DESIGN FOR QUALITY:

A MODEL-BASED PROBABILISTIC APPROACH

by

David Anthony Swan

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 1997

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

DESIGN FOR QUALITY:

A MODEL-BASED PROBABILISTIC APPROACH

Abstract

Detailed within this thesis is a method for determining and improving the quality of a system. An overall 'Design for Quality' method has previously not existed in the field of model-based design. Statistical experimental design has been used in 'off-line' quality control to determine the optimal settings for a system even when the mathematical model is known. Taguchi demonstrated how signal-to-noise ratios could be used to improve performance of a system through variance minimization. However, these statistical methods often don't use the full distribution information that may be available. Detailed within this thesis is an extension and complement to Taguchi's use of experimental design and signal-to-noise ratios for known system models. The use of a probability transformation method with the mathematical system model will allow designers to perform parameter and tolerance design simultaneously using a method of 'fast integration'. The result is a new method in the field of 'Quality by Design' that can handle both linear and non-linear systems, with components of any distribution type, with or without correlation of the variables, and with single or multiple responses. As an integral part of the method, an interpretation of Taguchi's classification of factors is given in context to this full distribution method. Through the examination of the gradients from the probability transformation method, the design variables can be classified into one of three types: neutral, adjustment, or control. In addition, two extensions to the design method are also detailed within the thesis. The first is the use of the probability transformation method to determine an approximate probability density function for the system responses, and the second is the use of the probability transformation method to perform a 'Worst-Case Analysis' on the system response. The former uses the information given about the system to 'profile' the response, while the latter uses uniform distributions for the system variables and an Interval Analysis type approach is performed to determine the upper and lower worst-case values for the response. Both of these extensions are important parts of the over-all 'Design for Quality' method.

All methods within this thesis require a mathematical model and gradient information of the system response. Graph-Theoretic Models (GTM) were used to model many of the systems since GTM provides many advantages. GTM can develop a system model from component models and connectivity equations, and in addition, easily find the required sensitivity information for design analysis and optimization.

Acknowledgements

The following dissertation was supported by various means for which the author is extremely grateful. These include:

- Numerous teaching assistantships under the supervision of Dr. Barry Wills, Prof. June Lowe, Dr. M. Chandrashekar, Prof. Kish Hahn, and Dr. Peter Roe, arranged by Dr. Mohammed Kamel and Dr. Paul Calamai.
- Research assistantships from Dr. Gordon Savage.
- Scholarships from the Ontario Graduate Scholarship, Faculty of Engineering, Institute for Computer Research, and the University of Waterloo.
- A three-time lecturer position at Wilfrid Laurier University – Physics and Computing Department under Dr. John Lit, Chair.
- A two-term lecturer position at the University of Waterloo – Systems Design Engineering under Dr. Keith Hipel, Chair, and Dr. Glenn Heppler, Undergraduate Chair.
- A three-time faculty position at Shad Valley – Waterloo under Dr. Ed Jernigan.
- A co-directorship and faculty position at Shad Valley – Bark Lake / Trent University through encouragement by Dr. Ed Jernigan and Dr. Jack Pal.
- Operating grants from the Bank of Mom and Dad (open 24 hours – next day service).

In addition to financial support, I would like to thank the following people for aiding me in my journey towards completion of this document:

- My parents – Sharon and Carl: I love you Mom and Dad. Your support has never wavered. This is for you. Now you'll have a 'doctor' in the family!
- My sister, brother-in-law, and niece – Sarah, Don, and Mary for their love and support and gifts of Star Wars toys.
- Dr. Gordon J. Savage for his knowledge and understanding throughout my good and bad times. I have been very lucky to have you as my mentor and guide into the academic world. Your sense of humour and sound advice has given me the privilege of completing both my Master's and Doctorate. Thank you very much, Gord.
- My comprehensive committee: Dr. Paul Calamai, Dr. Barry Wills, and Dr. Michael Hamada.
- My examination committee for their comments and suggestions: Dr. Paul Calamai, Dr. Barry Wills, Dr. Jerry Lawless, and Dr. Joe Pignatiello, Jr.
- Gord's research group: Tricia Cooper, Stephen Carr, and Andrew Row.
- Support Staff – Annette Dietrich, Sue Gooding, Fran Towner, Carol Kendrick, and Heather Hergott.
- Technical Staff – Dave Walsh, Lorus Rossi, Kevin Krauel, and Guenther Metzger
- To my wife – Barbara Kaylie. We have been through thick and thin. Thank you for the support and encouragement over these final few months. It's been a very long haul. I love you very much. Here's to our future.
- And finally to God...who again showed me the way and saved my data.

Dedicated to:

Carl and Sharon Swan – my father and mother
and
Dr. Gordon J. Savage – my supervisor and mentor

Table of Contents

List of Tables

List of Illustrations

Notation

$\varphi(u_i)/f_{V_i}(v_i)$	the density ratio relating the PDF $f(v_i)$ to the standard normal probability distribution
α	a unit vector parallel to the gradient of the vector of the trajectory at the most likely failure point, and is directed toward the failure set
β	the first-order approximation minimum distance from the failure surface to the origin
μ	mean value
σ	standard deviation
δ	a distance from the failure surface to the origin
δ^*	the minimum distance from the failure surface to the origin
∇	a gradient
$\Phi(\,)$	standard normal cumulative distribution function
$\varphi(\,)$	standard normal probability density function
$\nabla_D\, g(V)$	gradient of the margin in V-space with respect to the design variables
$\nabla_U\, g(V)$	gradient of the margin in V-space with respect to the U-space variables
$\nabla_V\, g(V)$	gradient of the margin in V-space with respect to the V-space variables
C	covariance matrix
CDF	cumulative distribution function
$f(V)$	a response function

Failure Surface	divides the state of success and failure, $m = 0$, also known as the limit-state function
FORM	first-order reliability method, also known as the transformation method
FOSM	first-order second moment
$g(U)$	margin in U-space variables
$g(V)$	margin in V-space variables
GTM	graph-theoretic models
HLRF	the Hasofer-Lind-Rackwitz-Fiessler algorithm
Limit-State Function	defined when the margin equals zero, also known as the failure surface
LSL	lower specification limit
m, margin	a function of random variables including design parameters, empirical parameters, uncontrolled random variables, and specifications. Success and failure events are defined by convention as $m > 0$ and $m \leq 0$, respectively.
MLFP	most likely failure point, also written as U^*
PDF	probability density function
Pr(*Failure*)	the probability of failure, $\Pr(m \leq 0)$
Pr(*Success*)	the probability of success, $\Pr(m > 0)$, also known as yield
SNR	signal-to-noise ratio
SL	specification limit
U^*	the most likely failure point in U-space, also known as MLFP
USL	upper specification limit
U-space	the standard normal probability space
V	a vector of random variables
V-space	parameter space, potentially correlated

Chapter 1

Introduction

In the quest for system measures, such as quality and reliability, development of engineering procedures has been rapid. In the literature, this field of research has been referred to as off-line quality control, parameter and tolerance design, sensitivity analysis, and a host of other names. There now exists a great need for the understanding of uncertainty and its effects on quality, reliability, system responses, parameters, and inputs.

Recent literature (Bagchi and Templeton 1994) (Belavendram 1995) (Bounou, Lefebvre, and Do 1992), (D'Errico and Zaino 1988) (Phadke 1986) (Taguchi and Phadke 1984) documents that experimental design has been used in industry to improve quality. The common link in these references lies in the fact that all of these studies had a mathematical relationship/model for the system under study and all used statistical experimental design to improve the response of the system. Most of the studies did not use the full distribution information available, only selected points of the distribution, often $(\mu - \Delta\sigma \quad \mu \quad \mu + \Delta\sigma)$ for experimental design, or its mean and variance for variance reduction methods. Recent advances now allow the use of arbitrary full distribution information in the evaluation and design of systems.

The method that is proposed in this thesis is new to the field of quality and system design. The method is an adaptation from the field of civil engineering that has used 'structural reliability theory' to get a probability of failure. This area of reliability has developed extensively over the past twenty years and has expanded from the simple use of safety factors into the sophisticated methods of probability-based structural design (Madsen, Krenk, and Lind 1986) (Melchers 1987). These sophisticated methods allow full distribution information to be used in the analysis and design of systems.

Previous work in quality design was done using discrete points, either deterministically or probabilistically with an assumption of normality while the method within this thesis is a continuous probabilistic approach. If any of the previous work had used distribution types other than normal, then the current methods to solve them could not be used. However, the method detailed within this thesis, although potentially more computationally expensive for small systems using normal distributions, is capable of handling arbitrary distributions of variables, with or without correlation.

In addition, this thesis will show the commonality between the commonly used 'Signal-to-Noise' ratio in 'off-line' quality control, and the probability-based 'transformation method' that easily computes the system's success/failure probability. Where signal-to-noise ratios aim to achieve the separability of design factors into control

factors and signal factors (Leon, Shoemaker, and Kackar 1987) (Phadke 1989) (Song, Mathur, and Pattipati 1995), the probability-based method does not require separability. In addition, the factor types need not be predetermined since they can mathematically be determined by the method and thus aid in the choosing of appropriate factors for design optimization.

To facilitate the development of the general approach to quality improvement a consistent modelling methodology will be used. Graph-Theoretic Modelling (GTM) has been used in the numerical calculation of system measures for the last thirty years (Koenig and Blackwell 1960). Recently, it has been shown that these measures can be developed and solved symbolically (Savage 1993). This development is the direct result of the creation of advanced symbolic computational software, such as MAPLE™ (Waterloo Maple Software 1996). The ability to solve systems algebraically allows the designer to determine the relationship between the system responses and the input and component values. This relationship is known as Sensitivity Analysis (Chandrashekar and Kesavan 1974) and provides useful quantitative information about the individual effects of parameter uncertainty.

Ultimately, there are two final issues that relate to quality: Analysis and Design. The first deals with a given system where quality levels are to be determined, and the second determines the optimal system given desired quality specifications. These issues can be handled by the methods developed within this thesis and provide a basis for an overall procedure for quality analysis and design.

The remaining sections of this chapter detail what I have accomplished (What), how I did it (How), and where this method fits in within the bigger systems picture (Where). Finally, an overview of the contents of this thesis is given in the last section.

1.1 What?

There are two aspects to this section. The first is the types of systems that I can handle with my research, and the second is the types of problems I can address. The first falls under the heading Models, and the second, Quality Problems.

1.1.1 Models

My methods currently are based on Linear Graph Models (Chandrashekar and Savage 1997) (Chandrashekar, Roe, and Savage 1992) due to the simplicity of modelling. However, as will be shown within this document, it is not necessary to use graph models. If any mathematical model exists for the system, my methods will still work. They are based solely on the mathematical model and not on the modelling mechanism. My methods can handle:

(a) linear systems
(b) non-linear systems
(c) frequency domain
(d) steady-state
(e) time domain with certain restrictions

Time domain systems have restrictions in that we can only determine the quality of a system at a given specified time. We cannot determine the 'global' quality level of a system (e.g., determining how close a response gets to the upper and lower specification limits at any given time). This will require a substantial amount of work and is recommended for further study.

In addition, the branch of Linear Graph Models which will be used is Graph-Theoretic Modelling (GTM), which can model all the above types of systems and easily get sensitivity information required for improving quality.

1.1.2 Quality Problems

There are two aspects to the quality problems issue: Analysis and Design. In the area of analysis, my methods can determine the quality of a system response regardless of whether there are one or two specifications on the response (i.e., either an upper or lower specification limit, or both). This quality level is denoted as the probability of conformance and is a function of the system variables and the system constraints or specifications. For multiple responses, the analysis is carried out on each response to determine its quality level.

For design, my method is formulated to improve the probability of conformance with respect to the system specifications. As in analysis, both one-sided and two-sided specifications on each response can be handled. In addition, both deterministic and probabilistic design variables can be incorporated into my methods, and both parameter (nominal) and tolerance (variance) design can be done simultaneously. This latter point is important since from recent research it has been found that this is a necessity for proper and accurate design (Bisgaard and Ankenman 1996). Also, improvement of quality on multiple responses is handled through an extension of work by Antreich, Graeb, and Wieser (1994).

1.2 How?

Having stated what types of systems can be handled and what types of quality problems can be addressed, this section deals with how it can be done. Again, there are two aspects: the models used and the quality problems addressed. The first deals with the types of probabilistic models that can be used in the determination and improvement of quality, and the second is the method by which we determine quality.

1.2.1 Probability Models

One of the unique aspects of the methods described in this document is the fact that any distribution type can be used as long as the transformation between it and standard normal distribution can be found. Using Rosenblatt's transformation (Rosenblatt 1952), a list of transformations for normal, lognormal, truncated normal, normal with center

missing, triangular, and uniform is given in Appendix A. In addition, the Rosenblatt transformation can handle correlated variables and distributions (see Appendix A).

1.2.2 Quality Problems

Within my method is an algorithm to find the most likely failure point in a system. This algorithm searches the parameter space efficiently to determine which combination of variables is 'most likely' to occur such that the result would be a system response not meeting its specifications. This method formulates the quality problem into a set of margins, where a margin is a function of a specification and its corresponding system response, and then calculates the failure probability using the above algorithm. For a system with one response and an upper and lower specification, there are two margins:

$$\text{Margin 1} = \text{Upper Specification} - \text{System Response}$$
$$\text{Margin 2} = \text{System Response} - \text{Lower Specification}$$

where if a margin is positive, the system response is said to be conforming to the associated specification. If the margin is negative, then the system response has either risen above or fallen below the specification and is no longer meeting specification. When the margin is equal to zero, then the system response lies on the specification boundary and is said to be on the failure surface. The point that has the highest probability of occurring on this surface is the Most Likely Failure Point.

The Hasofer-Lind-Rackwitz-Fiessler algorithm (Madsen, Krenk, and Lind 1986) described in Chapter 3 of this thesis is the algorithm used to find this point. A requirement of the algorithm is gradient information of the margin with respect to the random variables. This information can be obtained from the model's sensitivity information generated by GTM. For systems with only a mathematical model for the system response, sensitivity information needs to be found using other methods (partial differentiation using MAPLE, Matlab, etc.). In summary, the methods for design for quality proposed in this thesis are:

(a) Discipline independent. All types of systems can be handled.

(b) In need of a mathematical model for the response to be studied. The quality design methods are not dependent on GTM, but GTM adds ease of generation of equations.

(c) Independent of the method to find the Most Likely Failure point. The Hasofer-Lind-Rackwitz-Fiessler algorithm is used, but any method could be substituted (Melchers 1987).

(d) Not dependent on a specific optimization routine used to find the optimal design point. The routines need only accept the objective function, gradient of objective function with respect to design variables, constraints, and gradients of the constraints with respect to design variables. These may be non-linear. The methods currently use Matlab, but are not restricted to these. Other optimization routines are currently being reviewed for incorporation into the code.

From all of this, it can be summarized that the following is new to the field of quality or graph-theoretic modelling:

(a) An overall methodology for determining and improving quality. The method is a culmination of work previously done in many fields – Structural (Madsen, Krenk, and Lind 1986), Electrical – (Spence and Soin 1988) (Antreich, Graeb, and Wieser 1994), Reliability and Robust Design – (Carr 1990, 1992), (Carr and Savage 1996), Systems – (Swan, Savage, Cooper, and Carr 1997). However, none of this previous research fully developed a method to handle multiple responses, arbitrary distributions, and parameter and tolerance design all in one methodology.

(b) The methods in this document are discipline independent. Prior to this, most design for quality work was focused on either circuit design (Spence and Soin 1988), or structural engineering, where they referred to quality as performance reliability (Madsen, Krenk, and Lind 1986). Due to this, assumptions were often made that influenced the design methodology. No prior work has been all encompassing to include many types of systems.

(c) Probabilistic/mathematical description of Taguchi's separation of factors into Design-Adjustment, Design-Control, and Neutral variables (Taguchi and Phadke 1984) (Taguchi 1978, 1992, 1993) is achieved through study of the gradient information obtained within the methodology – specifically gradient information within both the parameter space and the transformed standard-normal space.

From this new information, extensions and derivations have been determined and noted. They include:

(a) The mathematical link to Taguchi's Factors is similar to Ben-Haim's "Hyperplane Separation" (Ben-Haim 1996).

(b) The ability to profile the system response using the Hasofer-Lind-Rackwitz-Fiessler algorithm. Instead of determining the probability of failure at only one specification, the specification is moved along the performance spectrum, thereby giving corresponding probabilities associated with each specification (see Chapter 6). Although Wu and Wirsching (1987) have done similar work, the purpose in this case is to determine the extent to which a distribution is not normal.

(c) Worst-Case analysis – the methods in this thesis can be modified to use uniform distributions for all random variables to give a "worst-case" approximation for the system response. Antreich et al. (1994) uses a first-order second-moment method to find "worst-case distances" from specifications. However, his assumption of normality leads to the question of whether they are truly finding "worst-case distances." By using uniform distributions instead of normal distributions, the problem becomes a non-linear interval method. The Hasofer-Lind-Rackwitz-Fiessler algorithm searches along the uniform distribution starting at the center/nominal/mean outward until it gets to the failure surface. Searching along this "uniform" distribution is equivalent to searching along all points of an interval (Dong and Wong 1986). The only assumption made is the

Principle of Indifference so that uniform distributions can approximate the unknown distribution on the interval and give the least biased approximation. Current worst-case methods, like Monte Carlo, are based on uniform distributions (Wojciechowski, Vlach, and Opalski 1997) (Schjaer-Jacobsen and Madsen 1979) and are computationally intensive, whereas the use of uniform distributions in the Hasofer-Lind-Rackwitz-Fiessler algorithm appears to be computationally more efficient (deatiled in Chapter 7).

1.3 Where?

In the bigger context of design, the question becomes, where do my methods fit in? In a recent paper, our research group developed definitions for quality and reliability that are discipline independent (Swan, Savage, Cooper, and Carr 1997). Within this, we have also proposed an overall framework that incorporates both process and product, with quality and reliability as measures to determine how a system is performing. In the following figure (Figure 1-1), this framework is depicted. My contributions fit in this framework as a method to quantify the quality of a system when component parameters are given, and as a means to improve quality when a specified level of quality is desired.

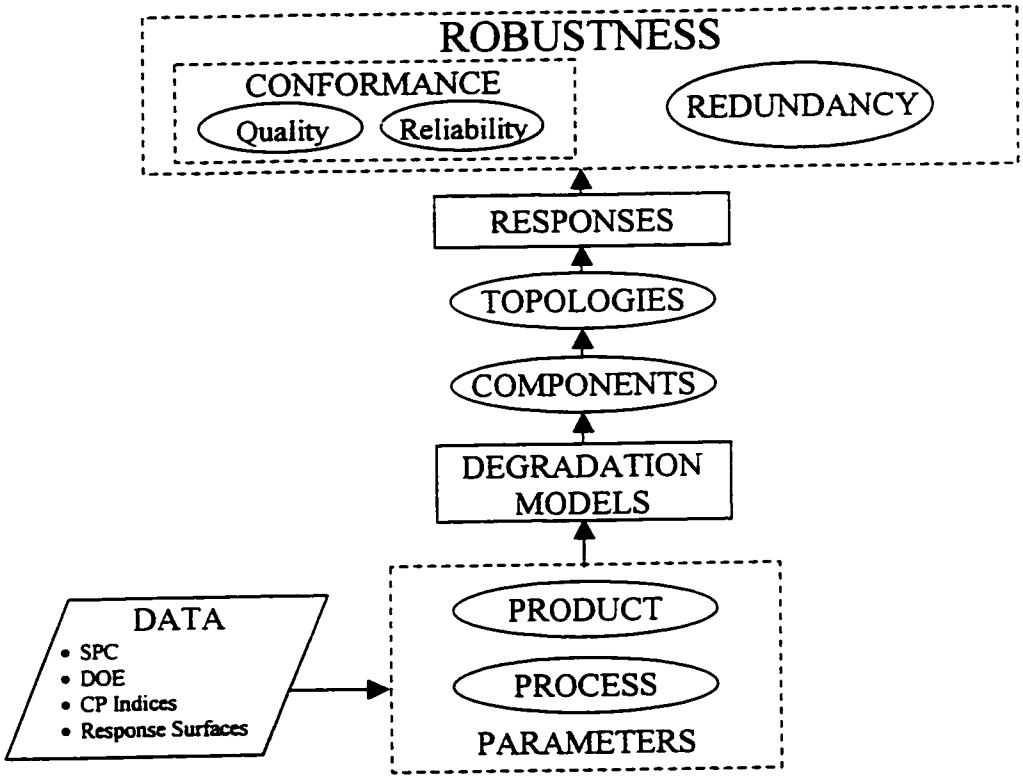


Figure 1-1: Framework for Robust Design

In addition, my methods helps to bridge the gap between current model-based approaches to design for quality using experimental design (Belavendram 1995) (Taguchi and Phadke 1984) or response surfaces (Khattree 1996) (Myers, Khuri, and Vining 1992), which use discrete values for variables in their experimentation, and the probability approach, which uses arbitrary distribution information.

Further work is required and ongoing to complete all aspects of the framework. When completed, it will present the designer with a toolbox to fully design and determine the quality and reliability of a system, given product and manufacturing constraints.

1.4 Overview

Within this thesis are eight chapters. Chapter 1 has been an introduction and overview of my research and methods. Chapter 2 will detail the current methods of quality in the area of circuit design, since most other fields adopted these methods. Chapter 3 details the probability approach, which includes the Hasofer-Lind-Rackwitz-Fiessler algorithm, the use of Rosenblatt's transformation, and two examples to help explain how the algorithm works. The extension into design is shown in Chapter 4 and details the link between Taguchi's Methods and factor determination and the probability approach. Chapter 5 lists the proofs and limitations of the algorithm and method. The next two chapters describe two of the extensions to my research. The first, the ability to profile a response distribution, is in Chapter 6, and the second, a proposed new approach to worst-case analysis, is in Chapter 7. Finally, Chapter 8 concludes and notes further research areas. The appendix includes the Rosenblatt transformations for a variety of distributions and the code for MAPLE and Matlab that I have used in my research.

Chapter 2

Current Approaches to Quality Design

In the field of 'Design for Quality' / 'Quality by Design' / 'Quality Design' in engineering, there have been historically two approaches: Electrical Engineering and Statistically Based Methods. In this chapter, both divisions are documented. Through explanation of each, it will be shown that Electrical Engineering has used aspects of the Statistically Based Methods, such as Monte Carlo Simulation, but in general, the two have remained very separate. Electrical Engineering developed the field of Tolerance Analysis and Design for circuits. The Statistically Based Methods branch into three approaches: Model-based Experimental Design, Response Surface Modelling, and Monte Carlo Simulation.

This chapter will show that the two approaches have recently become strongly inter-dependent. Electrical Engineering now, in its current approaches, uses many of the methods described in the Statistical approaches. This unification has allowed new methods to be used across all engineering disciplines. The result is the method detailed in the remainder of this thesis.

2.1 Electrical Engineering's Approach

The field of Electrical Engineering has been designing for quality using the same methods for the past twenty-five years (Wojciechowski and Vlach 1993) (Spence and Soin 1988). It is only in recent years that new methods, based on and extended from other disciplines, have been introduced to circuit design (Wojciechowski, Vlach, and Opalski 1997) (Antreich, Graeb, and Wieser 1994). In this section, both historical and current approaches will be detailed.

2.1.1 Historical Approach

Since the early 1970's, the field of electrical circuit design has been using a method for quality improvement called Tolerance Analysis, Design, and Assignment. This section will overview both the principles and practices of Tolerance Analysis and Tolerance Design and the methods associated with them. These methods were developed for electrical circuit design but clearly have the potential for broad application. Thus, the word "electrical circuit" was replaced with "product" since a general design methodology for quality is sought. For further information, please see Spence and Soin (1988) for in-depth description of the following methods.

Tolerance Analysis

Tolerance Analysis aids in answering the question "What effect will the component/parameter tolerances have on product performance?" (Spence and Soin 1988) Thus, the objective is to predict product performance given variation in the product's parameters. This prediction will provide an estimate of the fraction of manufactured products that will satisfy the specifications provided by the customer and becomes known as the 'manufacturing yield.' A number of approaches exist in Tolerance Analysis. They can be defined into the hierarchy shown in Figure 2-1 (adapted from Spence and Soin 1988):

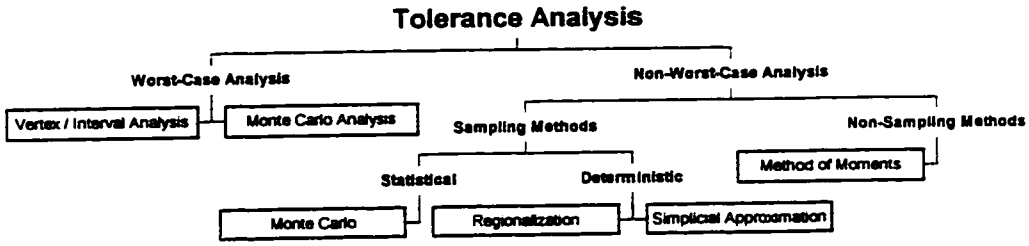


Figure 2-1: Tolerance Analysis Approaches

The first level divides Tolerance Analysis into two approaches: Worst-Case and Non-Worst-Case Approaches.

Worst-Case Analysis

The basis of Worst-Case Analysis is the identification of the extreme values of performance resulting from the variations in parameter values. In finding the extreme values, the objective is to compare the worst-case values (upper and lower) to the product's specifications. If the values exceed the specifications, then modification of the product will be necessary. In other words, the goal is to ensure 100% compliance. In Worst-Case Analysis, no information about the distributions associated with parameters is necessary. The only information required is the extreme values of each parameter.

The primary method used in Worst-Case Analysis is Vertex Analysis, or more correctly called Interval Analysis. Interval Analysis is the application of interval mathematics to problem solving (Hansen 1969) (Moore 1966, 1979) and involves combining all the upper and lower values of each component within the performance function. This gives 2^k combinations for a product with k parameters. For a product with a large number of parameters, the Worst-Case Analysis calculation using Interval Analysis can be extremely computationally intense, and thus Interval Analysis suffers from the curse of dimensionality. As well, this method assumes that the extreme values of the product's performance will occur at one of the 2^k combinations or vertices. This will always occur for linear and bilinear functions. Spence and Soin (1988) discuss alternative approaches to searching all combinations. The addition of sensitivity information allows for intelligent searches of the 2^k combinations.

An exhaustive search using Monte Carlo Analysis can also be performed on the function, where uniform distributions can be substituted for the intervals. This method too suffers from the intense computational effort needed to keep track of the worst-cases

within the tolerance region (See Section 2.2.1 for Monte Carlo Simulation and Chapter 7 for a Worst-Case Approach).

The main difficulty associated with worst-case approaches is that of identifying the component values at the worst-case. There currently does not exist a sure method that will apply to all functions. Another difficulty associated with worst-case approaches is the selection of the extreme values of the parameters. A thoughtful treatment of the process by which to select them is necessary. Nevertheless, useful results can be obtained from Worst-Case Analysis. Chapter 7 details the expansion of the transformation method (see Appendix A) into Worst-Case Analysis with most of these concerns eliminated.

Non-Worst-Case Analysis

Non-Worst-Case methods are applicable to the more general case where the compliance can be less than 100%. In addition, Non-Worst-Case Analysis requires full probability density functions for all parameters of the product and thus branch into two sections: Non-Sampling Methods and Sampling Methods.

Non-Sampling Methods

The only method worthy of consideration in the Non-Sampling Methods branch is the Method of Moments. The Method of Moments is based on a family of mathematical expressions known as the transmission of moments formulae. The formulae are based on Taylor series representations of the performance function.

Consider one quality characteristic and its corresponding performance function, f. The Taylor series representation of the mean and variance would be:

$$\mu_f = E\{f(x)\} = f(\bar{x}) + \tfrac{1}{2}\left\{vec\left(Hess(x)\right)^T vec\left(\mathrm{cov}(x)\right)\right\} \tag{2.1}$$

$$\sigma_x^2 = Var\left(f(x)\right) = \nabla_x f(x)^T \left[\mathrm{cov}(x)\right] \nabla_x f(x) \tag{2.2}$$

where (2.1), the mean, is calculated by a second-order Taylor series expansion and (2.2), the variance, a first-order Taylor series expansion. In the representations, $f(\bar{x})$ is the evaluation of the performance function at the response's nominal parameter values, $\nabla_x f(x)$ the gradient vector (vector of sensitivities of the function to the response parameters), $Hess(x)$ the Hessian of $f(x)$ with respect to the parameters, $\mathrm{cov}(x)$ the covariance matrix of the parameters and *vec* a function which stacks the columns of the matrix into one vector of length *nxm*, where *n* is the number of rows and *m* the number of columns in the respective matrix.

Using the method of moments and the transmission of moments formulae, one can relate the set of moments of the response's parameters to the moments of the response's function. From this point we can proceed in two directions, either by assuming a Gaussian probability density function (PDF) or by making use of methods to handle arbitrary PDF's. Regardless of the approach used, moment analysis is an extremely useful tool and has a great edge over Monte Carlo Analysis in computational effort required (See Section 2.2.1).

Sampling Methods

Sampling methods are based on performance function analysis at sample points in the response's parameter space. The sample points may be selected in a systematic/deterministic manner or in a pseudo-random/statistical fashion. In the deterministic approach, two methods exist: Regionalization and Simplicial Approximation.

(a) Regionalization

The Regionalization Method lays a regular grid over the tolerance region and performs function analysis at one representative point within each sub-region so formed. Often, the center of each region is chosen to represent the product's performance and compared against the specifications. All points within the region are assumed to pass or fail depending upon whether the product, defined by the central point, is a pass or fail. The result will be an approximation of the region of acceptability that falls within the tolerance region. The region of acceptability represents the combinations of parameter values that result in a performance function within specification.

Modifications of this method exist to reduce the number of calculations required and handle arbitrary PDF's. However, this method suffers from dimensionality. For a product with four parameters, the regionalization method using six intervals per parameter would result in $6^4 = 1296$ analyses. Changing the number of parameters to ten increases the number of analyses by 46 655% to $6^{10} = 60\ 466\ 176$. The yield can then be determined as the proportion of sub-regions containing a pass over the total number of sub-regions.

(b) Simplicial Approximation

An alternative deterministic approach in the sampling methods division is the concept of Simplicial Approximation. It involves a piecewise linear approximation to the boundary of the region of acceptability in multidimensional parameter space. The computation entails the determination of sufficient points on the boundary of the region to allow a polyhedral approximation to be developed.

The approach uses an algorithmic search pattern to find the approximation. By starting with a point that satisfies the specifications, one parameter is adjusted at a time. By adjusting one parameter, the extent of movement from the starting point to the failure region can be found. The next parameter is then adjusted and so on. The major computational effort involved in the simplicial approximation approach to tolerance analysis is associated with the function analyses involved in obtaining the approximation to the region of acceptability. Upon obtaining the approximation, generation of sample points from the tolerance region is performed, but no analysis of the product function is done. The points generated are then compared to the approximated region and it is determined if they lie in or out of this region.

The yield is then calculated as the number of sample points within the approximated region divided by the total number of sample points generated. In higher dimensions the sides of the approximation are hyperplanes and the polygon itself is referred to as a simplex, hence the name simplicial approximation. Unfortunately, the problem of dimensionality occurs within this approach too. In Spence and Soin (1988), a simple example with four parameters involved 307 hyperplanes and 455 function analyses to give only a moderately accurate approximation to the region of acceptability. In addition, the method requires the region of acceptability to be convex and simply connected, e.g., no 'black holes' in the region.

In the pseudo-random/statistical approach, there is only one method of significance. It is the Monte Carlo Analysis (See Section 2.2.1).

Tolerance Design

Tolerance Design aids in answering the question "What can be done to reduce the unwanted effect of component/parameter tolerances?" The principal objective in tolerance design is to maximize the manufacturing yield. Tolerance design is a systematic method of identifying the necessary adjustments to parameter nominal values and tolerances to improve product performance and yield. The methods of tolerance design are based upon the approaches to tolerance analysis. There are two main approaches to tolerance design and they are both based on sampling methods. They are Deterministic and Statistical.

Deterministic Approach

Within the field of electrical circuit design, the deterministic approaches are Regionalization and Simplicial Approximation. As discussed in the previous section, neither of these approaches proves worthy for design problems with numerous variables.

Statistical Approach

Of the sampling methods mentioned in the previous section, only the approaches based on statistical/Monte Carlo approaches were worthy of note. In fact for electrical circuit design, until the recent introduction of Taguchi Methods (see Section 2.2.3), the primary tools used in design for quality were based on Monte Carlo approaches.

There exist two methods for tolerance design in electrical circuit design. They are Yield Maximization and Worst-Case Design.

Yield Maximization

The problem common to the two approaches to be described in this section is that of maximization of the manufacturing yield of a mass-produced product. The maximization is to be achieved by adjusting the nominal values of the parameters while the parameter tolerances remain fixed. This assumption of fixed tolerances in widely held in the field of electrical circuit design. Thus, the problem can be formalized as:

$$\begin{aligned} &\text{Maximize } Y\left(f\left(x,x^0\right)\right) \\ &\quad \text{by choice of } \mathrm{X}^0 = x_1^0, x_2^0, \cdots, x_k^0 \\ &\quad \text{for fixed tolerances } \mathrm{t} = \mathrm{t}_1, \mathrm{t}_2, \cdots, \mathrm{t}_k \end{aligned} \tag{2.3}$$

where f is the K-dimensional parameter probability density function, with x^0 representing the nominal values. To accomplish this maximization, there are two principle approaches used in electrical engineering. They are the Simplex Method and Design Centering.

(a) Simplex Method

One straightforward approach to yield maximization is the Simplex Method. As illustrated in Figure 2-2, the method begins with three points in two dimensions and $k + 1$ points in k dimensions. Beginning with A, B and C, we have defined an equilateral triangle in parameter space. Estimating the yield at each point with its appropriate tolerances, we find which point gives us the smallest yield. In the figure below (adapted from Spence and Soin 1988), this point was B.

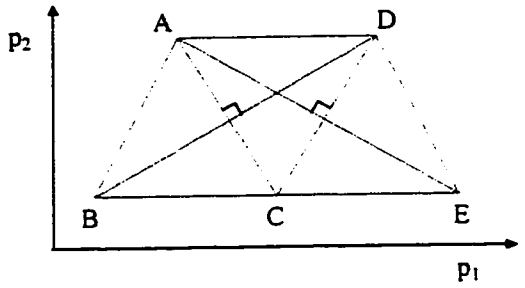


Figure 2-2: Simplex Method of Yield Maximization

In the next step, B is discarded for another exploratory point D, obtained by reflecting B in the opposite face AC of the equilateral triangle. ACD now forms the current triangle and D is evaluated for its yield. Discarding the point with the lowest yield, A, point E is now found. The search stops when the new point is the worst yield point. One way to proceed further is to reduce the size of the exploratory triangle and continue. There are several variants of this strategy. However, the simplex method appears most unattractive for yield maximization since an estimate of yield for each point is required prior to the determining the first movement. For a product with 50 parameters, a yield evaluation must occur at the 51 initial points before the worst yield point can be discarded and a new point found. Additionally, each yield evaluation itself requires analysis. In the case of using Monte Carlo, 100 simulations must be done to calculate just one yield evaluation. So for 50 parameters, the number of random numbers generated per iteration would be:

100 simulations x 50 parameters x 51 starting points = 255000

According to Section 2.2.1, this would provide an accuracy of ±10% for each yield evaluation.

(b) Design Centering

Spence and Soin (1988) state that one successful method of design centering is based upon the fact that a Monte Carlo analysis performs two useful functions:

1. The estimation of manufacturing yield.
2. The provision of spatial information about the region of acceptability by virtues of its pseudo-random sampling of the parameter space.

The first point is explained in the previous section. The second point corresponds to the fact that associated with each of the Monte Carlo sample points, we know whether the resulting performance function evaluation was acceptable (pass) or unacceptable (fail). The Centers of Gravity method takes a top-down view of the information presented by the Monte Carlo analysis (Figure 2-3 – adapted from Spence and Soin 1988). As illustrated in Figure 2-3 (a), we can see the tolerance region associated with the initial product design so positioned with respect to the region of acceptability that further yield increases are possible. Figure 2-3 (b) shows the result of Monte Carlo Analysis. R_A represents the acceptable region and R_T represents the tolerance region. Ideally, one would like to have the R_T contained within R_A. This would represent a system with 100% yield.

