-base integer
b	Base
m	Integer part of $\log_R p=[(\ln p)/(\ln R)]$
ϕ_b, ϕ_R	van der Corput base b/Radix inverse function
R	Prime number
Z_k	Halton Sequence

D^*_n	Star discrepancy
u	Uniformly distributed sample
z	Normally distributed sample
$\mathbf{z}$	Vector of normally distributed samples

Greek letters

θ	Uncertain parameter
$\mathbf{\Phi}$	Matrix of uncertain parameters
μ_θ	Mean of the uncertain parameter
σ_θ	Standard deviation of uncertain parameter
$\boldsymbol{\mu}$	Vector of means of uncertain parameters
$\mathbf{\Sigma}$	Variance-covariance matrix of uncertain parameters
ε	Error of integration
Γ	Correlation matrix of uncertain parameters
Θ	Domain of uncertain parameters

Set theory

$\forall$	For all
$\in$	Element of
$\exists$	There exist
D	Set of design variables
X	Set of state variables
Z	Set of control variables

Optimization acronyms

min	Minimize
w.r.t	With respect to
s.t.	Subject to

Subscripts

i Realization point i

References

Acevedo, J., and Pistikopoulos, E. N., "Stochastic optimization based algorithms for process synthesis under uncertainty", Computers and Chemical Engineering, Vol. 22, 647-671, 1998.

Bernardo F., P., Pistikopoulos E., N. and Saraiva P. M., "Integration and computational issues in stochastic design and planning optimization problems", Industrial and Engineering Chemistry Research, Vol. 38, 3056, 1999.

Bernardo, F., P., Pistikopoulos E. N. and Saraiva P. M., "Quality cost and robustness criteria in chemical process design optimization", Computers and Chemical Engineering, Vol 21, pp.27-40, 2001.

Brook, A., Kendrick, D., Meeraus, A. and Raman, R., *GAMS* — a user's guide, Washington, DC, 2005.

Bratley, P., and Fox, B., "Algorithm 659: Implementing Sobol's quasirandom sequence generator", ACM Transactions on Mathematical Software, Vol. 14, No. 1, 88-100, 1988.

Carnahan, B., Luther, H. A., and Wilkes. J.O., "Applied numerical methods", Wiley, New York, 1969.

Diwekar, U. M. "A novel sampling approach to combinatorial optimization under uncertainty", Computational Optimization and Applications, Vol. 24, 337-371, 2003.

Diwekar, U. M., and Rubin, E. S. "Parameter design methodology for chemical processes using a simulator", Industrial Engineering Chemistry Research, 33, 292-298, 1994.

Diwekar, U. M., and Kalagnanam, R. J., “Efficient sampling technique for optimization under uncertainty”, AIChE Journal, vol. 43, No. 2, 1997.

Dominguez-Palomeque, L. F., and McLean, D., “Investigation of robust techniques for the design of chemical processes under uncertainty”, M.A.Sc. thesis, University of Ottawa, 2006.

Faure, H., “Discrepance de suites associees a un systeme de numeration (en dimension s)“, Acta Arithmetica, Vol. 16, 337-351, 1982.

Fox, B., “Algorithm 647: Implementation and relative efficiency of quasirandom sequence generators”, ACM Transactions on Mathematical Software, Vol. 12, No. 4, 1986.

Gentle, J. E., “Random number generation and Monte Carlo methods”, Statistics and Computing series, Springer, Second edition, 2003.

Giunta, A. A., Wojtkiewicz Jr., S. F., and Eldred, M. S., “Overview of modern design of experiments methods for computational simulations”, American Institute of aeronautics and Astronautics, AIAA2003-0649, 2003.

Halton, J., “On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals”, Numerische Mathematik, Vol. 2, 84-90, 1960.

Hammersley, J., “Monte Carlo methods for solving multivariable problems”, Proceedings of the New York Academy of Science, Vol. 86, 844-874, 1960.

Hlawka, E., “Functionen von beschranker Variation in deter Therie der Gleichverteilung”. Ann. Math. Pura Appl., No., 54, 1961.

Iman R. L., and Conover W. J. "Small sensitivity analysis techniques for computer models, with an application to risk assessment", Communications in Statistics, Vol. A17, 1749-1842, 1982.

Kocis, L., and Whiten, W., "Computational investigations of low-discrepancy sequences", ACM Transactions on Mathematical Software, Vol. 23, No. 2, 266-294, 1997.

Lemieux and P. L'Ecuyer, "On the use of quasi-Monte-Carlo methods in computational finance", Computational Science, Lecture Notes in Computer Science, Springer-Verlag, 607-616, 2001.

Morokoff, W. J., and Caflisch R. E., "Quasi-random sequences and their discrepancies", SIAM J. Sci. Comput., Vol. 15, No. 6, 1251-1279, 1994.

Morgan, G, and Henrion, M. Uncertainty: "A guide to dealing with uncertainty in quantitative risk and policy analysis", Cambridge University press: New York, 1990.

Niederreiter, H., "Low-discrepancy and low-dispersion sequences", Journal of Number Theory, Vol. 30, 51–70, 1988.

Niederreiter, H., "Random number generation and quasi-Monte-Carlo methods", CBMS-NSF Regional Conference Series in Applied Mathematics, Philadelphia, Pennsylvania, 1992.

Novak, E., ''Deterministic and stochastic error bounds in numerical analysis,'' Lecture Notes in Mathematics, Vol. 1349, Springer-Verlag, New York/Berlin, 1988.

Papageorgiou, A., and Wasilkowski, G. W. "Average case complexity of multivariate problems", J. of Complexity, Vol. 6, No. 1, 1990.

