A SECOND-ORDER METHOD FOR ASSEMBLY

TOLERANCE ANALYSIS

A Thesis

Presented to the

Department of Mechanical Engineering

Brigham Young University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Charles G. Glancy

December 1994

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express appreciation to my research advisor Dr. Ken Chase for his help with this thesis. His guidance, feedback and, of course, research funding were important to the development of this new tolerance analysis method. Most of all, I thank Dr. Chase for his encouragement of my research and the confidence he placed in my abilities.

I express special appreciation to my wife Erika for her devoted support and understanding. I thank my wife for the sacrifices she made and for caring for our family during the months of research. Also, I am grateful for wonderful parents who always encouraged me to do my best. Last of all, I thank God, for truly He is the source of all inspiration and knowledge.

Chapter 1

INTRODUCTION

Tolerance analysis is increasingly becoming an important tool for mechanical design. This seemingly arbitrary task of assigning dimension tolerances can have a large effect on the cost and performance of manufactured products. With the increase in competition in today's marketplace, small savings in cost or small increases in performance may determine the success of a product.

Currently there are two prominent methods for mechanical assembly tolerance analysis: Monte Carlo simulation and the Linearized Method. There is quite a difference in capabilities between these two methods. Both Monte Carlo simulation and the Linearized Method have desirable features. The Linearized Method is well suited for design iteration because of its speed, and Monte Carlo simulation can provide very accurate results when large sample sizes are used. Large sample sizes, however, are computationally intensive. A fast *and* accurate tolerance analysis method would be ideal.

1.1 Research Goal

The goal of this research was to develop a statistical, second-order tolerance analysis method in an attempt to combine the accuracy of Monte Carlo simulation and the speed of the Linearized Method. Before stating the challenge that existed in implementing such a method, there are a few terms that must be defined. In addition, the details of Monte Carlo simulation and the Linearized Method need to be explained.

1.2 Definition of Terms

1.2.1 Tolerance Analysis vs. Tolerance Allocation

The two basic tasks in variation analysis are tolerance analysis and tolerance allocation. *Tolerance analysis* estimates the effect of manufacturing variation on assembly dimensions. Assembly dimensions are lengths and angles involving two or more different component parts. Tolerance analysis is sometimes called "tolerance stackup analysis" because variations in the component dimensions accumulate statistically, contributing to increased variation in the assembly dimensions.

Tolerance analysis provides an important link between design and manufacturing. The variation of an assembly will affect both the performance as well as the cost of manufacture. For example, tight component tolerances may yield excellent performance but high cost, while loose tolerances may reduce production cost yet cause unsatisfactory product performance. Tolerance analysis is a tool that quantifies the effect of assembly component variation on performance specifications.

Tolerance allocation is a closely related tool to tolerance analysis, but it works in the opposite direction. Tolerance allocation determines an appropriate set of component tolerances which meets the specified assembly variation limits. To further contrast the two, *tolerance analysis* estimates the percent rejects resulting from the specified component variations, while *tolerance allocation* determines a set of component tolerances which will meet the specified allowable reject limits. Both tolerance analysis and allocation empower the mechanical designer to incorporate the effects of manufacturing variation during the early stages of product design.

1.2.2 Monte Carlo Simulation

Monte Carlo simulation is an established method for doing assembly tolerance analysis. This method is based on random number generation. The manufacture of an assembly is simulated, for example, by creating a set of component dimensions with small random changes to simulate natural process variations. Next, the resulting assembly dimensions are calculated from the simulated set of component dimensions. These two steps, the creation of random component dimensions and the calculation of the resulting assembly dimensions, are repeated between 5,000 to 100,000 times based on the required accuracy of the simulation. The accuracy of Monte Carlo simulation is theoretically unlimited, however, high levels of accuracy require large sample sizes. For instance, to accurately predict the number of rejects for high assembly quality levels requires sample sizes of 100,000 or more. Obviously, the computational effort of such a simulation can be large, but Monte Carlo offers many advantages because of its flexibility. Monte Carlo simulation allows any component distribution to be specified and will output the resulting assembly distribution.

1.2.3 Vector-loop Assembly Model

Vector loops can be used to model manufactured assemblies. Figure 1.1 shows an example of a two-dimensional assembly described by three vector loops. A vector-loop model of an assembly mathematically establishes how the manufactured lengths and angles of each component combine in order to properly assemble together. The vector loops are able to model dimensional, form and kinematic variations.

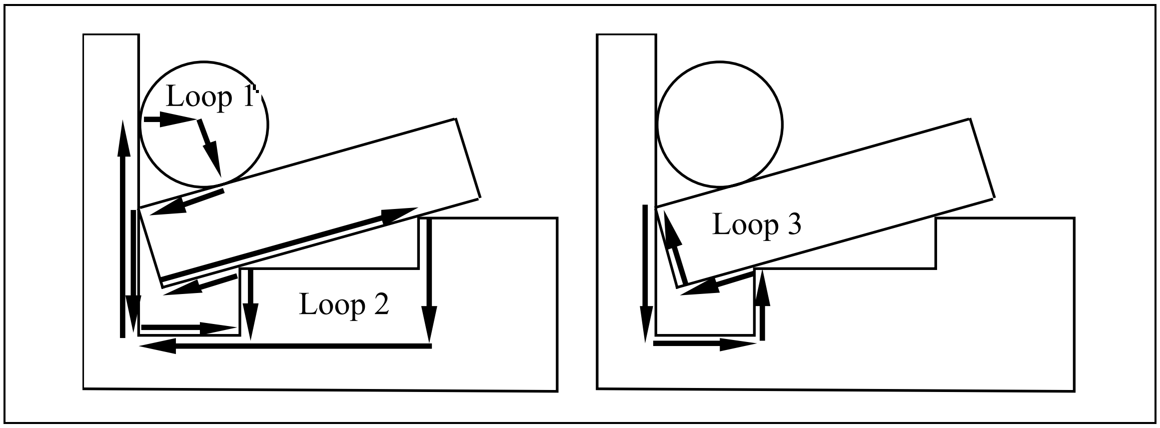

Figure 1.1: Vector loops for a 2-D assembly

Vector-loop closure is an important condition for assembly tolerance analysis. Closure simply refers to the condition when the beginning of the vector loop is the same position and orientation as the end of the loop. Loop closure is the mathematical equivalent of an assembly fitting together. The loop closure condition can be written as the system of nonlinear equations

$$\mathbf{h}(\mathbf{x},\mathbf{u}) = \mathbf{0} \tag{1.1}$$

where **h** is the system of loop equations, **x** is the set of vectors representing manufactured component dimensions, and **u** is the set of vectors representing unknown assembly lengths and angles. The unknown assembly lengths and angles are the kinematic assembly dimensions which change as a function of the component dimensions.

An example of kinematic assembly dimensions is shown in Figure 1.1 above. The assembly consists of a block, resting on two steps, and a cylinder resting against the block and upright wall. The corners of the two steps make contact at two points on the bottom surface of the block. The location of the two contact points relative to the end of the block are not component dimensions. The location of these two points are assembly dimensions which depend on the height and length of the steps. These assembly dimensions will change to accommodate small manufacturing variations in the step dimensions. All such kinematic dimensions are called the dependent variables. The component dimensions are called independent variables.

1.2.4 Linearized Method

The *Linearized Method* is a vector-loop-based method of assembly tolerance analysis. The method's name comes from the fact that the nonlinear equations of the vector-loop model are linearized for the analysis. The linearized equations determine how small changes of the component dimensions, form and contact affect an assembly. For this method only one assembly needs to be analyzed statistically. Linear analysis is extremely fast and allows for tolerance allocation and design iteration. It is, however, limited to normal component distributions and cannot be applied to non-normal assembly distributions.

When tolerances are small compared to the nominal dimension, on the order of 1/100 to 1/1000, the Linearized Method gives excellent results. A comparison [Gao 1993] between the Linearized Method and Monte Carlo simulation found that the accuracy of the Linearized Method corresponded to Monte Carlo simulation with a sample size of 30,000, for quality levels near three sigma. For four sigma quality levels, Monte Carlo simulation would require a sample size 100 times larger. However, for highly nonlinear assemblies or highly skewed distributions, the Linearized Method would lose accuracy.

The Linearized Method expands the loop closure equation, Equation 1.1, for small variations about the nominal by Taylor's series expansion, retaining first order derivatives. This expansion yields:

$$dh_i = \sum_{j=1}^{n} \frac{\partial h_i}{\partial x_j} dx_j + \sum_{j=1}^{m} \frac{\partial h_i}{\partial u_j} du_j = 0 \tag{1.2}$$

where dx_j are the specified tolerances of the component dimensions and du_j are the resultant variations in the dependent assembly dimensions. This expression can be put in vector form by forming the matrix **A** of partial derivatives $\frac{\partial h_i}{\partial x_j}$ and the matrix **B** of the partial derivatives $\frac{\partial h_i}{\partial u_j}$.

$$[\ \mathbf{A}\ ]\ \{\mathbf{dx}\} + [\ \mathbf{B}\ ]\ \{\mathbf{du}\} = \{\ \mathbf{0}\ \} \tag{1.3}$$

Solving for **du**:

$$\{\mathbf{du}\} = -[\ \mathbf{B}^{-1}]\ [\ \mathbf{A}\ ]\ \{\mathbf{dx}\} \tag{1.4}$$

Therefore, the product of the matrices $-\mathbf{B}^{-1}\mathbf{A}$ gives the sensitivities of the dependent assembly dimension with respect to the component dimensions. Having established this relationship, the statistical effect of the component dimension tolerances on the dependent assembly dimension variations may be estimated by the root sum squares expression:

$$du_i = \sqrt{\sum \left(\frac{\partial u_i}{\partial x_j} dx_j \right)^2} \tag{1.5}$$

where $\frac{\partial u_i}{\partial x_j}$ are the elements of the $-\mathbf{B}^{-1}\mathbf{A}$ matrix.

The formulation of the Linearized Method allows the implicit assembly dimensions in the loop equations to be expressed as an explicit, statistical function of the component dimensions.

1.3 Statement of the Thesis

The Linearized Method and Monte Carlo simulation provide different capabilities. The Linearized Method can perform an analysis and a tolerance allocation quickly, so it is suitable for design iteration. Linear analysis is limited in that it cannot output non-normal distributions or handle non-normal component distributions. Also, the Linearized Method will not be accurate for highly nonlinear assemblies. Monte Carlo simulation allows non-normal input distributions and a nonlinear analysis. However, Monte Carlo simulation is computationally expensive and does not accommodate design iteration, closed-loop constraints or tolerance allocation. Table 1.1 summarizes the features of the Linearized Method, Monte Carlo simulation and the Second-Order Tolerance Analysis (SOTA) method proposed in this thesis.

Table 1.1: Features of Three Tolerance Analysis Methods

Features	Linear Method	Monte Carlo Simulation	SOTA Method
1. Speed	✓		✓
2. Tolerance allocation	✓		✓
3. Closed-loop constraints	✓		✓
4. Nonlinear model		✓	✓
5. Non-normal input distributions		✓	✓
6. Non-normal output distributions		✓	✓

The ideal assembly tolerance analysis method would include the desirable features of both the Linearized Method and Monte Carlo simulation and at the same time avoid the drawbacks of either method. A statistical, second-order tolerance analysis method has been developed which incorporates the features of speed, tolerance allocation, closed-loop constraints, nonlinear model, non-normal input distributions and non-normal output distributions.

A logical starting point for the SOTA method was to extend the formulation of the Linearized Method to include second-order terms. The resulting method from this extension still used a vector-loop-based tolerance model; however, it required a second-order statistical technique. A major challenge in implementing such a statistical technique is described in the following section.

1.4 Research Challenge

Second-order statistical techniques calculate mean, variance and two higher statistical moments using second-order assembly information. For example, the Method of System Moments [Cox 1984] requires partial derivatives of the dependent variables with respect to the independent variables, $\frac{\partial u_i}{\partial x_j}$, $\frac{\partial^2 u_i}{\partial x_j^2}$, and $\frac{\partial^2 u_i}{\partial x_j \partial x_k}$. Much of the computational effort of a tolerance analysis method can be spent in finding these derivatives. The Linearized Method is able to take advantage of the fact that partial derivatives of the loop equations can be found analytically. The only partial derivatives required for the linearized analysis are $\frac{\partial u_i}{\partial x_j}$. These values are found by multiplying two matrices of analytic derivatives, the **A** and **B** matrices shown in Section 1.2.4. Computational effort could be minimized if the SOTA method were also able to use analytic derivatives to construct the required partial derivatives.