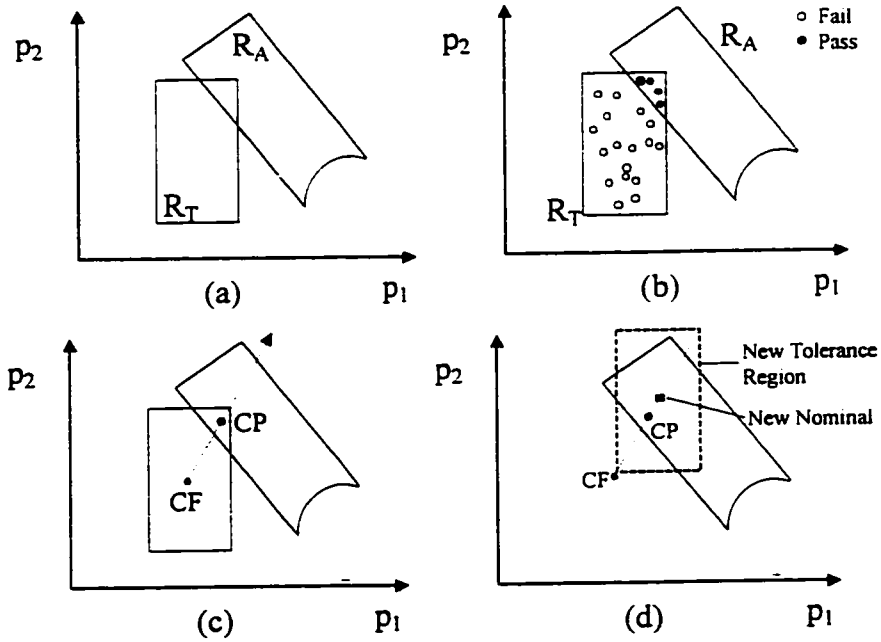


Figure 2-3: Design Centering - Centers of Gravity Method

By determining the center of gravity for each of the pass and fail regions and interpolating a direction vector from the fail center of gravity to the pass center of gravity, Figure 2-3 (c), we can determine the path to take to potentially increase the yield. Upon selecting a new nominal point on this

path equal to or greater than the distance from the fail center to the pass center, Figure 2-3 (d), the procedure begins again. Thus the Centre of Gravity method is iterative and forms the basis of the procedure shown in Figure 2-3.

However, there are a number of considerations associated with this simple method. Two considerations, determination of the step length on the path and the correct ranking of yield estimates, are discussed at length in Spence and Soin (1988).

In addition, there are numbers of variations on the Centre of Gravity approach. They include Correlated Sampling and the Common Points Scheme. Each of these is described in Spence and Soin (1988).

Worst-Case Design

Worst-Case Design is the selection of a design such that all manufactured products meet the specifications with 100% compliance. Within this design, only one approach exists: The Cut Method. The cut method is illustrated in Figure 2-4, adapted from Spence and Soin (1988).

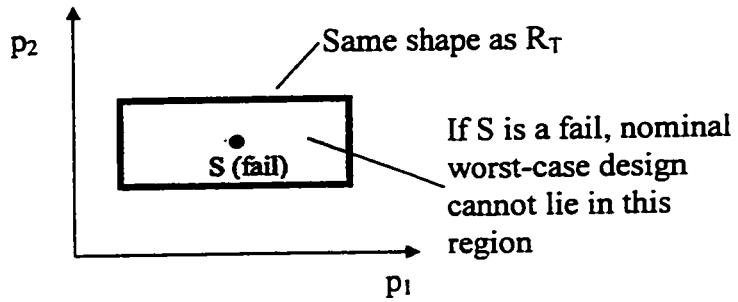


Figure 2-4: The Cut Method

Assume that the tolerances of the parameters are fixed. Therefore the size of R_T is invariant regardless of which design is considered. If, for whatever reason, a point S is found that violates the specifications, the design cannot lie anywhere within the region having the same shape as R_T and centered on the fail point S.

This rectangle is called the 'cut' because that region is removed from consideration as a possible location for the design. That the cut region is exactly the same size as the tolerance region R_T can be appreciated by imagining an experiment in which R_T is moved around S in such a way that S lies just outside R_T. The assumption of fixed tolerances allows for easy and explicit identification of the cut region.

The essence of the cut method is the gradual accumulation of cuts while at the same time gathering information about potentially successful regions in which the design can be located. Thus the cut method is iterative. Each iteration begins with the selection of a new trial design and the initiation of a Monte Carlo analysis. As soon as a fail point occurs in the Monte Carlo analysis, the iteration is halted and the next iteration begins. The outcome of the exploration of parameter space within each iteration is twofold:

1. The single fail sample enables the corresponding cut region to be established. This region is then amalgamated with all previously established cuts to identify the area in parameter space in which a successful nominal design cannot exist.
2. The pass samples generated prior to the single fail point collectively indicate how successful the location of a design nearby might be. The more pass samples encountered, the more likely that a simple shift of the nominal design will result in improved response.

Thus, the primary objective of the exploration is not to estimate the yield associated with the current design but to find the regions where one cannot place the nominal design. Methods to determine the new trial design are described in Spence and Soin (1988).

Tolerance Assignment

The aim of Tolerance Assignment is to minimize the cost of the manufactured products that meet specifications. It is therefore necessary to define a satisfactory cost model. There are many cost models, most of which would be satisfactory for tolerance assignment. Spence and Soin (1988) propose:

$$C(t) = b + \sum_{i=1}^{k} \frac{a_i}{t_i} \tag{2.4}$$

where b is the manufacturing constant that might represent labour costs, factory overheads, etc., a_i the cost of the i^{th} component/parameter – a constant, t_i the tolerance associated with the i^{th} component/parameter – the design variables, and k the number of parameters.

The model implies that the cost of a parameter/component is inversely proportional to its tolerance. In other words, the greater the tolerance, the cheaper the component. However, to recoup the cost of the products that don't meet specification, the unit product cost C for each satisfactory product is:

$$C_{Failures}(t) = \frac{1}{Y}\left[b + \sum_{i=1}^{k} \frac{a_i}{t_i}\right] \tag{2.5}$$

However, it is the quantity $C(t)$ that must be minimized, and in so doing, $C_{Failures}(t)$ will be minimized. Therefore, the optimization can be set up as:

$$\begin{aligned} &\text{minimize: } C = \sum_{i=1}^{k} \frac{a_i}{t_i} \\ &\text{such that: } Y \geq Y_{min} \end{aligned} \tag{2.6}$$

where Y, the yield, is a constrained to be greater than Y_{min}, a minimally acceptable level of yield. Note that b is a constant and can be removed from the problem.

2.1.2 Current Approaches

In the past five years, there have been significant advances in quality design of circuits. Since Spence and Soin's 1988 book, advances have improved the previously mentioned methods. Some advances include:

1. An ellipsoidal method for design centering and yield estimation by Wojciechowski and Vlach (1993),
2. A linearized performance penalty method for yield maximization has been developed by Kirshna and Director (1995),
3. Worst-case analysis and optimization of circuit performances using response surface modelling (see Section 2.2.2) by Dharchoudhury and Kang (1995),
4. The use of nonsymmetrical statistical distributions in circuit design by Wojciechowski, Vlach, and Opalski (1997),
5. Circuit analysis and optimization using worst-case distances by Antreich, Graeb, and Wieser (1994), and
6. Probabilistic design of integrated circuits with correlated input parameters – draft copy, by Seifi, Ponnambalam, and Vlach (1997).

The last three are examined below due to their relevance to this thesis.

Nonsymmetrical Statistical Distributions

This paper by Wojciechowski, Vlach, and Opalski (1997) is an extension to the earlier work mentioned in Wojciechowski and Vlach (1993). In electrical circuit design, it is commonplace to assume Gaussian distributions for circuit parameters. This paper extends circuit design to use more arbitrary distributions such as uniform, binomial, truncated normal, and triangular in design centering. However, their paper does not address correlation of distributions.

Worst-Case Distances and Probabilistic Design

The paper by Antreich, Graeb, and Wieser (1994) proposed a new methodology for integrated circuit design considering tolerances. The method was deterministic, used standard circuit simulators, and did not have the problem of dimensionality. The method detailed in this paper is similar to design methods presented by Hasofer and Lind in 1974. It is a first-order second moment method capable of handling correlation and is based on normal distributions. Skew distributions are approximated with lognormal distributions, which are then converted into normal distributions.

The first-order second moment (FOSM) method, a precursor to the transformation method in Section 3.3, has been well documented in the field of structural reliability

(Hasofer and Lind 1974) (Madsen, Krenk, and Lind 1986) (Melchers 1987). The method needs only the mean and variance of the parameters, and through exploitation of the properties of the normal distribution, accurate estimates of multidimensional integrals can be determined.

Antreich, Graeb, and Wieser establish the connection of the FOSM method to the traditional design methods in the previous section. FOSM provides a forum to accomplish nominal design, worst-case analysis, yield optimization, and design centering in one method.

Seifi, Ponnambalam, and Vlach (unpublished 1997) attempt to extend the work by Antreich et al. (1994) by detailing the FOSM method and showing examples of correlation. They have introduced cost as the objective function while including the yield maximization problem as a constraint. In addition, their method does not transform the problem into a standard-normal space, but rather leaves it in the parameter space. Since this paper only deals with normal distributions, the results are easily abtained in parameter space. Unfortunately, the method they present does not take into account the nonsymmetrical distributions dealt with in Vlach's earlier paper (Wojciechowski, Vlach, and Opalski 1997).

2.2 Statistical Methods – Off-line Quality Control

Currently there are three distinct statistical approaches to quality design. Monte Carlo Simulation is probably the most well-known and used method. The second is Response Surfaces and is an extension of regression analysis, and the third is Model-based Experimental Design, often referred to as Parameter Design. The second two methods encompass the majority of research currently devoted to improving quality (Khattree 1996) (Knottmaier 1993) (Myers, Khuri, and Vining 1992) (Logothetis and Wynn 1989). Each of the three statistically based methods is briefly described below.

2.2.1 Monte Carlo Simulation

The Monte Carlo Simulation approach (Logothetis and Wynn 1989) (Spence and Soin 1988) is the generation of sample points in parameter space in a pseudo-random manner to simulate the actual process. However, unlike the simplicial approximation from the Electrical Engineering approach, no attempt is made to obtain an approximation of the region of acceptability.

The Monte Carlo method directly mimics the process of random parameter value selection (including the correlation encountered between parameters) by generating values according to the known parameter probability density functions. The N sample points generated are then simulated and checked against the specifications. Thus, Monte Carlo Analysis is similar to measurements made on N actual responses. The yield for the N samples can be calculated as the fraction of acceptable performance samples to N, the total number of samples.

According to Spence and Soin (1988), Monte Carlo Analysis differs from the other methods in a number of distinctive ways:

1. It is conceptually simple and relatively easily programmed.
2. Although the Monte Carlo estimate of yield is known to be liable to error. the extent of this error can be estimated with known confidence. One can contrast this situation with the simplicial approximation approach for which it is very difficult to obtain an estimate of the error.
3. A mathematical analysis of the Monte Carlo method shows that the relationship between the accuracy of yield estimation and the number of samples N is independent of the number of parameters subject to random variation. Thus if the yield is, say, 60%, and one wishes to be 95% confident of the actual yield lying within ±10% of the estimate, then 100 Monte Carlo simulations are required whether the function contains two or as many as 2000 parameters. It is a property that allows Monte Carlo analysis to be employed for all levels of evaluation.
4. For a given confidence level in the estimate of *yield*, the confidence interval is proportional to the inverse of the square root of the number of samples N. Thus, if 100 Monte Carlo samples led to a confidence interval of ±10% about the estimate, then four times the number of samples, i.e., 400, would be necessary to halve the confidence interval, i.e., to ±5%. # of Simulations = 40000/(Confidence Width)^2

Confidence	#Simulations
±10.00%	100
±5.00%	400
±1.00%	10000
±0.50%	40000
±0.10%	1000000

Monte Carlo Analysis offers an attractive method for analysis and may result in useful clues as to how parameters can be changed to improve yield in the methods of design. In addition, arbitrary statistical distributions can be used in the evaluation of the functions. However, as the number of parameters increases, so does the computational expense.

The Monte Carlo methods used to estimate yield may also be used to estimate the gradients of yield with respect to parameter nominals and tolerances. This additional estimation does not require any further simulations and hence only incurs minimal cost. Two approaches exist in using yield gradients. The first is a general approach where at each iteration and for a particular set of nominals and tolerances, the yield and its gradients are estimated via Monte Carlo analysis and are subsequently used to choose a new trial design. When compared to the Centre of Gravity Method, it simply entails replacing the algorithm for choosing new nominals with a more sophisticated method employing yield gradients.

The simplest gradient-based optimization procedure is that of steepest ascent. However, the Monte Carlo analysis can also provide useful information on the second-

order yield gradients and thus a more effective gradient-based method (e.g., Fletcher and Powell, 1963) may be used.

The second approach starts with the estimation of the yield and its gradients for an initial starting point. These estimates are then substituted into an approximating analytic relationship referred to as a yield prediction formula (2.7), which relates yield to nominal points in the vicinity of the trial design,

$$Y(P^0) = Y(P_1^0) + \sum_{i=1}^{k} \Delta P_i^0 \frac{\partial Y}{\partial P_i^0} + \sum_{i=1}^{k}\sum_{j=1}^{k} \Delta P_i^0 \frac{\partial^2 Y}{\partial P_i^0 \partial P_j^0} \Delta P_j^0 \tag{2.7}$$

where $P_1^0 = p_{11}^0, p_{12}^0, p_{13}^0, \cdots p_{1k}^0$ is the nominal design at which the yield has been estimated by Monte Carlo analysis, and P^0 is a general point in the vicinity. The change in nominal points is denoted by $\Delta P^0 = P^0 - P_1^0$. The yield prediction formula predicts the yield that will be obtained if the nominal point is moved from P_1^0 to another point P_2^0.

If yield maximization is the objective, then P_2^0 could be found such that the yield predicted is maximized. Then a new Monte Carlo analysis would be performed with P_2^0 as the nominal, the yield prediction formula updated and the procedure continue until the design specifications are met.

2.2.2 Response Surfaces

Response Surface Methods (Khattree 1996) (Knottmaier 1993) (Myers, Khuri, and Vining 1992) are a combination of experiment design and regression analysis methods, which aim at the following:

- Determining local optima for quality characteristics.
- Showing how quality characteristics change when parameter levels are changed.

The quantities to be optimized can be, on the one hand, function results of the product/process; on the other hand, the experiment design can be superimposed by means of an error parameter matrix. In the latter case the level to be optimized is a quality measure. Regardless of the quantities to be optimized, there are two applications for response surface methodology; if the performance function of the quality characteristic is known or not known. It is the former case that will be studied.

With the performance function known, response surface methods seek to approximate the quality measure:

$$Y = Y(x) = f(x, \Theta) + \varepsilon \tag{2.8}$$

where $\Theta = (\theta_1, \theta_2, \cdots, \theta_n)$ are the coefficients of the average response to be estimated and ε is the error, thereby giving:

$$\hat{Y} = \hat{Y}(x) = f\left(x, \hat{\Theta}\right) \tag{2.9}$$

where $\hat{Y}$ represents the approximated response surface and $\hat{\Theta}$ represents the estimated coefficients. Very often two mathematical models are used for the layout of the response surface, depending on the behaviour of the factors to be optimized.

$$\hat{Y} = \hat{\theta}_0 + \hat{\theta}_1 x_1 + \hat{\theta}_2 x_2 + \cdots + \hat{\theta}_n x_n \tag{2.10}$$

$$\hat{Y} = \hat{\theta}_0 + \sum_{i=1}^{n} \hat{\theta}_i x_i + \sum_{i=1}^{n} \hat{\theta}_{ii} x_i^2 + \sum_{i<j} \sum \hat{\theta}_{ij} x_i x_j \tag{2.11}$$

Equation (2.10) is the model used if the response is linear and with no interactions, and equation (2.11) is the general quadratic model used for most other cases. It is important to note that the response surface only applies to a relatively small area, within which an approximation of second order does not obscure the reality. The first step for a response surface is to make a linear approximation and, if the linear model does not meet the requirements, to use a quadratic model, whereby the concerned sector can be shifted towards the optimum. Higher order models can be obtained if the quadratic is unacceptable. In these general settings, the response surface methodology does more than estimate and test the effects of parameter changes on the response, it also attempts to:

1. Fit $\hat{Y}$ to Y accurately.
2. Find the maximum or minimum of $\hat{Y}$.
3. Find the direction of maximum increase in $\hat{Y}$.
4. Find the x that keeps $\hat{Y}$ on target.
5. Find the general shape of Y.
6. Eliminate variables x_i which do not affect Y.

In addition to these general aims, response surface methods reflect more than the elementary factorial methods that they incorporate. They incorporate the desire to proceed sequentially. Several methods exist on how to move the response surface design. Knottmaier (1993) details methods using Steepest Gradient, Gauss-Seidel, and combinations of the two.

The response surface methodology proceeds as follows. Suppose we take observations $Y_1, Y_2, \cdots, Y_n$ at the points $(x_{11}, x_{21}, \cdots, x_{k1}), \ldots, (x_{1n}, x_{2n}, \cdots, x_{kn})$ respectively, where n is the number of points and k is the number of parameters. Then Least Squares analysis would seek to minimize:

$$\sum_{i=1}^{n} \left\{ Y_i - f(x_{1i}, x_{2i}, \cdots, x_{ki}), \theta \right\}^2 \tag{2.12}$$

For the quadratic case, the solution can be expressed in matrix notation as:

$$\hat{\Theta} = \left(X^T X\right)^{-1} X^T Y \tag{2.13}$$

where

$$\begin{aligned} \hat{\Theta} &= \left(\theta_1, \theta_2, \cdots, \theta_n\right)^T \\ Y &= \left(Y_1, Y_2, \cdots, Y_n\right)^T \end{aligned} \tag{2.14}$$

and

$$X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} & x_{11}^2 & x_{21}^2 & \cdots & x_{k1}^2 & x_{11}x_{21} & x_{11}x_{31} & \cdots & x_{(k-1)1}x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} & x_{12}^2 & x_{22}^2 & \cdots & x_{k2}^2 & x_{12}x_{22} & x_{12}x_{32} & \cdots & x_{(k-1)2}x_{k2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} & x_{1n}^2 & x_{2n}^2 & \cdots & x_{kn}^2 & x_{1n}x_{2n} & x_{1n}x_{3n} & \cdots & x_{(k-1)n}x_{kn} \end{bmatrix} \tag{2.15}$$

Thus, the response surface is estimated by:

$$\hat{Y} = \hat{\Theta} X \tag{2.16}$$

In addition, the covariance matrix of the parameter estimates is:

$$\operatorname{cov}\left(\hat{\Theta}\right) = \sigma^2 \left(X^T X\right)^{-1} \tag{2.17}$$

where

$$\sigma^2 \equiv s^2 = \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{n-k} \tag{2.18}$$

Upon estimation of the response surface, optimization can be done to determine the optimal point within the region. Of special importance in response surface methods is the variance of the estimated response $\operatorname{var}\left(\hat{Y}\right)$, which tells us how well we are predicting or interpolating $E(Y)$ over the whole region R. It is:

$$\operatorname{var}\left(\hat{Y}(x)\right) = \sigma^2 f(x)^T \left(X^T X\right)^{-1} f(x) \tag{2.19}$$

where $f(x)$ is the vector of functions in the model. For the case of k parameters:

$$f(x) = \left(1, x_{11}, x_{21}, \cdots, x_{k1}, x_{11}^2, x_{21}^2, \cdots, x_{k1}^2, x_{11}x_{21}, x_{11}x_{31}, \cdots, x_{(k-1)1}x_{k1}\right) \tag{2.20}$$

Thus, the response surface methodology is a useful approach for determining the relationship between the performance function and the quality measure. In this

application, the assumption is made that the parameter tolerance remains constant. This is necessary for calculation of the quality measure.

Special response surface designs exist to reduce the 2^k or 3^k factorial designs. These consist of fractional factorial designs with additional symmetrically placed points. The most common such additions are center points and star points. Logothetis and Wynn (1989) describe response surface methods using these fractional designs.

2.2.3 Model-Based Experimental Design (Parameter Design)

The model-based experimental design approach is generally attributed to Taguchi (1978). Experimental Design (Belavendram 1995) (Dehnad 1989) is a method of experimentation with a process in a controlled manner rather than random interference as may happen in uncontrolled procedures. It is a procedure used to determine the best/optimal levels at which to operate a process. The extension to model-based became a method by which computer experimentation could be performed on a design prior to implementation.

The basic problem of experimental design is deciding what pattern of design points will best reveal aspects of the problem of interest. Traditionally, the extremes of the process parameters were taken as starting points for the design. Given that the function of the inputs/parameters to the output/response is known, the experimental design systematically combines all possible combinations of the parameters to determine the respective output. Analyzing the output from the design, the parameters that have a greater effect of the response can be determined (Welch et al. 1992).

Using model-based experimental design allows the designer/engineer to find the best or optimal settings that put the required output on target with low variation and cost. This is accomplished by what Taguchi calls *Parameter Design*. Parameter Design exploits the nonlinear characteristics of the parameters and interactions among parameters and the noise or environmental factors to make the system objective function more robust. By robust we mean insensitive to change, where change can be from controllable factors such as variation of parameters or from uncontrollable factors such as noise.

Taguchi's Parameter Design is carried out through a sequence of steps. Taguchi (1992) suggests the following methodology:

1. Define the Problem
    - Provide a clear statement of the problem to be solved. It is important to establish just what the experiment is intended to achieve.
2. Determine the Objective
    - Identify the quality characteristics (responses) to be studied and eventually optimized. Determine the method of measurement.

3. Set Up a Brainstorming Session

   - Determine the controllable (inputs, excitations and parameters) and uncontrollable (noise and environmental) factors, the experimental range and the appropriate factor levels.

4. Design the Experiment

   - Select appropriate designs and construct the inner array (control factor array) and the outer array (noise factor array) (Coleman and Montgomery 1993).

5. Conduct the Experiment and Collect the Data

   - Use the function of the quality characteristic to determine the responses (Belavendram 1995) (Bagchi and Templeton 1994) (Welch and Sacks 1991) (D'Errico and Zaino 1988) (Taguchi and Phadke 1984).

6. Analyze the Data

   - There are many methods for analyzing experimental data. The Analysis of Variance (ANOVA) and Signal-to-Noise Analysis of Variance (S/N ANOVA) are the methods applied by Taguchi. Determination of the performance measures; Taguchi's use of S/N Ratios.

7. Interpret the Results

   - Upon analysis of the S/N data and the ANOVA, the control factors may be put into four classes:

     Class I: Factors that affect both average and variation (Control)

     Class II: Factors that affect only variation (unnamed)

     Class III: Factors that affect only average (Adjustment)

     Class IV: Factors that affect neither average nor variation (Neutral)

8. Selection of Optimal Levels

   - Proper levels of Class I and II control factors are determined to reduce variation. Proper levels of Class III factors are chosen to bring the average to target. Control factors of Class IV can be set at the most cost-efficient level.

9. Run a Confirmatory Experiment

   - Upon selection of the optimal levels, a confirmatory experiment is conducted to ensure and demonstrate that the new levels chosen do provide the desired results.

If the predicted results from step 8 are not confirmed, or if the results are otherwise unsatisfactory, additional experiments may be required and a reiteration of steps 3 to 9 might be necessary.

Chapter 3

The Probability Approach

One of the most common problems in engineering is determining the probability that a response is above/below some specified value. Given that the response has some arbitrary probability density function, the problem can be graphically depicted as (Figure 3-1):

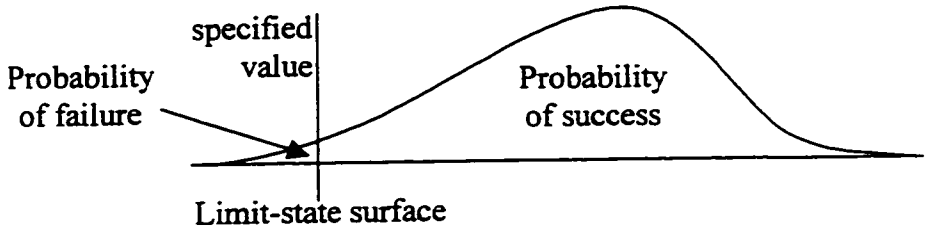


Figure 3-1: One-sided Probability Problem

where in order to find the probability of success or failure, we must integrate over the region of interest. For multidimensional problems this now involves a multidimensional integral. This problem type in the past has been solved through direct analytical methods (Kalbfleisch 1979) (Larson 1969) based on methods such as the Chebychev inequality or Monte Carlo (Spence and Soin 1988).

3.1 Margin and Limit-State Function

The margin, m, is a function of V, which is a vector of random variables that may include design parameters, empirical parameters, uncontrolled random variables, and specifications, and can be shown as:

$$m = g(\mathrm{V}) \tag{3.1}$$

The line that divides the state of success and failure is called the limit-state function or failure surface. Mathematically, the limit-state function is determined when the margin, m, is equal to zero.

$$m = g(\mathrm{V}) = 0 \tag{3.2}$$

Success and failure events are then defined by convention as $m > 0$ and $m \leq 0$, respectively and the value of the margin indicates the extent to which a particular limit-state has been successfully avoided or exceeded. For example, a system response, $f(\mathrm{V})$,

has a specified value of SL that it should be less than. Note that SL can be either probabilistic or deterministic. In most cases, as in the examples in this thesis, the specification is simply a deterministic value. In this case, the margin is:

$$m = SL - f(\mathrm{V})$$

and the limit-state function or failure surface is defined as:

$$m = SL - f(\mathrm{V}) = 0 \rightarrow f(\mathrm{V}) = SL$$

So, the margin is positive for $f(\mathrm{V}) < SL$ (the success region) and negative for $f(\mathrm{V}) > SL$ (the failure region). The associated probability of success is defined as the multidimensional integral of the joint probability density function (PDF) of the random variables over the success region of the probability space.

$$\Pr(m > 0) = \Pr(g(\mathrm{V}) > 0) \tag{3.3}$$

However, integration over arbitrary regions of multivariate distribution functions is generally not practical, and thus numerous methods have been found to approximate the integration over the failure region.

3.2 One-Sided and Two-Sided Probability Problems

If the response can be approximated by a normal distribution, such as in Figure 3-2.

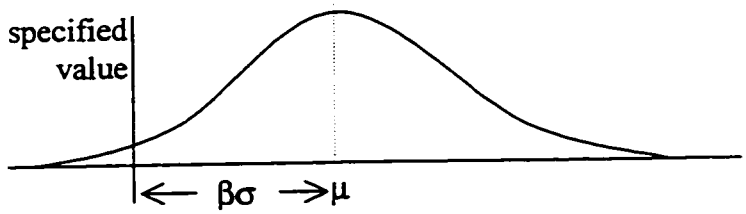


Figure 3-2: One-sided Approximate-Normal Probability Problem

then β represents the number of standard deviations the mean is away from the specified value. If the specified value is zero, then this formulation is known as Cornell's (1969) index, β_C, or more commonly, the inverse of the coefficient of variation:

$$\beta_C = \frac{\mu_R}{\sigma_R} \tag{3.4}$$

and only requires the mean and variance of the response R. Using the standard normal tables, the probability of μ_R having a value greater than zero is represented in equation (3.5) and falling below zero, equation (3.6).

$$\Pr(\mu_R > 0) = \Pr(Success) = \Phi(\beta_C) \tag{3.5}$$

$$\Pr(\mu_R \leq 0) = 1 - \Pr(Success) = \Pr(Failure) = 1 - \Phi(\beta_C) = \Phi(-\beta_C) \tag{3.6}$$

However, in practice, the specified value or specification will not be zero. In this case, the index becomes:

$$\beta = \begin{cases} \dfrac{\mu_R - spec}{\sigma_R}, & \text{for } \mu_R \geq spec \\ \dfrac{spec - \mu_R}{\sigma_R}, & \text{for } \mu_R < spec \end{cases} \tag{3.7}$$

where the first case β is often referred to as β_{LSL}, the number of standard deviations μ_R is above a lower specification limit (*spec*), and the second β_{USL}, the number of standard deviations μ_R is below an upper specification limit. If β_{LSL} and β_{USL} are applied to the same response, a two-sided probability problem is created and is graphically shown in Figure 3-3. If $\mu_R = spec$, then $\beta = 0$ since the distance between the mean and the failure surface is 0.

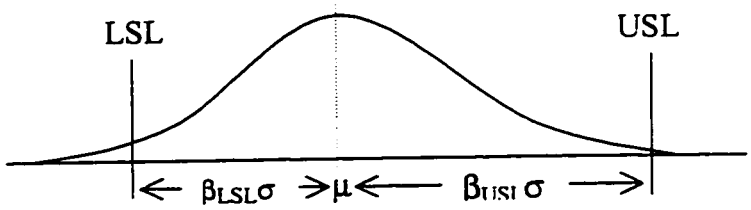


Figure 3-3: Two-sided Approximate-Normal Probability Problem

Two probabilities are associated with the above problem. Mathematically they are:

$$\Pr(Success) = \Phi(\beta_{LSL}) + \Phi(\beta_{USL}) - 1 \tag{3.8}$$

$$\Pr(Failure) = \Phi(-\beta_{LSL}) + \Phi(-\beta_{USL}) \tag{3.9}$$

Thus, for any normally distributed response, we can estimate the probability of success or failure. However, many one-sided and two-sided problems do not have normal distributions. If the response distribution is highly skewed, as in Figure 3-4,

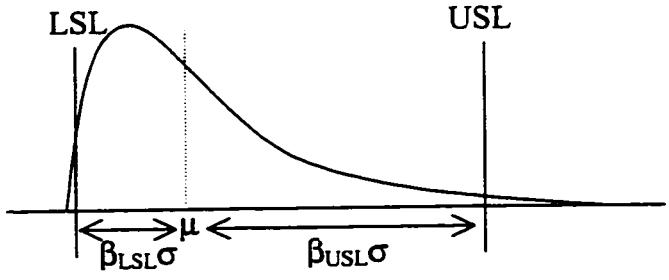


Figure 3-4: Incorrect Calculation of Non-Normal Probability Problem

then applying equation (3.8) or (3.9) would result in an extremely inaccurate calculation. To eliminate this problem, another method for calculating β is needed. This is discussed and detailed in the next section.

3.3 The Transformation Method

One of the best solutions to the above-mentioned problem comes from structural reliability theory. Hasofer and Lind (1974) developed a procedure for approximating multidimensional integrals by transforming the problem into the standard normal probability space ($\mu = 0$, $\sigma = 1$) and exploiting the geometrical properties of this space. Figure 3-5 (adapted from Carr 1992) shows a transformation from V-space, the design variable space, to U-space, the uncorrelated standard normal representation of the problem. The method is based on normal-tail approximation by Ditlevsen (1981) and the Rosenblatt (1952) transformation. Through this transformation to standard normal probability space, the previous approximations for Pr(*Success*) and Pr(*Failure*) can still be used since they are based on the assumption of normality. In Appendix A, the transformation of various distributions in V-Space to the uncorrelated standard normal probability space, U-space, are discussed in detail.

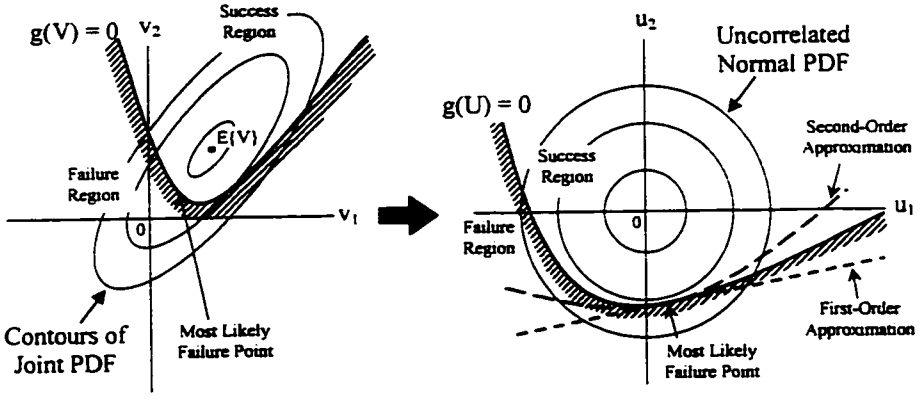


Figure 3-5: V to U Space Transformation

Exploiting the Standard Normal Space Properties

Within the standard normal probability space, the maximum probability density is located at the origin and is rotationally symmetrical about this point. The point on the transformed limit-state surface that is closest to the origin is called the most likely failure point (MLFP). This point has the highest probability density of all those in the failure region (Carr 1992) (Melchers 1987) (Madsen, Krenk, and Lind 1986). Mathematically, in the U-space, the probability density function decreases very quickly, as $\exp(-r^2/2)$, where r is the distance from the failure surface to the origin. The transformation method finds the most-likely failure point, U^*, and determines the shortest distance from it to the origin and is denoted as δ^*. The failure probability is calculated by approximating the failure surface at U^* by a hyperplane and exploiting the properties of the U-space. The probability of failure at the MLFP can then be approximated by $\Phi(-\delta^*)$, a linear Taylor series expansion about the failure point (Carr and Savage 1996).

So using the properties of the standard normal space, the approximate failure probability is given by:

$$\Pr(m_i \le 0) = 1 - \Pr(m_i > 0) = \Phi(-\delta^*) \tag{3.10}$$

The first-order approximation is fairly accurate in the region of the checking point as long as the principal curvatures of the limit-state surface are not too large. A second-order method is also available which uses a quadratic approximation of the limit-state surface. It is more accurate, but computationally more complex. For this thesis, only the first-order approximation will be used. Thus, the index is given as:

$$\beta = \delta^* \tag{3.11}$$

where β now represents the minimum distance to the hyperplane.

Efficient methods for determining the minimum distance, β, require the gradient of the limit-state function (Rackwitz and Fiessler 1978). An algorithm to find δ^* was developed by Hasofer, Lind, Rackwitz, and Fiessler (HLRF) and is detailed in Madsen, Krenk, and Lind (1986) and Melchers (1987). The following adaptation of the HLRF algorithm was detailed in Carr and Savage (1996), and has been expanded upon for readers of this thesis. It is:

1. Convert U_k to V_k using the inverse probability transformation.

   V_k is the vector of basic random variables (design and noise) that in general, may not be normally distributed. U_k is the associated vector of uncorrelated and standardized normally distributed variables. V_k is transformed into U_k by a probability transformation. Typically the starting point for the algorithm is U = 0.

2. Calculate $g(V_k)$ and $\nabla_U g(V_k)$,

   where $g(V_k)$ is the value of the margin at the k^{th} iteration, and $\nabla_U g(V_k)$ is the gradient of the margin with respect to the random U-space variables at the k^{th} iteration.

3. Calculate U_{k+1} as:

$$U_{k+1} = \frac{U_k^T \nabla_U g(V_k) - g(V_k)}{\nabla_U g(V_k)^T \nabla_U g(V_k)} \nabla_U g(V_k) \tag{3.12}$$

   In Chapter 5, the problems associated with this method are detailed. However, it should be noted here that if the limit-state surface has more than one stationary point, this algorithm might find different values for δ^* depending on the starting point, U_0. In addition, note that if $U_k^T \nabla_U g(V_k) = g(V_k)$, the algorithm terminates with U = 0. If this case occurs, a different starting point should be tried.

4. Calculate δ_{k+1}, the distance to the origin in U-space, as the Euclidean norm of U_{k+1}.
5. Stop when $|\delta_{k+1} - \delta_k| \leq \varepsilon$.

Typically, the Hasofer-Lind-Rackwitz-Fiessler algorithm converges in less than 12 iterations (Madsen, Krenk, and Lind 1986) to U^*, the MLFP to find:

$$\beta = \delta^* = \|U^*\| = \sqrt{\left(U^{*T} \cdot U^*\right)} \tag{3.13}$$

Madsen, Krenk, and Lind (1986) do note that the failure surface may contain multiple stationary points, and therefore it is necessary to use several starting points to find the corresponding δ^*'s. If there is only one stationary point, convergence is guaranteed and the global minimum will be found; otherwise the algorithm may converge toward another stationary point. Thus, the general convergence in the algorithm is not guaranteed and the algorithm should be modified to terminate after a pre-determined number of iterations. If this occurs, a trace of the δ values should help to determine if convergence is occurring.

Within the HLRF algorithm, the gradient of the margin with respect to the random variables, $\nabla_U g(V_k)$, is required and this can be calculated as follows:

$$\nabla_U g(V_k) = \begin{bmatrix} \dfrac{\partial g(V_k)}{\partial u_1} \\ \vdots \\ \dfrac{\partial g(V_k)}{\partial u_n} \end{bmatrix}, \text{ where } \frac{\partial g(V_k)}{\partial u_i} = \frac{\partial g(V_k)}{\partial v_i}\frac{\partial v_i}{\partial u_i} = \frac{\partial g(V_k)}{\partial v_i}\frac{\varphi(u_i)}{f_{V_i}(v_i)} \tag{3.14}$$

where $\varphi(\,)$ is the standard normal probability density function, and $f_{V_i}(v_i)$ is the marginal probability distribution of the variable v_i. Note that $\partial v_i / \partial u_i$ is the derivative of the basic variable with respect to its corresponding uncorrelated and standardized normal variable. This derivative, also stated as $\varphi(u_i)/f_{V_i}(v_i)$ is commonly referred to as a density ratio (the standard normal distribution to the marginal distribution of the variable) and is examined in detail for various distributions in Appendix A. For example, if the marginal distribution is a normal distribution with mean, μ, and standard deviation, σ, then the ratio is σ – the standard deviation of the marginal.