Pistikopoulos, E. N., "Uncertainty in process design and operations", Computers and Chemical Engineering, Vol. 19, Supplement, s553-s563, 1995.

Pistikopoulos, E. N., Ierapetritou, M. G. Novel approach for optimal process design under uncertainty, Computers and Chemical Engineering Vol. 19, No. 10, 1089, 1995.

Press, W. H., and Teukolsy, S. A., "Quasi- (that is, sub-) random numbers", Comput. Phys., No. 3, 76-70, 1989.

Press, W. H., Teukolsy, S. A., Vetterling W. T., and Flannery, B. P., "Numerical recipes in Fortran – The art of scientific computing", Second Edition, Cambridge University Press, 1992.

Sobol, I., and Levitan, Y. L, "The production of points uniformly distributed in a multidimensional cube", Preprint IPM Akad. Nauk SSSR, No. 40, Moscow, 1976.

Straub, D. A. and Grossman I. E., "Integrated stochastic metric of flexibility for systems with discrete state and continuous parameter uncertainties", Computers and Chemical Engineering, Vol. 14, No.9, pp. 967-985, 1990.

Thiémard, E., "An algorithm to compute bounds for the star discrepancy", Journal of Complexity, Vol 17, 850-880, 2001.

van der Corput, J. "Verteilungsfunktionen I & II", Nederl. Akad. Wetensch. Proc., Vol. 38, 813-820, 1935.

Wasilkowski, G. W., "Average case complexity of multivariate integration and function approximation: An overview", J. Complexity, 12(4): 257-272, 1996.

Wozniakowski, H., "Average case complexity of multivariate integration", Bulletin of the American Mathematical Society, Vol. 24, 185-194, 1991.

Chapter 6

Conclusions, Contributions and Recommendations

This chapter presents the conclusions, contributions and recommendations derived from this work. This work focused on the robust design of chemical processes when uncertainty in the values of model parameters and variability in process inputs are present.

6.1 Conclusions

With respect to the formulation of the objective function in the robust design of chemical processes the following conclusions can be drawn:

1) Careful consideration should be exercised when formulating the objective function to be optimized. Minimization of quality costs only by means of reducing the variability in quality variables can lead to over-designs which may increase the overall cost of the process or to under-designs which may compromise the safe operation of a process.

2) When closeness-to-target of quality characteristics is of paramount importance for the designer, the approach of Vining and Myers (1990) should be employed. On the other hand, when neither closeness-to-target nor variability of quality characteristics are favored and minimum quality costs are sought, the approach of Lin and Tu (1995) should be used. When exact information about the relative importance about both aspects of the quality characteristic is known, a more sophisticated approach such as that of Ding *et al.* (2004) should be employed. This approach provides a trade-off between closeness-to-target with low variability and minimum quality costs.

With respect to the effect of uncertainties in process design it is concluded that:

3) The overall expected cost of designing a process can be further reduced by eliminating the uncertainty of key parameters. Identification of parameters, whose uncertainty contributes the most to the overall expected cost, can be obtained by performing a sensitivity analysis. This is performed by eliminating the uncertainties of each uncertain parameter one at a time. This is a straightforward approach and, though interactions between uncertain parameters are not accounted for, it provides information for cost reduction.

4) The elimination of the uncertainty in key parameters highly depends on the objective sought. While the elimination of the uncertainty of one key parameter leads to a further reduction of the overall cost, for other design objectives (e.g., objective function with additional constraints) further reduction of the overall cost may not be achieved by eliminating the uncertainty in the same key uncertain parameter.

5) The magnitude of the mean values of the Taguchi loss constant had a significant impact on the final cost of a design. Processes having large quality costs are expected to have high capital and operating costs. On the other hand, for the range of values of the Taguchi loss constant studied in this work, the uncertainty in k had no significant impact on the overall cost of a design.

With respect to the sampling techniques as an approach to numerical integration of multidimensional integrals derived from probability distribution functions, one can conclude that:

6) Quasi-Monte-Carlo sampling techniques presented better performance than Monte Carlo Sampling Techniques. HSS showed better performance than MCS and LHS for the solution of stochastic optimization problems.

7) Despite the fact that cubatures require a low number of points for the estimation of expected values of objective functions, they did not provide accurate estimates of the objective functions for the problems that were solved in this work. The author is unaware of any formula capable of providing estimates of standard deviations and/or variances. This may represent a limitation, for instance, in planning problems where robustness in operating costs has to be ensured and quantified (Ahmed and Sahidinis, 1998).

8) The newly proposed quasi-Monte-Carlo sampling techniques, showed the potential of reducing the computational time involved in the solution of stochastic optimization problems. Among the quasi-Monte-Carlo sampling techniques proposed, SSS presented the lowest discrepancy. In one of the problems solved in this work, SSS required the fewest number of points. In the other, it required the same number of points as HSS. SSS seem to compete with HSS in the estimation of objective functions with the fewest number of function evaluations.

6.2 Contributions

This section summarizes the main contributions of this work in the field of chemical process design with mathematical programming techniques. The main contributions of this thesis are in the area of robust formulations for the minimization of quality costs and in the area of sampling techniques for numerical integration in the solution of stochastic optimization problems.

First, this work represents one of the first studies in providing a comparison of the current formulations for the minimization of quality cost in the field of chemical engineering. While most of the work has been in the field of quality engineering, with standard nonlinear programming techniques, this work focused on minimizing quality costs at the design stage of a chemical process within a stochastic optimization framework.