To begin formulating a second-order method, the Taylor's series of the loop equations was expanded about the nominal again, this time retaining second-order derivatives:

$$\begin{aligned} dh_i = \sum_{j=1}^{n} \frac{\partial h_i}{\partial x_j} dx_j + \frac{1}{2} \sum_{j=1}^{n} \frac{\partial^2 h_i}{\partial x_j^2} dx_j^2 + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \frac{\partial^2 h_i}{\partial x_j \partial x_k} dx_j dx_k \\ + \sum_{j=1}^{m} \frac{\partial h_i}{\partial u_j} du_j + \frac{1}{2} \sum_{j=1}^{m} \frac{\partial^2 h_i}{\partial u_j^2} du_j^2 + \sum_{j=1}^{m-1} \sum_{k=j+1}^{m} \frac{\partial^2 h_i}{\partial u_j \partial u_k} du_j du_k = 0 \end{aligned} \tag{1.6}$$

In this Taylor's series expression, all the derivatives are partial derivatives of the loop equation, so they can be found analytically. Ideally, we would like to manipulate the Taylor's series expression into the form:

$$du_i = \sum_{j=1}^{n} \frac{\partial u_i}{\partial x_j} dx_j + \frac{1}{2} \sum_{j=1}^{n} \frac{\partial^2 u_i}{\partial x_j^2} dx_j^2 + \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} \frac{\partial^2 u_i}{\partial x_j \partial x_k} dx_j dx_k \tag{1.7}$$

However, by inspection, it can be seen that manipulating Equation 1.6 to obtain a general, explicit expression for du_i as a function of dx_j would be difficult. Since a general, explicit expression for du_i could not be found by the same algebraic procedure used for the Linearized Method, a different procedure needed to be formulated to obtain the required partial derivatives.

To analytically extract the second-order assembly information from a vector-loop model is difficult because of the implicit nature of the assembly dimensions in the loop equations. This fact suggested the use of a more costly numerical method in order to obtain the required information from the loop equations.

1.5 Research Objectives

The goal of this research, restated, was to develop a second-order tolerance analysis method with the specific benefits of speed, nonlinear effects, tolerance allocation and non-normal distribution capability. The objectives that had to be completed in order to reach this goal are listed in Table 1.2:

Table 1.2: Research Objectives for the SOTA method

1. Develop numerical difference formulas to obtain the required partial derivatives for the second-order statistical technique.
2. Develop a nonlinear system solver for the loop equations to solve for the assembly dimensions for a given set of component dimensions.
3. Select a second-order statistical analysis technique for estimating the first four moments of the assembly dimensions.
4. Select a family of distributions to match the moments of the assembly dimensions in order to estimate the assembly reject fraction.
5. Integrate objectives 1-4 into a computer program to perform the SOTA method on a wide range of tolerance analysis problems.
6. Benchmark the SOTA method against Monte Carlo simulation and the Linearized Method.

1.6 Research Scope

The SOTA method was developed to analyze the closed-loop, kinematic variations of two-dimensional assemblies. The method does not consider open-loop variations, three-dimensional assemblies, or geometric tolerances. Expanding tolerance allocation methodology from the linear model to a second-order model was not investigated.

1.7 Overview of Thesis

The presentation of this thesis will proceed as follows. Chapter 2 will review and analyze previous work relevant to a second-order tolerance analysis method. Chapter 3 will describe the SOTA method and Chapter 4 will present five case studies comparing the SOTA method with Monte Carlo simulation and the Linearized Method. Finally, Chapter 5 will state the conclusions and recommendations concerning this research.

Chapter 2

REVIEW AND ANALYSIS OF PREVIOUS WORK

This chapter will first review related work to this thesis in the areas of the Linearized Method and Monte Carlo Simulation. Next, two statistical techniques used for second-order tolerances analysis will be presented. Finally, three distribution families used for empirical fits will be discussed.

2.1 The Linearized Method

The Linearized Method, explained in Section 1.2.4, provides a quick way to perform nonlinear tolerance analysis for both explicit and implicit assembly dimensions of a vector-loop tolerance model. Because of its speed, the Linearized Method is ideal for design iteration and tolerance allocation. Multiple research studies have continued to refine the Linearized Method, making it more general and accurate.

The Direct Linearization Method (DLM) [Marler 1988] prescribed a systematic approach to vector-loop model tolerance analysis. DLM has enabled the Linearized Method to be applied to a broad range of tolerance problems. Most importantly, DLM has allowed a general tolerance analysis methodology to be incorporated into a computer program suitable for integration with a CAD system.

More recently, the Global Coordinate Method [Gao 1993] for determining the partial derivatives of the loop equations was developed. This method simplified the calculations of these derivatives. In the same paper, Gao benchmarked the Linearized Method against a comparable Monte Carlo simulation system. The benchmark results showed that the accuracy of the Linearized Method corresponds to Monte Carlo simulation with a sample size of 30,000 for quality levels of three sigma.

The Linearized Method has demonstrated its usefulness as a design tool. However, the method is inadequate for highly nonlinear tolerance problems and non-normal input distributions.

2.2 Monte Carlo Simulation

There are two papers which deserve special attention because of their correlation with this research. Both papers deal with incorporating aspects of the Linearized Method into a nonlinear tolerance analysis method, Monte Carlo simulation. The most recent paper incorporates the kinematic constraints of a vector-loop tolerance model into the Monte Carlo procedure. The second paper introduces the Optimal Design System, a system to add tolerance allocation to Monte Carlo simulation.

2.2.1 Monte Carlo with Kinematic Constraints

Generally, Monte Carlo simulation is applied to an *explicit* function of random variables. The variables of interest in the equations of a vector-loop tolerance model are inherently *implicit*. McCATS, a Monte Carlo based tolerance analysis method developed recently [Gao 1993], is able to adequately handle the implicit equations of a vector-loop tolerance model.

The McCATS system starts by generating random variates for the assembly variables. These random variates are sent to an assembly function which then solves the nonlinear system of loop equations iteratively for the dependent assembly dimensions. The assembly dimensions are stored, new random variates are generated, and the assembly function is called again. This procedure is continued until the desired number of assemblies have been simulated. Solving the loop equations iteratively for each assembly simulation is critical to the accuracy of the tolerance model.

Including the capability for kinematic constraints in Monte Carlo simulation enabled Monte Carlo methods to be applied to a much broader range of design problems. However, the required iterative assembly function does add more calculations to an already computationally intense method.

2.2.2 Monte Carlo with Tolerance Allocation

The two major drawbacks to using Monte Carlo simulation as a design tool are computation time and lack of tolerance allocation capability. The Optimal Design System (ODS) [Larsen 1989] overcomes the latter drawback. The ODS method is able to allocate tolerances for linear problems by incorporating aspects of the Method of System Moments.

The ODS method uses Monte Carlo simulation to calculate the first four statistical moments of the assembly distribution in terms of a set of polynomial functions of the component tolerances. The application of the Method of System Moments enables the component tolerances to be uncoupled from the component dimensions. As a result, after an initial simulation is run, the component tolerances can be changed and a new assembly distribution calculated without repeating the simulation. The results showed that the ODS method was very capable for tolerance allocation; however, this method was only applied to linear assembly functions. Still, this added feature of tolerance allocation greatly increases the usefulness of Monte Carlo simulation as a design tool.

The ODS method and McCATS show that features can be incorporated into Monte Carlo simulation which make it a more useful tool for design. However, because of the steps of random number generation and then assembly function calculation, which must be repeated over and over, it is still very inefficient for design.

The next section will discuss two alternatives to Monte Carlo simulation for nonlinear assemblies and non-normal distributions.

2.3 Second-Order Statistical Techniques

There are two well-developed, second-order statistical techniques for estimating moments of functions of random variables. The two techniques are Moments by Quadrature and the Method of System Moments. Generally, these two techniques are used to estimate the first four moments of a function of random variables. These four moments are described in Figure 2.1. Moments by Quadrature and the Method of System Moments are formulated for explicit functions of random variables, and therefore, cannot be applied directly to the implicit assembly dimensions of a vector-loop tolerance model.

The general procedure for using Moments by Quadrature and the Method of System Moments would be to

1. Determine the functional relationship between the random component dimensions and the resultant assembly dimensions.
2. Obtain the expressions for the first four moments of the assembly function.
3. Insert the distribution moments of the component dimensions and evaluate the moments of the assembly dimensions.
4. Fit an empirical distribution to the calculated moments of the assembly dimensions.

The empirical distribution fit is useful for approximating the density or estimating the fraction of acceptable assemblies. Both Moments by Quadrature and the Method of System Moments, which will be described in detail below, are good candidates for use in the SOTA method.

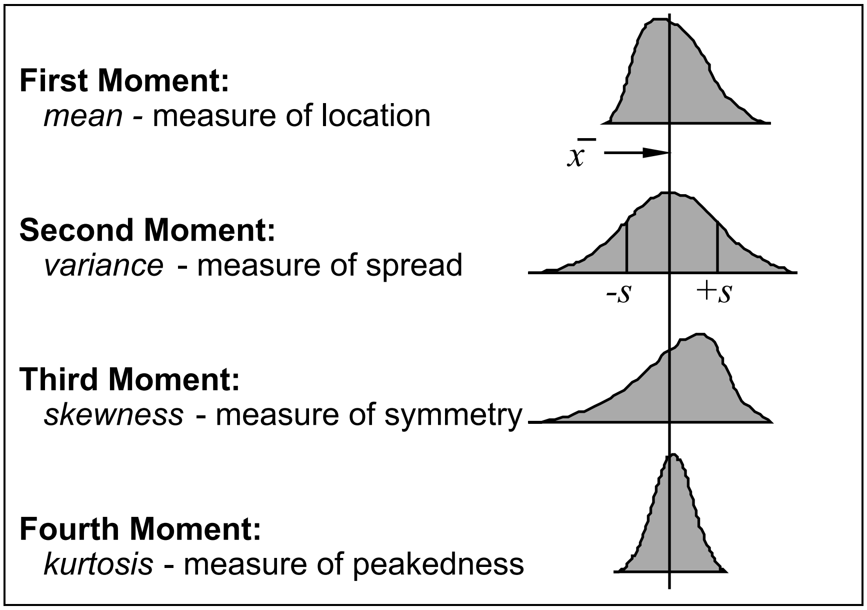

Figure 2.1: First Four Statistical Moments

2.3.1 Moments by Quadrature

The Moments by Quadrature [Evans 1975] technique for tolerance analysis is a way of estimating the moments of a function of random variables given a functional relationship between the inputs and the output and the distribution characteristics of the input variables. This method is useful where the analytic form of the function is not known. As the word quadrature would imply, numerical integration is the basis for this second-order method. The expected value for any function $f(x_1, x_2, \ldots, x_n)$ may be expressed in the integral form

$$I\{f\} = \int_{-\infty}^{+\infty} \cdots \int f(x_1, x_2, \ldots, x_n) \prod_k w_k(x_k)\, dx_k \tag{2.1}$$

where x_k are independent random variables with know densities $w_k(x_k)$, $k = 1,2,\ldots,n$. The integral expression can be approximated by the quadrature expression

$$Q\{f\} = C \cdot f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) + \sum_{j=1}^{n} H_j \left[\frac{f(\bar{x}_j + a_j^+ \sigma_j)}{a_j^+} - \frac{f(\bar{x}_j + a_j^- \sigma_j)}{a_j^-} \right] + \\ \sum_{j=1}^{n-1} \sum_{k=j+1}^{n} P_{jk} \left[\begin{array}{l} \dfrac{f(\bar{x}_j + b_j^+ \sigma_j, \bar{x}_k + b_k^+ \sigma_k)}{b_j^+ b_k^+} - \dfrac{f(\bar{x}_j + b_j^- \sigma_j, \bar{x}_k + b_k^+ \sigma_k)}{b_j^- b_k^+} - \\ \dfrac{f(\bar{x}_j + b_j^+ \sigma_j, \bar{x}_k + b_k^- \sigma_k)}{b_j^+ b_k^-} + \dfrac{f(\bar{x}_j + b_j^- \sigma_j, \bar{x}_k + b_k^- \sigma_k)}{b_j^- b_k^-} \end{array} \right] \tag{2.2}$$

where $C, H_j, P_{jk}, a_j^{\pm}, b_j^{\pm}$ are constants. The arguments of *f* not explicitly shown are at their nominal values, $x_i = \overline{x}_i$. This quadrature expression may be used to calculate the first four moments of an assembly function.

Moments by Quadrature requires *2n²+1* function evaluations to approximate the four moments of an assembly function, where *n* is the number of component dimensions. If the distribution of one assembly component is changed, two additional function evaluations must be made. If the nominal value of any component is changed, all *2n²+1* function evaluations must be performed once again.