There are two reasons why the problem should be converted into U-space. The first reason is the ability to exploit the properties of the standard normal space and the second is that the transformation allows for use of arbitrary distributions.

Using U-space versus V-space

A number of authors (Melchers 1987) (Antreich et al. 1994) (Seifi et al. unpublished 1997) do not transform the problem into U-space, rather they all solve for

the most likely failure point in V-space. These practitioners determine the most likely failure point in V-space by minimizing:

$$\beta = \sqrt{\left(\left(v_i - \mu_{v_i}\right)^T \cdot C^{-1} \cdot \left(v_i - \mu_{v_i}\right)\right)} \tag{3.15}$$

where v_i is the value of the i^{th} variable at the most likely failure point, μ_{v_i} is the mean of the i^{th} variable, and C is the covariance matrix. Note that if all the distributions are standard normal, and the covariance matrix is the identity matrix (no correlation), then this is the same as equation (3.13). They have simply eliminated the transformation of each variable and consolidated it into one equation. This is from first-order second moment theory (Hasofer and Lind 1974). This approach can handle normal PDF's, lognormal PDF's which are transformed into normal PDF's, and skew distributions, which are approximated by lognormal (then transformed).

There is only one reason to use equation (3.15) instead of transforming the problem into U-space and that is if all the probabilistic variables in the mathematical model are normally distributed AND correlated (i.e., the covariance matrix, C, is dense). In this case, the Rosenblatt transformation to convert the variables to and from U-space is computationally intensive due to the correlation. This special case allows the problem to be solved in V-space and avoids the Rosenblatt transformation, by modifying the minimum distance equation to calculate the deviation from the mean values with respect to the variances for normally distributed variables. In effect, the problem has been transformed prior to solution, while U-space transforms during the solution. If, however, there are any distributions in the model that are not normally distributed or that cannot be accurately approximated by one, then the problem must be solved in U-space. For example, the Rosenblatt transformation into U-space for a normally distributed variable is:

$$u_i = \left(v_i - \mu_{v_i}\right) / \sigma_{v_i}$$

If we substitute this into equation (3.13), we get equation (3.15). However, substituting the Rosenblatt transformation for a uniform or any other non-normal distribution into equation (3.13) does not yield equation (3.15). For this reason, conversion into U-space must be done during the solution as performed in the transformation method and not solved in V-space as Melchers (1987), Antreich et al. (1994), and Seifi et al. (unpublished 1997) suggest.

For arbitrary PDF's, the Rosenblatt transformation transforms the variable's PDF from V-space into U-space, resulting in a one-to-one mapping of the arbitrary PDF to the standard normal distribution (detailed for numerous distributions in Appendix A). Transforming all the variables into U-space allows the method to solve the problem with the same ease as the special case mentioned above. If there is no correlation among the variables, then the problem is even simpler and there is no difficulty in transforming the problem into U-space since each variable is independent and can be transformed separately.

In conclusion, we can take advantage of both methods within the 'Design for Quality' methodology since determination of the minimum distance is not restricted to either method.

3.4 Example of Finding δ^*

Let's define a function, $f = d^2$, which represents the response of a system. We wish to determine the probability that f is greater than 1 given that the variable d is normally distributed with μ_d = 2 and σ_d = 1. Therefore, the margin becomes $g(\mathrm{V}) = d^2 - 1$. The transformation from U-space to V-space for any normally distributed uncorrelated variable is (See Appendix A):

$$v = \mu + \sigma u \tag{3.16}$$

where in this example, v represents d, such that $d = \mu_d + \sigma_d u$ or $d = 2 + u$. In addition, the gradient of the margin with respect to the random U-space variable, u, is needed. It is defined as:

$$\nabla_{\mathrm{U}} g(\mathrm{V}) = \frac{\partial g(\mathrm{V})}{\partial u} = \frac{\partial g(\mathrm{V})}{\partial v}\frac{\partial v}{\partial u} = \frac{\partial g(\mathrm{V})}{\partial v}\frac{\varphi(u)}{f_{\mathrm{V}}(v)} \tag{3.17}$$

and for our example becomes:

$$\nabla_{\mathrm{U}} g(\mathrm{V}) = 2d\frac{\varphi(u)}{f_{\mathrm{V}}(v)} = 2d\sigma_d \tag{3.18}$$

Using the Hasofer-Lind-Rackwitz-Fiessler algorithm, we can iterate to find δ^*, the minimum distance to the limit-state surface (line). The iterations are detailed in the following table, Table 3-1.

Table 3-1: Determining minimum distance to limit-state surface (Normal Case)

The u → v transformation	$U_{k+1} = \frac{U_k^T \nabla_U g(V_k) - g(V_k)}{\nabla_U g(V_k)^T \nabla_U g(V_k)} \nabla_U g(V_k)$
$U_0 = 0.0000 \rightarrow d_0 = 2.0000$	$U_1 = \frac{U_0^T (2d_0\sigma_d) - (d_0^2 - 1)}{(2d_0\sigma_d)(2d_0\sigma_d)} (2d_0\sigma_d)$
	$= \frac{0 \cdot (2 \cdot 2 \cdot 1) - (2^2 - 1)}{(2 \cdot 2 \cdot 1)(2 \cdot 2 \cdot 1)} (2 \cdot 2 \cdot 1)$
	$= \frac{0-3}{4} = -0.7500$
$U_1 = -0.7500 \rightarrow d_1 = 1.250$	$U_2 = \frac{-0.75 \cdot (2 \cdot 1.25 \cdot 1) - (1.25^2 - 1)}{(2 \cdot 1.25 \cdot 1)(2 \cdot 1.25 \cdot 1)} (2 \cdot 1.25 \cdot 1)$
	$= \frac{-1.875 - 0.5625}{2.5} = -0.9750$
$U_2 = -0.9750 \rightarrow d_2 = 1.0250$	$U_3 = -0.9997$
$U_3 = -0.9997 \rightarrow d_3 = 1.0003$	$U_4 = -1.0000$
$U_4 = -1.0000 \rightarrow d_4 = 1.0000$	$U_5 = -1.0000$

The algorithm is terminated and $\beta = \delta^* = \sqrt{(U^T \cdot U)} = 1$ is found using equation (3.13), giving $\Pr(f > 1) = \Pr(Success) = \Phi(\beta) = 0.8413$. Checking this with basic probability, we find:

$$\Pr(d^2 > 1) = \Pr(d < -1 \text{ or } d > 1) = 0.0013 + 0.8413 = 0.8426 \qquad (3.19)$$

which is the exact answer. Our answer of 0.8413 only approximates the most likely failure point $(d > 1)$, and not the second failure point, which is over 3σ away $(d < -1)$. As β, the estimate of the most likely failure point grows larger (i.e., $\beta \rightarrow \infty$), then the Hasofer-Lind-Rackwitz-Fiessler algorithm approximation approaches the basic probability calculation.

The same example can be done with d having a uniform distribution, in which case,

$$v = (b-a)\Phi(u) + a \qquad (3.20)$$

is used as the transformation from U-space to V-space, and the gradient becomes:

$$\nabla_U g(V) = 2d \frac{\varphi(u)}{f_V(v)} = 2d \frac{\varphi(u)}{1/\Delta_d} = 2d\Delta_d \varphi(u) \qquad (3.21)$$

where $1/\Delta$ is the probability density function with Δ representing the width of the uniform interval and $\varphi(\ )$ the standard normal probability density function. Now, the gradient changes with respect to the value of $\varphi(\ )$, unlike the normal case where it was a constant equal to σ (See Appendix A).

Changing the random variable d to uniform [-1..5], points representing the $\pm 3\sigma$ if d was normally distributed. Solving for U, we find:

Table 3-2: Determining minimum distance to limit-state surface (Uniform Case)

$U_0 = 0.0000 \rightarrow d_0 = 2.000$	$U_1 = \dfrac{U_0^T(2d_0\Delta_d\varphi(u)) - (d_0^2 - 1)}{(2d_0\Delta_d\varphi(u))(2d_0\Delta_d\varphi(u))}(2d_0\Delta_d\varphi(u))$ $= \dfrac{0\cdot(24\cdot 0.39894) - (2^2 - 1)}{(24\cdot 0.39894)(24\cdot 0.39894)}(24\cdot 0.39894)$ $= \dfrac{0-3}{9.57456} = -0.3133$
$U_1 = -0.3133 \rightarrow d_1 = 1.2621$	$U_2 = -0.4164$
$U_2 = -0.4164 \rightarrow d_2 = 1.0314$	$U_3 = -0.4305$
$U_3 = -0.4305 \rightarrow d_3 = 1.0005$	$U_4 = -0.4307$
$U_4 = -0.4307 \rightarrow d_4 = 1.0000$	$U_5 = -0.4307$

And we quit, with

$$\delta^* = 0.43073,\ \beta = 0.43073, \text{ and } \Pr(f > 1) = \Pr(\textit{Success}) = \Phi(\beta) = 0.6667$$

which makes intuitive sense since one-third of the uniform distribution lies beyond the limit-state surface; the range [–1..1] is in the failure region and (1..5] is in the success region. The point, 1, lies on the limit-state surface and is considered a failure point.

Chapter 4

Extension to Design

The transformation method enabies designers to determine the probability of success or failure of a response at its nominal conditions. The gradient information developed in Chapter 3 can enable designers to optimize the variables such that maximum success probability or minimum failure probability can be achieved. It is in this chapter that we extend the transformation methods and its gradients into design.

There are two types of design to be considered. The first is where only one quality characteristic is to be optimized. Other responses in the system may be important and can exist as constraints, but the design variables will be adjusted to maximize the probability of success or minimize the probability of failure of one response. The second is multi-objective, where multiple quality characteristics are to be optimized. This implies that there may be tradeoffs required in order to optimize all responses of interest. In addition, numerous examples are given to show that the approach is discipline independent.

4.1 Single Quality Characteristic Optimization

The first and simplest case of design deals with single response problems. Here, one quality characteristic is important and the system design variables are adjusted so as to improve the quality of the system. However, there exist many issues within this one problem. This section will show that the probability approach from Chapter 3 is an extension of Taguchi's Signal-to-Noise ratios. The extension being that the approach uses full distribution information while signal-to-noise ratios are limited to first and second order information, typically the mean and variance of the response. Thus, with the approach deemed an extension of signal-to-noise ratios, improvement of quality is accomplished through maximization of this measure, as discussed in Section 4.1.1.

In addition, the transformation method and its gradient information supply the designer with knowledge about the variables. This information can be used to determine which variables should be used as design variables, and also classifies them according to Taguchi's factor types. This classification is detailed in Section 4.1.2.

After design variables have been selected, the methodology by which we can improve the system is presented in Section 4.1.3 with examples following.

4.1.1 Taguchi's 'Quality by Design'

Taguchi's definition of quality stems from the loss incurred to society if the product/systems/response does not meet the specifications set by the designers and engineers. In the extension of this definition to design optimization, Taguchi developed the Signal-to-Noise Ratio (Knottmaier 1993) (Barker 1990) (Maghsoodloo 1990) (Logothetis and Wynn 1989) (Leon, Shoemaker, and Kackar, 1987). The use of signal-to-noise ratios in system analysis provides a mathematical value for response variation comparison.

The Signal-to-Noise Ratio (SNR) is a concurrent statistic that combines two characteristics of a distribution into a single number (e.g., the coefficient of variation). In the application of signal-to-noise ratios to design optimization, maximization of the SNR results in minimization of the response variation. The recommended signal-to-noise ratio by Taguchi when there is a target value to be achieved is:

$$SNR = 10\log_{10}\left(\frac{\mu_R^2}{\sigma_R^2}\right) \tag{4.1}$$

Note that the Cornell index, equation (3.4), is similar to the above SNR. Different SNR's have been developed for 'smaller is best', 'larger is best', and 'target is best' conditions to aid designers in the minimization of variance. In a like manner, the probability approach seeks to minimize the variance of the response by maximizing the probability of success, equation (3.8), or minimizing the probability of failure, equation (3.9). By using the transformation method, arbitrary full distribution information can be used in the calculation of β and accurate approximations of Pr(*Success*) and Pr(*Failure*) can be found. For the two-sided probability problem, as in Figure 3-3, if the response distribution is symmetrical, we seek to make the probability of failure beyond each specification equal. This is equivalent to minimizing variance for the target-is-best condition, where the target is the center of the specifications. For non-symmetrical distributions, we require the gradient of the probability at the two specifications. Each of these values is weighted by the probability of being at the respective specification. When these two weighted values are equal, minimization of the variation of the non-symmetrical response is achieved. 'Design for Quality' can now be done by using the composite signal-to-noise ratios, Pr(*Success*) and Pr(*Failure*). In order to do this, two gradients are required: $\nabla_V g(\mathrm{V})$ and $\nabla_U g(\mathrm{V})$, where the subscripts refer to differentiation by V, the variables, and U, the variation of the variables. Derivation of $\nabla_U g(\mathrm{V})$ is found in equation (3.14).

4.1.2 Selection of Design Variables

In order to improve an existing system, design variables must be determined or specified. Taguchi often refers to the fact that the variables of a system can be classified into six factor types: Control, Signal, Adjustment, Variance-Affecting, Noise, and Neutral. It is through exploitation of the interactions of these different factor types that the signal-to-noise ratios can be improved (Dehnad 1988). By examining the partial

gradient of $g(\mathrm{V})$ with respect to v_i and u_i, one can determine the effect of the nominal value (in V-space) and of the variation (in U-space) that v_i has on the margin, and thus, how it affects the failure probability. Specifically, the probability approach aids in the determination of neutral, design–adjustment, design–control, and variance-affecting only factors. Noise factors and signal factors, like Taguchi's approach, must be user specified.

By examining the numerical values for $\partial g(\mathrm{V})/\partial v_i$ and $\partial g(\mathrm{V})/\partial u_i$ for each variable, there are two ways to use the information:

1. Evaluate the $\partial g(\mathrm{V})/\partial v_i$ and $\partial g(\mathrm{V})/\partial u_i$ at the MLFP, indicated as $\partial g(\mathrm{V})/\partial v_{i\,|MLFP}$ and $\partial g(\mathrm{V})/\partial u_{i\,|MLFP}$, so as to determine the effect of v_i and u_i on the failure probability. The results are only applicable in the neighbourhood of the MLFP.
2. Evaluate the $\partial g(\mathrm{V})/\partial v_i$ and $\partial g(\mathrm{V})/\partial u_i$ for all values of v_i and u_i so that conclusions can be drawn about global effect on the failure probability.

By studying the gradients only at the MLFP, if one classifies the variables into one of the four factor types possible (neutral, adjustment, control, variance), then these classifications will only be applicable in the neighbourhood of the current design. To draw conclusions on a global scale, more information is required. By evaluating the gradients at multiple failure points, a more accurate assessment of the different factor types is possible.

In studying the two gradients, $\partial g(\mathrm{V})/\partial v_i$ and $\partial g(\mathrm{V})/\partial u_i$, four cases can be enumerated so that factor types can be determined. They are detailed below and will be called Case A, B, C, and D.

CASE A. $|\partial g(\mathrm{V})/\partial v_i| \le \varepsilon_V$ where ε_V is very small. There are four conditions that satisfy this. The first two are special conditions where the partial gradient is equal to zero at the MLFP (Condition 1) or for all values of v_i (Condition 2), and the remaining two are if the partial gradient is less than ε_V at the MLFP (Condition 3) or for all values of v_i (Condition 4).

Condition 1: If $\partial g(\mathrm{V})/\partial v_{i\,|MLFP} = 0$, then $g(\mathrm{V})$ is not a function of v_i at MLFP, since the linearized failure surface is parallel to the v_i axis (see Figure 4-1).

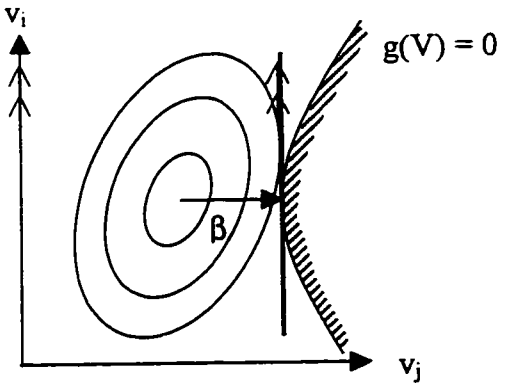


Figure 4-1: Linearized failure surface parallel to v_i axis

Thus, small perturbations in nominal value of v_i will have little effect on the value of the margin and the probability of failure. However, if v_i is moved significantly, $\partial g(\mathrm{V})/\partial v_{i|MLFP} = 0$ may not hold true for new values of v_i. Evaluation at the MLFP only provides information for the respective nominal value of v_i and its effect in its immediate neighbourhood.

For example, if the margin is defined as:

$$g(\mathrm{V}) = 5 - v_1 + \frac{1}{2}(v_2 - 4)^2 \tag{4.2}$$

then the gradient of the margin with respect to v_2, evaluated at the MLFP (5. 4), is zero. However, a small change away from the MLFP produces a non-zero value for the gradient.

Condition 2: If $\partial g(\mathrm{V})/\partial v_i = 0$ for all values of v_i, then $g(\mathrm{V})$ is not a function of v_i. This implies that the failure surface is parallel to the v_i axis at all values of v_i (see Figure 4-2). In this case, the value of the margin and the failure probability are independent of v_i.

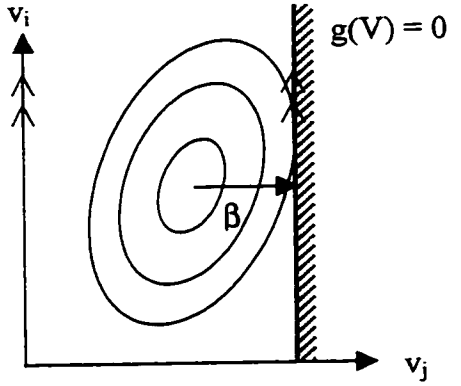


Figure 4-2: $g(\mathrm{V}) = 0$ parallel to $\mathrm{v_i}$ axis

Condition 3: If $\left|\partial g(\mathrm{V})/\partial v_{i|MLFP}\right| \le \varepsilon_v$, the linearized failure surface is approximately parallel to the v_i axis in the region of the MLFP (see Figure 4-3).

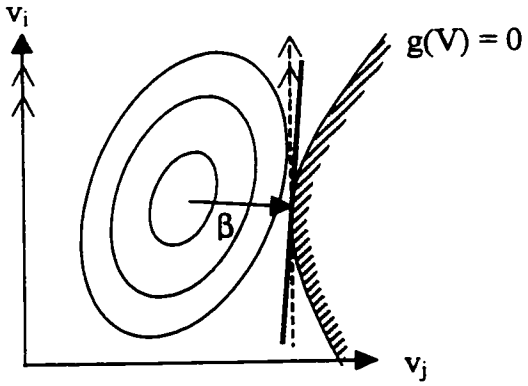


Figure 4-3: Linearized Failure Surface approximately parallel to $\mathrm{v_i}$ axis

Thus, a small change in the nominal value of v_i has little effect on the value of the margin, but is only true in the neighbourhood of the MLFP.

Condition 4: If $|\partial g(\mathrm{V})/\partial v_i| \le \varepsilon_v$ for all values of v_i, then the failure surface, $g(\mathrm{V}) = 0$. is approximately parallel to the v_i axis (see Figure 4-4).

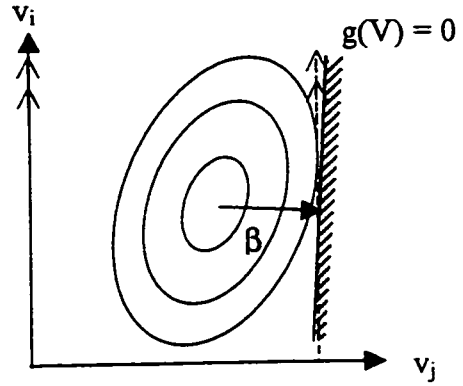


Figure 4-4: $g(\mathrm{V}) = 0$ approximately parallel to v_2 axis

This implies that any change in the nominal value of v_i has little effect on the value of the margin or failure probability

For example, if the margin is defined as:

$$g(\mathrm{V}) = 5 - v_1 + \frac{v_2}{10000} \tag{4.3}$$

then the gradient of the margin with respect to v_2 is very small (0.0001). Thus, the value of the margin and the failure probability are relatively insensitive to changes in the nominal value of v_2. Any change in v_2 would need to be large in order to cause any significant value change in $g(\mathrm{V})$ and ultimately Pr(*Failure*).

CASE B. $|\partial g(\mathrm{V})/\partial v_i| > \varepsilon_v$ for any value of v_i

Whether evaluating at the MLFP or at any value for v_i, if the partial gradient is greater than ε_v, then it implies that a change in the nominal value of v_i has a greater effect on the value of the margin and failure probability than in case A above. Using this gradient information, we can change v_i and move away from the failure surface to keep $g(\mathrm{V})$ positive. By doing this, β is increased.

Before looking at the gradients in U-space cases, which are C and D, a few notes must be made. The purpose of converting the quality problem into U-space is to remove any correlation and to be able to take advantage of the properties of standard normal space (see Section 3.3). Recall equation (3.14). The partial gradient of the margin with respect to the U-space variable, $\partial g(\mathrm{V})/\partial u_i$, is a two-part product; the gradient of the

margin in V-space and the density ratio. By studying the interaction between these parts, we can determine how any variation will affect the margin and the failure probability.

CASE C. $|\partial g(\mathrm{V})/\partial u_i| \le \varepsilon_U$ where ε_U is very small. There are four conditions that satisfy this. The first two are special conditions where the partial gradient is equal to zero at the MLFP (Condition 1) or for all values of u_i (Condition 2), and the remaining two deal with the partial gradient less than ε_U at the MLFP (Condition 3) or for all values of u_i (Condition 4). Conditions 1-3 are additionally sub-divided into two parts; (a) the effect of the partial gradient of the margin in V-space on the condition, and (b) the effect of the density ratio on the condition.

Condition 1: $\partial g(\mathrm{V})/\partial u_{i\,|MLFP} = 0$. There are two possibilities that this will occur; (a) if the partial gradient of the margin in V-space evaluated at the MLFP is zero, or (b) the density ratio evaluated at the MLFP is zero. Each possibility is discussed below.

(a) If CASE A, Condition 1 or 2 is true, then the transformed failure surface or its linearized counterpart is be parallel to the u_i axis (Figure 4-5).

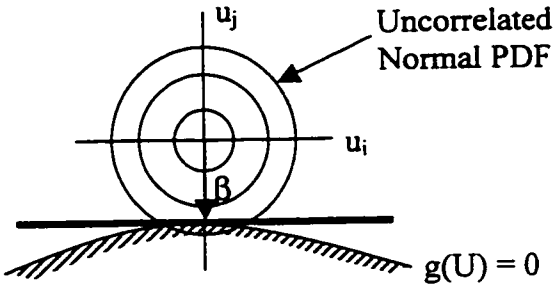


Figure 4-5: The linearized failure surface in U-space is parallel to the u_i axis since the V-space linearized failure surface or actual failure surface is parallel to the v_i axis

In this case, the gradient of the margin in U-space may only be zero when the variable v_i is at its nominal value. If the variable has a value other than its nominal (i.e., any other point on its PDF), then the gradient may not be zero, depending on the shape of the transformed failure surface. As such, the assumption of $\partial g(\mathrm{V})/\partial u_{i\,|MLFP} = 0$ should only be considered accurate within the neighbourhood of variable v_i's nominal value.

(b) $\varphi(u_i)/f_{\mathrm{V}_i}(v_i)_{|MLFP} = 0$ which implies the variance of variable v_i is zero. If the variance is zero, then $\varphi(u_i)/f_{\mathrm{V}_i}(v_i) = 0$ for all values of u_i and the limit-state function is not a function of u_i. For example, if variable v_i is a point with no variance, then

$$\varphi(u_i)/f_{\mathrm{V}_i}(v_i) = \partial v_i/\partial u_i = 0$$

and the limit-state function is independent of u_i.

Condition 2: $\partial g(\mathrm{V})/\partial u_i = 0$ for all values of u_i – There are two possibilities that this will occur: (a) if the partial gradient of the margin in V-space is zero for all v_i, or (b) the density ratio is zero. Each possibility is discussed below.

(a) CASE A – Condition 2 is true implying that the transformed failure surface is parallel to the u_i axis and is independent of u_i (see Figure 4-6).

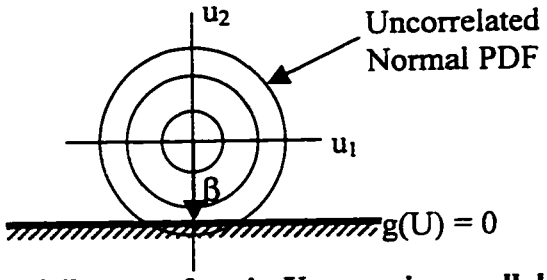


Figure 4-6: The failure surface in U-space is parallel to the u_i axis because the V-space failure surface is parallel to the v_i axis

(b) $\varphi(u_i)/f_{V_i}(v_i) = 0$ which implies the variance of variable v_i is zero making the limit-state function is independent of u_i.

Note that for CASE C – Condition 1 or 2, if $\varphi(u_i)/f_{V_i}(v_i) = 0$ at the MLFP or for all values of u_i, it implies that there is no variance in variable v_i.

Condition 3: $\left|\partial g(\mathrm{V})/\partial u_{i\,|MLFP}\right| \le \varepsilon_U$ – By splitting up the partial gradient into its two parts and dividing through by the density ratio, the following relationship can be found:

$$\left|\frac{\partial g(\mathrm{V})}{\partial v_i}_{|MLFP}\right| \le \varepsilon_U \frac{1}{\varphi(u_i)/f_{V_i}(v_i)_{|MLFP}} \tag{4.4}$$

By examining each part of the relationship, (a) the partial gradient of the margin in V-space with respect to v_i, and (b) the density ratio, one can determine why $\left|\partial g(\mathrm{V})/\partial u_{i\,|MLFP}\right| \le \varepsilon_U$ is true.

(a) The left-hand side equation (4.4) is the partial gradient of the margin in V-space with respect to v_i evaluated at the MLFP, $\left|\partial g(\mathrm{V})/\partial v_{i\,|MLFP}\right| \le \varepsilon_V$. If this partial gradient is small (CASE A, Condition 1), it can obviate the effect from the density ratio, $\varphi(u_i)/f_{V_i}(v_i)$. This implies that the linearized failure surface is approximately parallel to the v_i and u_i axes, thereby obviating the effect any variation in v_i (determined by the density ratio) has in U-space about the nominal. In this particular case where the partial gradient is evaluated at the MLFP, the obviation may only apply within the neighbourhood of the nominal value of v_i (see Figure 4-7).

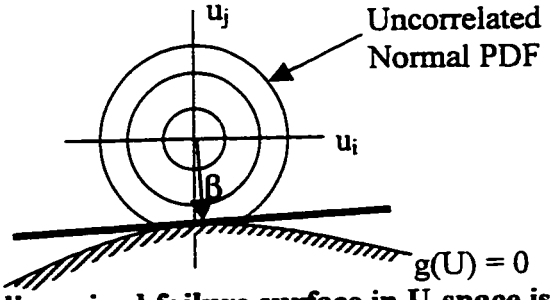


Figure 4-7: The linearized failure surface in U-space is approximately parallel to the u_i axis ($\left|\partial g(\mathrm{V})/\partial u_{i\,|MLFP}\right| \le \varepsilon_U$) because the V-space linearized failure surface or actual failure surface was approximately parallel to the v_i axis ($\left|\partial g(\mathrm{V})/\partial v_{i\,|MLFP}\right| \le \varepsilon_V$)

(b) The density ratio represents the relationship of the distribution in V-space to the standard normal distribution in U-space. By showing the density ratio for both normal-like and non-normal distributions, one can determine when $\varphi(u_i)/f_{V_i}(v_i)$ is sufficiently small as to obviate the effect from the partial gradient of the margin in V-space.

For a normally distributed variable with variance σ_{vi}, equation (4.4) becomes:

$$\left|\frac{\partial g(\mathrm{V})}{\partial v_{i\,|MLFP}}\right| \le \varepsilon_U \frac{1}{\sigma_{v_i}} \tag{4.5}$$

If the variance σ_{vi} is sufficiently small enough to make equation (4.5) true, then $\left|\partial g(\mathrm{V})/\partial u_{i\,|MLFP}\right| \le \varepsilon_U$. For example, using the previous margin function from Case A, equation (4.3), in U-space, if the parameters were normally distributed in V-space, becomes:

$$g(\mathrm{V}) = 5 - (\mu_1 + \sigma_1 u_1) + \frac{\mu_2 + \sigma_2 u_2}{10000} \tag{4.6}$$

and the gradients with respect to u_1 and u_2 are:

$$\frac{\partial g(\mathrm{U})}{\partial u_1} = \sigma_1 \text{ and } \frac{\partial g(\mathrm{U})}{\partial u_1} = \frac{\sigma_2}{10000}$$

in both cases, if σ_1 or σ_2 is very small, then $\partial g(\mathrm{V})/\partial u_{i\,|MLFP} \le \varepsilon_U$ for u_1 and u_2.

For non-normal distributions (e.g., uniform, triangular) if $\varphi(u_i)$ is small due to large u_i, it implies that the limit-state function passes through U-space sufficiently far away from the u_i axis such that $\varphi(u_i) \cong 0$ (see Figure 4-8).

For example, evaluating (4.4) for a uniform distribution,

$$\left|\frac{\partial g(\mathrm{V})}{\partial v_i}_{|MLFP}\right| \le \varepsilon_U \frac{1}{(b-a)\varphi(u_i)} \tag{4.7}$$

where a and b are constants represent the lower and upper values of the uniform distribution, if $\varphi(u_i)$ is sufficiently small for (4.7) to be true, then $\left|\partial g(\mathrm{V})/\partial u_{i_{|MLFP}}\right| \le \varepsilon_U$.

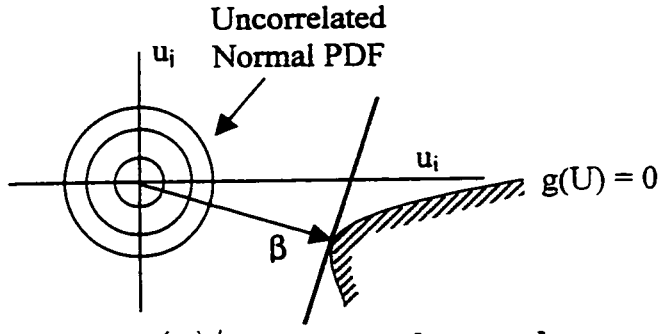


Figure 4-8: $\partial g(\mathrm{V})/\partial u_{i_{|MLFP}} \le \varepsilon_U$ **due to** u_i **large, making** $\varphi(u_i)/f_{V_i}(v_i)_{|MLFP} \cong 0$

Let us reconsider the limit-state function, equation (4.3), where the distributions in V-space are, for example, uniform. For this example, the margin and gradients would be as follows:

$$g(\mathrm{V}) = 5 - \left(\Delta_1 \Phi(u_1) + a_1\right) + \frac{\Delta_2 \Phi(u_2) + a_2}{10000} \tag{4.8}$$

$$\frac{\partial g(\mathrm{U})}{\partial u_1} = \Delta_1 \varphi(u_1) \text{ and } \frac{\partial g(\mathrm{U})}{\partial u_1} = \frac{\Delta_2 \varphi(u_2)}{10000}$$

where the partial gradients could be negligible for two reasons. For example, considering u_1, the gradient would be negligible if either Δ_1 is small implying a narrow width uniform distribution or $\varphi(u_1)$ is small because u_1 is large implying that the failure surface is a large number of standardized normal deviations away from the nominal value of v_1. If either or both is true, then the partial gradient can be considered negligible and $\left|\partial g(\mathrm{V})/\partial u_{i_{|MLFP}}\right| \le \varepsilon_U$.

Condition 4: If $\left|\partial g(\mathrm{V})/\partial u_i\right| \le \varepsilon_U$ for all values of u_i implies that:

$$\left|\frac{\partial g(\mathrm{V})}{\partial v_i}\right| \le \varepsilon_U \frac{1}{\max\left(\varphi(u_i)/f_{V_i}(v_i)\right)} \tag{4.9}$$

For this case, equation (4.4) is extended to all u_i and v_i. If equation (4.9) is true, then it implies that regardless of how large the variation is in variable v_i, it has little or no effect on the value of the margin or the failure probability (i.e., u_i varying from $-\infty$ to $+\infty$ has little or no effect on the failure probability).

Note: CASE C really applies to studying variables with non-normal distributions. Since the density ratio of a normal distribution is simply σ, the standard deviation, then unless sigma is zero or very small, the variable will always have an effect on the variation of the response. However, CASE C shows that there can be interaction between the products such that at certain values of the variable, the partial gradient is small thereby obviating the effect of the variance. This is the essence of robust design; it takes advantage of these interactions to reduce variation. Design tells us to move the nominal far away from the failure surface so that a large number of standard deviations exist between it and the failure surface. However, for non-normal distributions as CASE C shows, this does not have to be the case. Instead, we can move the design point such that one of the parts of the product obviates the other. At this point, the response is not very sensitive to changes in the parameter values.

CASE D. $|\partial g(V)/\partial u_i| > \varepsilon_U$ for all values of u_i

This implies that the variation in v_i (in U-space it is represented by u_i) has a greater effect on the value of the margin and failure probability than in CASE C.

Classification into Taguchi's Factor Types

Thus, using these four cases, each variable v_i can be classified into one of the factor types. The following truth table can be used to determine the factor type of a variable.

Table 4-1: Factor Determination using the Probability Approach
(T is True, F is False, X is Don't Care)

Factor Type	CASE A $\left\|\frac{\partial g(V)}{\partial v_i}\right\| \le \varepsilon_V$	CASE C $\left\|\frac{\partial g(V)}{\partial u_i}\right\| \le \varepsilon_U$	
Class I: Design–Control	F	F	
Class II: Unnamed (Variance only)	T	F	
Class III: Design–Adjustment	F	T	
Class IV: Neutral	T	T	
Signal	F	X	← User selected
Noise	X	X	← v_i uncontrollable

From the table, the last two factor types, as in Taguchi's method, must be user specified. If the variable is uncontrollable, then it is considered a noise factor. If the variable is considered to be a user input, then it is a signal factor. The remaining four cases are neutral, design–adjustment, design–control, and the unnamed factor which affects only variance. Depending on whether the partial gradients are evaluated only at

MLFP or at all v_i and u_i, the factor classifications will be applicable on either a local (MLFP) or global (for all v_i and u_i) scale. In some cases, a local classification of a variable factor type may be sufficient if the variable is not expected to vary much. However, for design, a global analysis should be done to ensure proper classification.

However, it must be remembered that the above classifications are for the purposes of comparison to Taguchi's method. Within the transformation method and the subsequent design for quality method, the factor types of the variables do not need to be enumerated. The algorithm by Hasofer, Lind, Rackwitz, and Fiessler automatically uses the gradient information correctly in its determination of the MLFP. Without specification of factor types, the optimization routines will still move the variables to the best optimal design using the gradient information. Indirectly, the optimization process classifies the factors at each iteration and adjusts the nominal values accordingly.

It is in design that we can use this classification as a preliminary step to determine factor types and ultimately reduce the number of design variables. Once classified, we wish to remove neutral factors (global level) from the optimization process and compare the results from the remaining "design" factor types to determine importance. From this information, we could choose the top 5 or 10 variables that may contribute most to the design optimization. An expert system could be developed to select the top design and noise variables and remove neutral variables from the optimization process. The result would be a computational savings over solving the entire system with all variables. This was what Taguchi did using analysis of variance (ANOVA). Based on results attained from ANOVA, he determined if variables were neutral, adjustment, or control.

The above mathematical link shows that Taguchi was indirectly finding out which variables did not have normal distributions, and which ones had little or no variance. In addition, his analysis was based on averages and from this determined the factor types based on multiple nominal points. His method agrees with the method presented in this section for determining factor types. Taguchi's use of a few nominal points lies somewhere between analysis at a single point, the MLFP, and analysis for all v_i and u_i. His use of a collection of nominal points would be equivalent to tracking the partial gradients in U- and V-space over a number of design iterations.

As an additional note, work by Ben-Haim (1996) coins the term "hyperplane separation" as a method to determine how much a variable can vary before the response exceeds some predefined limits. That work is similar to the above in the sense that hyperplane separation tries to determine to what degree a variable affects its response without using distributions. Instead, the work is based on intervals and uses the hyperplanes as a means to quantify the tangent plane between two expanding ellipses. The result is a mathematical estimation of the importance of each variable.