Second, three sampling techniques based on the quasi-Monte-Carlo sequences of Halton, Faure and Sobol were evaluated for the solution of stochastic optimization problems. Quasi-Monte-Carlo sequences are widely recognized and used in areas such as finance and economics. However, in the field of chemical engineering, these sequences have been overlooked. In this work, one of the first evaluations, if not the first one, of their performance for the solution of stochastic optimization problems for the design of chemical processes has been provided.

6.3 Recommendations

This section summarizes the recommendations derived from this work.

1) Throughout this work the uncertainty in process parameters was characterized by probability density functions (PDF's), specially, Gaussian. However other PDF's can also be incorporated such as Gamma, Poison or Binomial distributions. These distributions can be accounted for by modifying the procedure of transformation used by the sampling techniques. It is recommended that the uncertainties in form of other PDF's be studied and their effect assessed on the solution of stochastic optimization problems derived from process design.

2) It was assumed that the uncertainty level was fixed (i.e., standard deviation was assumed to be 10% of the mean, or it was already defined). For further work, other levels of uncertainty should be investigated over a wide level of uncertainty in order to assess their effect on the overall expected cost of a process.

3) In the sensitivity analysis performed in this work all the uncertainty in the parameters under consideration was eliminated. A more in-depth work should approach the optimal uncertainty reduction that would minimize the overall cost. This however is not an easy task, it requires of proper identification of the real cost involved in eliminating the uncertainties such as investment cost for laboratory facilities and

operating cost involved in performing the experiments in addition to a more complex formulation of the optimization problem. Recently, Bernardo *et al.*, (2000) has attempted to address this fact by formulating the problem a Mixed-Integer Nonlinear Programming (MINLP) problem.

References

Ahmed, S., and Sahinidis, N. V., "Robust process planning under uncertainty", Industrial and Engineering Chemistry Research, Vol. 37, 1883, 1998.

Bernardo, F. P, Saraiva, P., and Pistikopoulos, E. N., "Inclusion of information in process design optimization under uncertainty", Computers and Chemical Engineering, Vol. 24, 1695-1701, 2000.

Ding, R., Lin, D. K., and Wei, D., "Dual response surface optimization: A weighted MSE approach", Quality Engineering, Vol. 16, No. 3, 377-385, 2004.

Lin, D. K., and Tu, W., "Dual response surface optimization", Journal of Quality Technology, Vol. 27, No.1, 1995.

Vining, G. G., and Myers, R. H., "Combining Taguchi and response surface philosophies: A dual response approach", Journal of Quality Technology, Vol. 22, No.1, 1990.

Appendix A

Sampling Technique Codes

Sampling techniques were used throughout this work as an approach to numerical Integration for the solution of stochastic optimization problems. The following sampling techniques were coded in Matlab®.

MCS

The code MCS generates N_p samples derived from a multivariate normal probability density function with mean, μ, and standard deviation σ. All samples generated are assumed to be independent. It uses an inverse normal CDF code developed by Peter Acklam (1996) and the RAND function to generate uniformly distributed random numbers.

LHS

The code LHS generates N_p samples drawn from a multivariate normal probability density function with mean with mean, μ, and standard deviation σ. It uses the method of Stain (1987) to generate the samples. No correlation among input variables is assumed.

HSS

The code HSS generate N_p samples derived from a multivariate normal probability density function with mean, μ, and standard deviation σ. No correlation is imposed. It uses an inverse normal CDF code developed by Peter Acklam (1996) and a Hammersley sequence code developed by John Burkardt (2006).

HalSS

The code HalSS generates N_p samples derived from a multivariate normal probability density function with mean, μ, and standard deviation σ. No correlation is imposed. It uses an inverse normal CDF code developed by Peter Acklam (1996) and a Halton sequence code developed by John Burkardt (2006).

FSS

The code FSS generates N_p samples derived from a multivariate normal probability density function with mean, μ, and standard deviation σ. No correlation is imposed. It uses an inverse normal CDF code developed by Peter Acklam (1996) and a Faure sequence code developed by John Burkardt(2006).

SSS

The code SSS generates N_p samples derived from a multivariate normal probability density function with mean, μ, and standard deviation σ. No correlation is imposed. It uses an inverse normal CDF code developed by Peter Acklam (1996) and a Sobol sequence code developed by John Burkardt (2006).

Reference

Acklam, P. J., "Monte Carlo methods in state space estimation". Cand. Scient. thesis. University of Oslo, Norway, 1996.

Burkardt, J., Matlab Software resource, http://www.csit.fsu.edu/~burkardt/m_src.html

Stein, M., "Large Sample Properties of Simulations Using Latin Hypercube Sampling", Technometrics, Vol. 29, 143-151, 1987.

Appendix B

The HSS code

In this section only the HSS code is provided. Codes for the other sampling techniques can be easily obtained by modifying the procedure of generation of uniformly distributed sample numbers in the interval [0,1]. For instance, by replacing the code i_to_hammersley_sequence by i_to_halton_sequence, the HalSS code is obtained.

```
function s=hss(xmean,xsd,nsample)

%HSS generate nsamples derived from a multivariate normal probability
%density function. All samples generated are assumed to be
%independent.It uses Peter Acklam inverse normal CDF and a Hammersley
%sequence code developed by John Burkardt. %
%http://www.csit.fsu.edu/~burkardt/f_src/hammersley/hammersley.html
% Input:
%   xmean   :  mean of data (1,nvar)
%   xsd     :  std.dev of data (1,nvar)
%   nsample : no. of samples
%   nvar    : no. of variables
% Output:
%   s       : random sample (nsample,nvar)

% Luis Dominguez (2004)

nvar=length(xmean);
s=zeros(nsample,nvar);
    % Generation of Hammersley sequence
          step = 1;
          seed(1:nvar) = 0;
          leap(1:nvar) = 1;
          base(1) = nsample;
       for ( i = 2 : nvar )
          base(i) = prime ( i - 1 );
       end
    %I_TO_HAMMERSLEY computes an element of the Hammersley subsequence.
    r = i_to_hammersley_sequence (nvar,nsample,step,seed,leap,base);

    % Transformation
    HSS_matrix=r';
    s=zeros(nsample,nvar);
        for j=1: nvar
          s(:,j) = xmean(j) + ltqnorm(HSS_matrix(:,j)).* xsd(j);
       % ltqnorm is an algorithm for computing the inverse normal
       % cumulative distribution function. Author Peter J. Acklam
        end
%-------------------------------------------------------
```