2.3.2 Method of System Moments

The Method of System Moments (MSM) [Shapiro 1981, Cox 1984] is a technique of estimating system output based upon the relationship between input and output variables and information about the distribution of the inputs. MSM is also known as nonlinear propagation of error and propagation of moments. This method is formulated by expanding the function of interest in Taylor's series about its mean values. Retaining second-order terms the expansion yields:

$$u_i = u_i^0 + \sum_{j=1}^{n} \frac{\partial u_i}{\partial x_j}\left(x_j - \overline{x}_j\right) + \frac{1}{2}\sum_{j=1}^{n} \frac{\partial^2 u_i}{\partial x_j^2}\left(x_j - \overline{x}_j\right)^2 + \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} \frac{\partial^2 u_i}{\partial x_j \partial x_k}\left(x_j - \overline{x}_j\right)\left(x_k - \overline{x}_k\right) \qquad (2.3)$$

The term u_i^0 represents the nominal value of the assembly dimensions. The approximate system mean, the first distribution moment, is calculated by taking the expected value of the above expression, which gives:

$$E(u_i) = u_i^0 + \frac{1}{2}\sum_{j=1}^{n} \frac{\partial^2 u_i}{\partial x_j^2} \operatorname{var}(x_j) \qquad (2.4)$$

Equation 2.4 assumes that all component dimensions are mutually independent. Mutual independence of the input variables is a basic assumption of MSM.

In order to calculate the second distribution moment, the expected value of the Taylor's series in Equation 2.3 raised to the second power must be found. For the third and fourth moment the expected value of Equation 2.3 raised to the third and fourth power must be found. Obviously, the number of terms in the expressions for the higher moments increases dramatically.

The expected value operator must be applied to the total number of terms after raising Equation 2.3 to the correct power for the corresponding distribution moment. To illustrate how fast the number of terms increases for the higher moments, the number of terms resulting from raising Equation 2.3 to various powers may be calculated. The total number of terms resulting from raising the sum of *t* terms to the *p* power can be found from the expression

$$\frac{(p+t-1)!}{p!\,(t-1)!} \tag{2.5}$$

Since there are *(n (n + 3)) / 2* terms in Equation 2.3, where *n* is the number of component dimensions, the total number of terms for which the expected value must be found for MSM can be obtained from

$$\frac{\left(p+\left[\frac{n(n+3)}{2}\right]-1\right)!}{p!\left(\left[\frac{n(n+3)}{2}\right]-1\right)!} \tag{2.6}$$

For example, for an assembly with ten component dimensions the expected value is found from 65 terms for the first moment, 2,145 terms for the second moment, 47,905 terms for the third moment and 814,385 terms for the fourth moment. After the expected value operation is performed, the complexity of the four moment expressions is significantly reduced since covariance terms are assumed to be zero and the expected value of $(x_j - \bar{x}_j)$ is zero. However, even after the expected value operation, the expressions for the third and fourth moment are still tremendously long. Because of their length, only abbreviated expressions for the third and fourth moment will be given below.

The expressions for all first four moments are further simplified if the notation is modified, the origin shifted, and the *raw* moments calculated before calculating the moments about the mean.

For the new notation

$$b_j = \frac{\partial u_i}{\partial x_j} \quad b_{jj} = \frac{1}{2}\frac{\partial^2 u_i}{\partial x_j^2} \quad b_{jk} = \frac{\partial^2 u_i}{\partial x_j \partial x_k}$$

and the term $\mu_i(x_j)$ represents the *ith* distribution moment of the *jth* component dimension. The origin is shifted by introducing the variable y_i and setting $y_i = u_i - u_i^0$.

The moments of y_i, the raw moments, can be found directly as $E(y_i)$, $E(y_i^2)$, $E(y_i^3)$, and $E(y_i^4)$ and except for the first moment expression they are identical to the corresponding values for u_i.

The first four raw moments, listed in order from first to fourth, are

$$E(y_i) = \sum_{j=1}^{n} b_{jj}\mu_2(x_j) \tag{2.7}$$

$$\begin{aligned} E(y_i^2) &= \sum_{j=1}^{n}\left[b_j^2\mu_2(x_j) + 2b_jb_{jj}\mu_3(x_j) + b_{jj}^2\mu_4(x_j)\right] \\ &+ \sum_{j=1}^{n-1}\sum_{k=j+1}^{n}\left(2b_{jj}b_{kk} + b_{jk}^2\right)\mu_2(x_j)\mu_2(x_k) \end{aligned} \tag{2.8}$$

$$E(y_i^3) = \sum_{j=1}^{n} b_j^3\mu_3(x_j) \tag{2.9}$$

$$E(y_i^4) = \sum_{j=1}^{n} b_j^4\mu_4(x_j) + \sum_{j=1}^{n-1}\sum_{k=j+1}^{n} 6b_j^2b_k^2\mu_2(x_j)\mu_2(x_k) \tag{2.10}$$

Equations 2.9 and 2.10 have been truncated significantly in order to simplify the expressions. The complete third moment equation is lengthy, and the complete fourth moment equation is formidable. The complete equations for the raw third and fourth moments may be found in [Cox 1984].

After calculating the raw moments, the four moments of u_i about the mean may be found from

$$m_1(u_i) = E(y_i) + u_i^0 \tag{2.11}$$

$$m_2(u_i) = E(y_i^2) - \left[E(y_i)\right]^2 \tag{2.12}$$

$$m_3(u_i) = E(y_i^3) - 3E(y_i^2)E(y_i) + 2\left[E(y_i)\right]^3 \tag{2.13}$$

$$m_4(u_i) = E(y_i^4) - 4E(y_i^3)E(y_i) + 6E(y_i^2)\left[E(y_i)\right]^2 - 3\left[E(y_i)\right]^4 \tag{2.14}$$

To estimate the four moments of an assembly distribution using the full quadratic model requires the first eight moments of the component dimension distributions and the partial derivatives $\frac{\partial u_i}{\partial x_j}$, $\frac{\partial^2 u_i}{\partial x_j^2}$, and $\frac{\partial^2 u_i}{\partial x_j \partial x_k}$.

In a comparison of advanced tolerance analysis methods [Greenwood 1987], the Method of System Moments was recommended as the best method. The qualification for this recommendation was that MSM requires each controlled dimension to be modeled with a statistical distribution. Such information about the controlled dimensions is usually not known before production begins.

2.4 Empirical Distributions

Once the first four moments of a distribution are calculated, it is useful to fit an empirical distribution to these moments in order to describe the probability density or to estimate the fraction of assemblies outside of the specification limits. An empirical distribution is fit by the *matching of moments* of the calculated distribution to the moments of the empirical distribution. In general, the matching of moments requires the solution of a system of nonlinear equations to estimate the empirical distribution parameters. However, for many distributions the matching of moments is facilitated by tables.

Three empirical distributions are reviewed here: the Generalized Lambda Distribution, the Johnson System, and the Pearson System.

2.4.1 Generalized Lambda Distribution

The Generalized Lambda Distribution (GLD) [Hastings 1947, Dudewicz 1974] was developed as a simple function which could be used to approximate other distributions. GLD is defined in terms of its percentile function:

$$x_p = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2} \qquad 0 \le p \le 1 \tag{2.15}$$

The parameters λ_1, λ_2, λ_3, and λ_4 are the four parameters of GLD which relate to the first four moments of the distribution being approximated. This family of distributions is very flexible since, with one equation, it is able to cover most of the skewness and kurtosis plane. GLD may be fit to a distribution by both methods of matching of moments and matching of percentiles. The disadvantage of this distribution family is that an iterative algorithm must be used when calculating the quantile, the value of p when the value of x_p is given. The quantile is necessary for determining the fraction of assemblies outside the specification limits.

2.4.2 Johnson Distribution System

The Johnson System of distributions [Johnson 1949] is composed of three distribution families. These three families cover the entire allowable skewness and kurtosis plane. The basis for the Johnson System is that a distribution being approximated may be transformed in such a way that it can be considered to be normally distributed. Making this transformation allows normal distribution tables to be used. The disadvantage of the Johnson System is that the methods for fitting the distribution depend on which of the three families are appropriate.

Johnson S_L family:

$$f_1(x) = \frac{\eta e^{-\frac{1}{2}[\delta + \eta \ln(x-\varepsilon)]^2}}{\sqrt{2\pi}(x-\varepsilon)} \qquad \varepsilon < x < \infty \quad \eta > 0 \quad -\infty < \delta < \infty \quad -\infty < \varepsilon < \infty \qquad (2.16)$$

Johnson S_B family:

$$f_2(x) = \frac{\eta\lambda\, e^{-\frac{1}{2}\left\{\gamma + \eta \ln\left[\frac{(x-\varepsilon)}{(\lambda+\varepsilon-x)}\right]\right\}^2}}{\sqrt{2\pi}(x-\varepsilon)(\lambda+\varepsilon-x)} \qquad \lambda > 0 \quad \eta > 0 \quad -\infty < \gamma < \infty \quad -\infty < \varepsilon < \infty \qquad (2.17)$$

Johnson S_U family:

$$f_3(x) = \frac{\eta e^{-\frac{1}{2}\left\{\gamma + \eta \sinh^{-1}\left[\frac{(x-\varepsilon)}{\lambda}\right]\right\}^2}}{\sqrt{2\pi\left[\lambda^2 + (x-\varepsilon)^2\right]}} \qquad \lambda > 0 \quad \eta > 0 \quad -\infty < \gamma < \infty \quad -\infty < \varepsilon < \infty \qquad (2.18)$$

Both matching of moments and matching of percentiles may be used for each family, however, matching of moments for the Johnson S_B family is very complex. Once the parameters of the particular Johnson family are found, standard normal tables may be used to calculate quantiles or percentiles.

2.4.3 Pearson Distribution System

The Pearson System [Amos 1971] is composed of twelve families of distributions, all of which are solutions to the differential equation

$$\frac{df(x)}{dx} = \frac{(x-a)f(x)}{b+cx+dx^2} \tag{2.19}$$

The solutions differ in the values of the parameters *a*, *b*, *c*, and *d*. The Pearson system includes the normal, Gamma, and Beta distributions among the families. The twelve families cover the entire skewness and kurtosis plane. The advantage of this system is the tabulated percentile values indexed as a function of skewness and kurtosis. However, if the actual probability density or quantiles are needed, the Pearson System is not convenient. This inconvenience stems from the number of families included in the system and the logic needed to select a specific family and match distribution parameters. Unfortunately, the calculation of a quantile is precisely what is required by tolerance analysis applications. An empirical fit is needed by tolerance analysis to estimate the fraction of assemblies which meet the design limits.

2.4.4 Skewness-Kurtosis Plane

Figures 2.2 and 2.3 show the GLD distributions and the Johnson family of distributions on the skewness-kurtosis plane. The plane shown in both figures is only a partial view as values of skewness and kurtosis would continue to infinity. The plane is shown in coordinates of the standardized values of skewness and kurtosis. The standardized value of skewness, β_1, and the standardized value of kurtosis, β_2, are obtained from

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \tag{2.18}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \tag{2.19}$$

where μ_i represents the *ith* distribution moment. The normal distribution and uniform distribution are single points on the skewness-kurtosis plane at the β_1,β_2 coordinates 0,3 and 0,1.8 respectively. The skewness value of zero for the normal and uniform distributions indicate that these two distributions are symmetric. The exponential distribution is non-symmetric and is shown at the point 4,9.

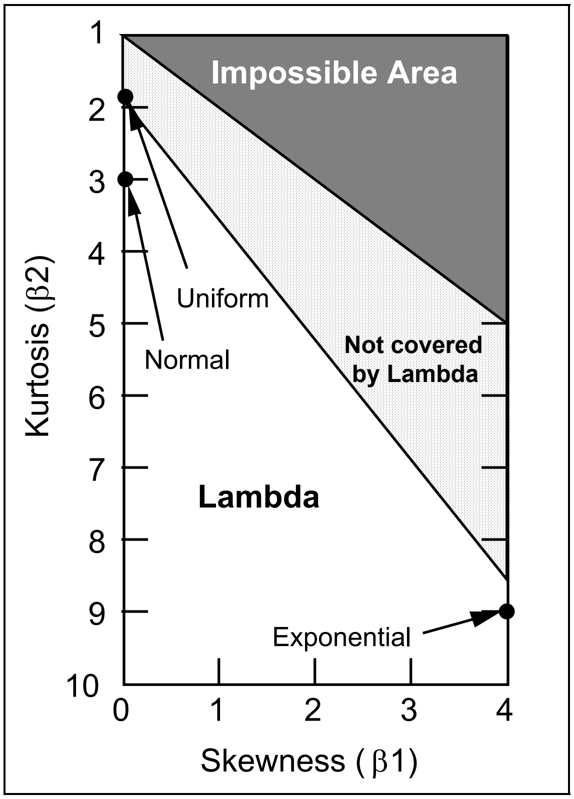

Figure 2.2: GLD Distribution

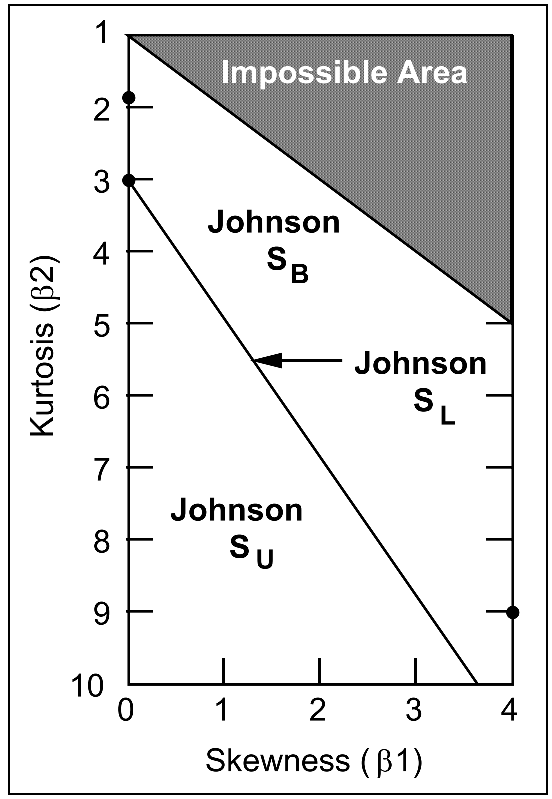

Figure 2.3: Johnson Distribution

Figure 2.2 and 2.3 compare the coverage of the GLD distributions and the Johnson system of distributions. Both GLD and the Johnson system include the normal, uniform and exponential distributions. Figure 2.2 illustrates a region corresponding to the Johnson S_B region that is not covered by GLD. However, GLD does include most practical distribution shapes likely to be encountered in mechanical assemblies. The Johnson families cover the entire skewness-kurtosis plane. The coverage of the Pearson system of distributions is identical to the Johnson diagram, but instead of two regions and one line to represent the Johnson sub-families, the Pearson sub-families would appear as multiple regions, lines and distinct points.