4.1.3 Methodology

Upon choosing design variables, $\nabla_V g(V)$ and $\nabla_U g(V)$ are reformulated such that,

$$\nabla_V g(V_k) = \begin{bmatrix} \nabla_D g(V) \\ \nabla_S g(V) \\ \nabla_N g(V) \end{bmatrix} \text{ and } \nabla_U g(V) = \begin{bmatrix} \frac{\partial g(V)}{\partial u_1} \\ \vdots \\ \frac{\partial g(V)}{\partial u_n} \end{bmatrix} \qquad (4.10) \text{ and } (4.11)$$

where $\nabla_V g(V)$ is found by stacking $\nabla_D g(V)$, $\nabla_S g(V)$, and $\nabla_N g(V)$ and the first d u_i's in $\nabla_U g(V)$ correspond to $\nabla_D g(V)$, the next s u_i's correspond to $\nabla_S g(V)$ and the remaining u_i's to $\nabla_N g(V)$, where D represents the design factors, S, the signal factors, and N, the noise factors. In addition, the gradient of our index β with respect to the design variables is found. Mathematically it is:

$$\nabla_D \beta = \begin{bmatrix} \frac{\partial \beta}{\partial d_1} \\ \vdots \\ \frac{\partial \beta}{\partial d_n} \end{bmatrix}, \text{ where } \frac{\partial \beta}{\partial d_i} = \frac{\frac{\partial g(V^*)}{\partial d_i}}{\left\| \nabla_U g(V^*) \right\|} \qquad (4.12)$$

where V^* is the value of the variables in V-space at the MLFP.

Using the algorithm by Hasofer, Lind, Rackwitz, and Fiessler, the point in U-space is found and then converted to V-space through the transformations specified in Appendix A. If multiple margins exist, then $\nabla_D \beta$ is n by m, where there are n design variables and m margins, and in order to maximize all β's, we need the gradient of the probabilities with respect to the design parameters. If we want to improve the probability of success, it is intuitive that we want to increase the β indices with respect to both the upper and the lower specification. However, it is obvious that we cannot change one without affecting the other. Madsen, Krenk, and Lind (1986) present the following derivation:

$$\frac{\partial \Pr(Success)}{\partial d} = \sum_{i=1}^{m} \varphi(\beta_i) \frac{\partial \beta_i}{\partial d} \qquad (4.13)$$

where to improve the overall probability of success, we need β and its corresponding gradient with respect to the design parameters. The resulting weighted summation of the gradients tells us which way to move the design variables to increase the total probability of success. Using the value of the standard normal PDF at β_i to weight its corresponding gradient causes margins with small β's to be weighted higher than margins with large β's. When the weighted gradients balance out and the sum equals zero, the point of maximum probability of success is found. Thus, for the two-sided problem in Figure 3-3, the formulations are:

$$\frac{\partial \Pr(Success)}{\partial d} = \varphi(\beta_{LSL})\frac{\partial \beta_{LSL}}{\partial d} + \varphi(\beta_{USL})\frac{\partial \beta_{USL}}{\partial d} \tag{4.14}$$

$$\frac{\partial \Pr(Failure)}{\partial d} = -\varphi(-\beta_{LSL})\frac{\partial \beta_{LSL}}{\partial d} - \varphi(-\beta_{USL})\frac{\partial \beta_{USL}}{\partial d} \tag{4.15}$$

In summary, the flow chart presented in Figure 4-9 captures the essence of the new method proposed.

4.1.4 Computational Requirements

The methodology presented in this chapter has three levels of computation (Carr and Savage 1996):

1. Number of design optimization iterations (N_{OPT})
   - At each iteration, the Hasofer-Lind-Rackwitz-Fiessler (HLRF) algorithm is solved for each limit-state
2. Number of HLRF algorithm iterations (N_{HLRF})
   - At each iteration, the margin, a function of the response, is evaluated – typically less than 12.
3. Number of iterations to find the solution of the system (N_S)
   - For a non-linear system, this is the number of Newton-Raphson algorithm iterations – typically less than 10.
   - For a linear system, the number of iterations is one.

Therefore, the average number of response evaluations can be described as:

$$\overline{N}_F = N_{OPT} \cdot N_{HLRF} \cdot N_S \cdot L \tag{4.16}$$

where L is the number of limit-state functions. By counting the total number of response evaluations in the optimization, this allows for comparison to the number of Monte Carlo simulations required to achieve the same improvement.

4.1.5 Examples

The following examples will compare previously solved problems to the results achieved through the method in this thesis to show that it achieves the same results or better. In addition, the examples are chosen from a wide variety of fields to show the universality of the method.

The first three examples are from publications that used experimental design and signal-to-noise ratios to improve quality. These examples all use experimental design on a mathematical model to improve the system. The first example is from Taguchi and Phadke (1984) and will be extensively detailed. The second example is from Bagchi and Templeton (1994), and the third example is from D'Errico and Zaino (1988). The fourth example compares a variance-transfer method (Xie 1994) to the probability approach.

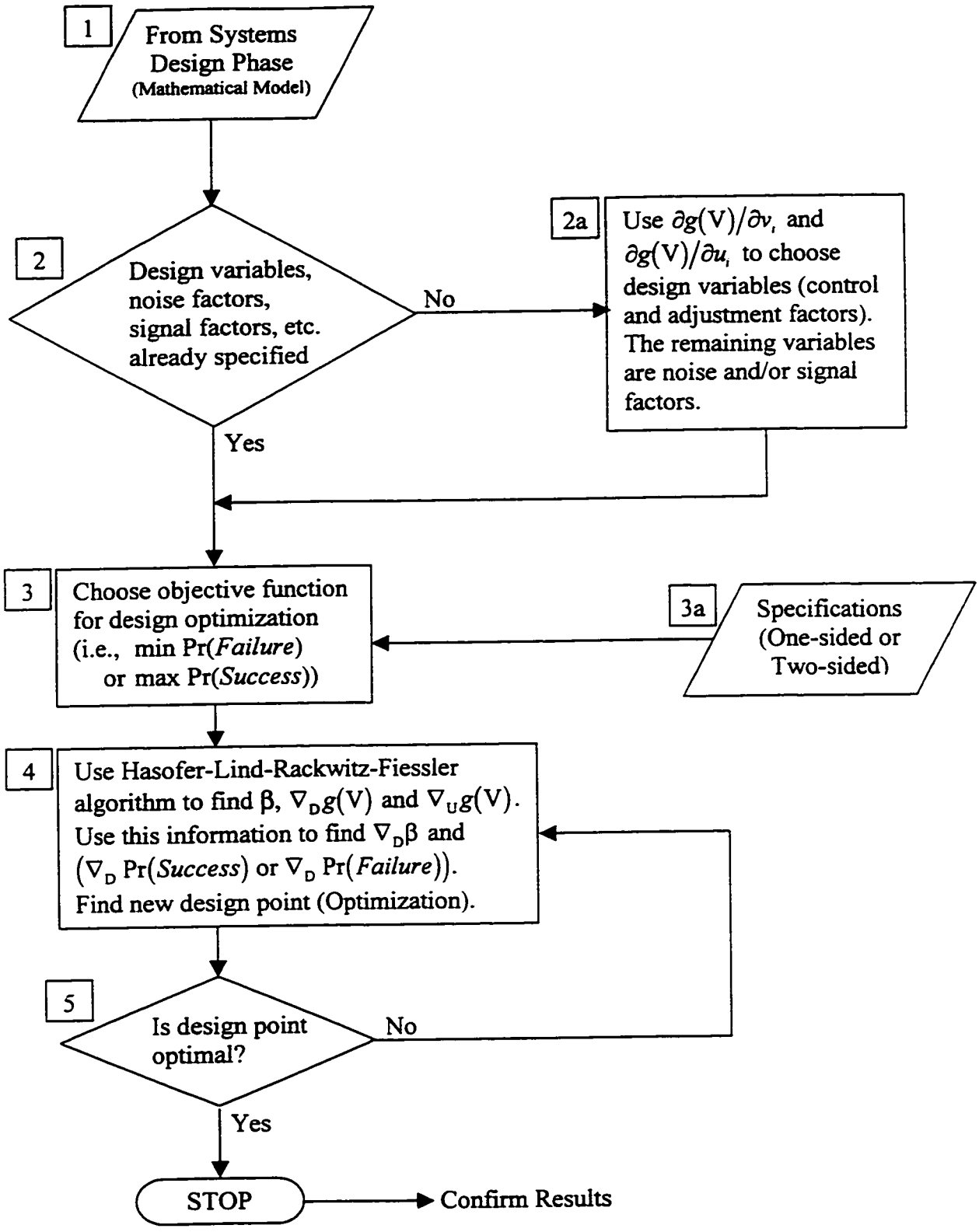


Figure 4-9: Design for Quality Flowchart

Design of a Temperature Controller Circuit

Taguchi and Phadke (1984) show off-line design for a temperature controller using experimental design. The goal was to find the highest signal-to-noise ratio for the resistive element $R_{T\text{-}ON}$, the resistive value of when the relay turns on, given a variable resistor R_3 used to set the desired temperature. The circuit is as follows:

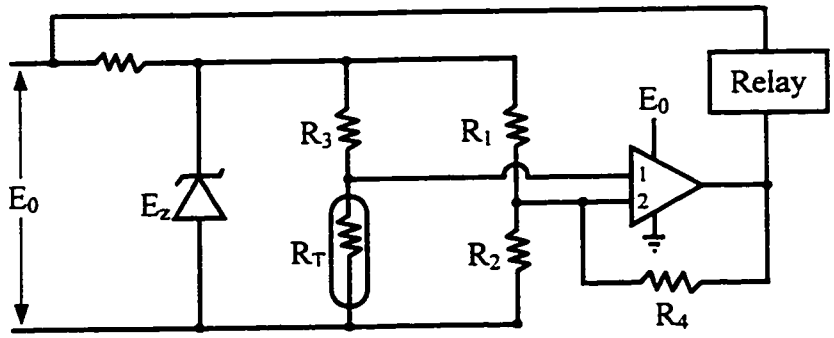


Figure 4-10: Temperature Controller Circuit

In the experimental design, R_3 was chosen as the signal factor and R_2 and R_4 were converted to scalars of R_1 ($\alpha = R_2/R_1$ and $\lambda = R_4/R_1$). The variables and their levels are given in Table 4-2 and noise factors in Table 4-3.

Table 4-2: Control Factors and Levels

Factor	Levels		
	1	2	3
R_1	2.67	4.0	6.0
α	1.33	2.0	3.0
λ	5.33	8.0	16.0
E_0	8.0	10.0	12.0
E_z	4.8	6.0	7.2

Table 4-3: Associated Tolerances for Control Factors

Factor	Tolerance (% of Nominal)	Levels (% of Nominal)		
		1	2	3
R_1	5	-2.04	0	2.04
α	5	-2.04	0	2.04
λ	5	-2.04	0	2.04
E_0	5	-2.04	0	2.04
E_z	5	-2.04	0	2.04

The resistance value of R_T at which the relay turns on is given as:

$$R_{T-ON} = \frac{R_3 R_2 (E_z R_4 + E_0 R_1)}{R_1 (E_z R_2 + E_z R_4 - E_0 R_2)} \tag{4.17}$$

As an aside, once the relay turns on, the resistance R_T must reach a value of:

$$R_{T-OFF} = \frac{R_2 R_3 R_4}{R_1 (R_2 + R_4)} \tag{4.18}$$

in order to turn the relay off. However, for this example Taguchi and Phadke only considered minimizing the variance of $R_{T\text{-}ON}$. For comparison purposes, only $R_{T\text{-}ON}$ will be considered here too.

Taguchi and Phadke compute 18 runs of the experimental design and find the optimal levels, i.e., highest signal-to-noise ratio, to be α – level 1, λ – level 3, E_0 – level 1, and E_z – level 3, and find that R_1 had little impact on the signal-to-noise ratio. The resistance R_T at the optimal settings had a value of 1.4356 Ohms.

To use the transformation method to solve this problem, the design variables have a normal distribution using the information given by Taguchi and Phadke. The starting values of the design variables (Box 2 in Figure 4-9) would correspond to the level 1's in the experimental design:

Table 4-4: Design Variables and Associated Information

Variable	Mean Value	Variation	Constraints
R_1	2.67	±5%	$2.67 < R_1 < 6.0$
α	1.33	±5%	$1.33 < \alpha < 3.0$
λ	5.33	±5%	$5.33 < \lambda < 16.0$
E_0	8	±5%	$8.0 < E_0 < 12.0$
E_z	4.8	±5%	$4.8 < E_z < 7.2$

and the signal factor, R_3, is assumed to have a uniform distribution ranging from [0.898...1.102] allowing the set-point represented by R_3 to vary anywhere along the range. This assumption was made for R_3 from the information given by Taguchi and Phadke – they stated that R_3 would be set at three levels, 0.898, 1, and 1.102 kilo-ohms, to allow for the estimation of the linear and quadratic effects. In the probability approach, the assumption of a uniform distribution is the least biased to this information.

These five design variable plus the signal factor are represented by the vector V, such that:

$$V = [R_1 \quad \alpha \quad \lambda \quad E_0 \quad E_z \quad R_3] \tag{4.19}$$

By setting some restrictions (i.e., upper and lower specifications) on the response, R_T, say [1.35...1.55] ohms (Box 3a in Figure 4-9), we can use the transformation method

and optimization to find an R_T within the specifications that has the minimum variance. The choice of these specifications is arbitrary in this case, and represents that one wishes to find a resistance value between 1.35 and 1.55 ohms. The objective function used will be the maximization of equation (3.7), Pr(*Success*) (Box 3 in Figure 4-9).

The two margins in this example become:

$$\begin{aligned} m_1 &= R_T - LSL = R_T - 1.35 \\ m_2 &= USL - R_T = 1.55 - R_T \end{aligned} \tag{4.20}$$

Using the above design variables and their respective constraints, the initial design point is evaluated with respect to both margins to find (Box 4 in Figure 4-9):

$$R_T = 2.0943,\ \beta = \begin{bmatrix} 10.8256 & -6.8086 \end{bmatrix}, \text{ and } \Pr(Success) = 0$$

$$\nabla_D g(V) = \begin{bmatrix} 0.0000 & 0.0000 \\ 1.5050 & -1.5721 \\ -0.0848 & 0.1062 \\ 0.1128 & -0.1327 \\ -0.1606 & 0.2002 \end{bmatrix} \text{ and } \nabla_U g(V) = \begin{bmatrix} 0.0000 & 0.0000 \\ 0.0334 & -0.0349 \\ -0.0075 & 0.0094 \\ 0.0150 & -0.0177 \\ -0.0128 & 0.0160 \\ 0.1289 & -0.1329 \end{bmatrix}$$

where the first five rows in the above gradients are with respect to the design variables and the sixth row in $\nabla_U g(V)$ corresponds to the random variable/signal factor R_3. Note that β is positive for margin 1 and negative for margin 2, indicating that the current design satisfies the lower constraint on R_T but not the upper constraint. In addition:

$$\nabla_D \beta = \begin{bmatrix} 0.0000 & 0.0000 \\ 11.1651 & -11.2459 \\ -0.6293 & 0.7600 \\ 0.8370 & -0.9492 \\ -1.1912 & 1.4320 \end{bmatrix},$$

$$\varphi(-\beta_{LSL}) \nabla_D \beta_{LSL} = 1.0e^{-24} \times \begin{bmatrix} 0.0000 \\ 0.1587 \\ -0.0089 \\ 0.0119 \\ -0.0169 \end{bmatrix}, \text{ and } \varphi(-\beta_{USL}) \nabla_D \beta_{USL} = 1.0e^{-9} \times \begin{bmatrix} 0.0000 \\ -0.3852 \\ 0.0261 \\ -0.0325 \\ 0.0490 \end{bmatrix},$$

combine to determine $\nabla_D \Pr(Success)$:

$$\nabla_D \Pr(Success) = \varphi(-\beta_{LSL})\nabla_D\beta_{LSL} + \varphi(-\beta_{USL})\nabla_D\beta_{USL} = 1.0e^{-9} \times \begin{bmatrix} 0.0000 \\ -0.3852 \\ 0.0260 \\ -0.0325 \\ 0.0490 \end{bmatrix}$$

Looking at $\nabla_D \Pr(Success)$, if the probability of success is to be increased, we need to decrease α, increase λ, decrease E_0 and increase E_z. Note that R_I has no effect on the probability. Taguchi and Phadke found the same in their analysis of variance. Since α and E_0 cannot be decreased, we can only increase λ and E_z. Using an optimization routine (Matlab 1994), a new design point is chosen. Since the design variables can be changed, the design is currently not optimal (Box 5 in Figure 4-9) – next iteration.

Following boxes 4 and 5 until the optimal design point is found, the exact answer from Taguchi and Phadke is found:

$$R_I = 2.67,\ \alpha = 1.33,\ \lambda = 16.0,\ E_0 = 8.0, \text{ and } E_z = 7.2$$

$$R_T = 1.4356,\ \beta = [0.7859 \quad 1.1278],\ \Pr(Success) = 0.6543, \text{ and}$$

$$\nabla_D \Pr(Success) = \varphi(-\beta_{LSL})\nabla_D\beta_{LSL} + \varphi(-\beta_{USL})\nabla_D\beta_{USL} = \begin{bmatrix} 0.0000 \\ -0.0513 \\ 0.0007 \\ -0.0029 \\ 0.0036 \end{bmatrix},$$

where a new design point cannot be found without the constraints being violated. The total number of iterations, N_{OPT}, required to find the optimal solution, using Matlab (1994) optimization package, is 12, which included a total of 238 function evaluations (N_F) of R_T and 226 Hasofer-Lind-Rackwitz-Fiessler algorithm iterations($\sum N_{HLRF}$).

The major difference between model-based experimental design and the probability approach is in solving using discrete values (levels) and continuous values (distributions with mean and variance). In this case, the answers are the same since all the design variables go to their maximum or minimum values. Had the maximum signal-to-noise ratio occurred with variable combinations not on their extremes (i.e., on the boundary of their constraints), say $\alpha = 1.6$, the first iteration of experimental design would have had to choose between α at level 1 ($\alpha = 1.33$) or level 2 ($\alpha = 2$). The transformation method, however, would have found this point since it is not restricted to choosing from discrete values. The use of the Hasofer-Lind-Rackwitz-Fiessler algorithm along with an optimization method allows for selection of values anywhere within the constrained region.

As an extension to this example, the equation for $R_{T\text{-}ON}$ can be developed using modelling methods (Chandrashekar and Savage 1997). The system is described in the following model:

$$\begin{bmatrix} \frac{1}{R_1}+\frac{1}{R_3} & 0 & -\frac{1}{R_3} & -\frac{1}{R_1} & 0 & 1 \\ 0 & \frac{1}{R_4} & 0 & -\frac{1}{R_4} & 1 & 0 \\ -\frac{1}{R_3} & 0 & \frac{1}{R_3}+\frac{1}{R_T} & 0 & 0 & 0 \\ -\frac{1}{R_1} & -\frac{1}{R_4} & 0 & \frac{1}{R_2}+\frac{1}{R_3}+\frac{1}{R_4} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_a \\ V_b \\ V_1 \\ V_2 \\ I_0 \\ I_= \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ E_0 \\ E_= \end{bmatrix} \tag{4.21}$$

The first four rows of the matrix represent the current sums at each node, while the remaining two rows describe the voltage sources applied to the system. Solving the set of equations using MAPLE (1996) and matrix calculus (Graham 1981), we find the voltages into 1 and 2 are:

$$V_1 = \frac{R_T}{R_T + R_3} E_= \tag{4.22}$$

$$V_2 = \frac{R_2(E_= R_4 + E_0 R_1)}{R_1 R_2 + R_2 R_4 + R_1 R_4} \tag{4.23}$$

For this circuit, the relay turns on when $V_1 - V_2 = 0$. Solving for this we find:

$$V_1 - V_2 = \Delta V = \frac{R_T}{R_T + R_3} E_= - \frac{R_2(E_= R_4 + E_0 R_1)}{R_1 R_2 + R_2 R_4 + R_1 R_4} \tag{4.24}$$

and simplifying we get:

$$\Delta V = \frac{E_=(R_1 R_2 R_T + R_1 R_4 R_T - R_2 R_3 R_4) - E_0 R_1 R_2 (R_T + R_3)}{(R_T + R_3)(R_1 R_2 + R_2 R_4 + R_1 R_4)} \tag{4.25}$$

If we set $\Delta V = 0$ (the point at which the relay turns on) and solve for R_T, we find:

$$R_{T\text{-}ON} = \frac{R_3 R_2 (E_= R_4 + E_0 R_1)}{R_1 (E_= R_2 + E_= R_4 - E_0 R_2)}$$

This is the same as equation (4.17). If we further simply this expression using the ratios suggest by Taguchi and Phadke, R_T simplifies to:

$$R_{T-ON} = \frac{R_3\alpha(E_z\lambda + E_0)}{E_z\alpha + E_z\lambda - E_0\alpha} \tag{4.26}$$

It is clear from this simplification why they found R_1 had little effect in the signal-to-noise ratio. It doesn't exist in this relationship. In addition, by solving for optimal R_T, the authors were simply trying to minimize the variation about the difference in voltages, ΔV.

A second method to solve this example is based on finding minimum variance about the point $V_1 - V_2 = 0$. For a fixed (given) R_T, the optimal values for α, λ, E_0, and E_z can be found. This is a better formulation to the problem since the manufacturer may specify the temperature sensor resistance and the remaining components of the system may be easier to modify or replace than the sensor. So given a fixed R_T and substituting $R_2 = \alpha * R_1$ and $R_4 = \lambda * R_1$, (substitutions from Taguchi and Phadke 1984), then V_1-V_2 is:

$$\Delta V = \frac{E_z(\alpha R_T + \lambda R_T - \alpha\lambda R_3) - E_0\alpha(R_T + R_3)}{(R_T + R_3)(\alpha + \alpha\lambda + \lambda)} \tag{4.27}$$

Setting an upper and lower bound of 0.0001 and –0.0001 Volts on the difference, we use the transformation method and optimization to find the optimum design, with R_3 = [0.898...1.102] and R_T = 1.4356, the optimal value found from Taguchi and Phadke's first design iteration. We find:

$$\alpha = 1.33, \lambda = 16.0, E_0 = 8.0, E_z = 7.2$$

after 7 iterations of the optimization routine, which included 215 function evaluations of ΔV, and 208 Hasofer-Lind-Rackwitz-Fiessler algorithm iterations. The probability of success associated with the design settings is Pr(*Success*) = 0.0061. Note that the optimal settings determined are the same as those found by Taguchi and Phadke.

The advantage of the modelling methods lies in the ability to choose any aspect of the system as a quality characteristic to be studied. In this one example, two different formulations, R_T and ΔV, were solved for. The problem could have been expanded so that resistors R_2 and R_4 are not simplified. This step was necessary only for comparison of answers with Taguchi and Phadke. The modified temperature controller circuit by Belavendram (1995) did not make this simplification. By allowing the resistor design variables to be modified independently, the design variables could be as in Table 4-5.

Table 4-5: Independent Resistor Design

Variable	Mean Value	Variation	Constraints
R_1	4	±5%	$2.67 < R_1 < 6.0$
R_2	5.32	±5%	$3.55 < R_2 < 18$
R_4	96	±5%	$14.23 < R_4 < 96$
E_0	8	±5%	$8.0 < E_0 < 12.0$
E_z	7.2	±5%	$4.8 < E_z < 7.2$

The mean starting values for this problem will be the results found from the previous formulation. Using the probability approach to optimize the design of equation (4.25) with constraints on the voltage difference of [–0.1…0.1], we find the optimal point at:

$$R_1 = 4.3306,\ R_2 = 5.5536,\ R_4 = 95.9996,\ E_0 = 8.0079, \text{ and } E_z = 4.8000$$

with Pr(*Success*) = 0.7715 and a nominal difference of 0.00098. Note that the optimal point suggested is different from the optimal point where the resistors where ratios. For comparison, that Pr(*Success*) using the constraints of [–0.1…0.1] is 0.5543. Therefore, by allowing the resistors to vary independently, we can find a design point with a higher probability of success.

To take this example one step further, the above resistors will be replaced with normal distributions with the center missing with t = 0.05 (i.e., the center 10% of the distribution is missing). Figure 4-11 shows a sample PDF of a resistor with the middle 10% missing. Appendix A.1.6 shows the mathematics associated with this distribution.

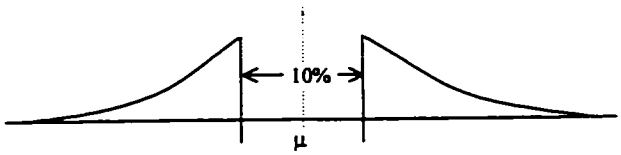


Figure 4-11: Distribution for a Resistor with the Center 10% Missing

Solving the problem again with R_1, R_2, and R_4 missing the center 10% of their distributions, the following solution is found:

$$R_1 = 4.3966,\ R_2 = 5.6149,\ R_4 = 95.9944,\ E_0 = 8.0766, \text{ and } E_z = 4.8000$$

with Pr(*Success*) = 0.7458 and a nominal difference of 0.0016. This was confirmed with a Monte Carlo Simulation that estimated the probability of success to be 0.7550. An interesting result of this approach is that removing the middle 10% of the three resistors only reduced the probability of success by 0.026; not a great amount considering the center 10% of the resistors' distributions were missing. Using the profiling method from Chapter 6, we can see the CDF for each of the results – Figure 4-12 (a) represents the approximated CDF with all resistors having normal distributions, and (b) with R_1, R_2, and R_4 having the normal with center missing. Note the slight leveling of the CDF around the mean in (b).

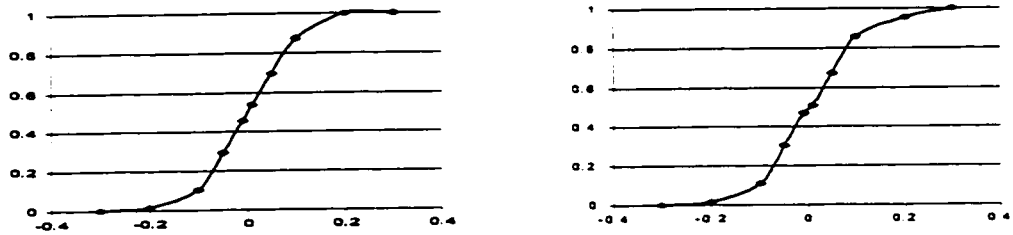


Figure 4-12: (a) CDF with all resistors having normal distributions (b) with all resistors having normal distributions with the center 10% missing

Design of a Thin Film Re-Distribution Layer

The example detailed by Bagchi and Templeton (1994) was a design problem encountered by IBM in Kingston, Ontario. Target circuit impedance of 85 ohms was required in addition to minimization of variability subject to manufacturing variations.

$$Z = 87\ln\left(5.98A/(0.8B+C)\right)/\sqrt{(\epsilon+1.41)} \tag{4.28}$$

The design variables are:

Table 4-6: Design Variables and Associated Information

Variable	Constraints	Variability
A – Insulator Thickness	20 < A < 30 (μm)	± 1 (μm)
B – Conductor Line Width	12.5 < B < 17.5 (μm)	± 0.67 (μm)
C – Line Height	4 < C < 6 (μm)	± 0.33 (μm)

with ϵ, the dielectric constant, equal to 3.094539.

The Bagchi and Templeton method was based on variance transmission through target constraints. In order to find the relationship of the output variance to the variables' variances, a procedure from Taylor (1991) was used to obtain:

$$\begin{aligned}\sigma_Z^2 &= \alpha_1\sigma_A^2 + \left(\alpha_2\sigma_B^2 + \alpha_3\sigma_C^2\right)/(0.8B+C) \\ &\quad + \alpha_4\sigma_A^4/A + \left(\alpha_5\sigma_B^4 + \alpha_6\sigma_C^4\right)/(0.8B+C)^2 \\ &\quad + \alpha_7\sigma_A^2\sigma_B^2/A + \alpha_8\sigma_A^2\sigma_C^2/A + \alpha_9\sigma_B^2\sigma_C^2/(0.8B+C)^2\end{aligned} \tag{4.29}$$

This equation was used to find the minimum variance of the response, Z. In addition, they noted that the constraint Z = 85 ohms simplifies the problem to $A = \text{constant} \times (0.8B+C)$ leaving only one design factor, A. Using this constraint to find B and C, the Z = 85 ohms specification. Their results were $[A \quad B \quad C] = [26.6 \quad 17.5 \quad 6]$.

Using the transformation method, this problem can be modelled two ways. The first is to assume that A, B, and C are normal with the variability representing ± 3σ. The second is to assume that A, B, and C are uniform with the variability representing Δ, the width of the interval. In this example, if the value 26.6 is used as the center point for A, it would have a width of 2 μm and range from 25.6 to 27.6 μm.

Regardless of which assumption made, if we sent an upper and lower bound of 84.999 and 85.001 ohms on equation (4.28) to constrain the function to be approximately 85, the transformation method and optimization converge to the same point as found by Bagchi and Templeton. For the normal assumption, we find Pr(*Success*) = 0.0119 after 7 iterations, 28 function evaluations, and 28 Hasofer-Lind-Rackwitz-Fiessler algorithm

iterations. For the uniform assumption, Pr(*Success*) = 0.0050 with 8 iterations, 32 function evaluations, and 32 Hasofer-Lind-Rackwitz-Fiessler algorithm iterations.

Design of a Hollow Cylinder

The design of a tube under a torsional load based on probability of failure was first discussed by Tribus (1969) and later solved using design of experiments and Taguchi's methods by D'Errico and Zaino (1988). It was a design problem considering the maximum stress caused by a twisting moment on a hollow cylinder (see Figure 4-13).

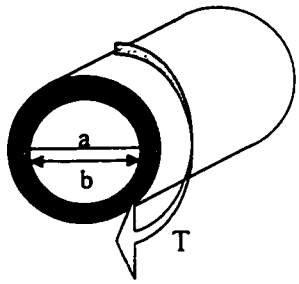


Figure 4-13: A Hollow Cylinder undergoing a twisting moment

Mathematically, the shear stress is defined as:

$$\tau = 16aM/\left(\pi\left(a^4 - b^4\right)\right) \tag{4.30}$$

where the design variables are shown in Table 4-7 and the noise factors in Table 4-8:

Table 4-7: Design Variables

Design Variable	Mean	Standard Deviation
a – outside diameter	2.4	0.02
b – inside diameter	2.0	0.02

Table 4-8: Noise Factors

Noise Factor	Mean	Standard Deviation
M – moment	1200	60
T – strength	900	90

D'Errico and Zaino develop a modification to Taguchi's method so as to approximate normal distributions using only three points in an experimental design. Their results allowed them to evaluate the probability of failure defined by them as:

$$\Pr(Failure) = \Pr(\tau > T) = \Phi\left(\frac{\mu_\tau - \mu_T}{\sqrt{\sigma_\tau^2 + \sigma_T^2}}\right) \tag{4.31}$$

where the expression in the brackets is identical to Cornell's index when the specification has a variance (see Madsen, Krenk, and Lind (1986), Melchers (1987) for further discussion), and is their approximation for β. So by approximating normal distribution through three points for each factor, the authors were able to evaluate the above probability.

To solve this problem with the transformation method, the only requirement is some constraints on the variables a and b. Referring to D'Errico and Zaino's paper, we can use the following constraints to aid on our comparison of answers:

Table 4-9: Constraints on Design Variables

Variable	Constraints
a – outside diameter	$2.376 < a < 2.424$
b – inside diameter	$1.976 < b < 2.024$

In the experimental design, they find the design factor combination that minimize shear stress, τ, given the variability in M and T, occurs at $a = 2.424$ and $b = 1.976$. Using the transformation method, we formulate the margin/limit-state surface as:

$$m = \tau - T = 16aM/\left(\pi\left(a^4 - b^4\right)\right) - T \tag{4.32}$$

Note in this example that the specification, T, is a random variable unlike the other examples where the specification was simply a value. Optimizing for minimum failure, we use the objective function, $\Pr(Failure) = \Phi(-\beta)$, and find the results in Table 4-10.

Table 4-10: Values from Probability Approach

a	2.424
b	1.976
τ	768.42
β	1.2034
$\Pr(Failure)$	0.1144

The design values are the same as D'Errico and Zaino and are found in 4 iterations, 31 function evaluations, and 23 Hasofer-Lind-Rackwitz-Fiessler algorithm iterations. The τ corresponds to the average of three values of τ found in the experimental design when $a = 2.424$ and $b = 1.976$ and $M = [1126.52, 1200, 1273.48]$. Where D'Errico and Zaino took 27 runs to calculate the optimal point, our method took 31. However, their method was only calculating 3 factors at 3 discrete levels while our method was calculating 4 factors, all continuous variables. If any or all of the factors in the example had been of another distribution type, the method presented by D'Errico and Zaino could not be used. However, as long as the probability transformation can be found, our method is capable of handling arbitrary distributions.

If we widen the constraints to $1.8 < a < 3.0$ and $0.5 < b < 2.5$, we can search for a combination of a and b that gives us Pr(*Failure*) ≈ 0. It occurs at $a = 3.0$ and $b = 0.5$, and $\tau = 226.5$ with $\beta = 7.4142$ and Pr(*Failure*) $= 0.6 \times 10^{-13}$, and is found in 58 iterations, 498 function evaluations, and 382 Hasofer-Lind-Rackwitz-Fiessler algorithm iterations. However, to conserve cross-sectional area, we could have specified a failure probability (e.g., $\beta = 4.2649$ or Pr(*Failure*) $= 0.00001$). This gives an answer of $a = 3.0$ and $b = 2.55507$, with $\tau = 477.71$, and a savings of 19.7 in^2 in the cross-sectional area.

Design of a Servo-Control System

Illustrated in Figure 4-14 is a multi-disciplinary system; a servo-control system to be optimized for two load types.

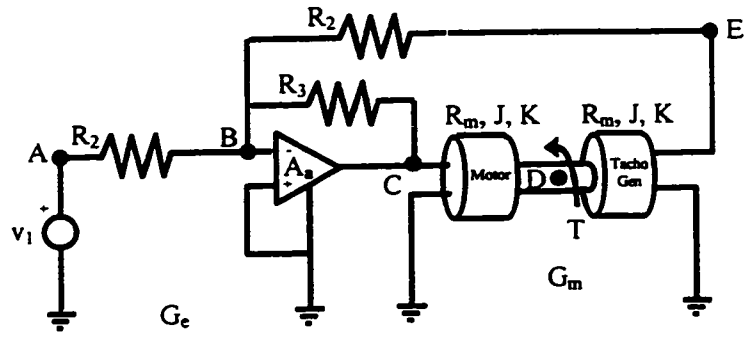


Figure 4-14: Servo-Control System

This servo-control system is comprised of both an electrical and mechanical sub-system. The corresponding system graphs can be seen in Figure 4-15.

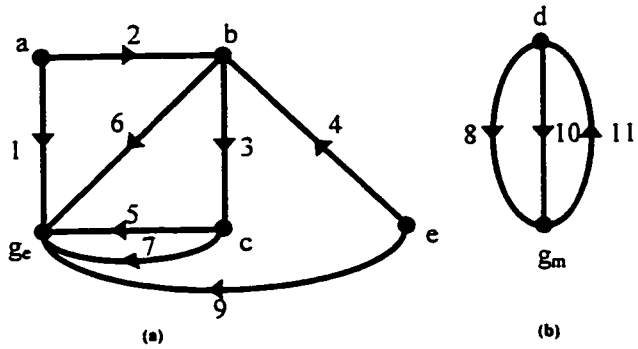


Figure 4-15: System Graph of Servo-Control System for (a) Electrical Sub-system and (b) Mechanical Sub-system

where the electrical representation of the system is depicted in (a) and the mechanical portion of the system in (b). Due to the interdisciplinary nature of the problem, there are two hybrid components in the system. The through variables are $y = [i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_9, T_8, T_{10}, T_{11}]^T$ and the across variables are $x = [v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_9, \omega_8, \omega_{10}, \omega_{11}]^T$, i.e., the currents and torques (through) and voltages and shaft speeds (across). Orientations of the through and across measurements

are represented by the system graphs shown in the figure above, the currents and voltages associated with (a) and the torques and speeds with (b).

Using the Mixed Nodal Formulation of GTM, the following set of equations is found (Xie 1994). They are given in matrix form for ease of understanding and use.

$$\begin{bmatrix} \frac{1}{R_2} & -\frac{1}{R_2} & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -\frac{1}{R_2} & \frac{1}{R_2}+\frac{1}{R_2}+\frac{1}{R_3} & -\frac{1}{R_2} & 0 & -\frac{1}{R_3} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{R_2} & \frac{1}{R_2} & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2J(s) & 0 & 0 & -K & -K & 0 \\ 0 & -\frac{1}{R_3} & 0 & 0 & \frac{1}{R_3} & 0 & 0 & 1 & 0 \\ 0 & A & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -K & 0 & 0 & -R_m & 0 & 0 \\ 0 & 0 & 0 & -K & 1 & 0 & 0 & -R_m & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_a(s) \\ v_b(s) \\ v_c(s) \\ \omega_d(s) \\ v_e(s) \\ i_5(s) \\ i_7(s) \\ i_9(s) \\ i_1(s) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ T_{11}(s) \\ 0 \\ 0 \\ 0 \\ 0 \\ v_1(s) \end{bmatrix} \quad (4.33)$$

The initial values of the system parameters can be found in Tables 4-11, 4-12, and 4-13.