Appendix C

Internal Functions used in Sampling Technique Codes

The sampling techniques used in this work mainly consist of two procedures which were implemented in Matlab®:

1. The generation of sample numbers in the interval [0,1] with uniform probability density function.
2. The transformation of the sample numbers to provide a set of samples for the variables of interest.

C.1 Generation uniformly distributed sample numbers

The sample numbers uniformly distributed in the interval [0,1] were generated with the following internal functions.

RAND

The rand function is an internal function of Matlab which generates arrays of random numbers whose elements are uniformly distributed in the interval (0,1). RAND is used in the Monte Carlo Sampling (MCS) code.

I_TO_HAMMERSLEY_SEQUENCE

The function i_to_hammersley_sequence is part of the Hammersley Sequence Sampling (HSS) code. It is a function that computes N_p elements of the Hammerlsey n-dimensional quasi-random sequence. Except for the first dimension, i_to_hammersley_sequence generates a 1-dimensional van der Corput sequence using bases of successive primes (2, 3, 5, 7, 11….). This function was modified so that the first base was $1/N_p$.

i_to_hammersley_sequence was developed by John Burkardt from the Florida State University and is available at:

http://www.csit.fsu.edu/~burkardt/f_src/hammersley/hammersley.html

I_TO_HALTON_SEQUENCE

The function i_to_halton_sequence is part of the Halton Sequence Sampling (HalSS) code. It is a function that computes N_p elements of the Halton k-dimensional quasi-random sequence. It is derived from the 1-dimensional van der Corput sequence. Each dimension uses a different prime number as the base of the sequence.
i_to_halton_sequence was developed by John Burkardt from the Florida State University and is available at http://www.csit.fsu.edu/~burkardt/m_src/halton/halton.html

FAURE

Fuare is implemented in the Faure Sequence Sampling (FSS) code. This function is Matlab implementation of the code ACM TOMS Algorithm 647 originally developed by Bennett Fox (1986). It is a merging and adaptation of the routines IFAUR and GOFAUR from the above mentioned codes. This implementation is due to John Burkardt form the Florida State University.
Faure is available at http://www.csit.fsu.edu/~burkardt/m_src/faure/faure.html

I4_SOBOL

i4_sobol is implemented in the Sobol Sequence Sampling (SSC) code and it is an implementation of the sobol quasi-random sequence. i4_Sobol is an improved adaptation of the ISOBL and GOSOBL routines presented in the papers of Bennett (1986) and Bratley and Fox, (1988). The implementation in Matlab is due to John Burkardt from the Florida State University.
i4_sobol is available at http://www.csit.fsu.edu/~burkardt/m_src/sobol/sobol.html.

C.2 Transformation of uniformly distributed sample numbers

The transformation of the uniformly distributed sample numbers was performed with the subroutine LTQNORM.

LTQNORM

This function returns the inverse of the standard normal distribution function. Performs the same operation as SQRT(2)*ERFINV(2*k-1) with standard functions of Matlab. The algorithm uses a minimax approximation by rational functions. It has a relative error less than 1.15e-9. This function is due to Peter Acklam (1996).
LTQNORM is available at http://home.online.no/~pjacklam/index.html.

References

Acklam, P. J., "Monte Carlo methods in state space estimation". Cand. Scient. thesis. University of Oslo, Norway, 1996.

Bratley, P., and Fox, B., "Algorithm 659: Implementing Sobol's quasirandom sequence generator", ACM Transactions on Mathematical Software, Vol. 14, No. 1, 88-100, 1988.

Fox, B, "Algorithm 647: Implementation and Relative Efficiency of Quasirandom Sequence Generators", ACM Transactions on Mathematical Software, Vol. 12, No. 4, 362-376, 1986.

Appendix D

Optimization software

D.1 Optimization using MATLAB

For the solution of stochastic optimization problems in Chapter three, the Optimization Toolbox of Matlab® was employed. The toolbox includes routines for different types of optimization including unconstraint nonlinear and constrained nonlinear minimization. In Chapter three, the robust design of a CSTR was posed a stochastic optimization problem. The problem involved equality and inequality constraints that were highly nonlinear. Because information about the first and second order derivatives was not available, the problem was solved using medium-scale algorithm. The first and second order derivatives were approximated by finite differences. Medium-scale algorithms employ variations of successive quadratic programming (SQP). The solver used to perform the optimization was *fmincon*.

FMINCON

The function *fmincon* attempts to find a constrained minimum of a scalar function of several variables starting at an initial estimate. It uses a sequential quadratic programming (SQP) method. In this method, the function solves a quadratic programming (QP) subproblem at each iteration. An estimate of the Hessian of the Lagrangian is updated at each iteration using the BFGS formula.

FZERO

This function is used in the computation of the quality variable, RB, in which the state variable, T, is calculated iteratively. *fzero* is an adaptation of a Fortran version developed by Forsythe et al (1976). It uses a combination of bisection, secant and inverse quadratic interpolation methods.