Even though the coverage of the GLD over the skewness-kurtosis plane is limited, GLD was determined to be the best empirical fit for use with the SOTA method. The region not covered by GLD includes U-shaped distributions and other distributions which are uncommon for mechanical assemblies. The Johnson and Pearson systems covered the entire skewness-kurtosis plane; however, the matching of moments required by MSM is extremely difficult to implement for these two systems. The ease of implementing GLD made GLD better suited for this research, and its range of coverage is still very acceptable.

2.5 Summary

The Linearized Method for tolerance analysis is quick, suitable for design iteration and tolerance allocation, and general for both implicit and explicit assembly functions. However, this method may be inaccurate for highly nonlinear assemblies or assemblies with non-normal input distributions.

Monte Carlo simulation is a nonlinear tolerance analysis method. This method allows normal and non-normal input distributions. While such features as kinematic constraints and tolerance allocation can be added for special cases, Monte Carlo simulation is too inefficient for design use.

Moments by Quadrature is a second-order statistical technique for estimating the first four moments of an explicit function of random variables. This technique includes the ability to handle non-normal distributions. Moments by Quadrature is based on numerical integration and requires $2n^2+1$ assembly function evaluations.

Another second-order statistical technique for estimating the first four moments of a function of random variables is the Method of System Moments. This technique is based upon Taylor series expansion of the assembly function. The Method of System Moment can handle non-normal input distributions. Thus far this method has only been applied to explicit functions.

Empirical distribution fits must be used by both Moments by Quadrature and the Method of System moments to match the moments of the assembly dimension of interest. The matching of moments is necessary to estimate the fraction of assemblies which meets a given design specification. The three empirical distributions reviewed were the Generalized Lambda Distribution, the Johnson System, and the Pearson System.

The next chapter describes a new second-order tolerance analysis method. This new method uses the Method of System Moments and the Generalized Lambda Distribution. MSM was chosen because it requires less computations than Moments by Quadrature in design iteration. The GLD system was selected because it was the easiest to implement and covers the useful region of the skewness-kurtosis plane with a single family of distributions.

Chapter 3

THE SOTA METHOD

The second-order tolerance analysis (SOTA) method is proposed as a general analysis method for vector-loop tolerance models. The SOTA method is comprised of three difference equations, a nonlinear system solver, the Method of System Moments (MSM), and a Generalized Lambda Distribution (GLD) empirical fit. The difference equations and nonlinear solver are used together to supply MSM with the required relationships between the component dimensions and the resultant assembly dimensions. MSM is then used to calculate the first four moments of the assembly dimensions. Finally, GLD is used to fit the calculated moments and approximate the distribution of the assembly dimensions. The SOTA method process is shown in Figure 3.1.

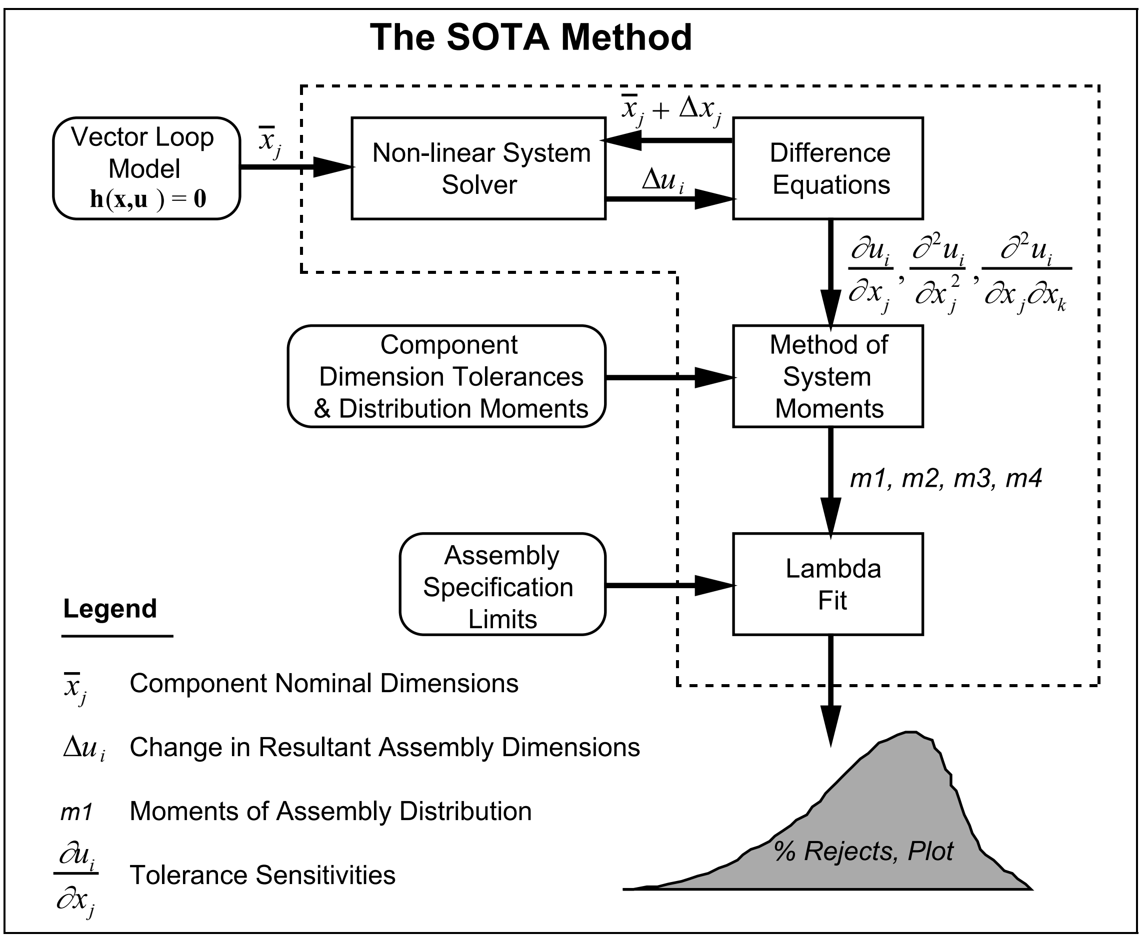

Figure 3.1: Functional Diagram of the SOTA method

The differentiation scheme, statistical technique, and empirical fit of the SOTA method will be presented in detail below.

3.1 Differentiation Scheme

MSM was selected as the best statistical technique for estimating the moments of the assembly dimensions in a vector-loop tolerance model. In order to apply the four moment equations by MSM, the partial derivatives $\frac{\partial u_i}{\partial x_j}$, $\frac{\partial^2 u_i}{\partial x_j^2}$, and $\frac{\partial^2 u_i}{\partial x_j \partial x_k}$ must be found. The research challenge involved finding these partial derivatives of implicit variables from a system of nonlinear loop equations. A numerical differentiation scheme using a nonlinear system solver and three difference formulas was developed to overcome this challenge.

3.1.1 Nonlinear System Solver

Since the assembly dimensions, u_i, are implicit variables, each function evaluation needed by the difference formulas requires an iterative solution . An iterative procedure is required to solve the nonlinear loop equations for a corresponding set of assembly dimensions, u_i, given a set of component dimensions, x_j. This procedure is identical to performing a position analysis in kinematics. In fact, significant computational effort may be saved by following well-developed position analysis procedures.

One particular position analysis procedure [Sandor 1984] suggests Newton's method for systems as the numerical method to solve the nonlinear system of equations. Newton's method [Burden 1993] has the best convergence of nonlinear system solvers, but it is can also be the most costly. The cost for Newton's method is incurred by the evaluation of the inverse Jacobian matrix at each iteration. The computational cost of Newton's method when applied to vector-loop equations is able to be reduced by the analytic derivatives of the loop equations. These analytic derivatives allow the Jacobian matrix to be obtained easily. The only major computational effort incurred is the inverting of the Jacobian.

For example, the linear equation that must be solved at each iteration of Newton's method is

$$[\mathbf{B}]\{\Delta\mathbf{u}\} = \{\mathbf{r}\} \qquad (3.1)$$

where the **B** matrix is the Jacobian matrix of the loop equations with respect to the unknown assembly dimensions. This **B** matrix is the same as shown in Section 1.2.4 and can be obtained analytically given the values of the component dimensions and the values of the assembly dimensions. The residual vector **r** is related to the closure error of the nonlinear loop equations. At each iteration **B** and **r** must be reevaluated and a new **Δu** found to solve Equation 3.1. A new estimate of the unknown assembly dimensions, **u**, is found by adding the vector **Δu** to the previous estimate of **u**.

$$\mathbf{u}_{\mathbf{new}} = \mathbf{u}_{\mathbf{old}} + \Delta\mathbf{u} \qquad (3.2)$$

Again, since the **B** matrix and **r** vector can be obtained from analytic expressions, the only significant effort at each Newton iteration is the inverting of the **B** matrix.

Another concern when using Newton's method is stability. Newton's method may diverge when initial values are not close to the actual solution. However, this will never be the case for the SOTA application. Since Newton's method is being used to find numerical derivatives, only one or two values of the component dimensions, x_j, will be perturbed at a time. These perturbations will be very small, which means the assembly dimensions, u_i, should always be close to their given value.

3.1.2 Difference Formulas

In order to approximate the three sets of partial derivatives, $\frac{\partial u_i}{\partial x_j}$, $\frac{\partial^2 u_i}{\partial x_j^2}$, and $\frac{\partial^2 u_i}{\partial x_j \partial x_k}$, three separate difference formulas are required. The linear partial derivatives are approximated by a central difference formula, Equation 3.3.

$$\frac{\partial u_i}{\partial x_j} \approx \frac{u_i(x_j + \Delta x_j, u_i) - u_i(x_j - \Delta x_j, u_i)}{2\Delta x_j} \qquad (3.3)$$

For this notation, $u_i(x_j + \Delta x_j, u_i)$ represents a function evaluation for the implicit assembly dimensions u_i, where the component dimensions, x_j, are at their nominal value except for the *jth* dimension which is perturbed by a value Δx_j. So, Equation 3.3 indicates two function evaluations for each component dimension x_j. Note that each function evaluation requires an iterative solution of a system of nonlinear equations.

A three point difference formula is used for the approximation of the quadratic partial derivatives [Burden 1993].

$$\frac{\partial^2 u_i}{\partial x_j^2} \approx \frac{u_i(x_j + \Delta x_j, u_i) - 2u_i(x_j, u_i) + u_i(x_j - \Delta x_j, u_i)}{\Delta x_j^{\ 2}} \tag{3.4}$$

Two of the function evaluations that appear in this three point difference formula also appear in the central difference formula, Equation 3.3. Therefore, if there are *n* component dimensions, *2n* function evaluations can be avoided if the same function evaluations are used for both difference Equations 3.3 and 3.4. If this is done the quadratic partial derivatives will require *2n* function evaluations plus one evaluation at the nominal and, without any further evaluations, the linear partials can be obtained.