Table 4-11: Design Variables

Design	Mean	$\pm 3\sigma$	Constraints
A	50000	$\pm$ 10%	$25000 \le A \le 75000$
R_2	15000	$\pm$ 10%	$R_2 \ge 10000$
R_3	1650	$\pm$ 10%	$R_3 \ge 1000$
R_m	10	$\pm$ 10%	$8 \le R_m \le 12$
K	1	$\pm$ 5%	$0.8 \le K \le 1.2$

and

Table 4-12: Noise Factor

Parameter	Mean	$\pm 3\sigma$
V_1	10	$\pm$ 10%

and two different load conditions (Signal factor):

Table 4-13: Signal Factor

Load	Distribution
(a) T_{11}	Normal(μ= 1000, σ = 66.667)
(b) T_{11}	Uniform(800...1200)

In addition, J is a constant equal to 0.000001.

The speed of the shaft at D is specified to be, at steady state, between [900...1100] with maximum quality. The equation for the shaft speed at steady state is:

$$\Omega_d = \frac{2R_m^2T + 2R_3R_mT_{11} + R_2R_mT_{11} + AR_m^2T_{11} + AR_3R_mT_{11} + v_1KR_m - v_1AKR_m - v_1AKR_3}{K^2(R_2 + 2R_3 + 4R_m + AR_2 + AR_3 + 2AR_m)} \quad (4.34)$$

Note that it is not dependent on J. Solving the initial system and determining the probability of success, we find for condition (a), a normally distributed load, that Pr(*Success*) = 0.723875. For condition (b), a uniformly distributed load, the Pr(*Success*) = 0.457919.

Using Matlab to find the optimal design point, we find:

Table 4-14: Optimal Design Points for Servo-Control System

For Normal T_{11}	For Uniform T_{11}
Pr(*Success*) = 0.723953	Pr(*Success*) = 0.462586
A = 49999.9999	A = 49999.999
R_2 = 15000.0000	R_2 = 15000.0000
R_3 = 1650.0106	R_3 = 1650.00000
R_m = 12.0000	R_m = 10.0066
K = 1.0966659	K = 1.017336
Ω_d = 993.889	Ω_d = 962.003
N_F = 1283	N_F = 542
N_{OPT} = 142	N_{OPT} = 45
ΣN_{HLRF} = 1141	ΣN_{HLRF} = 497

The corresponding V and U vectors are:

$$V = \begin{bmatrix} A \\ R_2 \\ R_3 \\ R_m \\ K \end{bmatrix} \text{ and } U = \begin{bmatrix} u_A \\ u_{R_2} \\ u_{R_3} \\ u_{R_m} \\ u_K \end{bmatrix}$$

For this system, by looking at the $\nabla_D\beta$ for each limit-state under each load, it is clear that the solution is extremely sensitive to any perturbations in K.

$$\nabla_D\beta_{Normal} = \begin{bmatrix} -0.0000 & 0.0000 \\ -0.0006 & 0.0007 \\ 0.0058 & -0.0060 \\ 0.8953 & -0.9240 \\ -19.1040 & 20.5170 \end{bmatrix}, \quad \nabla_D\beta_{Unform} = \begin{bmatrix} -0.0000 & 0.0000 \\ -0.0004 & 0.0005 \\ 0.0030 & -0.0043 \\ 0.5895 & -0.8002 \\ -11.4919 & 15.8452 \end{bmatrix}$$

In fact, the gradient information shows that K's effect on β is two orders of magnitude greater than R_m and several for the remainder of the design variables.

Had we studied the $\partial g(V)/\partial v_i$ and $\partial g(V)/\partial u_i$ for the design variables at each margin, we would have found:

$$\nabla_V g_{LSL}(V) = \begin{bmatrix} -0.0000 \\ -0.0538 \\ 0.4902 \\ 90.8998 \\ -1797.4049 \end{bmatrix}, \quad \nabla_U g_{LSL}(V) = \begin{bmatrix} -0.0060 \\ -26.9121 \\ 26.9603 \\ 30.2999 \\ -29.9567 \end{bmatrix} \text{ and}$$

$$\nabla_V g_{USL}(V) = \begin{bmatrix} 0.0000 \\ 0.0662 \\ -0.5934 \\ -110.0789 \\ 2206.5466 \end{bmatrix}, \quad \nabla_U g_{USL}(V) = \begin{bmatrix} 0.0072 \\ 33.1066 \\ -32.6349 \\ -36.6930 \\ 36.7758 \end{bmatrix}$$

showing that R_m and K have the greatest effect in both V and U space and thus should be control factors. A is a neutral factor and R_2 and R_3 could be consider noise because they only effect the variance of the system. If R_2 and R_3 were to remain as design variables, their variances should be considered the control factors, not their mean, since adjustment of their mean has little or no effect on the value of the function.

Xie's Variation of the Servo-Control Example

Using the same servo-control system, Xie (1994) formulated the problem to deal with the mechanical time constant, τ. In addition, Xie's set did not distinguish R_2 from R_3. She simply set $R_2 = R_3$ and called it R.

The system of equations can be solved to find the shaft speed, $\Omega_d(s)$, in Laplace (s) domain. The inverse-Laplace transformation of the equation the shaft speed equation in the time (t) domain. It is:

$$\Omega_d(t) = A + Be^{-Ct} \tag{4.35}$$

where the information of interest is contained in C. The mechanical time constant, τ, is equal to the inverse of C. Thus, τ is:

$$\tau = \frac{1}{C} = \frac{2JR_m(AR + AR_m + 2R + 2R_m)}{K^2(2AR + 2AR_m + 3R + 4R_m)} \tag{4.36}$$

This design exercise seeks to find the highest probability of success that the mechanical time constant, τ, lies between 0.0622 and 0.0628. In addition, a target value of 0.0625 is set. Using the design variables as [A, R, J, K, R_m], with constraints on the parameters:

$$10000 \le A \le 50000$$
$$1000 \le R \le 10000$$
$$0.00000010 \le J \le 0.00000200$$
$$0.01 \le K \le 0.05$$
$$1 \le R_m \le 10$$

Using Xie's optimum point found,

$$\left.\begin{aligned} A &= 50000 \\ R &= 1200 \\ J &= 0.00000066 \\ K &= 0.01028 \\ R_m &= 10 \end{aligned}\right\} \text{ with } \tau = 0.062455 \text{ and } \Pr(Success) = 0.0468766$$

as the starting point and [LSL..USL] as [0.0622..0.0628], Matlab optimization finds the optimal design point after 22 iterations as:

$$\left.\begin{aligned} A &= 50000 \\ R &= 1200 \\ J &= 0.000002 \\ K &= 0.017879 \\ R_m &= 10 \end{aligned}\right\} \text{ with } \tau = 0.062565 \text{ and } \Pr(Success) = 0.046888$$

This is intriguing since this is not the answer that Xie received. In fact, the problem is completely insensitive to any change in A (changes in the order of 10^{-6}) or R (changes in the order 10^{-9}) over the entire range of each design variable. Thus, they can be considered neutral and can be chosen at any value suitable to the design (or to cost).

To understand why this value was picked as optimum, we consider the methods used to choose the optimum point. Xie used a weighted function of variances and minimized the function, $S_{wc}(\mu_{wc})$.

$$S_{wc}(\mu_{wc}) = 2S_A^2(\mu_A) + S_R^2(\mu_R) + 5S_J^2(\mu_J) + 4S_K^2(\mu_K) + 3S_{R_m}^2(\mu_{R_m}) \tag{4.37}$$

where $S_D^2(\mu_d)$ represents the partial derivative of function, τ, with respect to D evaluated at the nominal values, μ_d, and where D is the respective design variable.

The inherent problem with this method lies in choosing the weightings for each derivative. In Xie's case, the weighting were based on results obtained by Hajdukiewicz (1993).

In the probability-based design method however, the optimization routine maximizes a function, in this case the probability of success, using supplemental gradient information. This method requires no prior knowledge of the variance importance levels, interactions, or weightings. The changes to the optimum point are found through summation of the normalized gradients at each specification limit, i.e., equation (4.14), the gradient of the probability of success.

In this example, at the two differing points, the probability gradients are:

Table 4-15: Comparison of Xie's and Probability Gradient found by Matlab

Probability Gradient information at Xie's optimum point		*Probability Gradient information at Matlab's optimum point*	
A	0	A	0
R	0	R	0
J	31087.8411	J	4051.2989
K	0.5625	K	1.7126
R_m	0.0021	R_m	0.0008

Note that the design variable J has a large effect at the chosen optimum point. However, the point chosen by Matlab is less sensitive to changes in variable values than the point chosen by Xie. In fact, there is an order of magnitude difference. In the point selected by the Matlab optimization routine, J is taken to its upper limit and according to the probability gradient given above, a more optimum point could be found by relaxing the upper limit and increasing J.

4.2 Multiple Quality Characteristic Optimization

Having established the method for single response design optimization using the probability approach and the transformation method, an extension to multi-response design is necessary to approach real-world applications. The problem becomes:

$$\max \Pr(\textit{Success}) \text{ for all responses simultaneously}$$

Section 4.1 showed that for two limit-state surfaces on a quality characteristic, the optimization objective becomes the probability of success or failure, equation (3.7) or (3.8), for that response. In the case of multiple responses, the probability of success can be found for each response using the associated β's for its respective response. There can only be one or two β's associated with a response, representing the distance from an

upper or lower specification (in the first case), or both (in the second case). So the problem becomes

$$\max\ \beta_i \text{ for all } i = 1,\cdots, L \text{ simultaneously} \tag{4.38}$$

where L is the number of limit-state functions. Given that β_i can be calculated using the transformation method, a method is required that considers all β's simultaneously.

There have been numerous papers and methods for optimizing multiple responses. In a paper by Pignatiello, Jr. (1993), several strategies are detailed for handling systems with multiple responses. Pignatiello discusses:

1. A Loss Function for Multiple Quality Characteristics
2. A Response Model-Based Strategy
3. Priority-Based Approaches
4. Computer Graphics-Based Approaches

Within these approaches, the second and third strategies are applicable to the model-based approach in this thesis. The third strategy, however, relies on choosing a primary response to be optimized with all other responses constrained. This approach, while common (Myers and Carter 1973) (Biles 1975), is not suitable for the methodology within this thesis. It requires assumptions to be made about importance of responses that cannot be made without user input, or a restructuring of the problem.

The second strategy is based on a response surface being obtained and then using a least squares approach to determine the expected loss (squared deviation from target). A similar approach is used by Antreich, Graeb, and Wieser (1994), Low and Director (1989), and Lightner, Trick, and Zug (1987). They use the following weighted function:

$$\min \sum_{i=1}^{L} k_i \cdot \exp(-a \cdot \beta_i) \tag{4.39}$$

where $a = 2$ and $k_i = 1$, giving the least squares formulation:

$$\min \gamma^T \gamma, \text{ with } \gamma = \left[\exp(-\beta_1) \quad \cdots \quad \exp(-\beta_i) \quad \cdots \quad \exp(-\beta_L)\right] \tag{4.40}$$

From this formulation, if all specifications are satisfied, i.e., all β's are positive, then the corresponding component of γ, i.e., $\exp(-\beta_i)$, has a value between 0 (if $\beta_i = \infty$) and 1 (if $\beta_i = 0$). If a specification is violated, i.e., $\beta_i < 0$, the component has a value greater than 1. This formulation handles nominal design, yield optimization, and design centering (Antreich, Graeb, and Wieser 1994). Their paper compares (4.40) to these three methods and shows:

1. It ensures that a feasible solution is found if present (Nominal Design).
2. The probability of success for each quality characteristic is maximized. The method is a form of Yield Maximization (see Section 2.1.1, equation (2.3)).

3. The design is centered in the acceptable region. This is identical to Design Centering (see Section 2.1.1).

Figure 4-16 shows a system in U-space with two responses to be optimized. The above method seeks to move the nominal point in U-space to the geometrical center of the constraint region (indicated by the arrow). Doing this satisfies the above three points.

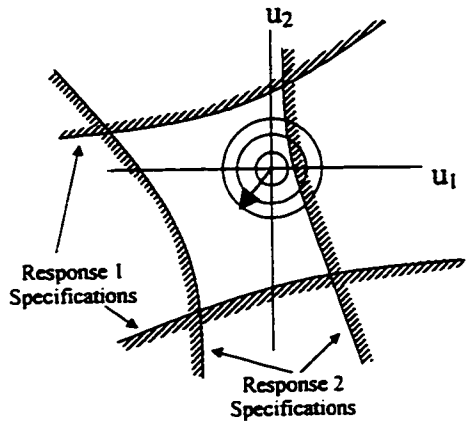


Figure 4-16: Nominal Design, Yield Maximization, and Design Centering

Since this method has been proven to work for this class of problems, it will be used in this thesis as a means to optimize multiple response problems. This does not imply that other methods cannot be used or explored. However, in using this optimization approach to multiple responses, gradient information of the objective function with respect to the design variables is needed. Taking the derivative of equation (4.40) with respect to the design variables, it yields:

$$\nabla\left[\gamma^T\gamma\right] = \sum_{i=1}^{L} -2\nabla_D\beta_i \cdot \exp(-2\beta_i) \tag{4.41}$$

In order to illustrate the multiple quality characteristic optimization, three examples will be shown. The first and the last problem attempt to maximize the probability of success of two quality characteristics, while the second problem is formulated differently. The problem is to minimize cost of the system while meeting or exceeding a specified level of quality for each response.

4.2.1 Examples

The first is a detailed example of a simple non-linear two-pipe, one source, two-demand problem and the second example is a complex non-linear eight-pipe, one source, six-demand problem from Lansey, Duan, Mays, and Tung (1989), and the third is the optimization of a teacup design suggested by Savage.

Non-Linear Two-Pipe Problem

The following figure depicts a non-linear two-pipe system (Carr and Savage 1996):

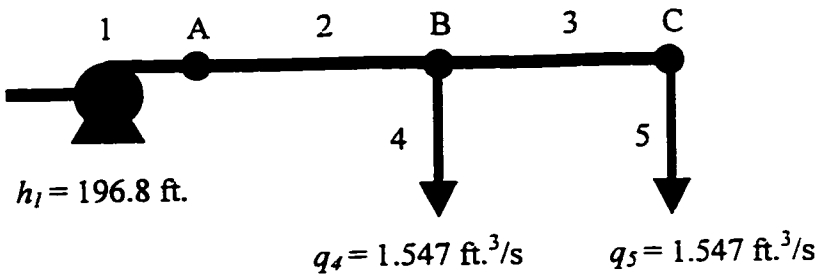


Figure 4-17: Simple Two-Pipe, One Source, Two-Demand System

The system is characterized by a source and two demands, all random normally distributed. The specifications for the system are:

$$90 \text{ psi} < h_b < 130 \text{ psi}$$

$$90 \text{ psi} < h_c < 130 \text{ psi}$$

where the heads at each of the two nodes are constrained to lay between a lower specification limit, e.g., minimum pressure in the system is required to ensure supply, and an upper specification limit, e.g., safety design limit for the pipe so that it doesn't burst. The purpose of this example is to choose the diameters of the pipes such that the probability of being within the specifications at each node (i.e., both h_b and h_c lie between 90–130 psi) is maximized.

For this problem, the following information is given. The non-linear constitutive equation for the pipes is the Hazen-Williams equation:

$$q - \frac{Cd^{2.63}}{(kl)^{0.54}} h^{0.54} = 0 \tag{4.42}$$

where q and h represent the flow and head in the pipe, C is the pipe roughness, d is the diameter of the pipe in inches, and the variables k and l are constants representing a units conversion constant and the pipe length in feet, 852000 and 3280 respectively.

The design and noise factors are given in the following tables:

Table 4-16: Noise Factors

Factors	Mean	$\pm 3\sigma$
C_2 – Coefficient of Roughness	100	± 20%
C_3 – Coefficient of Roughness	100	± 20%
h_1 – Source Head	196.8	± 20%
q_4 – Demand at B	1.547	± 20%
q_5 – Demand at C	1.547	± 20%

Table 4-17: Deterministic Design Factors

Factors	Initial	Constraints
d_2	9.5	$8 \le d_2 \le 20$
d_3	11.7	$8 \le d_3 \le 20$

Solving the system using the optimal points found by Carr and Savage through minimization of cost, we find:

$$\begin{bmatrix} h_a \\ h_b \\ h_c \end{bmatrix} = \begin{bmatrix} 196.8000 \\ 119.3352 \\ 111.5538 \end{bmatrix}$$

and

$$\Pr(Success) \text{ at } h_b = 0.6746$$

$$\Pr(Success) \text{ at } h_c = 0.7323$$

Using (4.40) as the objective function and (4.41) as the gradient of the objective function, the initial evaluation in Matlab gives:

$$\gamma^T\gamma = 0.5523 \text{ and } \nabla[\gamma^T\gamma] = \begin{bmatrix} 1.1262 \\ 0.0090 \end{bmatrix}$$

The gradient information indicates that we need to decrease both design variables to find the next design point. Note that d_2 is to be decreased more than d_3. The next design point is:

Table 4-18: Second Design Point

Factors	Value
d_2	9.2315
d_3	11.6910

giving

$$\gamma^T\gamma = 0.6455 \text{ and } \nabla[\gamma^T\gamma] = \begin{bmatrix} -2.0746 \\ -0.1087 \end{bmatrix}$$

indicating that the design variables should be increased. After 20 iterations, 1501 function evaluations (through Newton-Raphson), and 334 HLRF algorithm iterations, the optimal design point is found:

Table 4-19: Optimal Design Point

Factors	Optimal
d_2	9.3206
d_3	16.6904

with

$$\begin{bmatrix} h_a \\ h_b \\ h_c \end{bmatrix} = \begin{bmatrix} 196.8000 \\ 111.7986 \\ 110.4193 \end{bmatrix}$$

$$\Pr(Success) \text{ at } h_b = 0.7190$$

$$\Pr(Success) \text{ at } h_c = 0.7208$$

$$\gamma^T\gamma = 0.4621 \text{ and } \nabla\left[\gamma^T\gamma\right] = \begin{bmatrix} 0.00000009 \\ 0.00000001 \end{bmatrix}$$

Note that the objective function used, equation (4.40), approximately centered the head at B and C in the specification [90..130] and approximately equalized the probabilities of success.

Lansey, Duan, Mays, and Tung Water Distribution Example

The following Figure 4-18 depicts an eight-pipe system from Lansey, Duan, Mays, and Tung (1989). The system is characterized by one probabilistic source and six probabilistic demands. In addition, the equations for the pipes (4.42) are non-linear with probabilistic elements.

The design of the water distribution system concerns the ability of the system to supply the demands at the nodes within the system at required minimum pressures. Lansey et al. (1989) wish to design the system for minimum cost for a required probability of success. The required minimum pressure at each node is 100 ft.

The paper by Lansey et al. (1989) defines a methodology for finding minimum cost given constraints. The constraints are related to the probability of success at each node (i.e., $\Pr(h_i \geq 100 \text{ ft}) \geq 0.99$). The cost function used was:

$$COST = 0.331 LD^{1.51} \tag{4.43}$$

and the method was chance-constrained (Charnes and Cooper 1963), (Charnes and Sterdy 1966). Starting with their optimal solution for minimum Pr(*Success*) > 0.99, the noise and design factors are found in Tables 4-21 and 4-22:

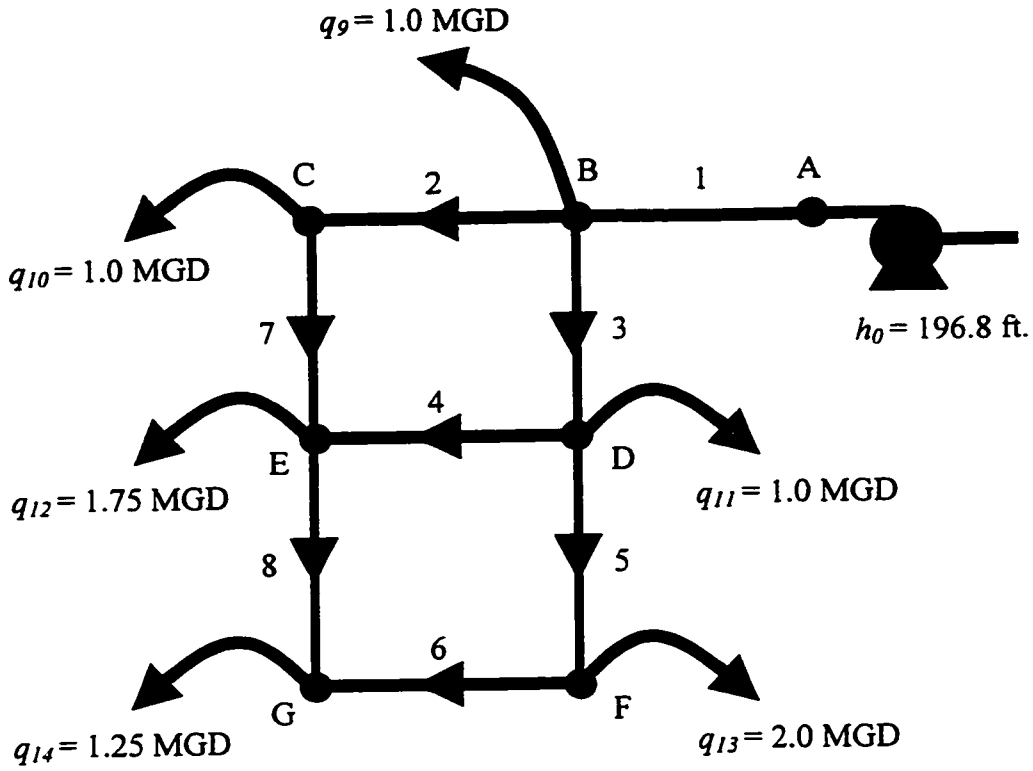


Figure 4-18: Eight-Pipe, One Source, Six-Demand System

Table 4-20: Noise Factors

Factors	Mean	σ
C_1 – Coefficient of Roughness	100	10
C_2 – Coefficient of Roughness	100	10
C_3 – Coefficient of Roughness	100	10
C_4 – Coefficient of Roughness	100	10
C_5 – Coefficient of Roughness	100	10
C_6 – Coefficient of Roughness	100	10
C_7 – Coefficient of Roughness	100	10
C_8 – Coefficient of Roughness	100	10
h_0 – Source Head	196.8	10
q_9 – Demand at B	1.00	0.25
q_{10} – Demand at C	1.00	0.25
q_{11} – Demand at D	1.00	0.25
q_{12} – Demand at E	1.75	0.25
q_{13} – Demand at F	2.00	0.25
q_{14} – Demand at G	1.25	0.25

Table 4-21: Lansey et al.'s Optimal Design Point

Control	Optimal
d_1	23.0
d_2	8.4
d_3	20.1
d_4	10.9
d_5	15.6
d_6	11.9
d_7	0.0
d_8	0.0

and again, the variables k and l are constants of values 852000 and 3280 respectively. Using the probability approach and the transformation method, we can optimize for maximum probability of success for a given cost. Setting the cost constraint to the minimum cost found in Lansey et al., after 10 iterations, 1755 function evaluations, and 289 HLRF algorithm iterations, we find the optimal design point in Table 4-22.

The design points give $\gamma^T\gamma = 0.00001295$, compared to $\gamma^T\gamma = 0.00043183$ from Lansey et al. In addition, the solution is very close to that found by Lansey et al. Note that pipes 7 and 8 are removed from the system in the optimal design. Since the problem is focused on both probability and cost, pipes 7 and 8 are redundant. According to the gradient information (Section 4.1.2), they are very close to being neutral factors, Therefore, they are the first to be reduced in size for cost reasons.

Table 4-22: Optimal Design Point found by Swan's method

Control	Optimal
d_1	22.9
d_2	8.9
d_3	20.0
d_4	10.8
d_5	15.8
d_6	11.8
d_7	0.0
d_8	0.0

Optimizing the Design of a Teacup:

This example comes from recent work done by Chandrashekar and Savage (1997) and relates the lumped parameters we normally design with to the physical components of the system. The system under study is a teacup, shown in Figure 4-19. However, there is one distinct difference between this teacup and a regular teacup. The jacket of this teacup has a water-core to help absorb the excess heat from the tea and retain it to keep the tea warmer longer. In the analysis by Chandrashekar and Savage, only nominal design was done. To further improve his design, we can use the probabilistic design methods from this thesis. By assigning "quality" specifications to the response, e.g.,

target time to cool and minimum hold time, optimization can be performed using these specifications as constraints, and the probability that they are met as the overall function to be maximized.

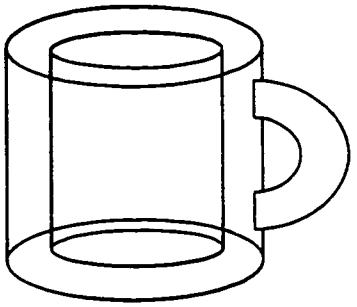


Figure 4-19: A New Teacup

So the problem can be formulated as follows: The variables of the system are found in Table 4-23.

Table 4-23: Teacup Variables

Variable	Type
z, Height of Cup (m)	Design
r, Radius of Tea (m)	Design
dS, Thickness of Separator (m)	Design
dJ, Thickness of Jacket Cavity (m)	Design
dI, Thickness of Insulation (m)	Design
kS, Conductivity of Separator (W/m-°C)	Noise
kI, Conductivity of Insulation (W/m-°C)	Noise
hT, Convection Coefficient of Tea Interface (W/m-°C)	Noise
hW, Convection Coefficient of Water Interface (W/m-°C)	Noise
den, Density of Water (kg/m^3)	Constant
cW, Specific Heat of Water (J/kg-°C)	Constant

Ranges are given for each of the design variables and nominal values for the noise variables. Both variable types will be assumed to be normally distributed with $\pm 3\sigma = \pm$ 5% of the nominal value.

Table 4-24: Teacup Design Variable Ranges

Design Variables	Design Range
z, Height of Cup (m)	0.05–0.15
r, Radius of Tea (m)	0.015–0.035
dS, Thickness of Separator (m)	0.0005–0.0020
dJ, Thickness of Jacket Cavity (m)	0.005-0.020
dI, Thickness of Insulation (m)	0.005-0.020

Table 4-25: Teacup Noise Variables and their Variation

Noise Variables	Value
kS, Conductivity of Separator (W/m-°C)	0.70 ± 5%
kI, Conductivity of Insulation (W/m-°C)	0.05 ± 5%
hT, Convection Coefficient of Tea Interface (W/m-°C)	39.5 ± 5%
hW, Convection Coefficient of Water Interface (W/m-°C)	36.9 ± 5%

Table 4-26: Teacup Constants

Constants	Value
den, Density of Water (kg/m^3)	4200
cW, Specific Heat of Water (J/kg-°C)	1000

From the work done by Chandrashekar and Savage, we can find the symbolic time/temperature relationship of the tea as a function of the design and noise variables:

$$Temperature = f(z, r, dS, dJ, dI, kS, kI, hT, hW, den, cW, t)$$

where t is time in seconds. The function is of the form:

$$Temperature = T_0 + k_1 e^{\frac{-t}{\tau_1}} + k_2 e^{\frac{-t}{\tau_2}}$$

where k_1 and k_2 are coefficients composed of the different variables and τ_1 and τ_2 are the time constants, which are be found from the characteristic polynomial. They are:

$$\tau_{1,2} = \frac{-2}{(a_{11} + a_{22}) \pm \sqrt{a_{22}^2 - 2a_{11}a_{22} + a_{11}^2 + 4a_{12}a_{21}}}$$

Chandrashekar and Savage (1997) give further details on the development of this model. Also specified were constraints on the temperature/time relationship. They are:

(a) temp = 70°C (± 2°C) at target time of 600 seconds (10 minutes). This is the time to steep and cool.

(b) temp ≥ 45°C at time greater than 3000 seconds (50 minutes). This allows for 40 minutes of drinking time.

These constraints can be formulated into Figure 4-20. The temperature response must fall between the specifications at t = 10 minutes and above the specification at t = 50 minutes. Since we know that the temperature of the tea will only decline, we can optimize the design by checking the temperature at t = 10 and t = 50 only. If the response is within specification at these two points, then the design will be accepted.

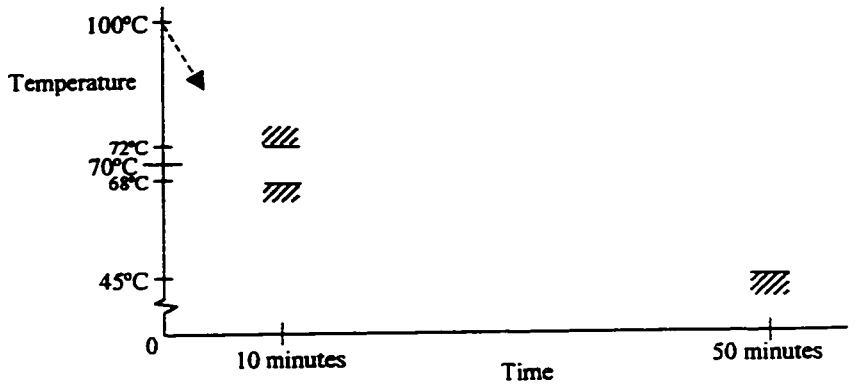


Figure 4-20: Specification on Teacup Design Properties

Using Maple to generate the derivatives of the temperature/time function with respect to the design variables and noise variables, the gradients required for the Matlab routines are found. The result is the following design:

Table 4-27: Teacup Optimal Design Point

Design Variables	Value
z, Height of Teacup (m)	0.0843
r, Radius of Teacup (m)	0.0150
dS, Thickness of Separator (m)	0.0005
dJ, Thickness of Jacket Cavity (m)	0.0050
dI, Thickness of Insulation (m)	0.0200

with the following temperatures:

Table 4-28: Teacup Design Temperatures and Probabilities

Time	Temperature	Probability of Success
10 minutes	72.94 °C	0.51%
50 minutes	44.42 °C	7.08 %

Note that the probabilities of success are very small, even with the temperatures very close to the specifications. This is because the response function variance is very small. By looking at the above results, we can deduce that at the first specification 49.49% of the response lies between 72–72.94 °C, a one-degree spread. At the second specification almost the same amount of the distribution, 42.92%, now lies between 44.42–45 °C, 0.6 of a degree. The variance of the response is reducing as time increases.

However, the specifications set by Chandrashekar and Savage have not been met. Looking at the design suggested by Matlab, it is clear that four of the five variables are at the limits of the ranges imposed in the design. In order to get a better design, either the design ranges need to be modified, or the conductivity properties could be changed by using different materials.

4.3 Modelling Complex Systems

Often today, the system design phase now provides a quantitative description of the system and its responses. Through the use of Graph-Theoretic Modelling (Chandrashekar, Roe, and Savage 1992), Bond Graph Theory (Paynter 1968), System Dynamics (Rowell and Wormley 1997), Response Surfaces (Myers, Khuri, and Vining 1992), a mathematical relationship can be found that relates all inputs, components, and outputs. This 'model' can then be used in the design phase.

Full system models can be developed using Graph-Theoretic Modelling (GTM). The advantage of GTM is its ability to model at levels A to D in Taguchi's description of system design (Taguchi 1993, 1992) detailed in Table 4-29. Taguchi stated that the system design development of new products or processes could be subdivided into the following levels:

Table 4-29: Taguchi's System Design Development Steps

Level	Description
A	System Design
B	Subsystem Design
C	Element or Component Design
D	Development of Elements or Components
E	Development of Raw Materials

Level D, Development of Elements or Components, and Level C, Element or Component Design, can all be achieved through recent advances in graph-theoretic modelling (Gupta 1994) (Gupta and Chandrashekar 1995) (Carr and Savage 1996). Level B, Subsystem Design, can be aided through work done by Savage (1997) and Savage and Row (1992) on subsystem modelling, and Level A, System Design, encompasses all the techniques used in the previously mentioned references.

In addition, higher-order derivative information can be found with little computational effort by using Graph-Theoretic Modelling (Wills 1969) (Savage 1993). With the work in electrical engineering by Wojciechowski, Vlach, and Opalski (1997) and the transformations detailed in Appendix A, non-symmetrical statistical distribution information for components can be used in addition to typical normal approximations to system and component parameters.

Chapter 5

Proofs and Limitations

Having detailed the probability approach with the Hasofer-Lind-Rackwitz-Fiessler algorithm, and our extensions into multiple responses, this chapter will show both the proofs behind the methods and some limitations. There are four aspects that need to be addressed.

All four aspects deal with the Most Likely Failure Point (MLFP). The first point addresses inconsistencies between β and the probability of failure. The second describes problems when using β as a means of comparing systems. The third details what can occur if there are multiple minima on a limit-state surface, and the fourth explains that there is a potential for overlap of failure regions when multiple margins exist on one response.

5.1 The Most Likely Failure Point (MLFP)

Described in Chapter 3 is an algorithm from Hasofer-Lind-Rackwitz-Fiessler (HLRF) for finding the most likely failure point for a response and its corresponding specification. By definition, the Most Likely Failure Point is the point on the limit-state / failure surface, $g(V)$, that has the highest probability of occurring. Madsen, Krenk, and Lind (1986) and Melchers (1987) show that if we replace the failure surface with a hyperplane that is tangent to the failure surface at this point, then the result is an approximation of the true failure probability.

Figure 5-1 illustrates the approximate failure calculation (Madsen at al. 1986).

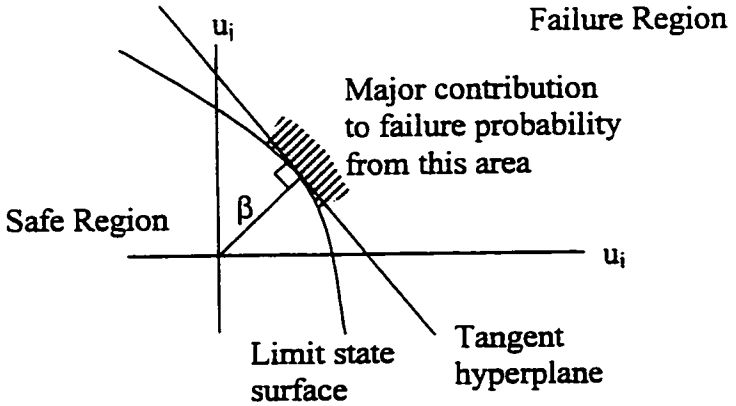


Figure 5-1: Approximation of Failure Probability

Again, according to Madsen, Krenk, and Lind (1986), β can provide a good approximation to the failure probability if the point on the limit-state (failure) surface closest to the origin is the only stationary point of the function and if the principal curvatures of the limit-state surface at this point are not too large in magnitude (Madsen, Krenk, and Lind 1986) (Melchers 1987). Mathematically, in the U-space / standard normal space, the probability density function decreases very quickly, namely as $\exp(-r^2/2)$, with the distance r from the origin. Using the rotational symmetry of the standard normal space, the distance r becomes β, and the probability of failure at the MLFP is approximated by $\Phi(-\beta)$. In addition, if the stationary point is the closest failure point to the origin, then this point represents the point of maximum likelihood of failure.

Inconsistency between β and the Probability of Failure

Since the HLRF algorithm is based on a first-order approximation of the failure surface, when the failure surface is non-linear, the probability is estimated at the most likely failure point. The following diagram depicts how different forms of the margin or limit-state function can have an effect on β and the estimate of the failure probability.

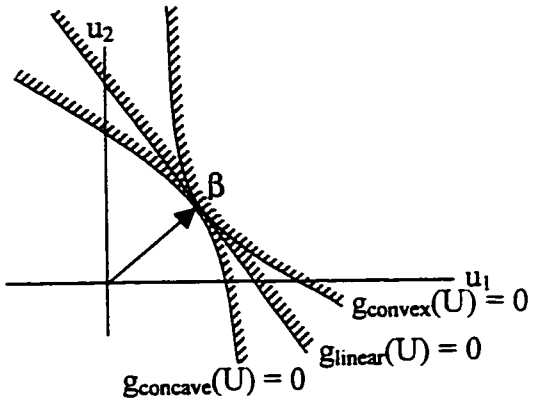


Figure 5-2: Inconsistency between β and the Probability of Failure

From the diagram, it can be seen that all three limit-state functions have the same β value. However, depending on the curvature of the function, β will either under or over-estimate the true probability of failure. If the limit-state function is $g_{linear}(U)$, then the tangent hyperplane is also $g_{linear}(U)$ and the estimated probability of failure is equal to the actual probability of failure. However, if the limit-state surface is $g_{convex}(U)$, then the tangent hyperplane will over-estimate the failure probability and for $g_{concave}(U)$, the hyperplane will under-estimate the failure probability.

Again, however, as shown in Melchers (1987) and Madsen et al. (1986), the Hasofer-Lind β is a good approximation as long as the principal curvatures of the limit-state functions are not too large. For a function with one stationary point, as the estimated distance to the most-likely failure point approaches infinity, the error between the estimate and the true failure probability approaches zero. However, in the field of design, when $\beta > 6$, the Pr(Failure) is less than 0.000000001 or 1-in-a-billion, and is often assumed to satisfactory (Harry and Lawson 1992).