D.2 Optimization using GAMS

GAMS stands for the General Algebraic Modelling System (Brook et al, 2005). It is a high-level modeling system suitable for mathematical programming and optimization. GAMS was design for modeling nonlinear complex models. Currently, there are three types of algorithms available in GAMS, CONOPT, MINOS and SNOPT. CONOPT was used in the solution of one-stage stochastic optimization problems in Chapter Four. A brief description of this solver is presented here.

CONOPT

Conopt is a well suited solver for models with very nonlinear constraints. It has a fast method for finding a first feasible solution for models with few degrees of freedom. CONOPT make use of the GRG algorithm. CONOPT searches through the feasible region for optimal solution points. Each point tested in the optimization process is feasible and the value of the objective function improves at each iteration. Since each point generated in the search procedure is feasible, the final value of the objective function is also feasible.

References

Forsythe, G. E., M. A. Malcolm, and C. B. Moler, "Computer Methods for Mathematical Computations", Prentice-Hall, 1976.

Brook, A., Kendrick, D., Meeraus, A. and Raman, R., *GAMS* — a user's guide, Washington, DC, 2005.

Appendix E

Optimization Codes

E.1 Optimization code of the CSTR system

```
function stochasticsim
clear variables
tiempo=cputime;
global Ka Kb Hra Hrb Ea Eb Cp R rho
%global mean_Rb

Ka=8.4e5;
Kb=7.6e4;
Hra=-2.12e4;
Hrb=-6.36e4;
Ea=3.64e4;
Eb=3.46e4;
Cp=3.2e3;
R=8.314;
rho=1180;

%Cai=3118;  <----- reference value
%Cbi=342;   <----- reference value
%Ti=300;    <----- reference value
%Q=1.71e6;  <----- reference value
%V=0.0391;  <----- reference value
%F=0.0781;  <----- reference value
%T=314.05;  <----- reference value


% Execution of Stochastic Optimization
%initialguess;
    weight=0.5
    [x,fval] = optimvarHSS (weight);
    objective=20*(variance_Rb + bias);
    fprintf('\n The Objective function is = %6.6f \n',objective);
    fprintf('\n %6.6f',x);
    fprintf('\n        ');
    fprintf('\n (mean_Rb-60)^2 = %6.6f',bias);
    fprintf('\n mean_Rb= %6.6f',mean_Rb);
    fprintf('\n Variance = %6.6f',variance_Rb);
    fprintf('\n w = %6.6f',weight);

tiempo2=cputime;
CPU_time=(tiempo2-tiempo)/60

%-------------------------------------------------------------------------
function [x,fval] = optimvarHSS(weight)
objval = []; % Initialize shared variable
```

```
  s=[];
  mean_Rb=[];
  initialguess =[3118;342;300;0.0391;0.078;2.71e6];%Make a starting guess
  options =
optimset('Display','iter','LargeScale','on','MaxFunEvals',600);
  [x, fval] = ...
  fmincon(@objfun,initialguess,[],[],[],[],...
   [0;0;0;0;0;0],...
   [inf;inf;inf;inf;inf;inf],...
   @constrfun,options);

%-----------------------------------------------------------------------------
 function f = objfun(x)
  % Error level 10 %
  error_level=10; %Error level of input variables for simulation
        mean_input(1)=x(1);
        mean_input(2)=x(2);
        mean_input(3)=x(3);
        mean_input(4)=x(4);
        mean_input(5)=x(5);
        mean_input(6)=x(6);
       %
         sigma_input(1)= error_level* x(1) /100;  % x(7);
         sigma_input(2)= error_level* x(2) /100;  % x(8);
         sigma_input(3)= error_level* x(3) /100;  % x(9);
         sigma_input(4)= error_level* x(4) /100;  % x(10);
         sigma_input(5)= error_level* x(5) /100;  % x(11);
         sigma_input(6)= error_level* x(6) /100;  % x(12);

         %Number of samples
         %LOOP for the different sammples sizes
          n=550;
          s=hss2(mean_input,sigma_input,n);

          for i=1:n
            Cai_prime= s(i,1); %mean_input(1);
            Cbi_prime=s(i,2);  %mean_input(2);
            Ti_prime= s(i,3);  %mean_input(3);
            V_prime= s(i,4);   %mean_input(4);
            F_prime= s(i,5);   %mean_input(5);
            Q_prime= s(i,6);   %mean_input(6);
            Thao_prime=V_prime/F_prime;
            % Sub routine that computes the state and quality variables
            [Ca_prime,Cb_prime,ra_prime,rb_prime,T] = calc(Cai_prime,...
                Cbi_prime,Q_prime,F_prime,V_prime,Ti_prime,Thao_prime);

            Ca_primex(i,1)=Ca_prime;
            Cb_primex(i,1)=Cb_prime;
            ra_primex(i,1)=ra_prime;
            rb_primex(i,1)=rb_prime;
            T_primex(i,1)=T;
            Rb(i,1)=V_prime*rb_primex(i,1);
          end

       %Calculation of mean's quality variable
       Rb;
```

```
        vector_size=length(Rb);
        mean_Rb=mean(Rb);
        %Calculation of sdev's quality variable
        standard_deviation_Rb=std(Rb);
        %Calculation of Variance
        variance_Rb=(standard_deviation_Rb)^2 ;
        bias= (mean_Rb-60)^2;
        % Nonlinear objective function
        % Variable objval shared with objfun and runsharedvalues
          objval = 20*((1-weight)*variance_Rb + weight*bias);
          %weight skipped from bias term --- loss function
          f = objval;
  end
%----------------------------------------------------------------------
     function [c,ceq] = constrfun(x)
        % Nonlinear inequality constraints
        % Variable objval shared with objfun and optimizingvarianceHSS
   for i=1:550
         c(i) = - s(i,1) + 1000;
   end
   for i=551:1100
       j=i-550;
         c(i) = - 5000 + s(j,1);
   end
   for i=1101:1650
         j=i-1100;
         c(i) = - s(j,2) + 100;
   end
   for i=1651:2200
       j=i-1650;
         c(i) = - 500 + s(j,2);
   end
   for i=2201:2750
         j=i-2200;
         c(i) = - s(j,3) + 210;
   end
   for i=2751:3300
       j=i-2750;
         c(i) = - 390 + s(j,3);
   end
   for i=3301:3850
         j=i-3300;
         c(i) = - s(j,4) + 0.01;
   end
   for i=3851:4400
       j=i-3850;
         c(i) = - 0.09 + s(j,4);
   end
 for i=4401:4950
         j=i-4400;
         c(i) = - s(j,5) + 0.01;
   end
   for i=4951:5500
       j=i-4950;
         c(i) = - 0.17 + s(j,5);
   end
 for i=5501:6050
```

```
        j=i-5500;
        c(i) = - s(j,6) + 1126000;
  end
  for i=6051:6600
      j=i-6050;
        c(i) = - 2394000 + s(j,6);
  end

      ceq = [];
    end
%-----------------------------------------------------------------------
end
end
```