The approximation of the cross-derivatives is more complicated. These partial derivatives are found by using the central difference of a central difference.

$$\frac{\partial^2 u_i}{\partial x_j \partial x_k} \approx \frac{\left[\begin{array}{c}\dfrac{u_i(x_j + \Delta x_j, x_k + \Delta x_k, u_i) - u_i(x_j - \Delta x_j, x_k + \Delta x_k, u_i)}{2\Delta x_j} - \\ \dfrac{u_i(x_j + \Delta x_j, x_k - \Delta x_k, u_i) - u_i(x_j - \Delta x_j, x_k - \Delta x_k, u_i)}{2\Delta x_j}\end{array}\right]}{2\Delta x_k} \tag{3.5}$$

Since for *n* component dimensions there are $(n^2-n)/2$ unique cross-derivatives, and four new function evaluations must be performed for each derivative, $2n^2-2n$ evaluations are required to obtain the cross-derivatives. Together with the $2n+1$ function evaluations for the linear and quadratic partial derivatives, the total becomes $2n^2+1$ function evaluations. Thus, with $2n^2+1$ function evaluations, the required partial derivatives for the SOTA method are obtained.

3.1.3 Difference Formula Error

Finite difference equations are subject to both truncation error and round-off error. Care must be used to select a step size, Δx, to minimize both of these error sources. In general, as the step size increases, the truncation error will increase, and as the step size decreases, the round-off error will increase. An optimal step size can be found which balances the effect of truncation error and round-off error and minimizes the total error. The optimal step sizes were calculated for Equations 3.3, 3.4 and 3.5 and appear in Appendix A.

The optimal step size equations may not seem to be valuable, since third and fourth derivatives are needed when first and second derivatives are being approximated.

However, there are some tolerance models for which closed form solutions exist, and the bounds on the third and fourth derivative may be calculated. Such closed form solutions help to establish a good estimate for the optimal step size over a wide range of problems. A step size of 0.001 was used for the five case studies analyzed in Chapter 4.

3.2 Statistical Technique

The Method of System Moments (MSM) was selected as the statistical method for estimating the moments for the tolerance model. MSM was selected over the Moments by Quadrature method because of computation requirements. Recall from Section 2.3.1 that quadrature requires $2n^2+1$ function evaluations to estimate the first four moments, as does MSM. The difference comes when one of the component tolerances or distributions is changed. When one component tolerance or distribution changes, Moments by Quadrature requires two additional function evaluations as well as the calculation of the four moments of the assembly dimensions. For MSM the partial derivatives do not change, only the four moments of the assembly dimensions need to be recomputed for the new values of the input moments when an input distribution changes. Of course, when any nominal dimension changes, the partial derivatives must be reevaluated, and all $2n^2+1$ function evaluations must be performed. This is true for both MSM and Moments by Quadrature.

The fact that for MSM the partial derivatives only need be calculated once is a desirable feature. This means that after the initial evaluation of the partial derivatives an unlimited number of "what if" studies can be done with different component tolerances without further assembly function evaluations. For example, during a tolerance allocation it is possible that all n component tolerances may change. For MSM only the four moments of the assembly dimension would need to be recalculated. For this same case using Moments by Quadrature, not only would the four moments need to be recalculated, but also $2n$ function evaluations would be required. For a large problem, $2n$ function evaluations may be costly. This subtle difference gives MSM a great advantage over Moments by Quadrature as a design tool.

The one criticism of MSM and other advanced tolerance analysis methods is the level of information required for each component distribution. The argument is that not enough information is known at the design stage to specify a distribution for each component. The ability of advanced methods to handle non-normal distributions should not be viewed as a drawback. When information about a component distribution is not known, the distribution should be assumed to be normal. Such an assumption is identical to the assumption that *must* be made when using the Linearized Method. In this way, added complexity is eliminated while, at the same time, the addition accuracy of advanced methods such as MSM is realized.

3.3 Empirical Distribution Fit

The Generalized Lambda Distribution was chosen as the best empirical model for fitting the statistical moments based on ease of implementation. The single form of the GLD make the method of matching of moments easily applied to the moments calculated by MSM, whereas the Johnson and Pearson systems require multiple distribution forms to cover a full range of moments. In addition, a GLD table, indexed by skewness and kurtosis values, was readily available for use in a computer program because of earlier research [Larsen 1989, Gao 1993]. The GLD's range of coverage is the smallest compared to the Johnson and Pearson systems, however, it does cover most practical distribution shapes likely to be encountered in mechanical assemblies.

The next chapter will compare analysis results of the SOTA method with results from the Linearized Method and Monte Carlo Simulation.

Chapter 4

CASE STUDIES

This chapter contains five case studies for which the SOTA method was used to analyze different manufactured assemblies. The case studies were selected as representative of real engineering problems. The objective of these case studies was to validate the SOTA method and compare the results obtained with those obtained using Monte Carlo simulation and the Linearized Method. Three different sample sizes were run for Monte Carlo simulation. The sample sizes were one million (1M), one hundred thousand (100k), and thirty thousand (30k).

It is important to note that the Monte Carlo simulator used in this comparison included the capability to handle closed loops (kinematic constraints) [Gao 1993]. This functionality was added to the Monte Carlo simulator by using the same nonlinear system solver used in the SOTA method. The pseudo-random-number generator used by the Monte Carlo simulator was the *ran1* function from [Press 1992]. The uniform deviates from the *ran1* function were transformed to the component dimension distributions using the Generalized Lambda Distribution.

The five case studies are 1.) the tape hub, 2.) the stacked blocks, 3.) the pawl and ratchet, 4.) the one-way clutch and 5.) the bicycle crank. The assembly for each case will be described, followed by the analysis results for that case. The analysis results for each case are the estimates of the first four distribution moments and the prediction of the parts-per-million rejects outside of the upper and lower specification limits. Specification limits were chosen for one assembly dimension out of each case study. The limits for each case were chosen to yield parts-per-million rejects at a quality level of approximately three sigma. The results for all the case studies will be summarized in Section 4.6.

The first derivatives for each case study obtained numerically using the SOTA method corresponded well with the first derivatives obtained using the Linearized Method. The Linearized Method obtains the first derivatives using analytic means coupled with linear algebra. Both the first and second-order derivatives obtained by the SOTA method are contained in Appendix B.

4.1 Tape Hub

4.1.1 Tape Hub Description

The tape hub is a mechanism used to secure computer tape reels in place. The plunger is pushed down to force the arm and pad outward. With the pad extended completely outward, an interference fit is established with the tape reel. The interference fit locks the tape reel in place and allows the hub to rotate the tape reel in either direction. The variation of the length from the center line to the pad is critical for this interference fit. This length, **H**, will be the dimension of interest for the tolerance analysis.

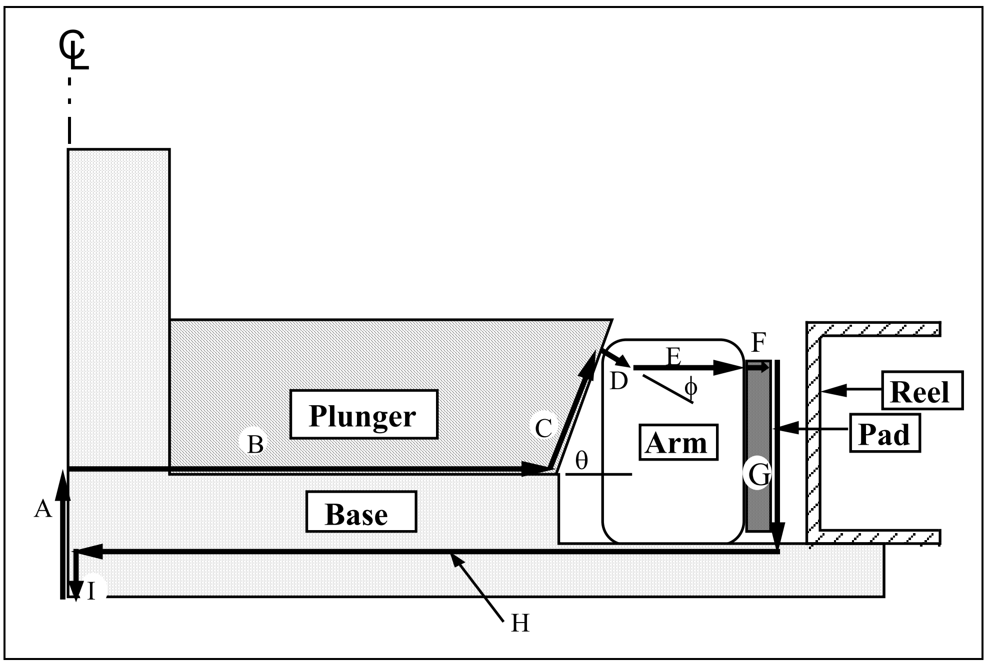

Figure 4.1: Tape Hub mechanism cross-section

Figure 4.1 shows a cross-section view of the tape hub. Overlaid on the cross-section is a diagram of the vector loop used for the analysis. The interference fit between the pad and the reel is shown as a clearance for clarity.

For this mechanism there are eight manufactured dimensions, including the angle θ. The nominal values and tolerances for these dimensions are listed in Table 4.1. Two separate analyses were performed using different assumptions about these dimensions. The first analysis used normally distributed dimensions, and the second analysis used non-normal distributions. The first analysis assumed all nine dimensions were normally distributed with means at the nominal dimensions and tolerance values equal to three standard deviations. In other words, the distribution of dimension **A** was normal with mean 0.39 and standard deviation 0.001. The second analysis assumed the nine dimensions had skewed distributions. The skewed distributions all had the following parameters:

Mean:	nominal value
Standard deviation:	tolerance/3.7
Skewness:	0.5
Kurtosis:	3.5

The value of 3.7 was chosen to make the number of rejects the same magnitude as the normal distributions case. An analysis using the Linearized Method could not be performed for the skewed input distributions.

Table 4.1: Tape Hub Manufactured Dimensions

Variable Name	Basic Size	Tolerance (±)
A	0.39 in	0.003 in
B	1.36 in	0.004 in
D	0.06 in	0.002 in
E	0.235 in	0.003 in
F	0.05 in	0.002 in
G	0.5275 in	0.004 in
I	0.2 in	0.003 in
θ	75°	0.5°

The assembly dimensions are listed in Table 4.2 with the given specification limits for dimension **H**.

Table 4.2: Tape Hub Assembly Dimensions

Variable Name	Basic Size	Upper Spec. Limit (USL)	Lower Spec. Limit (LSL)
C	0.3655 in	--	--
H	1.79755 in	1.80355 in	1.79155 in
ϕ	15°	--	--

4.1.2 Tape Hub Results

Table 4.3 and Table 4.4 show the results for the two analyses. The first four statistical moments are shown along with the parts-per-million (ppm) assembly rejects. The assembly rejects represent the fraction of assemblies for which the assembly dimension **H** is outside of the specification limits. For the first analysis, both the Linearized Method and the SOTA method predicted rejects very close to the Monte Carlo 1M analysis. For the non-normal distribution analysis, the SOTA method was, again, very close the Monte Carlo 1M analysis. The effect of the skewed input distributions is apparent from the slight positive skewness in the analyses of the second case.

Table 4.3: Tape Hub Results, normal Distributions

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	1.797550	0.002277	0.002612	3.118114	4231	4233
MC 100k	1.797552	0.002278	-0.000895	2.964939	4320	4400
MC 30k	1.797553	0.002262	-0.019353	2.975415	3867	4100
Linear	1.797550	0.002276	0	3	4196	4196
SOTA	1.797551	0.002276	0.002666	3.000036	4198	4192

Table 4.4: Tape Hub Results, Non-normal Distributions

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	1.797561	0.00196	0.059019	3.110862	1762	1133
MC 100k	1.797562	0.001962	0.055531	3.138308	1830	1170
MC 30k	1.797563	0.001949	0.03707	3.076703	1500	1000
SOTA	1.797551	0.001954	0.058369	3.118478	1643	1139

4.2 Stacked Blocks

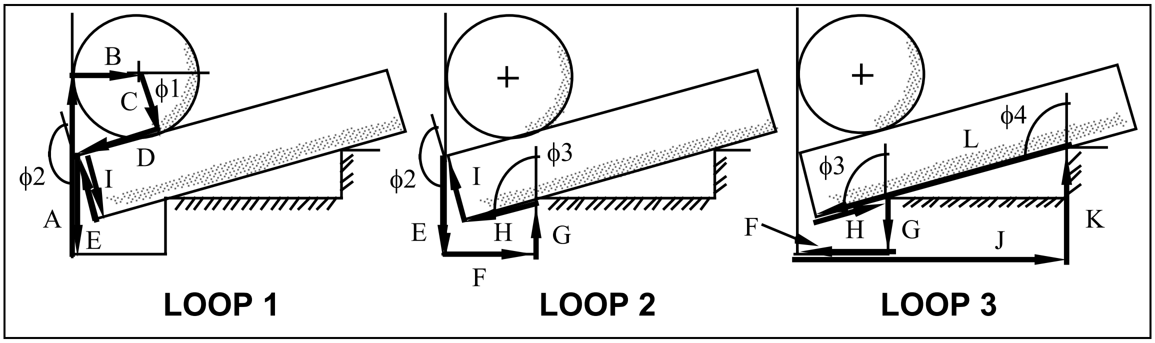

Figure 4.2: Stacked blocks

4.2.1 Stacked Blocks Description

The stacked blocks assembly is a tolerance analysis training problem. The assembly consists of a stepped base, a block, and a cylinder. Since there is no true function for this assembly, length **A** was arbitrarily selected as the dimension to analyze.