Comparison of Systems Using β

In the field of structures, the original purpose of β was to allow designers to compare the probabilities of failures of different systems under selection. From this, it is therefore necessary that the ordering of systems according to their β's is both consistent and at the same time not too crude (Madsen, Krenk, and Lind 1986). According to Madsen, Krenk, and Lind (1986), the Hasofer-Lind β, (as defined and used in this thesis), generally provides a satisfactory ordering, but examples can be constructed in which ordering is not reasonable (see Figure 5-3).

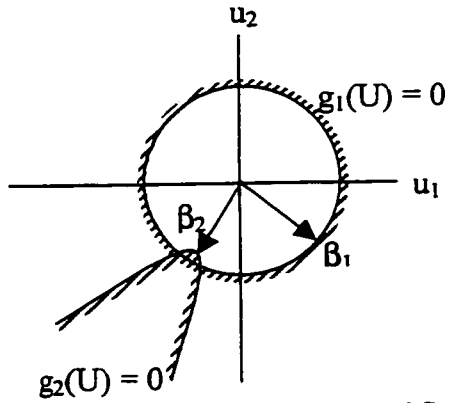


Figure 5-3: Improper Ordering of Systems

From the diagram, β_1 is greater than β_2, but it is unreasonable to state that structure 1 is safer than structure 2. This shortcoming is due to the first-order approximation Hasofer-Lind's β makes at the failure point. By using a higher-order approximation, this ordering problem would become less prevalent.

However, when using β to compare one response with different parameters, this improper ordering will not occur. Since the limit-state function is a mathematical combination of the system response and its associated specification, the values at which the function will equal zero will not change. For example, in a voltage divider, the voltage between the two resistors is equal to:

$$V_b = \frac{V_{in} R_3}{R_2 + R_3} \tag{5.1}$$

If V_{in} = 5 Volts, and we have a specification such that the circuit cannot exceed 3 Volts, then the limit-state function can be written as:

$$g(\mathrm{V}) = \text{Upper Specification} - \text{System Response} = 3 - \frac{5R_3}{R_2 + R_3} = 0 \tag{5.2}$$

Simplifying, we can see that the surface defined by $g(V) = 0$ is:

$$R_2 = \tfrac{2}{3} R_3 \tag{5.3}$$

In this case it is a line, but the proof is still applicable to non-linear surfaces. Since the mathematical structure of the limit-state function is fixed, then variations in the parameter values cannot change the shape of the function, only the distance between the nominal values and the MLFP. This distance, β, changes under two conditions; either a change in any mean value of any or all parameters (e.g., R_2 and/or R_3), or a change in the variance associated with any or all of the parameters. Since β is calculated as the distance from the nominals to the MLFP, then ordering will never be a problem. Figure 5-4 shows three different nominal design points and their corresponding β's.

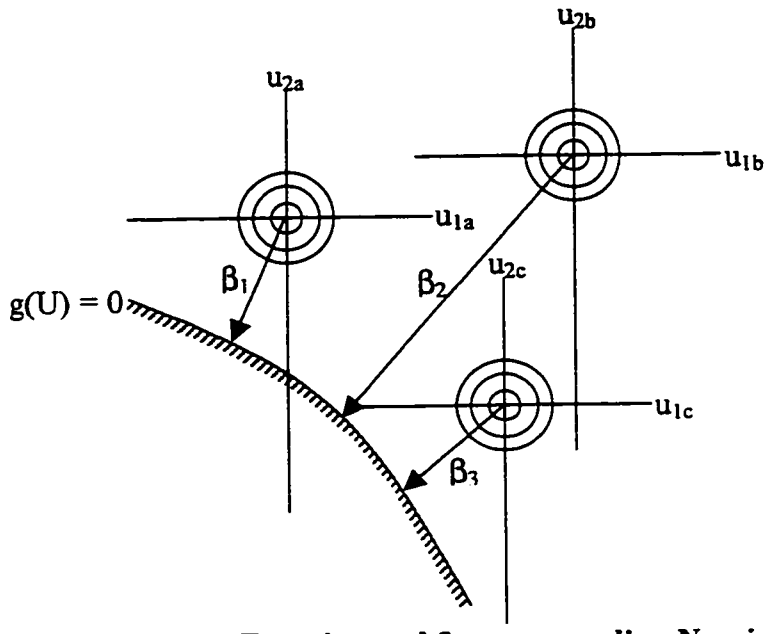


Figure 5-4: A Limit-State Function and 3 corresponding Nominal Designs

The set of parameters u_{1a} and u_{2a} result in β_1. The other two sets could result from, for example, a change in the nominal values of u_1 and/or u_2, or a change in the variance of u_1 and/or u_2, or a combination of both. In the above example, the ordering of the designs would be b(β_2), a(β_1), and c(β_3).

Multiple Minima

There are two cases associated with multiple minima. The first case results when a limit-state function has multiple parts, and the second occurs when the limit-state function has principal curvatures that are large in magnitude. For both cases, the condition that the MLFP is the only stationary point of the function is violated. Thus, there is a concern with where the HLRF algorithm starts. If the same initial condition is used for the HLRF algorithm on a response with multiple minima in two consecutive analysis, only one minima will be found. By randomly picking an initial condition in U-space, there is a greater potential for finding a different failure point.

Limit-States with Multiple Parts

Limit-states can have multiple minima if the limit-state function has multiple roots, thereby creating multiple parts to the failure surface. In the first example used in Chapter 3, the problem was posed such that we wished to find the probability of success associated with the function being greater than 1. In this example, there are two failure points. The first occurs at $d = 1$ and the second at $d = -1$. This can be depicted as:

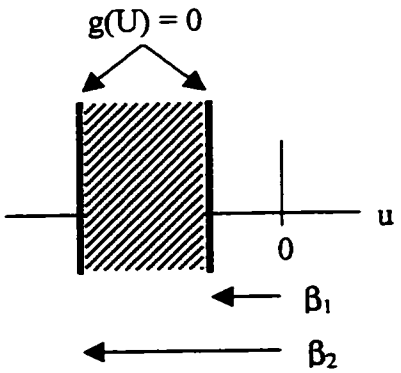


Figure 5-5: Chapter 3 Example In U-space

We find $\beta_1 = 1$ and $\beta_2 = 3$, and the hatched area represents the region of points that cause the function to be less than 1. So, for this example, if we started our search with $u < -3$, then we would find β to be 3 and the Pr(Failure) = 0.0013, an extremely inaccurate estimate of the failure probability. However, if we assume that all systems are designed such that the nominal values of the parameters are the starting conditions of the system, for this example, our algorithm searches and finds β to be 1 and the Pr(Failure) = 0.1587, a more accurate estimate of the failure probability.

Caution is warranted when using the Hasofer-Lind-Rackwitz-Fiessler algorithm for limit-states that may have multiple roots. By having multiple starting points and comparing the resulting β's, the existence of multiple minima may be determined and in so doing, can be handled. Determination of multiple minima is, however, not guaranteed.

Limit-States with Large Principal Curvatures

The second case involving multiple minima occurs with limit-states that have principal curvatures that are large in magnitude. For the example in Figure 5-6, depending on the starting point for U, there is the potential to pick β_{local} instead of β_{global}. In addition, by choosing either β, we are still greatly under-estimating the probability of failure with a simple first-order approximation to the failure surface.

Like the first case above, the potential for this error is also rare since this failure surface type does not occur very often. However, we cannot determine how non-linear the limit-state function is by simply studying the parameters. It can only be determined by directly solving for all values on the failure surface. Since this would require an extensive amount of work for even the smallest problem, it is a prohibitive course of action. However, fortunately it can be assumed that, like the examples in this thesis and

those in Madsen, Krenk, and Lind (1986), the limit-state function is fairly well approximated by the first-order method due to the rarity of this type of failure surface.

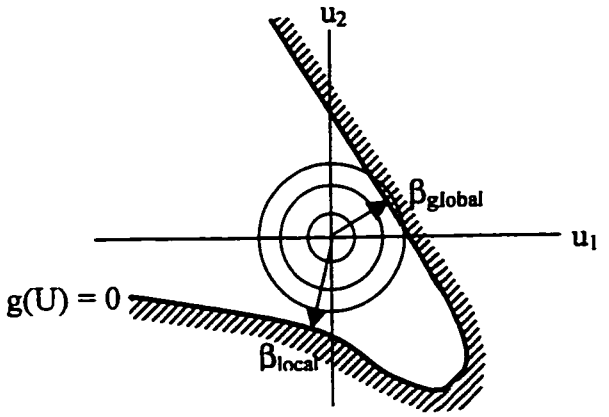


Figure 5-6: Multiple Minima

To help determine if multiple minima exist, multiple starting points are required. If multiple minima are found, then an approximation by a second-order method or even Monte Carlo can be done to better estimate the failure probability. However, in design the purpose of the HLRF algorithm is to find the gradient information of the margin with respect to the design variables so that we can move away from the failure surface. This fast integration method simply gives us an approximation of the probability of failure and a direction in which to move to decrease this probability.

Multiple minima, if they exist, can be found by the HLRF algorithm since it finds stationary points. For each of these points, the respective β must be checked in order to ensure that the smallest β is found (Dolinsky 1983). Melchers (1987) further elaborates on the problem by stating that this difficulty does not arise if a numerical algorithm for seeking the point of maximum likelihood is used instead of the iteration algorithm that identifies only stationary points. In section 4.3.4 of his book, he outlines such an approach. However, as a final thought, he also states that in most practical situations this difficulty will not arise.

To determine how non-linear the limit-state function is, and where the resulting β's lie on the function, the methods from the next section on Multiple Margins on One Response can be used. Within that section, the tools to help this problem will be detailed and referenced back to this section.

Multiple Margins on One Response

In the field of quality, it is very common to have both an upper and lower specification on a response. Since the Hasofer-Lind-Rackwitz-Fiessler algorithm only handles one specification within its margin function, care must be taken when we simply combine the probabilities from two separate HLRF results. There exists two problems: Under-estimating and Over-estimating the probability of success or failure.

The first problem, under-estimating the probability of success, is caused by the overlap of failure regions. The second problem is more serious and requires attention.

Both problems occur due to the non-linearity of the limit-state functions. Fortunately, the answer to determining if either problem exists lies in one approach, but the solution to correct for this under- and over-estimation still needs to be researched and implemented.

Determination If Overlap Exists

Overlap may occur such that, for example, the result is similar to Figure 5-7. In this example, the problem of over-estimating the probability of failure can occur when multiple margins exist on a single response. However, the error is typically negligible since the regions of overlap occur far away from the origin, and the area represented by the overlap represents a small probability when compared to the estimated probabilities at the two failure points.

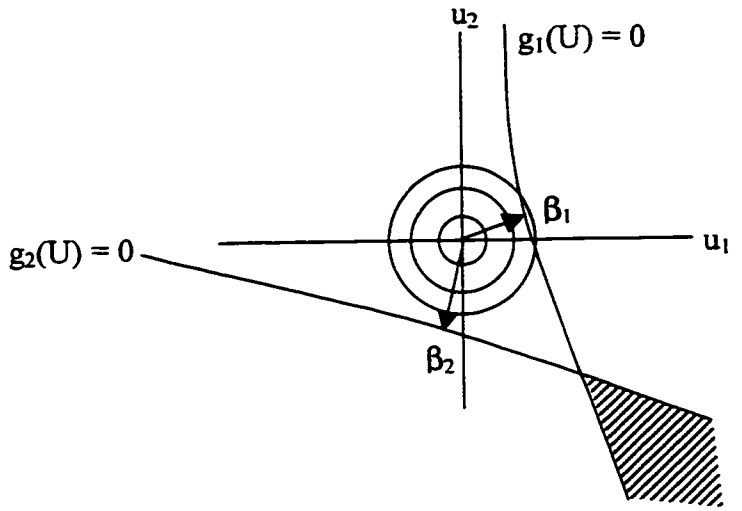


Figure 5-7: Intersection of failure probabilities

To help determine if there may be an overlap, we determine the vector α for each β. The unit vector α is parallel to the gradient of the vector of the trajectory at the most likely failure point, and is directed toward the failure set. It is defined as:

$$\alpha = -\frac{\nabla_U g(V)}{|\nabla_U g(V)|} \tag{5.4}$$

and can be shown to satisfy:

$$U^* = \beta\alpha \tag{5.5}$$

a measure of the sensitivity of β to inaccuracies in the value of u_i at the point of minimum distance U^*. If the α vectors are negative scalars of each other (i.e., $\alpha_1 = -k\alpha_2$), then this implies that the tangent hyperplanes associated with the most likely failure points and the β's must be parallel. An example may look like this:

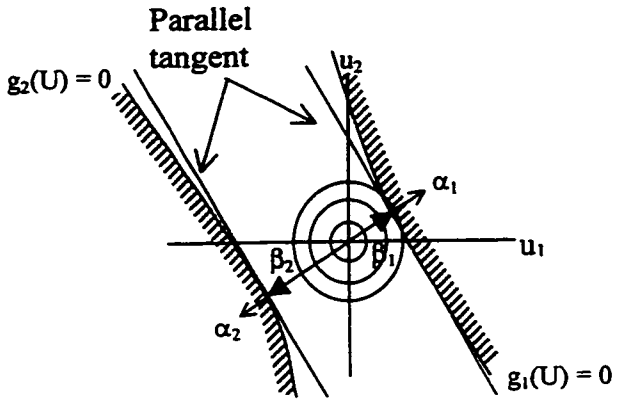


Figure 5-8: Two β's with α's that are negative scalars of each other

but tells nothing of the potential overlap of failure areas. The limit-state functions for an identical set of β's and α's could have been like Figure 5-9.

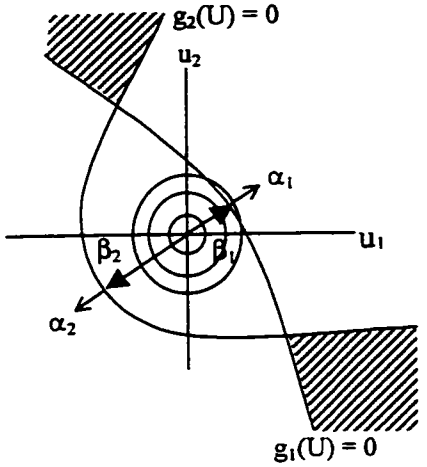


Figure 5-9: Another Two β's with α's that are negative scalars of each other

For example, depicted in Figure 5-10 is response with a normal distribution and one wishes to determine if there is any overlap in failure regions.

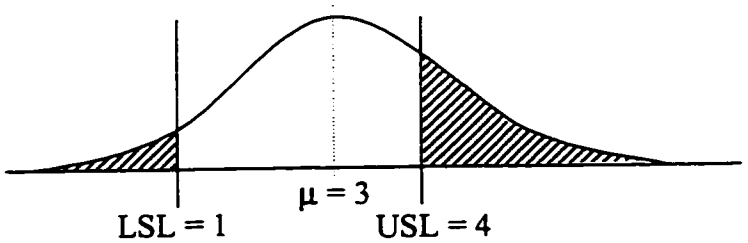


Figure 5-10: A Normal (μ = 3, σ = 1) Distribution

From the diagram, it is obvious that the failure regions do not overlap since they are on opposite ends of the parameter space. The α's are:

$$\alpha_{LSL} = \frac{-1 \cdot 1 \cdot 1}{|1 \cdot 1 \cdot 1|} = -1 \text{ and } \alpha_{USL} = \frac{-(-1) \cdot 4 \cdot 1}{|1 \cdot 4 \cdot 1|} = 1 \qquad (5.6) \text{ and } (5.7)$$

The first α corresponds to the margin, $g_{LSL}(V) = v - 1$, and the second α to $g_{USL}(V) = 4 - v$, where v is the value of the normal distribution. Since the α's are negative scalars of each other, then the associated hyperplanes of the most likely failure points (at 1 and 4) must be parallel to each other. With this knowledge, it can be assumed that there is likely little overlap, if any, of the failure regions.

Algorithm to Aid in Determining Potential Overlap of Failure Regions

So, in summary, the following approach should be taken to help determine if there is the potential for overlap of failure regions.

1. Compute β and α for each margin (upper and lower specification limits)
2. Determine if $\alpha_{LSL} = -k\alpha_{USL}$, where k is a scalar value.
   - If yes, then the tangent hyperplanes of the two margins are parallel, and if the principle curvatures are small (as assumed), then the overlap will be negligible. STOP
3. Determine if $\alpha_{LSL} = -k\alpha_{USL}$, where k is vector of length i.
   - If yes, and the i values of vector k are approximately equal, then the two tangent hyperplanes are almost parallel. In this case, if the principle curvatures are small (as assumed), then the overlap will be negligible.
   - If yes and the above does not hold, in this case, it can only be said that $\text{sign}(u_i)|_{LSL} \neq \text{sign}(u_i)|_{USL}$ for all i. If this is true and the principle curvatures are small (as assumed), and the β's are sufficiently large (greater than 1), then the overlap will be small.

     For example, depicted in Figure 5-11 is a two variable problem where the failure surfaces are linear and approximately perpendicular to each other where both β's are equal to one.

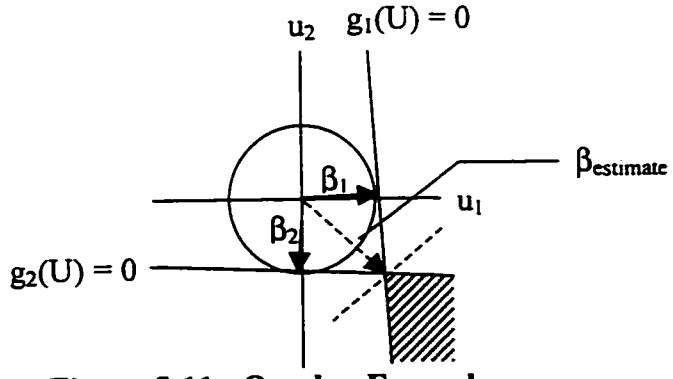


Figure 5-11: Overlap Example

In this case, the U vectors are:

$$U_1 = \begin{bmatrix} 1 \\ 0.00001 \end{bmatrix} \text{ and } U_2 = \begin{bmatrix} -0.00001 \\ -1 \end{bmatrix}$$

This satisfies the second case in this step (3), where k is:

$$k = [0.00001 \quad 100000]$$

and each element of k is not approximately equal. In this case, the overlap probability can be estimated by $\Phi(\beta_{estimate})$, which is $\Phi(\sqrt{2}) = 0.08$. Thus, in this example the probability of success was underestimated by 8%, a reasonably small error. However, as the β's increase, the resulting over- or under-estimation of the probability decreases.

4. If neither of the above are applicable, then further study of the problem is necessary to determine if overlap is a potential problem.

Determination of Limit-State Curvature

In addition to determining the α vectors, there has been some recent work done to extend the first-order method to an approximate second-order method by tracking the changes in curvature as the β point is found. Since the HLRF algorithm first finds the failure surface then moves along it until the minimum distance is found (Madsen, Krenk, and Lind 1986), this information could be used to determine if the curvature is convex or concave, and whether or not the first-order method will be an adequate approximation to the failure surface. For example, as the HLRF algorithm is converging, if we collect the points whose values satisfy $m = 0$ and fit a curve to these points, we can estimate the curvature of the failure surface. If the principal curvatures appear to be small, then the first-order approximation is sufficient. If not, then determine what order approximation is necessary. If the principal curvatures are large, then this will indicate that the approximations used within this thesis are inadequate.

Both of these approaches are applicable to the multiple minima problem. The determination of the α vectors can help determine where the minima are in relation to each other. The determination of the principal curvatures may also yield information about potential multiple minima. As both approaches are improved, there will be some overlap in use between these limitations of the Hasofer-Lind-Rackwitz-Fiessler algorithm.

Chapter 6

Profiling a Response

One of the classic problems in design for quality comes from the study of television sets from two different Sony plants (Phadke 1982). When colour density was examined and compared between plants, the production distribution was determined to be similar to the distributions depicted in Figure 6-1.

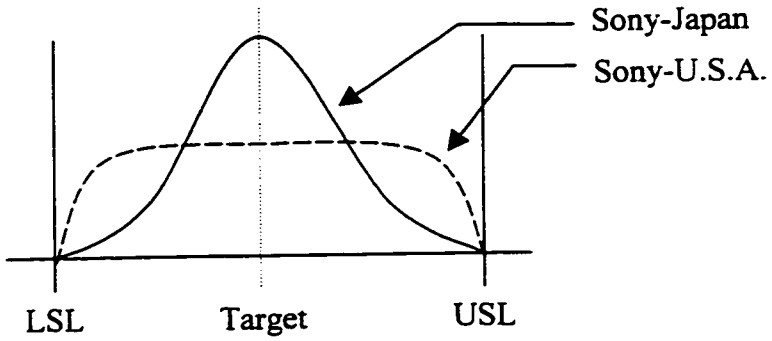


Figure 6-1: Comparison of Sony Television Set Production

The Sony-U.S.A. televisions were all to specification, but the response had an approximate uniform distribution shape. Sony-Japan's response was centered on target and had a normal-like distribution. The problem in comparing the probability of failure as a means of determining quality lies in the fact that in the above example, both distributions give similar values. However, if each response was to shift off-target, the Sony-U.S.A. probability of success would decrease faster than Sony-Japan's.

This was determined using empirical data and probability plots. However, in the examples used in this thesis, the mathematical model already exists for the response. Thus, it would be nice to be able to determine what the response looks like given the variation within the model variables. Monte Carlo methods (Section 2.2.1) are often used to determine the response distribution given variables with arbitrary distributions. If all the variables are normally distributed, D'Errico and Zaino (1988) provide a method by which the first six moments of the response can be estimated using design of experiments. However, a normal assumption may not always be the case.

The previous three chapters dealt with using the transformation method in the probability approach. A unique extension to this method exists that allows the profiling of a response with components having arbitrary distributions (Wu and Wirsching 1987). The method provides an approximate depiction of the response distribution while not being computationally expensive like Monte Carlo Simulation.

6.1 Extension of the Transformation Method

The probability approach enables designers to determine the probability of success or failure that a response is above/below some target value. It was shown that both one-sided and two-sided probability problems could be handled using the transformation method from Chapter 3. Since the transformation method calculates the probability of exceeding a specified value, we could use the method to profile the response by varying the specification.

Graphically, the extension is shown in U-space (Figure 6-2).

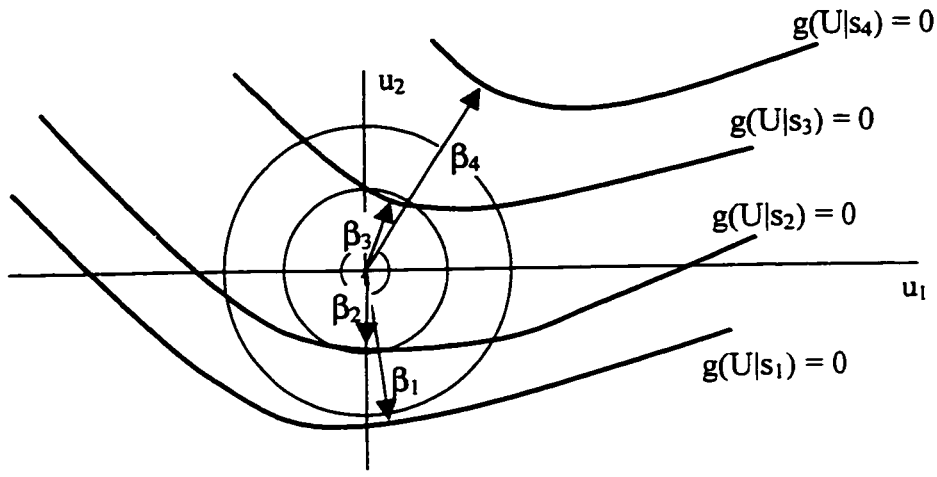


Figure 6-2: Profiling a Response in U-Space

In Figure 6-2, the limit-state surface moves as the specification, s_i, is adjusted. Remember that the limit-state surface, in its simplest form, is a function of the response and a specification:

$$g(\mathrm{V}) = \begin{cases} spec - f(\mathrm{V}) & \text{if } spec > f(\overline{\mathrm{V}}) \\ f(\mathrm{V}) - spec & \text{if } f(\overline{\mathrm{V}}) > spec \end{cases} \tag{6.1}$$

where $f(\overline{\mathrm{V}})$ is the function evaluated at its components' nominal values. For this extension, only the first case, $g(\mathrm{V}) = spec - f(\mathrm{V})$, is used. In using only this case of $g(\mathrm{V})$, the resulting values of β will be representative of the cumulative distribution function (CDF) of the response; when β is negative, the specification is less than the nominal value of the response and β is positive when the specification is greater. Therefore, as the specification is increased and decreased from the nominal value, probabilities will be associated with each specification. From the results attained, the CDF can be drawn and converted into a PDF.

There exists only one problem with the method. The transformation method assumes that the origin in U-space corresponds to the mean of the response distribution. This may not be the case. The origin corresponds to the response evaluated at the nominal values of the parameters. This inaccuracy is due to the fact that the normal-tail

approximation is not designed for use around the mean value, but was developed to approximate the tails of the distributions. However, this is negligible since the purpose of this method is to determine the overall shape of the response. In reality, we wish to distinguish between these two cases:

P(*Failure*) = 0.05 P(*Failure*) = 0.05

(a) (b)

Figure 6-3: Comparison of Two Responses

6.2 Examples

The first example of profiling a response by the transformation method is the servo-control system given in Chapter 4. For the first case where the load T_{11} was Normal, the CDF and PDF for the optimal settings of the shaft speed at D can be found using 10 specifications. The limit-state surfaces to be used to find the corresponding probabilities are:

$$g(\mathrm{V}) = 600 - \Omega_d,\ g(\mathrm{V}) = 700 - \Omega_d,$$
$$g(\mathrm{V}) = 800 - \Omega_d,\ g(\mathrm{V}) = 900 - \Omega_d,$$
$$g(\mathrm{V}) = 1000 - \Omega_d,\ g(\mathrm{V}) = 1100 - \Omega_d,$$
$$g(\mathrm{V}) = 1200 - \Omega_d,\ g(\mathrm{V}) = 1300 - \Omega_d,$$
$$g(\mathrm{V}) = 1400 - \Omega_d,\ g(\mathrm{V}) = 1500 - \Omega_d$$

The resulting CDF is:

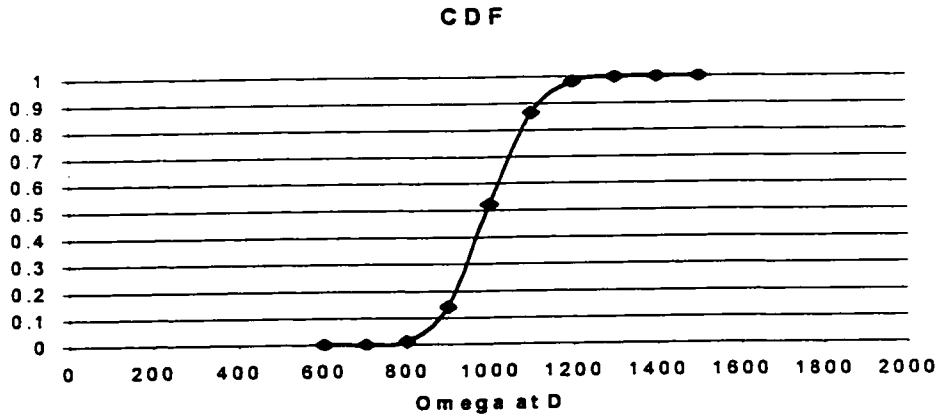


Figure 6-4: Estimated CDF of Shaft Speed at D

The corresponding PDF is:

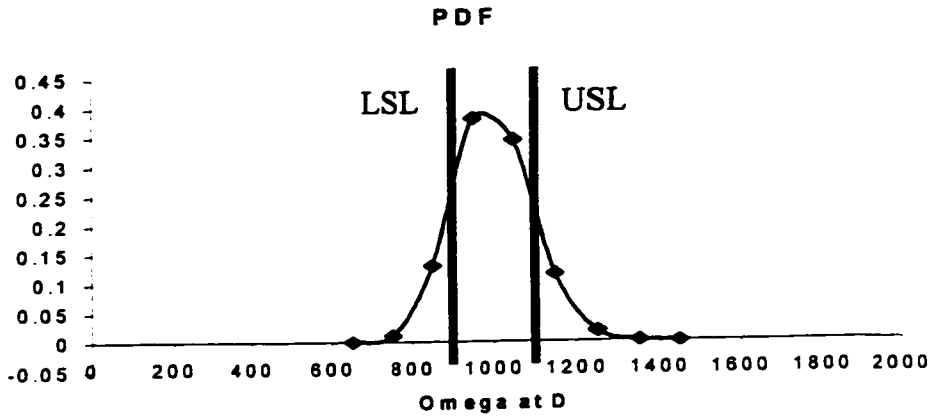


Figure 6-5: Estimated PDF of Shaft Speed at D

The two lines in the above figure represent the upper and lower specification limits given in the servo-control system example where LSL = 900 and USL = 1100. The resulting probability of success was 72.39%. To verify the above figure, a Monte Carlo simulation of 5000 evaluations of the function was performed. With seven random variables, 35000 random numbers were generated and the result is shown in Figure 6-6.

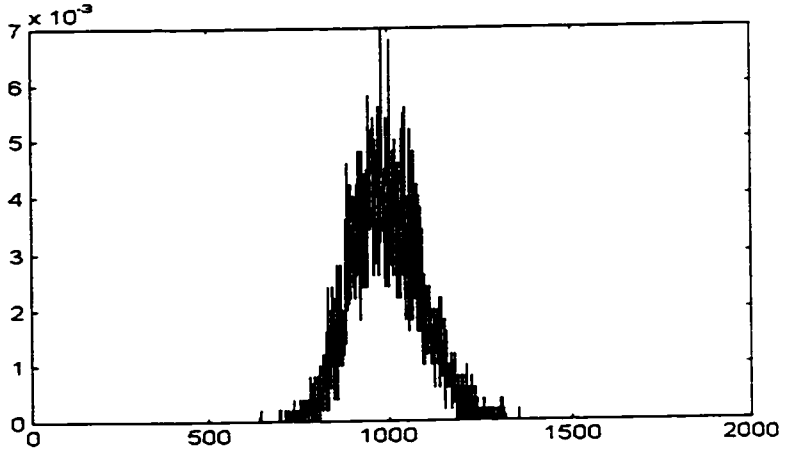


Figure 6-6: Estimated PDF of Shaft Speed at D by Monte Carlo Simulation

The Monte Carlo simulation estimated the probability of success as 72.30% and a mean value of 995.883. This compares very well with the results achieved through the transformation method. In addition, the Monte Carlo simulation estimated the standard deviation to be 98.1819. This currently cannot be done in the probability approach. Future research needs to address this problem.

The same procedure can be used for the second case where the load T_{11} was Uniform. The CDF and PDF using 10 margins are:

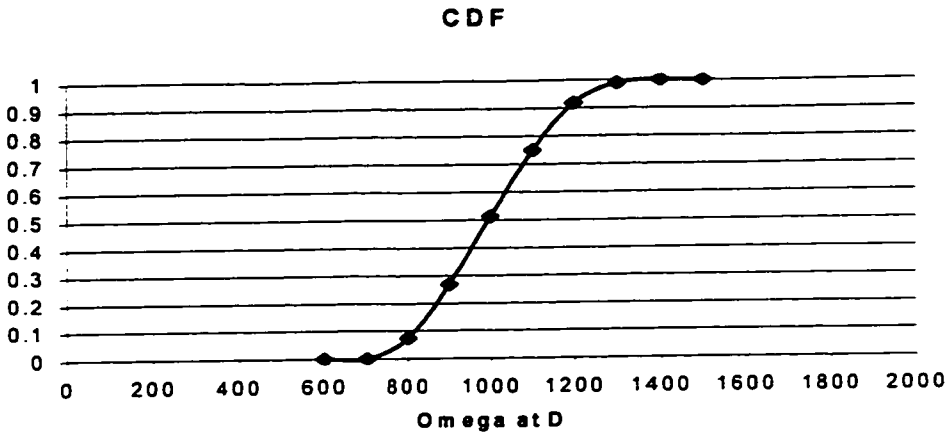


Figure 6-7: Estimated CDF of Shaft Speed at D (Uniform)

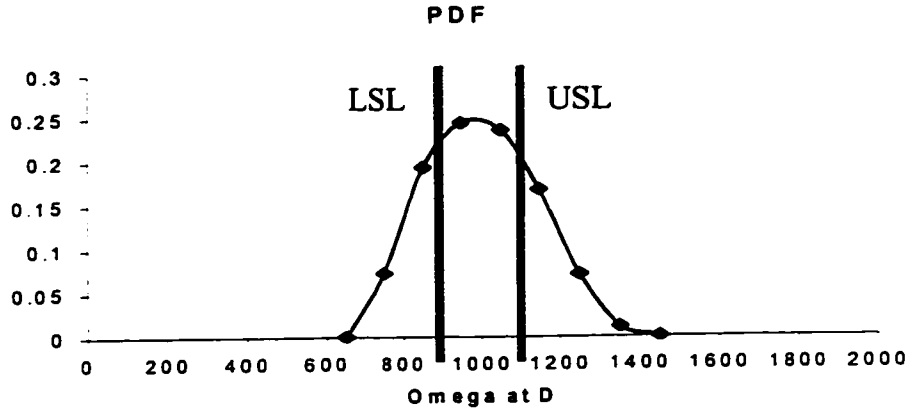


Figure 6-8: Estimated PDF of Shaft Speed at D (Uniform)

Again, a Monte Carlo simulation of 5000 evaluations was performed. Figure 6-9 shows the PDF.

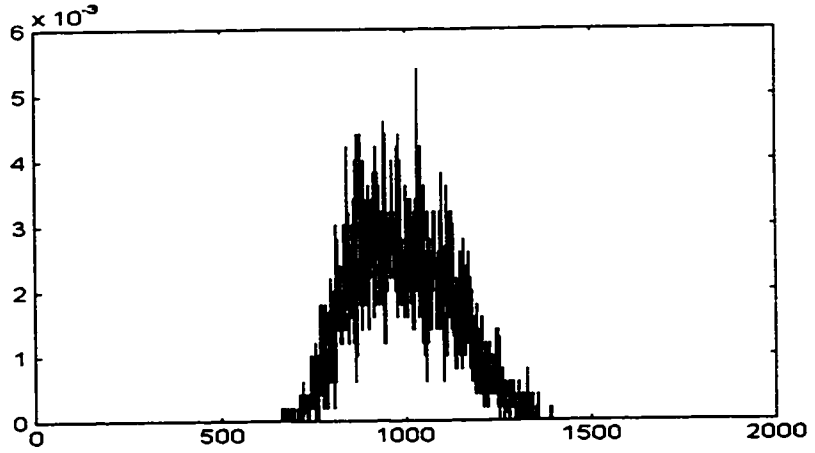


Figure 6-9: Estimated PDF of Shaft Speed at D by Monte Carlo (Uniform)

The Monte Carlo simulation results estimated the probability of success to be 49.92%, a mean of 996.4414, and a standard deviation of 130.8915. The first two values are close to those obtained in Chapter 4.

As can be seen in Figures 6-5 and 6-8, the probability approach centered the distributions such that the probability of success was maximized. In these two cases, since the distribution of the responses were approximately symmetrical, it also ensured that there was an equal probability of failure on both sides of the specifications. Therefore, a small shift in the mean would result in a similar loss to the probability of success regardless of the direction of shift.

The second example is from the non-linear two-pipe system in Chapter 4. For this example, using the initial point from Carr and Savage (1996), the CDF and PDF of the pressure head at B can be profiled using 16 margins. With the diameters of the pipes set at $d_2 = 9.5$ and $d_3 = 11.7$, the mean pressure head is approximately 119 ft. The probability of success was 67.46%.