E.2 Optimization code of the reactor-heat-exchanger system.

```
$title Reactor Heat exchanger system
set
  i /1, 2, 3, 4, 5, 6/
  j /F, Xa, Tw2, Fw, T, T2/
  s /A, V/
  k /1*10/

Parameters  Cai, ratio, Hr, Cp, Cpw, kloss;

Table input(k,i)
        1        2        3        4        5        6
1      39.43    333.17   292.59   11.45    1531.17  29.47
2      40.33    330.30   292.94   11.88    1587.43  30.40
3      40.92    336.12   292.47   12.36    1632.29  31.32
4      41.37    328.36   289.24   12.65    1670.86  31.85
5      41.74    334.55   296.40   10.71    1701.36  32.28
6      42.06    331.80   290.69   11.47    1747.98  32.36
7      42.34    338.30   296.96   12.04    1418.55  32.90
8      42.60    326.39   292.23   12.28    1539.60  33.40
9      42.84    333.85   289.04   12.80    1593.35  33.92
10     43.07    331.08   291.37   10.91    1610.42  34.92

scalar np /10/;

Variables deviation(k), expectcost, sdevobj  ;

Positive variables
des(s), meanXa, stdev, variance, descost, expopcost, sdevopcost, exppenalty,
sdevpenalty, y(k,j), z(k,j), z2(k,j), x2(k,j),  x(k,j), x(k,j),objective(k)
opcost(k), deltaXa(k), dTlm(k), penalty(k), denom(k), logarithm(k),
```

```
squared_dev(k);

Equations
obj(k), eq1, eq2(k), eq3, eq4, eq5(k), eq6(k), eq7(k), eq8(k), eq9(k), eq10(k),
eq11(k), eq12(k), eq13(k), eq14(k), eq15(k), eq16(k), eq17(k), eq18(k), eq19,
eq20(k), eq21(k, eq22, eq24, eq116, eq117, eq118,
objfunc;

obj(k).. objective(k) =e= descost + opcost(k) + penalty(k);

eq1.. descost =e= 691.2*des('V')**0.7 + 873.6*des('A')**0.6;
eq2(k).. opcost(k) =e= 1.76*z(k,'Fw') + 7.056*z2(k,'F');

eq3.. expopcost =e= (sum(k,opcost(k)))/np;
eq4.. exppenalty =e= (sum(k,penalty(k)))/np;

eq116.. sdevopcost =e=  sqrt((sum(k, power(opcost(k)- expopcost,2)))/(np-1));
eq117.. sdevpenalty =e=  sqrt((sum(k, power(penalty(k)- exppenalty,2)))/(np-1));
eq118.. sdevobj =e=  sqrt((sum(k, power(objective(k)- expectcost,2)))/(np-1));


* Reactor Material balance
eq5(k).. input(k,'1')*y(k,'Xa')-input(k,'4')*exp(-ratio/x(k,'T'))
      *input(k,'6')*(1-y(k,'Xa'))*des('V') =e= 0;

*Reactor energy balance
eq6(k).. input(k,'1')*Cp*(input(k,'2')-x(k,'T'))-z2(k,'F')*Cp*(x(k,'T')
      - x2(k,'T2'))-Hr*input(k,'1')*y(k,'Xa') =e= 0;

*Heat exchanger design equation
eq7(k).. dTlm(k) =e= ((x(k,'T') - x(k,'Tw2'))-( x2(k,'T2')
      - input(k,'3')))/logarithm(k);

eq8(k).. denom(k) =e= ((x(k,'T')-x(k,'Tw2'))/(x2(k,'T2')-input(k,'3')));
      denom.lo(k)=1;

eq9(k).. logarithm(k) =e= log(denom(k));
      logarithm.lo(k)=1e-3;

eq10(k).. z2(k,'F')*Cp*(x(k,'T')-x2(k,'T2')) =e= des('A')*input(k,'5')*dTlm(k);

*Heat exchanger Energy balance
eq11(k).. z2(k,'F')*Cp*(x(k,'T')-x2(k,'T2')) =e= z(k,'Fw')*Cpw*(x(k,'Tw2')
      - input(k,'3'));

*Additional constraints by modeler/designer
eq12(k).. x(k,'T') - x2(k,'T2') =g= 0;
```

```
eq13(k).. x(k,'Tw2') - input(k,'3') =g= 0;
eq14(k).. x(k,'T') - x(k,'Tw2') =g= 11.1;
eq15(k).. x2(k,'T2') - input(k,'3') =g= 11.1;

* Robustness criteria: Quality constraints are relaxed,
* with process performance f being penalised through a
* Taguchi loss function

eq16(k).. deltaXa(k) =g= 0.9-y(k,'Xa') ;
eq17(k).. deltaXa(k) =g= 0;
eq18(k).. penalty(k) =e= kloss*(deltaXa(k)**2);


*Additional robustness metric: second moment
eq19.. meanXa =e= (sum(k, y(k,'Xa')))/np;
eq20(k).. deviation(k) =e= y(k,'Xa')- meanXa;
eq21(k).. squared_dev(k) =e= power(deviation(k),2);

eq22.. variance =e= (sum(k, squared_dev(k)))/(np-1);

eq24.. stdev =e= sqrt(variance);
*eq25.. stdev =l= 0.002;

objfunc.. expectcost =e= (SUM(k, objective(k)))/np;
Model RHE /all/;

kloss=6.4e6;
*Taguchi loss constant
Cai=32.04;
*(Kmol/m3)- Concentration of A in the feed stream
ratio=555.6;
*K- Ratio (E/R) of activation energy to the perfect gas constant
Hr=-23260;
*KJ/mol - Molar heat of reaction
Cp=167.4;
*KJ/Kmol.K - Reactant heat capacity
Cpw=4.184;
*KJ/Kg.K - Cooling water heat capacity

* Boundary values for variables

x.lo(k,'T')= 311;
x.up(k,'T')= 389;
x2.lo(k,'T2')= 311;
x2.up(k,'T2')= 389;
x.lo(k,'Tw2')= 294;
x.up(k,'Tw2')= 323;
```

```
y.lo(k,'Xa') = 0;
y.up(k,'Xa') = 1;

*initial guess
des.l('A')= 1.045;
des.l('V')= 3.089;

z2.l(k,'F')=75.241;
y.l(k,'Xa')=0.89;
x.l(k,'Tw2')=323;
z.l(k,'Fw')=4177.37;
x.l(k,'T')=389;
x2.l(k,'T2')=347.37;


RHE.OPTFILE=1;

FILE OPT CONOPT option file / CONOPT.OPT /;
PUT OPT;
PUT
' set rtnwmi 1e-7'/
' set rtnwma 1e-6'/
PUTCLOSE OPT;

option nlp=examiner;
option nlp=conopt;

RHE.SCALEOPT=1;

solve RHE minimizing expectcost using nlp
```