This assembly requires three loops to model the three stacked shapes. A diagram of the three loops is shown in Figure 4.2. The six manufactured dimensions were assumed to be normally distributed with tolerance values equal to three standard deviations. These six dimensions, with their nominal sizes and tolerances, are listed in Table 4.5. The cylinder radius vectors, **B** and **C**, were considered the same source of variation.

Table 4.6 contains the nine assembly dimensions and their nominal values. The specification limit for the length **A** is also listed in this table.

Table 4.5: Stacked Blocks Manufactured Dimensions

Variable Name	Basic Size	Tolerance (±)
B,C	6.620 mm	0.200 mm
F	3.905 mm	0.125 mm
G	4.060 mm	0.150 mm
I	6.805 mm	0.075 mm
J	24.22 mm	0.350 mm
K	10.675 mm	0.125 mm

Table 4.6: Stacked Blocks Assembly Dimensions

Variable Name	Basic Size	Upper Spec. Limit (USL)	Lower Spec. Limit (LSL)
A	18.7182 mm	19.018 mm	18.418 mm
ϕ1	74.724°	--	--
D	8.6705 mm	--	--
ϕ2	164.724°	--	--
E	10.0477 mm	--	--
ϕ3	105.276°	--	--
H	2.1894 mm	--	--
ϕ4	105.275°	--	--
L	27.2965 mm	--	--

4.2.2 Stacked Blocks Results

The estimates of the first four assembly dimension moments as well as the parts-per-million (ppm) assembly rejects are listed in Table 4.7. All five methods provided very adequate predictions of the assembly rejects.

Table 4.7: Stacked Blocks Results

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	18.71849	0.099913	-0.002268	3.006475	1282	1407
MC 100k	18.71899	0.099956	-0.004995	3.020634	1340	1430
MC 30k	18.71888	0.100124	-0.008354	2.998437	1167	1467
Linear	18.71824	0.099923	0	3	1340	1340
SOTA	18.71847	0.099923	-0.002263	3.000063	1350	1329

4.3 Pawl and Ratchet

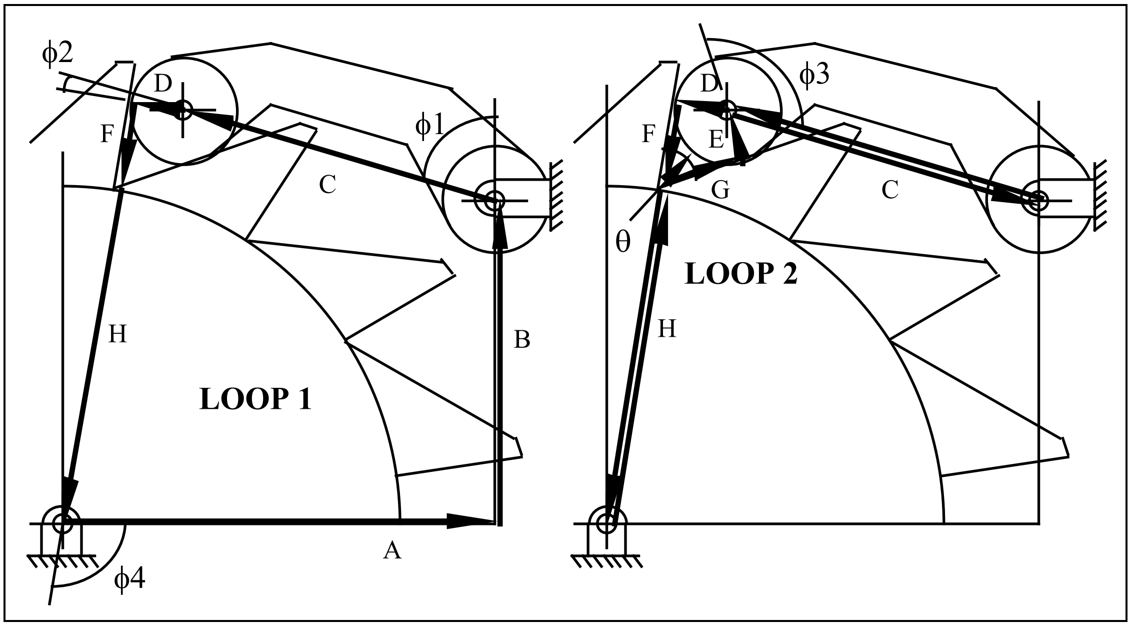

Figure 4.3: Pawl and Ratchet

4.3.1 Pawl and Ratchet Description

The pawl and ratchet assembly consists of a ratchet, pawl, and a housing. Both the pawl and ratchet are attached to the housing. The pawl arm has a cylindrical end that stops the clockwise rotation of the ratchet at increments equal to the spacing of the teeth. The assembly is used to rotationally position an object attached to the ratchet. The two vector loops used to model the pawl and ratchet assembly are shown in Figure 4.3.

The pawl and ratchet assembly has seven manufactured dimensions. These dimensions are listed in Table 4.8. The dimensions were assumed to be normally distributed with tolerance values equal to three standard deviations. The assembly dimension of interest is **ϕ4**, the angular orientation of the ratchet. The specification limits for **ϕ4** are given in Table 4.9.

Table 4.8: Pawl and Ratchet Manufactured Dimensions

Variable Name	Basic Size	Tolerance (±)
A	0.400 in	0.0015 in
B	0.300 in	0.0015 in
C	0.300 in	0.001 in
D	0.050 in	0.001 in
E	0.050 in	0.001 in
H	0.313 in	0.002 in
θ	62.0°	0.8°

Table 4.9: Pawl and Ratchet Assembly Dimensions

Variable Name	Basic Size	Upper Spec. Limit (USL)	Lower Spec. Limit (LSL)
F	0.0832 in	--	--
G	0.0832 in	--	--
ϕ1	73.863°	--	--
ϕ2	7.069°	--	--
ϕ3	125.069°	--	--
ϕ4	99.0679°	99.3679°	98.7679°

4.3.2 Pawl and Ratchet Results

Table 4.10 shows the estimate of the first four assembly dimension moments with the predicted assembly rejects shown in parts-per-million (ppm). For the pawl and ratchet case, all five analyses predicted the number of assembly rejects well. The Monte Carlo 30k analysis was the least accurate, over predicting by only 259 ppm.

Table 4.10: Pawl and Ratchet Results

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	99.06824	0.098336	-0.003276	3.11833	1080	1161
MC 100k	99.06792	0.098524	-0.005219	2.837427	1130	1140
MC 30k	99.06793	0.098864	-0.019827	2.969778	1067	1433
Linear	99.0679	0.098193	0	3	1124	1124
SOTA	99.07002	0.098191	-0.000644	3.000055	1208	1046

4.4 One-Way Clutch

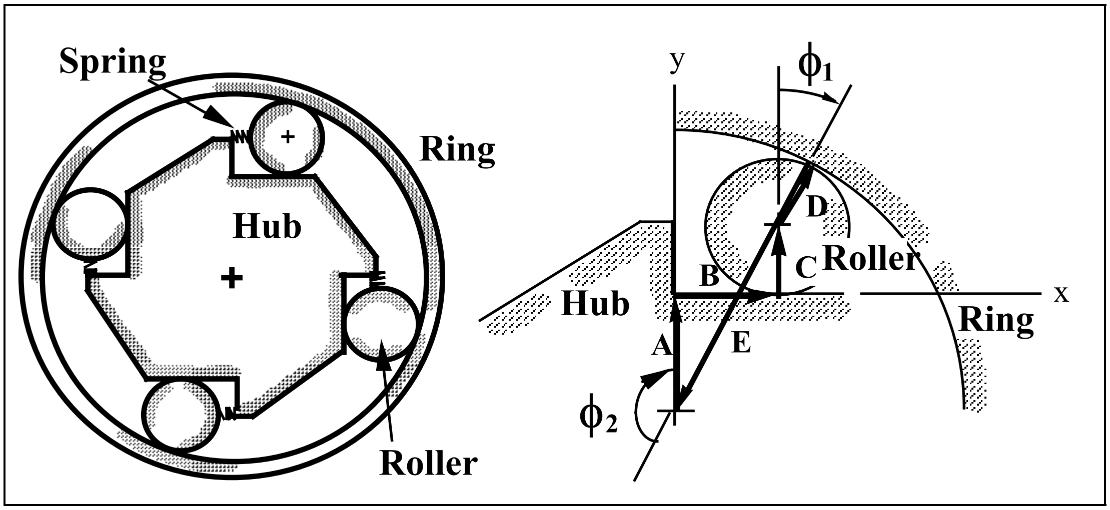

Figure 4.4: One-way Clutch

4.4.1 One-Way Clutch Description

A one-way clutch transmits torque in a single direction. This assembly consists of the following components: a hub, an outer ring, four rollers, and four springs. When the hub rotates in a counter-clockwise direction, the roller jams between it and the ring, locking them together. When the hub turns in a clockwise direction, the spring is compressed by the roller, the roller slips, and the hub is allowed to rotate freely. The one-way clutch assembly and the single vector loop used to model this assembly are shown in Figure 4.4.

The one-way clutch mechanism has three manufactured dimensions. The nominal values and tolerances of these three dimensions are shown in Table 4.11. The vectors representing the roller radius, vector **C** and **D**, were treated as the same variation source. The pressure angle, **ϕ1**, is critical to the function of the clutch, so **ϕ1** was given specification limits as shown in Table 4.12.

Two separate cases were analyzed for the one-way clutch assembly. The first case assumed all three component dimensions were normally distributed with tolerance values equal to three standard deviations. The second case assumed the components distributions were skewed. The skewed distributions had the following parameters:

Mean:	nominal value
Standard deviation:	tolerance/2.5

Skewness: 0.5
Kurtosis: 3.5

The value of 2.5 was chosen to make the number of rejects the same magnitude as the normal distributions case. An analysis using the Linearized Method could not be performed for the skewed input distributions.

Table 4.11: Clutch Manufactured Dimensions

Variable Name	Basic Size	Tolerance (±)
A	27.645 mm	0.050 mm
C,D	11.430 mm	0.010 mm
E	50.800 mm	0.0125 mm

Table 4.12: Clutch Assembly Dimensions

Variable Name	Basic Size	Upper Spec. Limit (USL)	Lower Spec. Limit (LSL)
ϕ1	7.0184°	7.6184°	6.4184°
ϕ2	172.9816°	--	--
B	4.8105 mm	--	--

4.4.2 One-Way Clutch Results

The results of the two clutch cases are shown in Table 4.13 and Table 4.14. The clutch problem exhibited nonlinear characteristics. Both analyses showed the clutch angle **ϕ1** to be negatively skewed and more peaked than a normal distribution. This nonlinear effect limited the ability of the Linearized Method to estimate rejects accurately for the clutch problem. The Linearized Method assumes a normally distributed assembly distribution, skewness of zero and kurtosis equal to three.