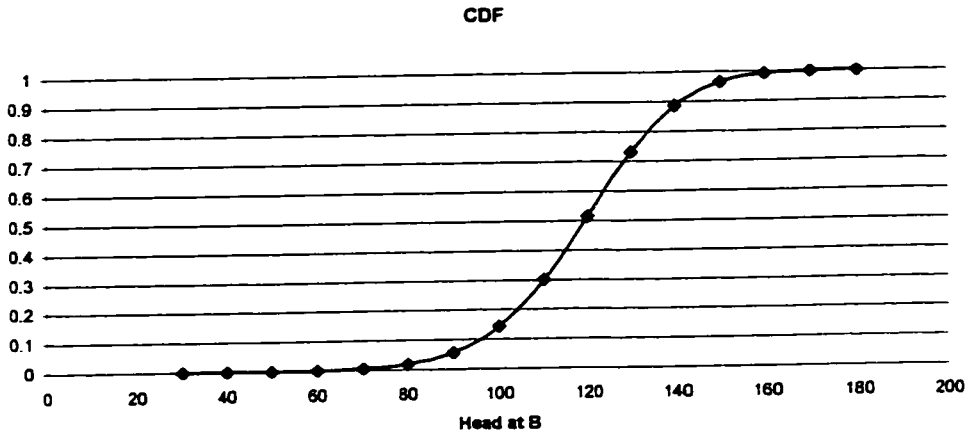


Figure 6-10: Estimated CDF of Pressure Head at B

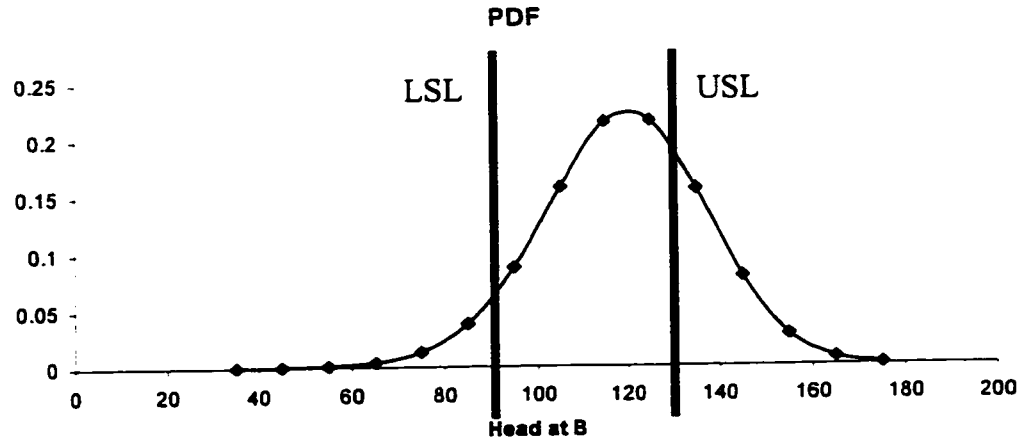


Figure 6-11: Estimated PDF of Pressure Head at B

A Monte Carlo Simulation of 5000 evaluations (20000 random numbers) was done for comparison.

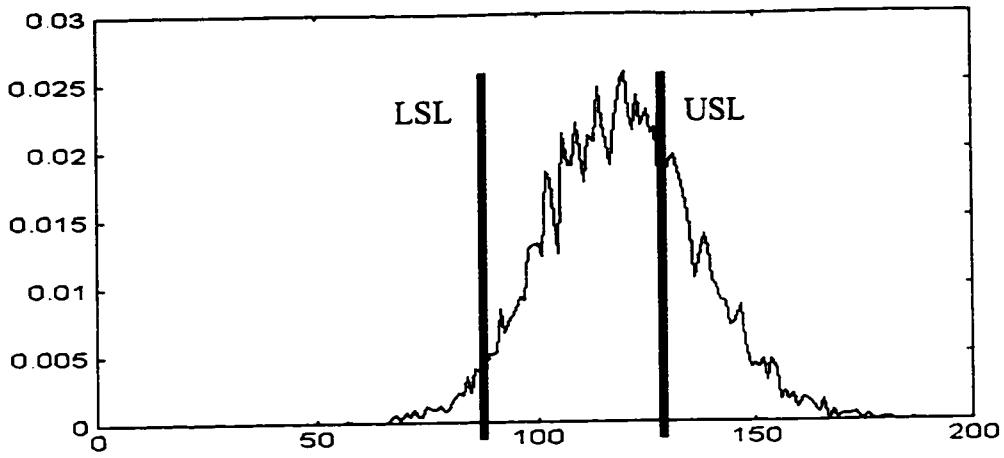


Figure 6-12: Estimated PDF of Pressure Head at B (Monte Carlo)

The estimated probability of success is 68.22%, with a mean of 119.12 and standard deviation of 17.42. Again, the first two values are similar to those determined by the method in Chapter 4.

In addition to obtaining similar results, it can be seen from the PDF that the distribution is negatively skewed. This is confirmed by the value of the mean (119), which is off-center.

Chapter 7

A Worst-Case Analysis Method using the Probability Approach

Worst-Case Analysis is the identification of the extreme values of performance resulting from the variations in parameter values (Spence and Soin 1988). By finding the extreme values, one can compare the worst-case values (upper and lower) to the product's specifications and determine if the expected worst-cases of the response are contained within the specification limits. If they are, then one can assume that the probability of success should be 100%. Therefore, Worst-Case is a special case of 'Design for Quality' where the goal is to have the probability of success be 100%. As such, it needs to be addressed within this thesis.

Detailed in this chapter is a proposed method for determining the worst-case. The method works for both linear and non-linear systems, and most importantly, easily gives the component values at the worst-case – the main difficulty associated with current worst-case approaches.

The proposed method was inspired by two papers: Antreich et al. (1994) and Wojciechowski et al. (1997). The first paper used a first-order second moment method to find the most likely failure point. They refer to β as the "worst-case distance" of the nominal (mean) value to the failure surface. However, their assumption of normality does not allow us to find true worst-case since there is always a probability, regardless of how small, that a value may occur. The second paper suggested the use of intervals in design centering. By using intervals, the best choice for the nominal point would be at the geometrical center of the design space / constraint region, equidistant from the respective failure surfaces. This suggestion of intervals, in combination with the work by Antreich et al., became the inspiration for this method.

The method in this chapter proposes the use of uniform distributions (intervals) instead of assuming normality. The use of uniform distributions also satisfies the constraint that only extreme values need to be given. Our assumption will be that all points between these extremes will be equally likely. By adapting our Hasofer-Lind-Rackwitz-Fiessler algorithm, we propose to find the true worst-case of the system.

7.1 Background/History

Currently, in Worst-Case Analysis, no information about the distributions associated with parameters is necessary. The only information required is the extreme

values of each parameter, given or estimated. The primary method used in Worst-Case Analysis is Vertex Analysis, or more correctly called Interval Analysis. Interval Analysis is the application of interval mathematics to problem solving. The field of interval mathematics was developed in the early 1960's to determine the error bounds in computer arithmetic (Hansen 1969) (Moore 1966, 1979). The field grew into a separate branch of Applied Mathematics dealing with computer arithmetic, mathematical software, linear and non-linear systems, optimization, and operator equations (Moore 1979).

The use of interval mathematics for solving sets of linear equations has been fully documented since the early 1960's (Moore 1966). Surprisingly, throughout the first twenty years, interval mathematics was not mentioned in the engineering journals. This is not to say however, that interval mathematics has not been used in engineering at all. In 1979, Skelboe showed that interval mathematics is a useful tool in the calculation of worst-case analysis for electrical circuits. However, the conclusions found that the development of special software was necessary to solve the system and thus this approach was not competitive with traditional methods for finding the worst-case (in 1979 computing power terms). Despite the obvious compatibility, interval mathematics and engineering would not cross paths in the literature until 1986. In 1986, Deif clearly showed that interval mathematics has applications for engineering. In his book, Deif works through a simple example of a circuit and how to find sensitivity factors. However, this six-page example is the only mention of engineering applications in his 224-page book.

Spence and Soin (1988) discuss Vertex Analysis; the combinatorial evaluation of all extreme values of parameters to aid in worst-case analysis. Their method is Interval Analysis despite its title. In 1990, Kolev and Mladenov discussed a method for finding all the operating points of a non-linear resistive circuit using interval mathematics. Although this work is similar to Skelboe (1979), it has been independently developed. Swan and Savage (1994) showed how intervals could be used in design calculations to quantify uncertainty. However, this method only applied to linear and bilinear models.

7.2 Applications/Limitations of Current Analysis

Interval Analysis methods involve combining all the upper and lower values of each component within the performance function. This gives 2^k combinations for a product with k parameters. For a product with a large number of parameters, the calculation of Worst-Case Analysis using Interval Analysis can be extremely computationally intense, and suffers from dimensionality. As well, this method assumes that the extreme values of the product's performance will occur at one of the 2^k combinations or vertices. This will always occur for linear and bilinear functions. For these functions, Spence and Soin (1988) discuss alternative approaches to searching all combinations. The addition of sensitivity information allows for intelligent searches of the 2^k vertices.

However, for both linear and non-linear functions, an exhaustive search using Monte Carlo Methods can be performed on the function, where uniform distributions can

be substituted for the intervals. However, this method too suffers from the intense computational effort needed to keep track of the worst-cases within the tolerance region. Until now, this has been the main difficulty associated with worst-case methods.

Another difficulty associated with worst-case approaches is the selection of the extreme values of the parameters. This difficulty is not dealt with in this thesis, but please refer to Tribus (1969) and Der Kiureghian (1986) for more information on the classification and representation of uncertainty. Nevertheless, useful results can be obtained from current worst-case analysis methods. This chapter deals with an extension of interval analysis into the probability approach to help determine the true worst-case of both linear and non-linear functions.

7.3 Extension of Probability Approach to Worst-Case

By using the probability approach to calculate the probability of success for two-sided quality problems, we can extend the method into finding the maximum and minimum of the function over a specified set of design parameters.

The problem with non-linear functions is that the maximum/minimums might not occur on the vertices. Interval Analysis can therefore not be used if the function is non-linear. Monte Carlo methods avoid this problem by using distribution information for the design variables. If we have the case where only the extreme values are known, then Monte Carlo assumes uniform distributions since any value between the extremes is equally possible. This is equivalent to checking all points on an interval. By doing this, we can ensure that the proper combination of variables necessary for minimization or maximization can be found. However, two difficulties lie with the Monte Carlo approach. There can be a large computational expense, even for a small number of design variables, and there is the difficulty in keeping track of component values at the worst-case.

By using uniform distributions for the variables of a response in the probability approach, the problem is transformed into the standard normal probability space and this provides a quick and easy check to find the most likely failure point (MLFP) with each check providing the combination of variables that comes closest to violating the specification limit. The only assumption is the use of uniform distributions. Numerous papers (Wojciechowski, Vlach, and Opalski 1997) (Cui and Blockley 1991) (Dong and Wong 1986) (Schjaer-Jacobsen and Madsen 1979) have all proposed and used uniform distributions in calculating worst-case, and so, uniform distributions will be used within this approach since it provides the least biased approximation for the variables (Principle of Indifference).

Given that the CDF for a uniform distribution is:

$$p = F(v|a,b) = \frac{v-a}{b-a} I_{[a,b]}(v) \tag{7.1}$$

and using the Rosenblatt transformation, the corresponding point for v in standard probability space is:

$$u = \Phi^{-1}\left(\frac{v-a}{b-a}\right) \tag{7.2}$$

where u is the uncorrelated standard probability space variable. From this transformation we can see:

$$\begin{aligned} &\text{if } v = a,\ \Phi^{-1}(0) \rightarrow u = -\infty \\ &\text{if } v = b,\ \Phi^{-1}(1) \rightarrow u = +\infty \\ &\text{if } v = \left(\frac{b+a}{2}\right),\ \Phi^{-1}(0.5) \rightarrow u = 0 \end{aligned} \tag{7.3}$$

Similarly, the u-space to v-space transformation is:

$$v = \Phi(u) \times (b-a) + a \tag{7.4}$$

To determine the worst-case, uniform distributions will be used for all design variables with no correlation, i.e., the covariance matrix is a unit-diagonal matrix. The Hasofer-Lind-Rackwitz-Fiessler algorithm converges to U^*, the MLFP. However, the use of uniform distributions creates a problem in that there is the potential that the margin, $g(V_k)$, may never be zero. If this occurs, we have successfully found a specification for which the margin will never be zero. This creates a problem in the calculation of U_{k+1} since it will not converge.

For example, to find the maximum value of a response, we set the margin equal to:

$$m = SL - f(V) \tag{7.5}$$

where SL is the estimated maximum worst-case value. If in the search for the MLFP $g(V_k)$ stays positive, then the value SL lies above the maximum worst-case value. Since the HLRF algorithm gives the point U^* which is the point on the limit-state surface closest to the origin, the termination criteria in the algorithm needs to be changed to minimize the absolute value of the margin, $|g(V_k)|$ and $\nabla_V g(V)$. This additional criteria of minimizing the absolute value of the margin now handles the non-zero margin case. If the absolute value of the margin is minimized, but not zero, and the last two iterations are within an error of ε, then the algorithm will terminate. In addition, one may wish to place an additional termination criteria on the algorithm such that it terminates if the number of iterations exceeds some specified value. This will aid if convergence is not achieved.

The proposed algorithm to find the worst-cases using a modified HLRF algorithm follows. It requires three termination criteria to be specified in advance (the maximum number of iterations, ε_g – the maximum allowable value for the margin, and ε_{wc} – the maximum allowable difference for SL).

1. Construct the margin.
   - g(V) = SL - Response
2. Choose a starting value for SL.
   - Choose a value such that $|g(V_k)| = 0$ is possible (i.e., the limit-state surface passes through the parameter space (in U and V).
3. Using the HLRF algorithm (Section 3.3), find the δ^* and minimum $|g(V_k)|$.
4. If the last two values for SL are such that the absolute value of the margin is zero for one but not for the other, (e.g., $|g(V_k)| = 0$ and $|g(V_{k-1})| \neq 0$) AND the absolute value of the margin for the non-zero case is less than some specified error, (i.e., $|g(V_{k-1})| < \varepsilon_g$) AND the difference between the last two values for USL is less than some specified error (i.e., $\Delta SL < \varepsilon_{wc}$) then STOP, else
   - Find minimum worst-case value
     - If $|g(V_k)| = 0$ then reduce SL
     - If $|g(V_k)| \neq 0$ then increase SL
   - Find maximum worst-case value
     - If $|g(V_k)| = 0$ then increase SL
     - If $|g(V_k)| \neq 0$ then reduce SL
5. Continue steps 3 and 4 until (k > Maximum Iterations Allowed) OR ($\Delta SL < \varepsilon_{wc}$ AND $|g(V_k)| \neq 0$ AND $|g(V_k)| < \varepsilon_g$)

Graphical Depiction of Worst-Case Analysis using the HLRF Algorithm

To show the process, Figures 7-2 through 7-4 have been constructed to depict the finding of the worst-cases. Figure 7-1 depicts a two-variable problem in V-space.

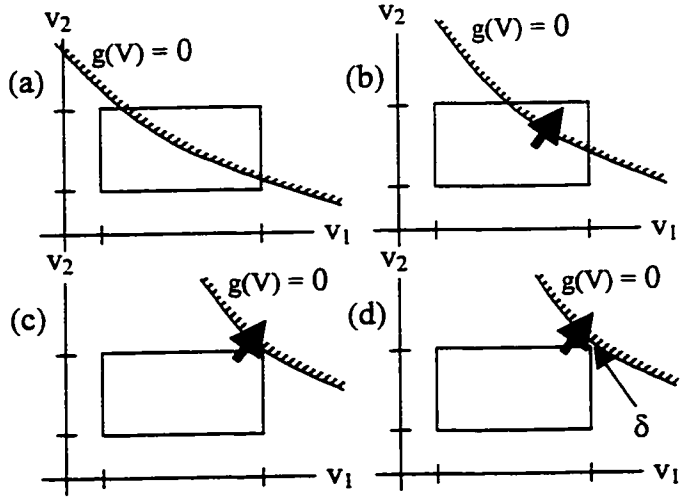


Figure 7-1: Finding the Worst-Cases in V-space

To determine the worst-case values, each variable is assigned a uniform distribution. The limit-state surface is shown as g(V) = 0, and defines the success / failure region. In (a), the limit-state surface passes through the variable space. As the limit-state surface is moved by changing the USL (Figure 7-1 (b)), the surface will eventually only be satisfied by one combination of design variables (Figure 7-1 (c)). As the specification is changed just a little more, there does not exist any combination of variables within the design space to satisfy g(V) = 0 (Figure 7-1 (d)), and the result is a minimum margin value of δ indicating that the failure surface can never be reached or crossed.

Figure 7-2 shows the corresponding U-space diagram and limit-state surface for (a). The remaining corresponding U-space diagrams are shown in Figure 7-3, and it can be seen that the limit-state does not exist in U-space if no combinations of variables satisfy g(V) = 0. This is because U-space only represents possible combinations of variables in V-space, and since the design variables cannot have values outside their uniform distributions, they are not represented in U-space. Mathematically, this can be seen in transformation, equations (7.2) and (7.4).

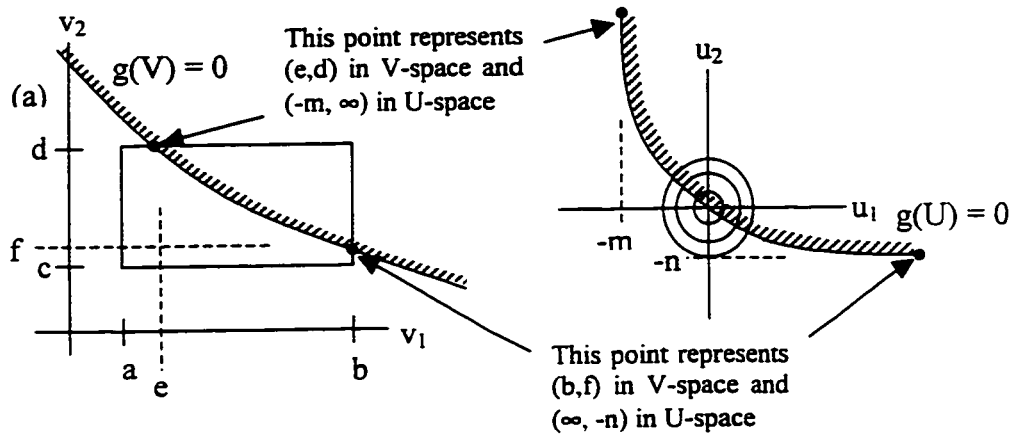


Figure 7-2: Corresponding U-space diagram for (a) in Figure 7-1

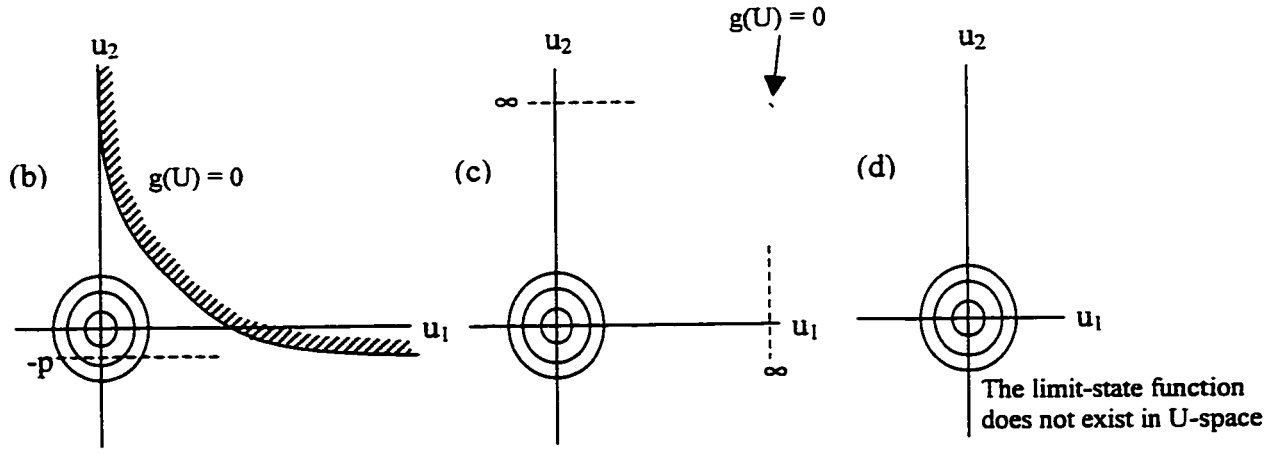


Figure 7-3: Corresponding U-space diagrams for (b), (c), and (d) in Figure 7-1

7.4 Examples:

Two examples are detailed below which show the differences between Interval Analysis, Monte Carlo, and the Probability Approach in worst-case analysis. The first, a linear voltage divider, is solved out for each step to show the process, and the second is a non-linear voltage gain problem.

7.4.1 Linear Systems

The following diagram, Figure 7-4, depicts a simple voltage divider. The objective is to determine the extreme values of the voltage at node B. The circuit has a voltage source, v_1, which is deterministic with a value of 5 Volts. Both resistors are normally distributed; resistor 2 is 10 Ω ± 10% (σ = 0.3333) and resistor 3 is 20 Ω ± 10% (σ = 0.6667). Note that $2(10\% \cdot \mu) = 6\sigma$.

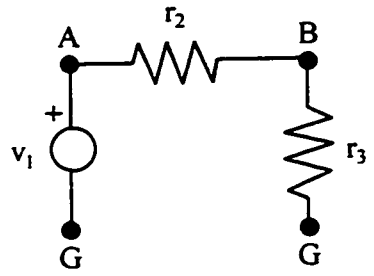


Figure 7-4: A Simple Voltage Divider

If the problem were solved using only mean values (i.e., deterministic solution), we would find that the voltage at B is:

$$v_b = v_1 \frac{R_3}{R_3 + R_2} = v_1 \frac{1}{1 + \frac{R_2}{R_3}} = 3.3333 \tag{7.6}$$

using nominal values. However, we cannot state anything about the range of the voltage. To do this, the problem can be solved using bounds/interval analysis.

Bounds/Interval Analysis

In engineering calculations, often the only points needed to determine the characteristics of a system are the nominal and the two end-points of the distribution. Since all normal distributions tend toward positive and negative infinity, it becomes useful to define a reference point for the end-points. In industry, these end-points are typically characterized by 3σ, three standard deviations away from the nominal value.

Thus, Interval Mathematics can be used to give quick and simple solutions to probabilistic problems. For the example, we can define the resistors as

$$r_2 = [9..11], \quad r_3 = [18..22] \tag{7.7}$$

with the nominal values, r_2 = 10 Ohms and r_3 = 20 Ohms. Using these values, we find the response to be

$$v_b = 5\frac{1}{1+\frac{[9..11]}{[18..22]}} \tag{7.8}$$

and following Interval Arithmetic, we can simplify the response to

$$v_b = 5\frac{1}{1+\left[\frac{9}{22}..\frac{11}{18}\right]} = 5\times\left[\frac{18}{29}..\frac{22}{31}\right] \tag{7.9}$$

which gives a final interval response vector of

$$v_b = [3.103..3.548] \tag{7.10}$$

From this answer, we can extract that the voltage at node b could range from 3.103 to 3.548 Volts given that the resistors may vary over their respective intervals. It should be noted that the interval response vector does contain the nominal response vector. Thus, Interval Mathematics can be used to give a preliminary estimate as to the extent of the distribution. However, it does not give us the combination of values for each parameter that result in the maximum and minimum.

Worst-Case using the Probability Approach

Using the algorithm in Section 7.3 to find worst case, we find the following solution:

$$v_b = [3.1034..3.5484] \tag{7.11}$$

With the values occurring at the following combinations:

Table 7-1: Worst-Case Values and their Respective Parameter Values

Parameter	Minimum (3.1034)	Maximum (3.5484)
V	*5*	*5*
R2	*11*	*9*
R3	*18*	*22*

In this case, the maximum and minimum lie on the vertices so both the bounds analysis and the algorithm in Section 7.3 give the same result. In fact, this is true for any linear or bilinear problem since the response will be monotonically increasing over each of the design variables.

7.4.2 Non-Linear Systems

The following equation represents the magnitude of the voltage at a current source in an RLC circuit excited by an oscillating current.

$$V(s) = \frac{I(s)}{\frac{1}{R} + sC + \frac{1}{sL}} \tag{7.12}$$

$$|V(jw)| = \sqrt{\mathrm{Re}^2 + \mathrm{Im}^2} = \frac{RwL}{\sqrt{\left(w^2L^2 + R^2w^4C^2L^2 - 2R^2w^2CL + R^2\right)}} \tag{7.13}$$

For $R = 1$, $L = 1$, $w = 0.75$, and $C = [1..2]$. we can find the graph showing |V| versus C.

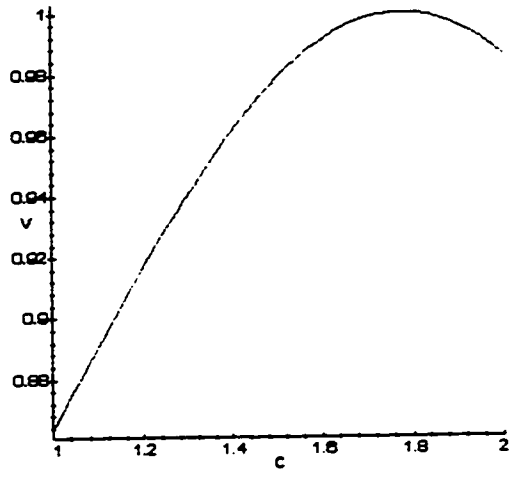


Figure 7-5: RLC Circuit Response

We can see that the maximum voltage does not occur at any of the vertices. In fact, the capacitance value of 1.77 gives the maximum voltage.

Bounds/Interval Analysis

Substitution of intervals into the above equation gives a result of:

$$|V| = [0.86377..0.98639] \tag{7.14}$$

occurring at C=1 and C=2. However, we know this to be incorrect by looking at Figure 7-5. This illustrates why Interval Analysis is not appropriate for Non-Linear Systems.

Worst-Case using the Probability Approach

Using the algorithm in Section 7.3 to find worst case, we find:

$$|V| = [0.86377..1.0] \tag{7.15}$$

occurring at C = 1 and C = 1.7737. This was achieved by setting the lower and upper specifications of V to 0.5 and 1.2. The algorithm finds these points as the best solutions to minimizing the margin and $|\delta_{k+1} - \delta_k| \leq \varepsilon$. The algorithm finds u in the first case (lower specification) to be approaching negative infinity. From equation (7.3), this implies that the lower bound is achieved when the lower extreme, C = 1, is approached. The upper bound is found at u = 0.7388, and the margin reaches a minimum of 0.2, indicating that the limit-state function does not intersect the probability space. Transforming this into v-space, we find the value of C is 1.7737. This solution was found in 11 iterations of the algorithm.

7.5 Comments

This adaptation of the probability approach appears to have two distinct advantages over current worst-case approaches. The first is that it is far less computationally expensive. For a few hundred function evaluations we can determine the worst-cases of a system. Even compared to efficient Monte Carlo methods, the proposed method appears to be better. The second is that the combination of parameters that gives the worst-case is automatically found with the algorithm. There is no need to keep track of combinations separately as in Monte Carlo worst-case analysis methods.

However, it must be noted that the extension may be computationally difficult for a large number of variables. Further study is necessary to ensure that true worst-case points are found in this case.

Chapter 8

Conclusions and Future Directions

Within this document, there are still a number of issues to be addressed. The purpose of this chapter is to list and conclude what has been accomplished, and almost more importantly, what is still to be done in future research.

8.1 Conclusions

Proposed and detailed within the previous chapters is a method for determining the quality of a system. It is efficient, accurate, and effective for both linear and non-linear systems. The method is based on a probability approach used within the field of structural engineering, and allows for the use of full arbitrary distribution information. This is a significant advance over the current first-order second moment methods widely used in other fields.

In addition, by using the probability approach and its requirement to transform the parameter distributions into the standard normal space, the mathematical relationships of how a parameter's mean and variance affects the response's probability of failure is found. Within this relationship, a link to Taguchi's factor types is established. By examination of the gradient information generated in both V- and U-space, a parameter's factor type (i.e., neutral, noise, design-adjustment, or design-control) can be determined due to the relationship found between the failure surface and the parameter.

Extensions of the method proposed in this thesis relate to the profiling of response distributions and the determination of worst-case. The first extension results from altering the specification value systematically and then re-evaluating β, the distance between the nominal values and the most likely failure point. The result is a quick method for determining the shape of a response distribution. The second extension, in a sense, results from the first extension in combination with work previously done by Antreich et al. (1994). By modifying the Hasofer-Lind-Rackwitz-Fiessler algorithm to search for the minimum margin value, we find the two points in U-space that come closest to violating the upper and lower specifications.

With respect to the models to be improved, the use of Graph-Theoretic Models as a method to model the system provides many advantages. The ability to stamp together a system and its sensitivities is important due to the requirement of gradient information. However, the use of Graph Theory is not a requirement. All the methods within this document only require a mathematical model of the response. From this model, all

gradient information must be obtained through other means (Matlab, MAPLE, etc.). As a final overview of the methodology for quality design in this thesis, Figure 8-1 is given.

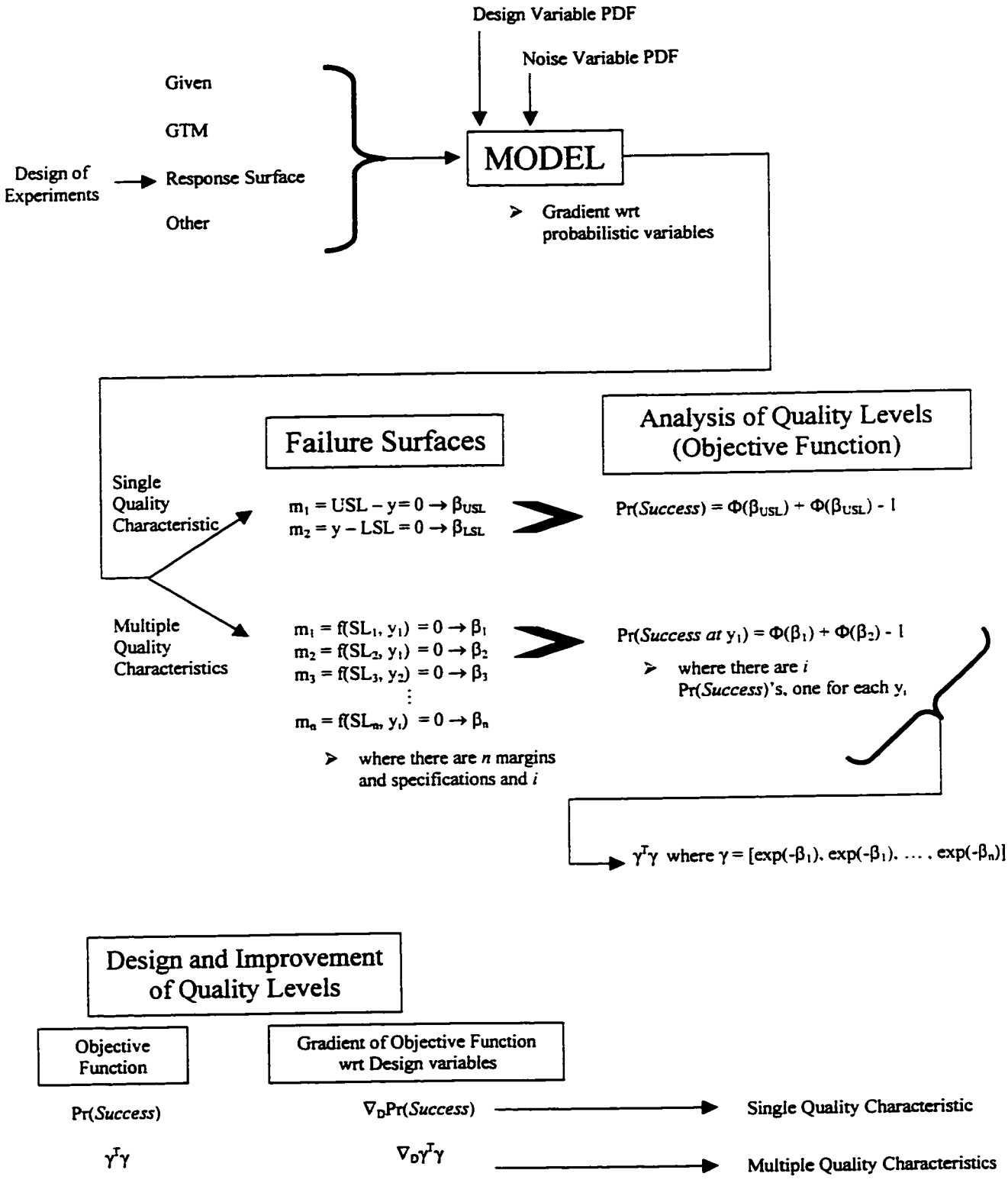


Figure 8-1: Overview of Design for Quality

8.2 Future Directions

The future of this research lies in six areas: Coding, Incorporation of Process Information, Incorporation of Reliability Information, Development of the Worst-Case Analysis Extension, Determination of Mean and Variance using the Probability Approach, and the Global Quality Level for Time Domain Systems.

8.2.1 Coding

Currently, all the computer code is in MAPLE and Matlab routines (see Appendix B for a list of the routines used). To further improve the efficiency and speed of the algorithms, they should be coded into an appropriate language and ultimately meshed with a graphical interface. Three approaches exist. The first is to simply enhance the existing code in Maple and Matlab so that it becomes a toolbox for this software. This has its advantages due to the prevalence of these programs within industry, but lacks the user-friendly graphical interface.

The second approach is to modify the code for use within PROSIM, a software package currently developed at the university. This program has a graphical interface and the ability to draw the system under study, input mathematical equations for components and systems, and solve the system. Unfortunately the program was written with little room for expansion. Extensive work may be necessary to modify this program to handle the types of problems currently solved using the Maple and Matlab routines.

The third option is to write a new program from scratch. The program should begin as a shell to the current routines, but could ultimately be expanded to encompass different aspects of the methodology (e.g., optimization routine, HLRF algorithm, etc.). This program would probably have a similar user interface to PROSIM, but with the ability to calculate quality and be expandable to include routines for reliability, redundancy, process models, etc.

In addition to the above coding, more efficient optimization routines should be researched to aid in the design optimization of the system. The routines need only accept the objective function, gradient of objective function with respect to design variables, constraints, and gradients of the constraints with respect to design variables. Any or all of these may be non-linear. Additionally, non-gradient based methods should be explored in the event that gradient information is not attainable.

8.2.2 Incorporating Process Information

Instead of designing with lumped parameters (e.g., resistance), the methods should be used to design for physical properties such as dimensions, densities, temperatures, etc., as in the teacup example from Chapter 4. This would add a dimension of realism to the design optimization process, instead of simply choosing arbitrary ranges for design parameters.

8.2.3 Incorporating Reliability Information

Of all the future directions, this one intrigues me the most. Current reliability methods determine the probability of a system being in a functional state. For systems with redundancy (e.g., structures, pipes, power distribution), this implies that there may be multiple topologies that could result in different quality levels. Analysis of these topologies is straightforward – simply use the methods in this thesis to analyze each topology for its quality level. Design, however, is very difficult.

In the eight-pipe, six-demand problem in Chapter 4, the objective was cost and thus the design removed extra pipes from the system. A method for improving design quality with redundancy is needed. No current method exists, but by adapting the methods in this thesis, an approach dubbed the "multiple topology, multiple response" method is being examined within our research group.

8.2.4 Development of the Worst-Case Extension

The work in Chapter 7 is an extension of the methods discussed in this thesis. Future work needs to examine the properties and limitations of that method. Improvement of the algorithm is necessary to find the worst-cases more efficiently.

8.2.5 Determination of Mean and Variance

Through conversations with Dr. Michael Hamada, it was suggested that the methods in this thesis should be able to determine the mean and variance of the response. From this suggestion, Chapter 6 was created as a step toward solving this problem. However, a solution is still needed.

This problem is unique since we wish to find the mean and variance of a response given only a handful of probabilities associated with points along the response. In effect, we have an accurate depiction of the tails, and we wish to find the center. Entropy-based methods may be applicable.

8.2.6 Global Quality Level of Time Domain Systems

It was stated in the introduction that there were restrictions on the determination of quality in systems that varied with time. The methods within this thesis can only determine the quality level for specified times. As in the teacup problem in Chapter 4, we can find the probability of success or failure associated with a time, but we cannot currently extend this to a global level. Future study to determine the quality level of a time varying system needs to be done.

In dynamic problems, the system is often constrained to stay within specified levels. To currently determine how close the response comes to exceeding the specifications, we would need to solve the system at every specified time. Only then could we estimate the quality by taking the minimum of the probabilities found.

References

Antreich, Kurt J., Helmut E. Graeb, and Claudia U. Wieser, "Circuit Analysis and Optimization Driven by Worst-Case Distances", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 13, No. 1, pp. 57-71, 1994.

Bagchi, Tapan P., and J. G. C. Templeton, "Robust Design Engineering using Variance Transfer," *Proceedings of the ISSAT International Conference on Reliability and Quality in Design*, Editor: Hoang Pham, Seattle, Washington, March 16-18, pp. 203-207, 1994.

Barker, Thomas B., *Engineering Quality by Design: Interpreting the Taguchi Approach*. ASQC Quality Press, Milwaukee, 1990.

Belavendram, Nicolo, *Quality by Design: Taguchi Techniques for Industrial Experimentation*, Prentice Hall, London, 1995.

Ben-Haim, Yakov, *Robust Reliability in the Mechanical Sciences*, Springer, New York, 1996.

Biles, W. E., "A Response Surface Method for Experimental Optimization of Multi-Response Processes," *Industrial and Engineering Chemistry, Process Design and Development*, Vol. 14, pp. 152-158, 1975.

Bisgaard, Søren, and Bruce Ankenman, "Analytic Parameter Design," *Quality Engineering*, Vol. 8, No. 1, pp. 75-91, 1995-96.