Appendix F

Specialized Cubatures

SC's were used in Chapter Four to compare their accuracy with that of HSS in the estimation of the overall expected cost. In this section we present the parameter values used in the construction of the grid of points with corresponding weights.

F.1 Specialized cubature of degree 3 (SC_3)

For the SC_3, the grid of points, u_i, are constructed by a fully symmetric permutation of the elements and signs of a set of points (r, 0,…,0) with corresponding weights B_0. The parameter, r, and weights, B_0, are given in Table E.1 and the grid of points is presented in Table F.2.

Table F.1Parameter values used to perform the transformation $u_i \in R_n \rightarrow \theta_i \in \Theta$ for SC_3

n =	6
V =	31.00628
N_p =	12
r^2 =	3
r =	1.732051
B_0 =	2.583856

Table F.2 Grid of points θ_i in Θ space generated with SC_3

$\theta_1(u)$	$\theta_2(u)$	$\theta_3(u)$	$\theta_4(u)$	$\theta_5(u)$	$\theta_6(u)$
52.55	333.00	293.00	12.00	1635.00	32.04
38.17	333.00	293.00	12.00	1635.00	32.04
45.36	343.56	301.29	12.00	1635.00	32.04
45.36	322.44	284.71	12.00	1635.00	32.04
45.36	341.29	302.29	12.00	1635.00	32.04
45.36	324.71	283.71	12.00	1635.00	32.04
45.36	333.00	293.00	13.90	1635.00	32.04
45.36	333.00	293.00	10.10	1635.00	32.04
45.36	333.00	293.00	12.00	1894.22	32.04
45.36	333.00	293.00	12.00	1375.78	32.04
45.36	333.00	293.00	12.00	1635.00	37.12
45.36	333.00	293.00	12.00	1635.00	26.96

B_i
2.583856
2.583856
2.583856
2.583856
2.583856
2.583856
2.583856
2.583856
2.583856
2.583856
2.583856
2.583856

F.2 Specialized cubature of degree 5 (SC_5)

For the SC_5 the grid of points, u_i, are constructed the same way as in the SC_3. SC_5 has an additional parameter, s. The grid of points u_i are given by the fully symmetric permutation of the points (r, 0, …, 0) with corresponding weight B_0 and (s, s, …, s) with corresponding weight B_1. The parameters r and s, along with corresponding weights, B_0 and B_1, are given in Table F.3. The grid of points is presented in Table F.4.