Table 4.13: Clutch Results, normal distributions

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	7.015373	0.219884	-0.094774	3.027695	2206	4467
MC 100k	7.015453	0.220172	-0.101679	3.021511	2080	4580
MC 30k	7.012982	0.220541	-0.09758	3.082748	2567	5000
Linear	7.018389	0.219292	0	3	3109	3109

SOTA	7.018523	0.219346	-0.093555	3.011671	2452	3998

Table 4.14: Clutch Results, Non-normal distributions

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	7.012377	0.266847	-0.525843	3.582245	91	4867
MC 100k	7.012311	0.267340	-0.535775	3.59812	80	4880
MC 30k	7.009413	0.268300	-0.541735	3.684566	133	5033
SOTA	7.017216	0.259859	-0.520600	3.305675	46	3118

4.5 Bicycle Crank

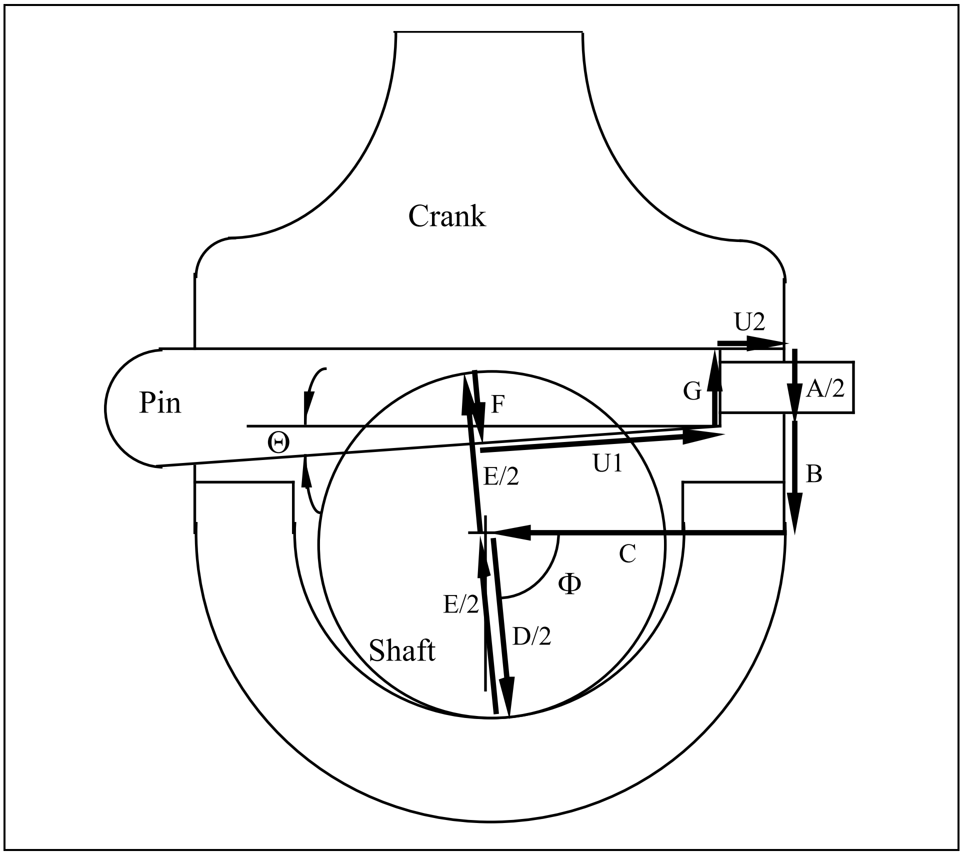

Figure 4.5: Bicycle Crank Loop Diagram

4.5.1 Bicycle Crank Description

The bicycle crank assembly is used to secure the pedal crank to the sprocket shaft. The assembly is composed of the crank, the shaft and a pin that holds them together. The pin is designed to wedge between a flat portion of the shaft and hole drilled through the crank. The threaded portion of the pin extends out beyond the surface of the crank to allow locking the pin with a nut. As the nut is tightened the pin is wedged tighter and the crank and the shaft are held together. A cross-section view of the assembly is shown in Figure 4.5.

The length of thread that remains in the crank hole, length **U2**, is important to the proper function of the assembly. If the resulting length is zero, the nut will not be able to fasten the crank to the shaft securely. On the other hand, if the resulting length is the entire length of the threaded portion of the pin, the nut cannot even be attached. This length, **U2**, will be the dimension of interest in the analysis.

There are eight manufactured dimensions involved in the bicycle crank assembly. The eight dimensions along with their nominal sizes and tolerances are shown in Table 4.15. These dimension were assumed to be normally distributed with means at the nominal sizes and tolerance values equal to three standard deviations.

After some preliminary analyses of the bike crank assembly, it was noted that the wedge angle, θ, had a significant nonlinear effect. This effect was first noted from the magnitude of the second derivative of the length **U2** with respect to the angle θ (see Appendix B). Furthermore, it was found that as the nominal value of the wedge angle decreased, this nonlinear effect increased. In order to study how well the Linearized Method and SOTA method perform with this type of nonlinear problem, the bike crank assembly was analyzed at three different wedge angle values: 4°, 3.5°, and 3°. Tables 4.15, 4.16 and 4.17 show the assembly dimension nominal values and the corresponding specification limits for the three angle sizes.

Table 4.15: Bicycle Crank Manufactured Dimensions

Variable Name	Basic Size	Tolerance (±)
A	9.520 mm	0.015 mm
B	7.650 mm	0.076 mm
C	13.550 mm	0.127 mm
D	15.700 mm	0.025 mm
E	15.660 mm	0.013 mm
F	4.500 mm	0.050 mm
G	8.500 mm	0.050 mm
θ	4.0°	0.5°

Table 4.16: Bicycle Crank Assembly Dimensions, Wedge Angle at 4°

Variable Name	Basic Size	Upper Spec. Limit (USL)	Lower Spec. Limit (LSL)
U1	8.7169 mm	--	--
U2	5.0852 mm	6.8852 mm	3.2852 mm
ϕ	86.0°	--	--

Table 4.17: Bicycle Crank Assembly Dimensions, Wedge Angle at 3.5°

Variable Name	Basic Size	Upper Spec. Limit (USL)	Lower Spec. Limit (LSL)
U1	9.9294 mm	--	--
U2	3.8412 mm	5.6412 mm	2.0412 mm
ϕ	86.5°	--	--

Table 4.18: Bicycle Crank Assembly Dimensions, Wedge Angle at 3°

Variable Name	Basic Size	Upper Spec. Limit (USL)	Lower Spec. Limit (LSL)
U1	11.5511 mm	--	--
U2	2.1880 mm	3.9880 mm	0.3880 mm
ϕ	87.0°	--	--

4.5.2 Bicycle Crank Results

The results of the three cases analyzed for the bicycle crank are shown in Tables 4.19, 4.20 and 4.21. The nonlinear assembly effect is illustrated well with these three cases. As the wedge angle decreased for each case the skewness and kurtosis values for the length **U2** increased. Also, as the nonlinear effect increased and the skewness and kurtosis values increased, the error of the Linearized Method increased. The error of the SOTA method also increased, but not a severely as the Linearized Method.

Table 4.19: Bicycle Crank Results, Wedge Angle at 4°

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	5.070325	0.625727	-0.143747	3.046986	986	3490
MC 100k	5.0687	0.626505	-0.150672	3.042868	1000	3400
MC 30k	5.065652	0.627186	-0.154475	3.044086	967	3533
Linear	5.085188	0.622715	0	3	1923	1923
SOTA	5.084174	0.623427	-0.14449	3.041235	1387	3029

Table 4.20: Bicycle Crank Results, Wedge Angle at 3.5°

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	3.818987	0.751315	-0.179491	3.071622	666	3464
MC 100k	3.816909	0.752402	-0.186888	3.069802	600	3550
MC 30k	3.813427	0.753469	-0.189744	3.065199	633	3700
Linear	3.841216	0.746559	0	3	1605	1605
SOTA	3.836815	0.747728	-0.179432	3.058847	959	2912

Table 4.21: Bicycle Crank Results, Wedge Angle at 3°

	Mean	Standard Deviation.	Skew-ness	Kurtosis	Upper Rejects (ppm)	Lower Rejects (ppm)
MC 1M	2.15259	0.93874	-0.23005	3.114978	373	3614
MC 100k	2.149822	0.940358	-0.238019	3.116546	290	3760
MC 30k	2.145769	0.942055	-0.240143	3.104797	300	3900
Linear	2.187993	0.930411	0	3	1309	1309
SOTA	2.176559	0.932506	-0.228017	3.088323	610	2898

4.6 Summary of Results

To compare the performance of the different tolerance analysis methods, a relative error measure for each method was calculated. Monte Carlo simulation using one million samples was used as the baseline of the comparison. Monte Carlo simulation with such a large sample size was assumed to be adequately accurate. In other words, the Monte Carlo 1M analysis was assumed to have zero error, and the four remaining analyses, the SOTA method, Linearized Method, Monte Carlo 100k, and Monte Carlo 30k, were compared relative to the Monte Carlo 1M analysis.

4.6.1 Error in Statistical Moments

The error in estimating the first four statistical moments of the assembly distribution was compared for the five case studies of the tape hub, pawl, stacked blocks, clutch, and the bike with the wedge angle at 4°. The relative error in estimating moments was calculated by taking the difference between the calculated moment from the Monte Carlo 1M method and the calculated moment from the method in question, divided by the Monte Carlo 1M moment. The only exception was the skewness, for which just the absolute difference from the Monte Carlo 1M analysis was used. This was done since the skewness value for most cases was very small.

4.6.1.1 Error in the Mean

Figure 4.6 shows the chart of the error in prediction of the first moment, the mean. The error for the tape hub, the pawl and ratchet, and the stacked blocks was extremely small. The error increased for the clutch and bicycle crank assemblies, though it did not exceed 0.3% for the SOTA method and the Linearized Method. It is interesting to note that the clutch and bicycle crank assemblies also yielded assembly distributions with the largest magnitudes of skewness and kurtosis among the five assemblies shown in the chart. Deviation from a normal distribution indicates a nonlinear assembly function. A nonlinear assembly function explains the increase in error for the clutch and bicycle crank cases.

4.6.1.2 Error in the Standard Deviation

Figure 4.7 shows the Monte Carlo simulation with 30k samples did not predict the standard deviation as well as the other tolerance analysis methods. Note also, as in the last plot, the SOTA method and the Linearized Method were not as accurate in predicting the standard deviation for the clutch and the bicycle crank assemblies as for the other three more linear cases.

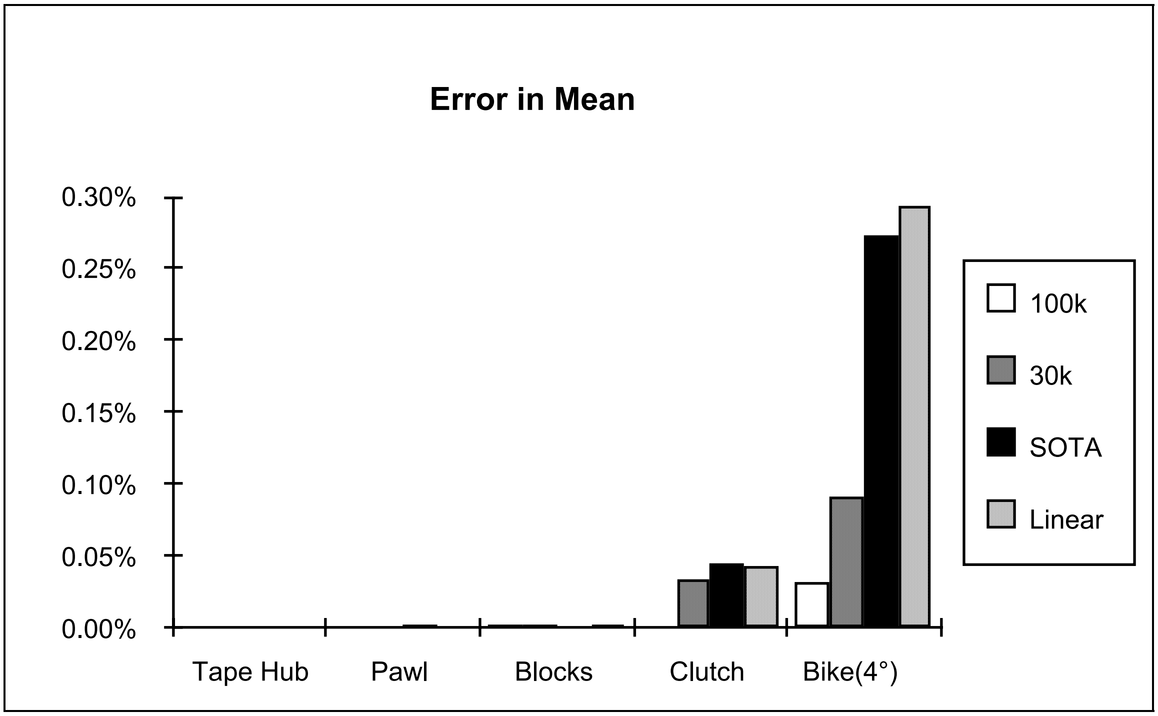

Figure 4.6: Error in Mean Chart

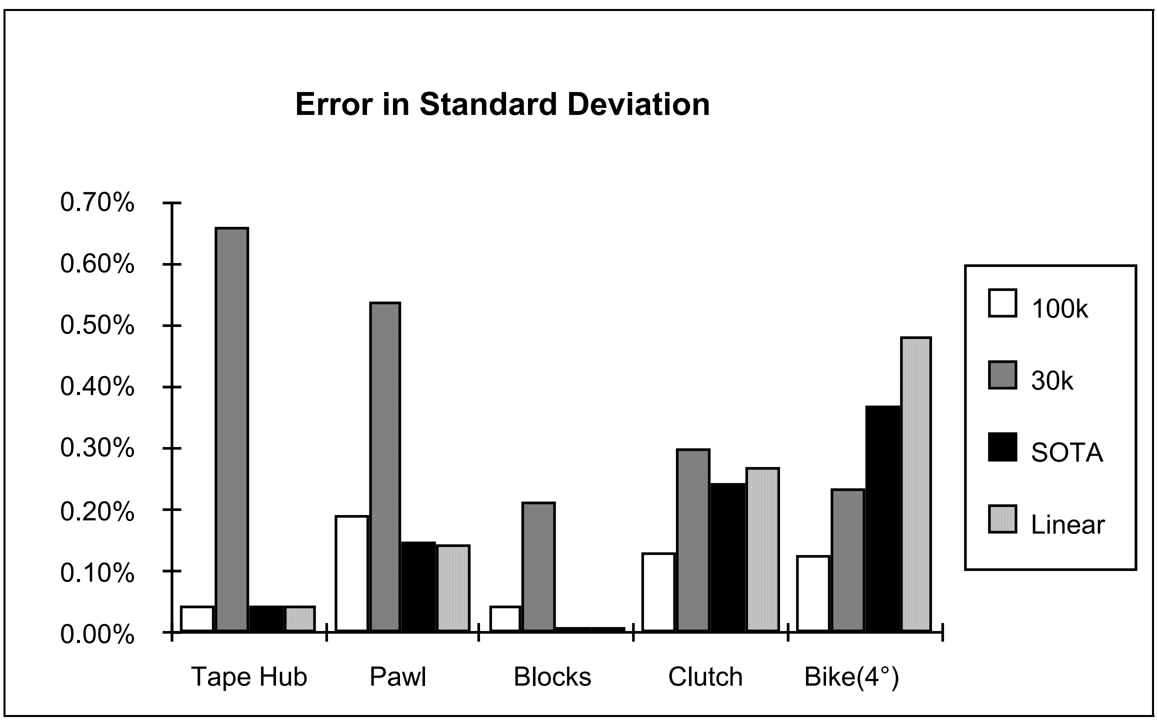

Figure 4.7: Error in Standard Deviation Chart

4.6.1.3 Error in Skewness

Figure 4.8 shows the SOTA method predicted the skewness extremely well, meeting or exceeding the accuracy of the 100k and 30k methods in every case. The Linearized Method cannot predict skewness, so the assumption of zero skewness must be made. A zero skewness assumption was not a good one for the nonlinear assembly functions of the clutch and the bicycle crank.