Carr, S. M., and G. J. Savage, "Reliability-Based Optimization on a Graph-Theoretic Foundation", *Invited paper for special session: Recent Advances in Graph-Theoretic Modelling and Analysis*, IEEE International Conference on Systems, Man, and Cybernetics, Beijing, China, Vol. 4, pp. 3072-3076, October 14-17, 1996.

Carr, Stephen M., *A Unified Methodology for Reliability-Based Design of Physical Systems*. Master's Thesis, Waterloo, University of Waterloo, 1990.

Carr, Stephen M., *Systems-Theoretic Formulation of Robust Design – Proposal for Doctoral Thesis*. Department of Systems Design Engineering, University of Waterloo, 1992.

Chandrashekar, M., and H.K. Kesavan, "Network Sensitivity Simplified," *Proceedings of the IEEE*, Vol. 62, No. 8, pp. 1179-1180, 1974.

Chandrashekar, M., and G. J. Savage, *Engineering Systems: Analysis, Design and Control*, In Preparation, Department of Systems Design Engineering, University of Waterloo, Waterloo, 1997.

Chandrashekar, M., P. Roe, and G. J. Savage, "Graph-Theoretic Models – A Unifying Modelling Approach," *Proceedings of the 23rd Annual Pittsburgh Conference on Modelling and Simulation*, Pittsburgh, Pa. Vol. 18, pp. 2217, 1992.

Charnes, A., and A. C. Sterdy, "A Chance-Constrained Model for Real-time Control in Research and Development Management," *Management Science*, Vol. 12, No. 8, pp. B353-B363, 1996.

Charnes, A., and W. Cooper, "Deterministic Equivalents for Optimizing and Satisfying under Chance Constraints," *Operations Research*, Vol. 11, No. 1, pp. 18-39, 1963.

Coleman, David E., and Douglas C. Montgomery, "A Systematic Approach to Planning for a Design Industrial Experiment," *Technometrics*, Vol. 35, No. 1, 1993.

Cornell, C. A., "A Probability-Based Structural Code," *Journal of the American Concrete Institute*, Vol. 66, No. 12, pp. 974-985, 1969.

Cui, Weicheng, and David I. Blockley, "On the Bounds for Structural System Reliability," *Structural Safety*, Vol. 9, pp. 247-259, 1991.

D'Errico, John R. and Nicholas A. Zaino, "Statistical Tolerancing Using a Modification of Taguchi's Method," *Technometrics*, Vol. 30, No. 4, 1988.

Dehnad, Khosrow, *Quality Control, Robust Design, and the Taguchi Method.* Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, California, 1988.

Deif, Assem, *Sensitivity Analysis in Linear Systems*. New York, Springer-Verlag, 1986.

Der Kiureghian, A., "Measures of Structural Safety under Imperfect States of Knowledge," *Journal of Structural Engineering*, Vol. 115, No. 5, pp. 1119-1140, 1989.

Dharchoudhury, Abhijit, and S. M. Kang, "Worst-Case Analysis and Optimization of VLSI Circuit Performances", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 14, No. 4, pp. 481-492, 1995.

Dickinson, R. R., "A Unified Approach to Second Moment Modelling of Probabilistic Physical Systems," Ph.D. Thesis, Department of Systems Design Engineering, University of Waterloo, 1987.

Dickinson, R. R., and G. J. Savage, "A Systems Theoretic Approach to Second Moment Modeling," *Civil Engineering Systems*, Vol. 5, pp. 8-16, 1988.

Dickinson, R. R., and G. J. Savage, "Automated Second Moment Analysis of Systems with Probabilistic Constitutive Components," *Civil Engineering Systems*, 1990.

Ditlevsen, O., "Principle of Normal Tail Approximation," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 107, pp. 1191-1208, 1981.

Dolinsky, K., "First Order Second-moment Approximation in Reliability of Structural Systems: Critical Review and Alternative Approach," *Structural Safety*, Vol. 1, No. 3, pp. 211-231, 1983.

Dong, Wei-Min, and Felix S. Wong, "From Uncertainty to Approximate Reasoning: Part 1: Conceptual Models and Engineering Interpretations," *Civil Engineering Systems*, Vol. 3, pp. 143-154, 1986.

Fletcher, R., and N. Powell, "A Rapidly Convergent Descent Method for Minimization," *Computer Journal*, Vol. 6, pp. 163-168, 1963.

Graham, Alexander, *Kronecker Products and Matrix Calculus with Applications*. Ellis Howard Ltd., Chichester, 1981.

Gupta, S., and M. Chandrashekar, "A Unified Approach to Modelling Photovoltaic Powered Systems," *Solar Energy*, Vol. 55, No. 4, pp. 267-285, 1995.

Gupta, Suneel, *Application of Non-linear Graph Theoretic Modeling for the Simulation of Directly Connected – Stand Alone Photovoltaic Systems*, Master's Thesis, University of Waterloo, Waterloo, Ontario, Canada, 1994.

Hajdukiewicz, J., "System Analysis of a Servo-Control System," Course Project of SD351 Systems Models I, Department of Systems Design Engineering, University of Waterloo, Waterloo, 1993.

Hansen, Eldon, *Topics in Interval Analysis*. Oxford University Press, London, 1969.

Harry, Mikel J., and J. Ronald Lawson, *Six Sigma Producibility Analysis and Process Characterization*. Addison-Wesley, Reading, Massachusetts, 1992.

Hasofer, A. M. and N. C. Lind, "Exact and Invariant Second-Moment Code Format," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 100, pp. 11-121, 1974.

Hohenbichler, M., and R. Rackwitz, "Nonnormal Dependent Vectors in Structural Reliability," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 107, pp. 1127-1238, 1981.

Kalbfleish, J. G., *Probability and Statistical Inference Volume 1: Probability*, Springer-Verlag, New York, 1979.

Khattree, Ravindra, "Robust Parameter Design: A Response Surface Approach", *Journal of Quality Technology*, Vol. 28, No. 2, pp. 187-198, 1996.

Knottmaier, J., *Optimizing Engineering Design*, McGraw-Hill, London, 1993.

Koenig, Herman E., and W.A. Blackwell, "Linear Graph Theory - A Fundamental Engineering Discipline," *IRE Transactions on Education*, Vol. 105, pp. 42-49, 1960.

Kolev, L.V., and V.M. Mladenov, "An Interval Method for Finding All Operating Points of Non-Linear Resistive Circuits," *International Journal of Circuit Theory and Applications*, Vol. 18, pp. 257-267, 1990.

Krishna, Kannan, and Stephen W. Director, "The Linearized Performance Penalty (LPP) Method for Optimization of Parametric Yield and Its Reliability", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 14, No. 12, 1995.

Lansey, Kevin, E., Ning Duan, Larry W. Mays, and Yeou-Koung Tung, "Water Distribution System Design Under Uncertainties," *Journal of Water Resources Planning and Management*, Vol. 115, No. 5., pp. 630-645, 1989.

Larson, H. J., *Introduction to Probability and Statistical Inference*, Wiley, Chichester, 1969.

Leon, Ramon V., Anne C Shoemaker, and Raghu N. Kackar, "Performance Measures Independent of Adjustment: An Explanation and Extension of Taguchi's Signal-to-Noise Ratios", *Technometrics*, Vol. 29, pp. 253-285, 1987.

Lightner, M., T. Trick, and R. Zug, "Circuit Optimization and Design," *Circuit Analysis, Simulation and Design, Part 2 (A. Ruehli). Advances in CAD for VLSI 3*, Amsterdam: North-Holland, pp. 333-391, 1987.

Logothetis, N., and A. Haigh, "Characterizing and Optimizing Multi-response Processes by the Taguchi Method," *Quality and Reliability Engineering International*, Vol. 4, pp. 159-169, 1988.

Logothetis, N., and H.P. Wynn, *Quality Through Design: Experimental Design, Off-line Quality Control, and Taguchi's Contributions*. Oxford Scientific, Oxford, 1989.

Low, K., and S. Director, "An Efficient Methodology for building macro-models of IC Fabrication Processes," *IEEE Transactions on Computer-Aided Design*, Vol. 8, pp. 1299-1313, 1989.

Madsen, H. O., S. Krenk, and N. C. Lind, *Methods of Structural Safety*. Prentice-Hall, Englewood Cliffs, NJ, 1986.

Maghsoodloo, Saeed, "The Exact Relationship of Taguchi's Signal-to-Noise Ratio to His Quality Loss Function," *Journal of Quality Technology*, Vol. 22, No. 1, pp. 57-67, 1990.

Matlab, Version 4, The Math Works Inc., Massachusetts, 1994.

Melchers, R. E., *Structural Reliability: Analysis and Prediction*, Ellis Horwood Limited, New York, 1987.

Moore, R. E., *Interval Analysis*, Prentice-Hall Inc., Englewood Cliffs, NJ, 1966.

Moore, R. E., *Methods and Applications of Interval Analysis*, Siam, Philadelphia, 1979.

Myers, Raymond H., Andre I. Khuri, and Geoffrey Vining, "Response Surface Alternatives to the Taguchi Robust Parameter Approach," *The American Statistician*, Vol. 46, No. 2, pp. 131-139, 1992.

Myers, R. H. and W. H. Carter, Jr., "Response Surface Techniques for Dual Response Systems," *Technometrics*, Vol. 15, pp. 301-317, 1973.

Paynter, H. M., *Analysis and Design of Engineering Systems*, M.I.T. Press, Cambridge, Mass., 1968.

Phadke, M., *Quality Engineering Using Robust Design*, Englewood Cliffs, NJ, Prentice-Hall, 1989.

Phadke, M. S., "Quality Engineering Using Design of Experiments," *Proceedings of the American Statistical Association, Section on Statistical Education*, Cincinnati, pp. 11-20, 1982.

Phadke, Madhav S., "Design Optimization Case Studies", *AT&T Technical Journal*, AT&T, 1986.

Pignatiello, Jr., Joseph J., "Strategies for Robust Multiresponse Quality Engineering," *IIE Transactions*, Vol. 25, No. 3, pp. 5-15, 1993.

Rackwitz, R., and B. Fiessler, "Structural Reliability under Combined Random Load Sequences," *Computers & Structures*, Vol. 9, pp. 489-494, 1978.

Rosenblatt, M., "Remarks on a Multivariate Transformation," *The Annals of Mathematical Statistics*, Vol. 23, pp. 470-472, 1952.

Rowell, Derek, and David N. Wormley, *Systems Dynamics: An Introduction*, Prentice Hall, NJ, 1997.

Savage, G. J., "Automatic Formulation of Higher-Order Sensitivity Models," *Civil Engineering Systems*, Vol. 10, No. 4, pp. 335-350, 1993.

Savage, G. J., "MAPLE as Design Tool: The Teacup Problem," *To be submitted to MAPLE Journal*, University of Waterloo, 1997.

Savage, G. J., "Solution of Non-linear Engineering Systems Through Subsystems," *Advances in Engineering Software*, Vol. 28, pp. 247-258, 1997.

Savage, G. J., and A. G. Row, "Uncertainty Modelling Through Subsystems," *Modeling and Simulation*, Vol. 23, No. 1, pp. 333-341, 1992.

Schjaer-Jacobsen, Hans, and Kaj Madsen, "Algorithms for Worst-Case Tolerance Optimization," *IEEE Transactions on Circuits and Systems*, Vol. CAS-26, No. 9, pp. 775-783, 1979.

Seifi, Abbas, K. Ponnambalam, and Jiri Vlach, "Probabilistic Design of Integrated Circuits with Correlated Input Parameters, *Submitted to IEEE Circuits and Systems*, University of Waterloo, January 28, 1997.

Skelboe, Stig, "True Worst-Case Analysis of Linear Electrical Circuits by Interval Arithmetic," *IEEE Transactions on Circuits and Systems*, Vol. 26, No. 10, pp. 874-879, 1979.

Song, Alan A., Amit Mathur, and Krishna R. Pattipati, "Design of Process Parameters Using Robust Design Techniques and Multiple Criteria Optimization", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 25, No. 11, 1437-1446, 1995.

Spence, Robert, and Randeep Singh Soin, *Tolerance Design of Electronic Circuits*, Addison Wesley, Wokingham, England, 1988.

Swan, D. A., *Bounds Analysis: The Application of Interval Mathematics to Graph-Theoretic Modelling.* Master's Thesis, Department of Systems Design, University of Waterloo, 1993.

Swan, D. A. and G. J. Savage, "The Use of Intervals in Design Calculations: The Application of Interval Mathematics to Graph-Theoretic Modelling," *Proceedings of the ISSAT International Conference on Reliability and Quality in Design* (Ed: Hoang Pham), pp.159-163, Seattle, Washington, March 16-18, 1994.

Swan, D. A., G. J. Savage, P. L. Cooper, and S. M. Carr, "Defining Quality and Reliability: A Linear Graph Model Perspective," *Invited Paper for Special Session on Recent Advances in Graph-Theoretic Modelling and Analysis*, IEEE Conference on Systems, Man, and Cybernetics, to be held in Orlando, Florida on October 12-15, 1997.

Taguchi, G., "Performance Analysis Design," *International Journal of Production Research*, Vol. 16, pp. 521-530, 1978.

Taguchi, G., and M. S. Phadke, "Quality Engineering Through Design Optimization," *IEEE Global Telecommunications Conference*, Globecom '84, Atlanta, GA, pp. 1106-1113, November 26-29, 1984.

Taguchi, Genichi, *Taguchi Methods: Research and Development.* ASI Press, Tokyo, 1992.

Taguchi, Genichi, *Taguchi on Robust Technology Development: Bringing Quality Engineering Upstream.* ASME Press, New York, 1993.

Taylor, Wayne A., *Optimization & Variation Reduction in Quality*. McGraw-Hill, New York, 1991.

Tribus, M., *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York, 1969.

Waterloo Maple Software 766884 Ontario Inc., *MAPLE-V for Windows 95*, 450 Phillip Street, Waterloo, Ontario, Canada, N2L 5J2, 1996.

Welch, William J., and Jerome Sacks, "A System for Quality Improvement Via Computer Experiments," *Communications in Statistics*, Vol. 20, No. 2, pp. 477-495, 1991.

Welch, William J., Robert J. Buck, Jerome Sacks, Henry P. Wynn, Toby J. Mitchell, and Max D. Morris, "Screening, Predicting, and Computer Experiments", *Technometrics*, Vol. 34, No. 1, pp. 15-25, 1992.

Wills, B. L., "Dynamic Sensitivity Co-efficients for RLC Networks," *Proceedings of the IEEE Conference on Circuits, Systems and Computers*, Houston, Texas, pp. 378-389, May 1969.

Wojciechowski, Jacek M., and Jiri Vlach, "Ellipsoidal Method for Design Centering and Yield Estimation", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 12, No. 10, pp. 1570-1579, 1993.

Wojciechowski, Jacek, Jiri Vlach, and Leszek Opalski, "Design for Nonsymmetrical Statistical Distributions," *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, Vol. 44, No. 1, pp. 29-37, 1997.

Wu, Y.-T. and P. H. Wirsching, "New Algorithm for Structural Reliability Estimation," *Journal of Engineering Mechanics*, Vol. 113, No. 9, pp. 1319-1335, 1987.

Xie, Cong-Rong, *Design for Quality: A System Theoretic Approach*, Master's Thesis, University of Waterloo, Waterloo, 1994.

Appendix A

U↔V Transformations

In general, variables are not normally distributed (Wojciechowski, Vlach, and Opalski, 1997). However, through transformation into the standard normal probability space, the probability contents of the non-normal distributions are easily approximated. In order to accomplish this, a one-to-one transformation is required:

$$\mathbf{T}: \mathbf{V} = (V_1, V_2, \ldots, V_n) \rightarrow \mathbf{U} = (U_1, U_2, \ldots, U_n) \tag{A.1}$$

where $U_1, U_2, \ldots, U_n$ are uncorrelated and standardized normally distributed. This chapter will detail the transformation of specified independent distributions (normal, lognormal, uniform, triangular, truncated normal, normal with center missing) and correlated (normal, lognormal) into the standard normal probability space. The correlated distributions dealt with will be restricted to normal and lognormal since they are most common.

A.1 Independent Distribution Transformation

The simplest definition of the transformation **T** occurs when the distributions of the variables are mutually independent, and therefore, each variable can then be transformed separately. Madsen, Krenk, and Lind (1986) and Melchers (1987) show the identity:

$$\Phi(u_i) = F_{V_i}(v_i), \quad i = 1 \text{ to } n \tag{A.2}$$

where $F_{V_i}(v_i)$ is the cumulative distribution function for variable, v_i. Thus, the transformations are:

$$\mathbf{T}: u_i = \Phi^{-1}\left(F_{V_i}(v_i)\right), \quad i = 1 \text{ to } n \tag{A.3}$$

$$\mathbf{T}^{-1}: v_i = F_{V_i}^{-1}\left(\Phi(u_i)\right), \quad i = 1 \text{ to } n \tag{A.4}$$

In addition, the density ratio is required for evaluation of the gradient of the margin with respect to u.

A.1.1 Normal

Independent normal distributions are the simplest to transform into standard normal space. A normal distribution has a probability density function (PDF):

$$f(v)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right\};\quad -\infty<v<\infty \tag{A.5}$$

where the distribution is characterized by its mean, μ, and standard deviation, σ. Therefore, $f(v)$ is the value of the PDF at point v. It is well known (Kalbfleisch 1979) that the transformation into standard normal space ($v \rightarrow u$) is accomplished by letting:

$$\mathbf{T}: u=\frac{v-\mu}{\sigma} \tag{A.6}$$

Through the use of the chain rule, we find:

$$\begin{aligned}\varphi(u)=f(v)\cdot\left|\frac{dv}{du}\right| &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right\}\cdot\sigma \\ &= \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}u^2\right\} \text{ for } -\infty<u<\infty\end{aligned} \tag{A.7}$$

which is the PDF of the standardized normal distribution. To accomplish the inverse transformation ($u \rightarrow v$), we need:

$$\mathbf{T}^{-1}: v=\mu+\sigma u \tag{A.8}$$

to find:

$$\begin{aligned}f(v)=\varphi(u)\cdot\left|\frac{du}{dv}\right| &= \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}u^2\right\}\cdot\frac{1}{\sigma} \\ &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{1}{2}\left(\frac{v-\mu}{\sigma}\right)^2\right\} \text{ for } -\infty<v<\infty\end{aligned} \tag{A.9}$$

Thus, for the probability method, equation (A.6) and (A.8) will be used to move between **V** and **U** space. The density ratio is:

$$\frac{\varphi(u)}{f(v)}=\sigma \tag{A.10}$$

A.1.2 Lognormal

For a lognormal distribution with mean, λ, and standard deviation, ξ, the PDF is:

$$f(v)=\frac{1}{\sqrt{2\pi}v\xi}\exp\left\{-\frac{1}{2}\left(\frac{\log v-\lambda}{\xi}\right)^2\right\} \text{ for } 0\le v<\infty \qquad (A.10)$$

where λ and ξ are found by transforming a normal distribution (μ, σ^2) into a lognormal distribution.

$$\lambda=\log\mu-\frac{1}{2}\log\left(\frac{\sigma^2}{\mu^2}+1\right) \qquad (A.12)$$

$$\xi^2=\log\left(\frac{\sigma^2}{\mu^2}+1\right) \qquad (A.13)$$

or inversely,

$$\mu=\exp\left(\lambda+\frac{1}{2}\xi^2\right) \qquad (A.14)$$

$$\sigma^2=\exp\left(2\cdot\left(\lambda+\xi^2\right)\right)-\exp\left(2\lambda+\xi^2\right)=\mu^2\left(e^{\xi^2}-1\right) \qquad (A.15)$$

Thus, the transformations to and from **U** space are based on (A.6) and (A.8) and are:

$$\mathbf{T}: u=\frac{\log v-\lambda}{\xi} \qquad (A.16)$$

$$\mathbf{T}^{-1}: \log v=\lambda+\xi u \rightarrow v=\exp(\lambda+\xi u) \qquad (A.17)$$

The corresponding density ratio is:

$$\frac{\varphi(u)}{f(v)}=v\xi \qquad (A.18)$$

A.1.3 Uniform

For a uniform distribution as in Figure A-1,

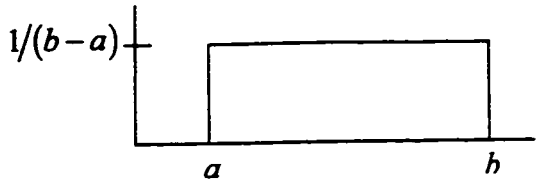


Figure A-1: Uniform Distribution

the PDF is:

$$f(v)=\begin{cases}1/(b-a) & \text{for } a \le v \le b \\ 0 & \text{otherwise}\end{cases} \tag{A.19}$$

the transformations are defined as:

$$\mathbf{T}: u=\Phi^{-1}\left(\frac{v-a}{b-a}\right) \tag{A.20}$$

$$\mathbf{T}^{-1}: v=(b-a)\Phi(u)+a \tag{A.21}$$

and the density ratio is:

$$\frac{\varphi(u)}{f(v)}=(b-a)\varphi(u) \tag{A.22}$$

A.1.4 Triangular

For a symmetrical triangular distribution as in Figure A-2,

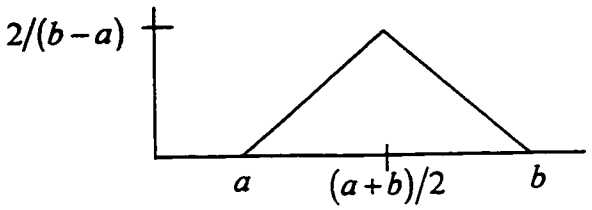


Figure A-2: Triangular Distribution

the PDF is:

$$f(v)=\begin{cases}\dfrac{4(v-a)}{(a-b)^2} & \text{for } a\le v\le (a+b)/2\\ \dfrac{4(b-v)}{(a-b)^2} & \text{for } (a+b)/2\le v\le b\\ 0 & \text{otherwise}\end{cases} \tag{A.23}$$

the transformations are defined as:

$$\mathbf{T}:\begin{cases}u=\Phi^{-1}\left(\dfrac{2(v-a)^2}{(a-b)^2}\right) & \text{if } v\le (a+b)/2\\ u=\Phi^{-1}\left(1-\dfrac{2(b-v)^2}{(a-b)^2}\right) & \text{if } v>(a+b)/2\end{cases} \tag{A.24}$$

$$\mathbf{T}^{-1}:\begin{cases}v=a+\dfrac{\sqrt{2}}{2}\sqrt{\Phi(u)(a-b)^2} & \text{if } u\le 0\\ v=b-\dfrac{\sqrt{2}}{2}\sqrt{(1-\Phi(u))(a-b)^2} & \text{if } u>0\end{cases} \tag{A.25}$$

and the density ratio is:

$$\frac{\varphi(u)}{f(v)}=\frac{1}{4}(a-b)^2\varphi(u)\cdot\begin{cases}1/(v-a) & \text{for } a\le v\le (a+b)/2\\ 1/(b-v) & \text{for } (a+b)/2\le v\le b\end{cases} \tag{A.26}$$

A.1.5 Truncated Normal

For a symmetrical truncated normal distribution, as in Figure A-3,

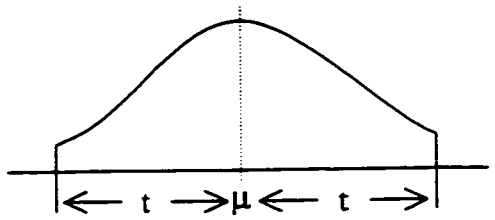


Figure A-3: Truncated Normal Distribution

the PDF is:

$$f(v)=\begin{cases}\dfrac{1}{2\sqrt{2\pi}\sigma t}\exp\left\{-\dfrac{1}{2}\left(\dfrac{v-\mu}{\sigma}\right)^2\right\} & \text{for } \mu(1-t)\le v\le\mu(1+t)\\ 0 & \text{otherwise}\end{cases} \tag{A.27}$$

where t is a tolerance ranging 0 to 0.5, where 0 implies an impulse at the mean, and 0.5 implies a full normal distribution. The transformations are:

$$\mathbf{T}: u=\Phi^{-1}\left(\frac{\Phi\left((v-\mu)/\sigma\right)-(0.5-t)}{2t}\right) \tag{A.28}$$

$$\mathbf{T}^{-1}: v=\mu+\sigma\left(\Phi^{-1}\left(2t\Phi(u)+(0.5-t)\right)\right) \tag{A.29}$$

and the density ratio is:

$$\frac{\varphi(u)}{f(v)}=2\sigma t \tag{A.30}$$

A.1.6 Normal with Center Missing

For a normal distribution with a symmetrical center portion missing, as in Figure A-4,

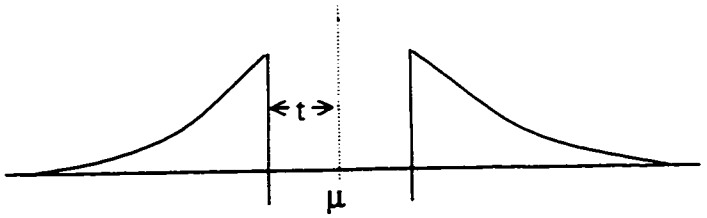


Figure A-4: Normal Distribution with Center Missing

the PDF is:

$$f(v)=\begin{cases}\dfrac{0.5}{(0.5-t)\sqrt{2\pi}\sigma}\exp\left\{-\dfrac{1}{2}\left(\dfrac{v-\mu}{\sigma}\right)^2\right\} & \text{for } v\le\mu(1-t) \\ & \text{or } v\ge\mu(1+t) \\ 0 & \text{otherwise}\end{cases} \tag{A.31}$$

where t is a tolerance ranging 0 to 0.5, where 0 implies a full normal distribution, and 0.5 implies two impulses, one at positive infinity and the other at negative infinity. The transformations are:

$$\mathbf{T}: u=-sign\left(\frac{v-\mu}{\sigma}\right)\Phi^{-1}\left(\Phi\left(-\left|\frac{v-\mu}{\sigma}\right|\right)\cdot\left(\frac{0.5}{0.5-t}\right)\right) \tag{A.32}$$

$$\mathbf{T}^{-1}: v=\mu+\sigma\cdot\left(-sign(u)\right)\cdot\Phi^{-1}\left(\Phi\left(-|u|\right)\cdot\left(\frac{0.5-t}{0.5}\right)\right) \tag{A.33}$$

and the density ratio is:

$$\frac{\varphi(u)}{f(v)}=\sigma\left(\frac{0.5}{0.5-t}\right) \tag{A.34}$$

A.2 Correlated Distribution Transformation

When correlation in variables is encountered, the Rosenblatt (1952) transformation is required to determine the transformation mechanism to and from **U** space. Correlation is normally restricted to variables with similar distributions. For example, the roughness of a pipe is often characterized by a normal distribution and is correlated with the roughness of other pipes in the system to ensure that the roughness throughout the system is similar. In this example, it is normal distributions correlated with normal distributions. It is rare, if at all, that correlation between different distribution types occurs. Therefore, in this thesis, only correlation between like distributions will be handled. In addition, it will be further restricted to correlation of normal-like distributions (normal, lognormal, truncated normal, normal with center missing).

A.2.1 General Case

When the variables are not mutually independent, Hohenbichler and Rackwitz (1981) suggested the Rosenblatt transformation. The transformation is defined similar to (A.3) and (A.4) and is:

$$\mathbf{T}:\begin{cases} u_1 = \Phi^{-1}\left(F_{V_1}(v_1)\right) \\ u_2 = \Phi^{-1}\left(F_{V_2}(v_2|v_1)\right) \\ \quad\vdots \\ u_i = \Phi^{-1}\left(F_{V_i}(v_i|v_1, v_2, \cdots, v_{i-1})\right) \\ \quad\vdots \\ u_n = \Phi^{-1}\left(F_{V_n}(v_n|v_1, v_2, \cdots, v_{n-1})\right) \end{cases} \tag{A.35}$$

$$\mathbf{T}^{-1}:\begin{cases} v_1 = F_{V_1}^{-1}\left(\Phi(u_1)\right) \\ v_2 = F_{V_2}^{-1}\left(\Phi(u_2)|u_1\right) \\ \quad\vdots \\ v_n = F_{V_n}^{-1}\left(\Phi(u_n)|u_1, u_2, \cdots, u_{n-1}\right) \end{cases} \tag{A.36}$$

In most cases, **T** and its inverse are determined numerically.

A.2.2 Correlated Normal-like Distributions

The mechanism through which correlation is achieved in normal-like distributions lies in the use of the covariance matrix and Cholesky factorization (Melchers 1987). For correlated normal distributions, the transformations are:

$$\mathbf{T}:\mathbf{U} = \left(Chol(\mathbf{C}_\sigma)^T\right)^{-1}(\mathbf{V} - \mu_\mathbf{V}) \tag{A.37}$$

$$\mathbf{T}^{-1}:\mathbf{V} = \mu_{\mathbf{V}} + Chol(\mathbf{C}_{\sigma})^{T}\mathbf{U} \tag{A.38}$$

where **U** and **V** are vectors of the transformation variables, $\mu_{\mathbf{V}}$ is a vector of the mean values corresponding to the **V** variables, $\mathbf{C}_{\sigma}$ is the covariance matrix, and *Chol* is the Cholesky factorization in upper triangular form.

This form can be extended into the other normal-like distributions such as the lognormal and truncated normal:

Table A-1: Correlated Lognormal and Truncated Normal Distributions

Lognormal	$\mathbf{T}:\mathbf{U} = \left(Chol\left(\mathbf{C}_{\xi}\right)^{T}\right)^{-1}\left(\log \mathbf{V} - \lambda_{\mathbf{V}}\right)$	(A.39)
	$\mathbf{T}^{-1}:\mathbf{V} = \exp\left(\lambda_{\mathbf{V}} + Chol\left(\mathbf{C}_{\xi}\right)^{T}\mathbf{U}\right)$	A.40)
Truncated Normal	$\mathbf{T}:\mathbf{U} = \Phi^{-1}\left(\dfrac{\Phi\left(\left(Chol(\mathbf{C}_{\sigma})^{T}\right)^{-1}\left(\mathbf{V} - \mu_{\mathbf{V}}\right)\right) - (0.5 - \mathrm{t})}{2\mathrm{t}}\right)$	(A.41)
	$\mathbf{T}^{-1}:\mathbf{V} = \mu_{\mathbf{V}} + Chol(\mathbf{C}_{\sigma})^{T}\left(\Phi^{-1}\left(2\mathrm{t}\Phi(u) + (0.5 - \mathrm{t})\right)\right)$	(A.42)

The normal with center missing distribution is a little more complicated and will not be shown here.

The density function is found as (Madsen, Krenk, and Lind 1986):

$$\frac{\partial v_j}{\partial u_i} = \begin{cases} 0 & i < j \\ \dfrac{\varphi(u_i)}{f_i\left(v_i \middle| v_1, \cdots, v_{i-1}\right)} & i = j \\ \dfrac{\varphi(u_i)}{\dfrac{\partial F_i}{\partial v_i}\left(v_i \middle| v_1, \cdots, v_{i-1}\right)} & i > j \end{cases} \tag{A.43}$$

Appendix B

Maple and Matlab Code

Steps to determining the performance of a system

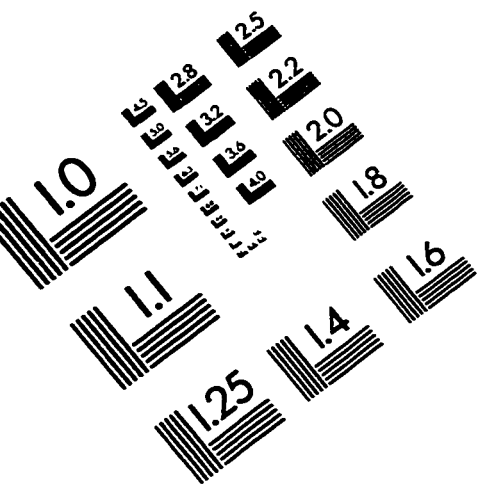
1. Open a Maple session
2. Find the system equations:
    - There are many ways to handle this. Steve Carr has written a number of files to do this.
        - Load file ('matcalc') which defines some matrix algebra,
        - Load file ('sysinfo'), the input file, which contains all the system information,
        - Load file ('system') that takes the input file, stamps up the model, and solves it.
    - The result is a symbolically solved system of equations.
3. Convert the design_parameters vector into a matrix
4. Define the U_V relationship
    - Load file ('xdata')
5. Convert the Maple code to Matlab code
    - Load files ('matlab', 'mproc', and 'gradm')
    - Use gradm function to convert code
    - The output is a file ('output')
6. Copy the Matlab output ('output') into the files:
    - 'gradm' – the gradient information of the function wrt design variables
    - 'znew' – the set of system equations
    - 'data' – the nominal values and tolerances of the parameters
7. Ensure that the following files are correct:
    - 'margins' – correct number
    - 'optfun' – correct number of constraints, correct objective function
    - 'optgrad' – correct gradient info for objective function and constraints
8. Open Matlab session
9. Load data file – the problem is solved

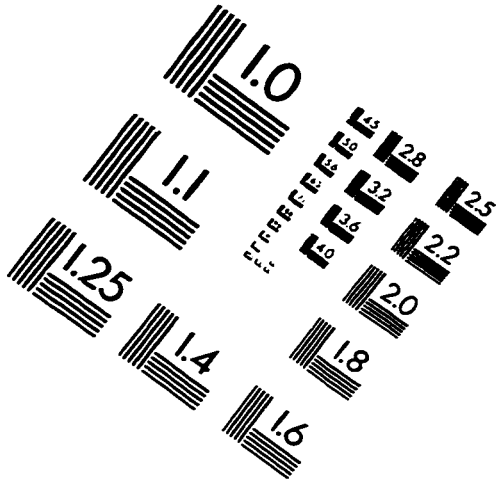
There are several other support files that are needed:

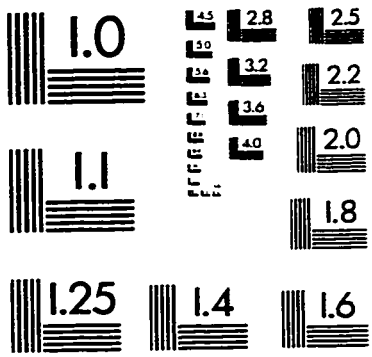
- 'phi', 'pdf', 'newton', 'rf', 'probdu', plus the optimization toolbox from Matlab

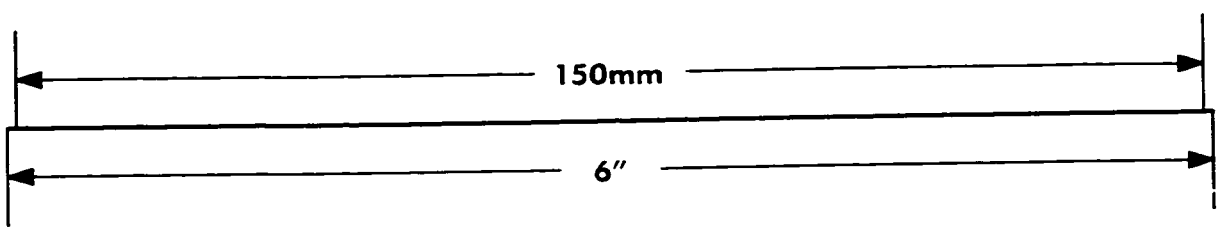
The only file that may have to be changed is 'probdu' depending on the distribution type used in the system.

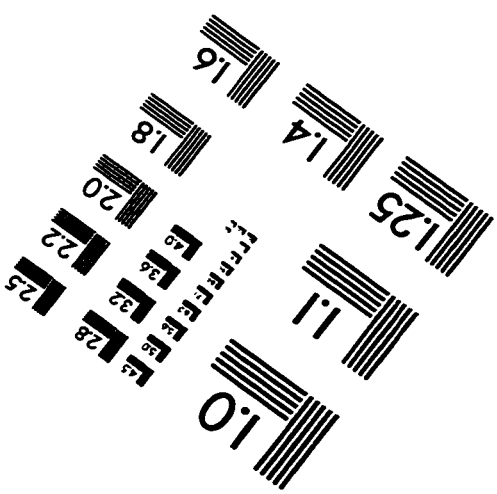
IMAGE EVALUATION
TEST TARGET (QA-3)

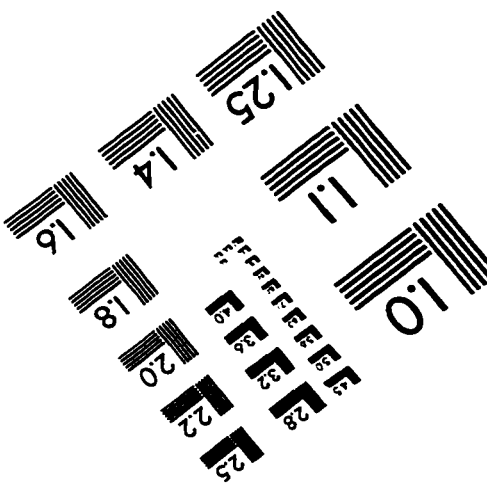












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989