Table F.3 Parameter values used to perform the transformation $u_i \in R_n \to \theta_i \in \Theta$ for SC_5

n	6
V	31.00628
r^2=	2
s^2=	1
r	1.414214
s	1
N_p=	76
B_0	1.937892
B_1	0.121118

Table F.4 Grid of points θ_i in Θ space generated with SC_5

$\theta_1(u)$	$\theta_2(u)$	$\theta_3(u)$	$\theta_4(u)$	$\theta_5(u)$	$\theta_6(u)$	B_i
51.23	333.00	293.00	12.00	1635.00	32.04	1.428034
39.49	333.00	293.00	12.00	1635.00	32.04	1.428034
45.36	341.62	299.77	12.00	1635.00	32.04	1.428034
45.36	324.38	286.23	12.00	1635.00	32.04	1.428034
45.36	339.77	300.59	12.00	1635.00	32.04	1.428034
45.36	326.23	285.41	12.00	1635.00	32.04	1.428034
45.36	333.00	293.00	13.55	1635.00	32.04	1.428034
45.36	333.00	293.00	10.45	1635.00	32.04	1.428034
45.36	333.00	293.00	12.00	1846.65	32.04	1.428034
45.36	333.00	293.00	12.00	1423.35	32.04	1.428034
45.36	333.00	293.00	12.00	1635.00	36.19	1.428034
45.36	333.00	293.00	12.00	1635.00	27.89	1.428034
49.51	343.88	303.15	13.10	1784.66	34.97	0.100409
49.51	343.88	303.15	13.10	1784.66	29.11	0.100409
49.51	343.88	303.15	13.10	1485.34	34.97	0.100409
49.51	343.88	303.15	13.10	1485.34	29.11	0.100409

49.51	343.88	303.15	10.90	1784.66	34.97	0.100409
49.51	343.88	303.15	10.90	1784.66	29.11	0.100409
49.51	343.88	303.15	10.90	1485.34	34.97	0.100409
49.51	343.88	303.15	10.90	1485.34	29.11	0.100409
49.51	334.31	292.42	13.10	1784.66	34.97	0.100409
49.51	334.31	292.42	13.10	1784.66	29.11	0.100409
49.51	334.31	292.42	13.10	1485.34	34.97	0.100409
49.51	334.31	292.42	13.10	1485.34	29.11	0.100409
49.51	334.31	292.42	10.90	1784.66	34.97	0.100409
49.51	334.31	292.42	10.90	1784.66	29.11	0.100409
49.51	334.31	292.42	10.90	1485.34	34.97	0.100409
49.51	334.31	292.42	10.90	1485.34	29.11	0.100409
49.51	331.69	293.58	13.10	1784.66	34.97	0.100409
49.51	331.69	293.58	13.10	1784.66	29.11	0.100409
49.51	331.69	293.58	13.10	1485.34	34.97	0.100409
49.51	331.69	293.58	13.10	1485.34	34.97	0.100409
49.51	331.69	293.58	10.90	1784.66	34.97	0.100409
49.51	331.69	293.58	10.90	1784.66	29.11	0.100409
49.51	331.69	293.58	10.90	1485.34	34.97	0.100409
49.51	331.69	293.58	10.90	1485.34	29.11	0.100409
49.51	322.12	282.85	13.10	1784.66	34.97	0.100409
49.51	322.12	282.85	13.10	1784.66	29.11	0.100409
49.51	322.12	282.85	13.10	1485.34	34.97	0.100409
49.51	322.12	282.85	13.10	1485.34	29.11	0.100409
49.51	322.12	282.85	10.90	1784.66	34.97	0.100409
49.51	322.12	282.85	10.90	1784.66	29.11	0.100409
49.51	322.12	282.85	10.90	1485.34	34.97	0.100409
49.51	322.12	282.85	10.90	1485.34	29.11	0.100409
41.21	343.88	303.15	13.10	1784.66	34.97	0.100409
41.21	343.88	303.15	13.10	1784.66	29.11	0.100409
41.21	343.88	303.15	13.10	1485.34	34.97	0.100409
41.21	343.88	303.15	13.10	1485.34	29.11	0.100409
41.21	343.88	303.15	10.90	1784.66	34.97	0.100409
41.21	343.88	303.15	10.90	1784.66	29.11	0.100409
41.21	343.88	303.15	10.90	1485.34	34.97	0.100409
41.21	343.88	303.15	10.90	1485.34	29.11	0.100409
41.21	334.31	292.42	13.10	1784.66	34.97	0.100409
41.21	334.31	292.42	13.10	1784.66	29.11	0.100409
41.21	334.31	292.42	13.10	1485.34	34.97	0.100409
41.21	334.31	292.42	13.10	1485.34	29.11	0.100409
41.21	334.31	292.42	10.90	1784.66	34.97	0.100409
41.21	334.31	292.42	10.90	1784.66	29.11	0.100409
41.21	334.31	292.42	10.90	1485.34	34.97	0.100409
41.21	334.31	292.42	10.90	1485.34	29.11	0.100409
41.21	331.69	293.58	13.10	1784.66	34.97	0.100409
41.21	331.69	293.58	13.10	1784.66	29.11	0.100409
41.21	331.69	293.58	13.10	1485.34	34.97	0.100409
41.21	331.69	293.58	13.10	1485.34	34.97	0.100409
41.21	331.69	293.58	10.90	1784.66	34.97	0.100409

41.21	331.69	293.58	10.90	1784.66	29.11		0.100409
41.21	331.69	293.58	10.90	1485.34	34.97		0.100409
41.21	331.69	293.58	10.90	1485.34	29.11		0.100409
41.21	322.12	282.85	13.10	1784.66	34.97		0.100409
41.21	322.12	282.85	13.10	1784.66	29.11		0.100409
41.21	322.12	282.85	13.10	1485.34	34.97		0.100409
41.21	322.12	282.85	13.10	1485.34	29.11		0.100409
41.21	322.12	282.85	10.90	1784.66	34.97		0.100409
41.21	322.12	282.85	10.90	1784.66	29.11		0.100409
41.21	322.12	282.85	10.90	1485.34	34.97		0.100409
41.21	322.12	282.85	10.90	1485.34	29.11		0.100409