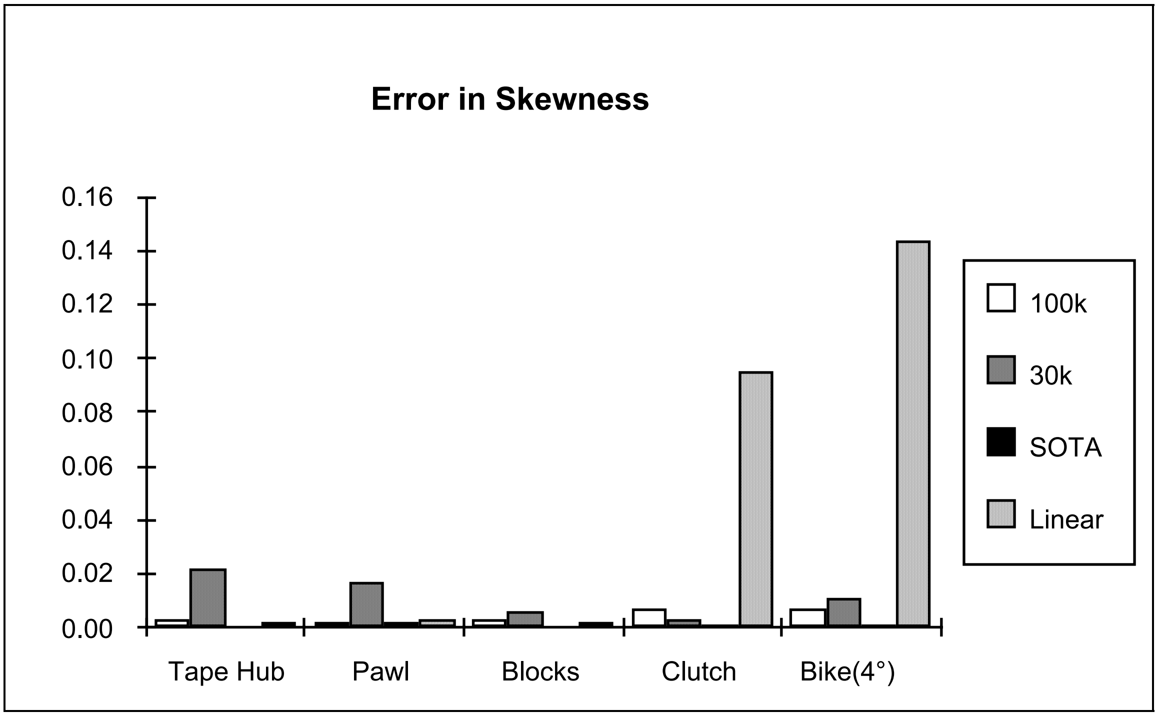

Figure 4.8: Error in Skewness Chart

4.6.1.4 Error in Kurtosis

The first obvious observation of Figure 4.9 is the magnitude of error. The greatest magnitude of the error is especially intriguing since it occurred with two linear assemblies, the tape hub and the pawl and ratchet. Furthermore, the SOTA method and the Linearized Method analyses had less error than both the Monte Carlo cases. No explanation has been found of why the magnitude of error for these two cases was so high.

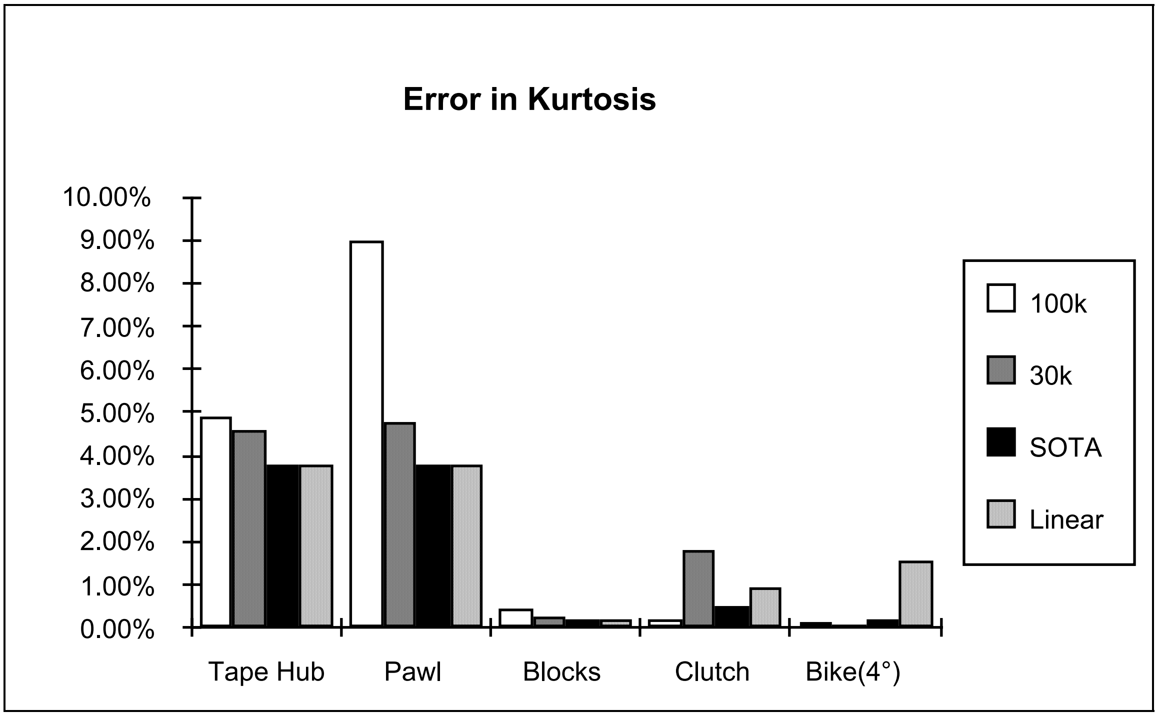

Figure 4.9: Error in Kurtosis Chart

4.6.2 Error in Prediction of Rejects

The prediction of total rejects is a good overall measure of the accuracy of a tolerance analysis method. For the SOTA method and the Linearized Method, estimating the statistical moments is an intermediate step toward the final approximation of the assembly distribution. The estimate of the fraction of rejects is the final result, and this result depends on all the intermediate steps of the analysis method.

The measure of error for the prediction of rejects was the absolute difference between the parts-per-million (ppm) rejects predicted by the Monte Carlo 1M analysis and the ppm rejects predicted by the particular analysis method.

Figure 4.10 shows the error in prediction of total rejects. For the two nonlinear cases, the clutch and the bike, the error for the Linearized Method was relatively high. For the same two cases, the error of the SOTA method was quite a bit less. Overall, the SOTA method performed very well, predicting rejects more accurately than the Monte Carlo 100k case 4 out of 5 times as well as the Monte Carlo 30k case 4 out of 5 times. The SOTA method equaled or improved the prediction of rejects over the Linearized Method in all cases.

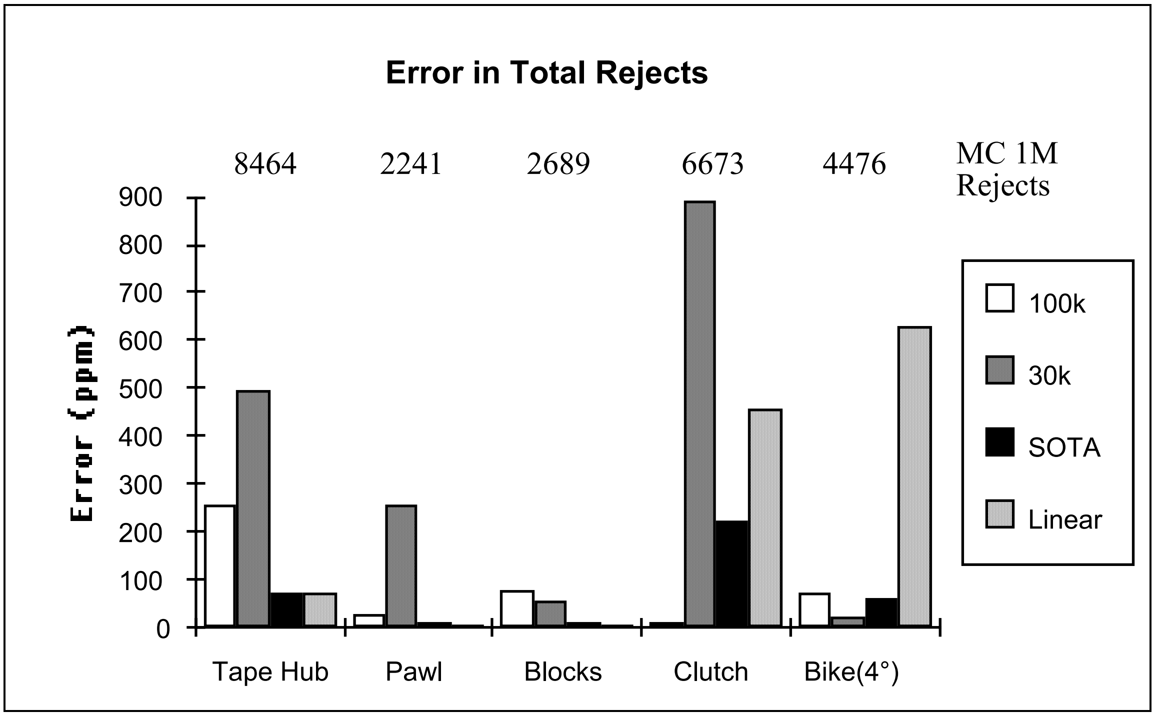

Figure 4.10: Error in Total Rejects Chart

4.6.3 Error vs. Time

There was a dramatic difference in computation time between the SOTA method and Monte Carlo simulation as shown in Table 4.21. All analyses were computed using a HP 9000 series workstation, model 715/64. On average the SOTA method was roughly 900 times faster than the Monte Carlo simulations with 100k samples and 280 times faster than the Monte Carlo simulations with 30k samples. Viewed another way, the computational effort of the SOTA method is equivalent to about 110 Monte Carlo simulations. The time difference between the Linearized Method and the SOTA method was less noticeable, since both analyses were completed in less than one second for all case studies.

Table 4.22: Analysis Time in Seconds

	Tape Hub	Pawl	Blocks	Clutch	Bike (4°)
MC 1M	3833	6206	9146	3119	3973
MC 100k	374	610	900	316	382
MC 30k	112	189	268	94	113
SOTA	0.65	0.44	0.80	0.22	0.65
Linear	0.073	0.071	0.073	0.070	0.058

Figure 4.11 was created by adding a second axis representing analysis time to Figure 4.10. This figure demonstrates a combined speed and accuracy measure of the different tolerances analysis methods. The impressive performance of the SOTA method is illustrated clearly in this chart. The SOTA method was able to perform each of the analyses in under one second and predict total rejects to within 250 parts-per-million for all the case studies. The speed and accuracy performance is a significant improvement over the Monte Carlo simulation analyses as well as a noteworthy accuracy improvement over the Linearized Method.

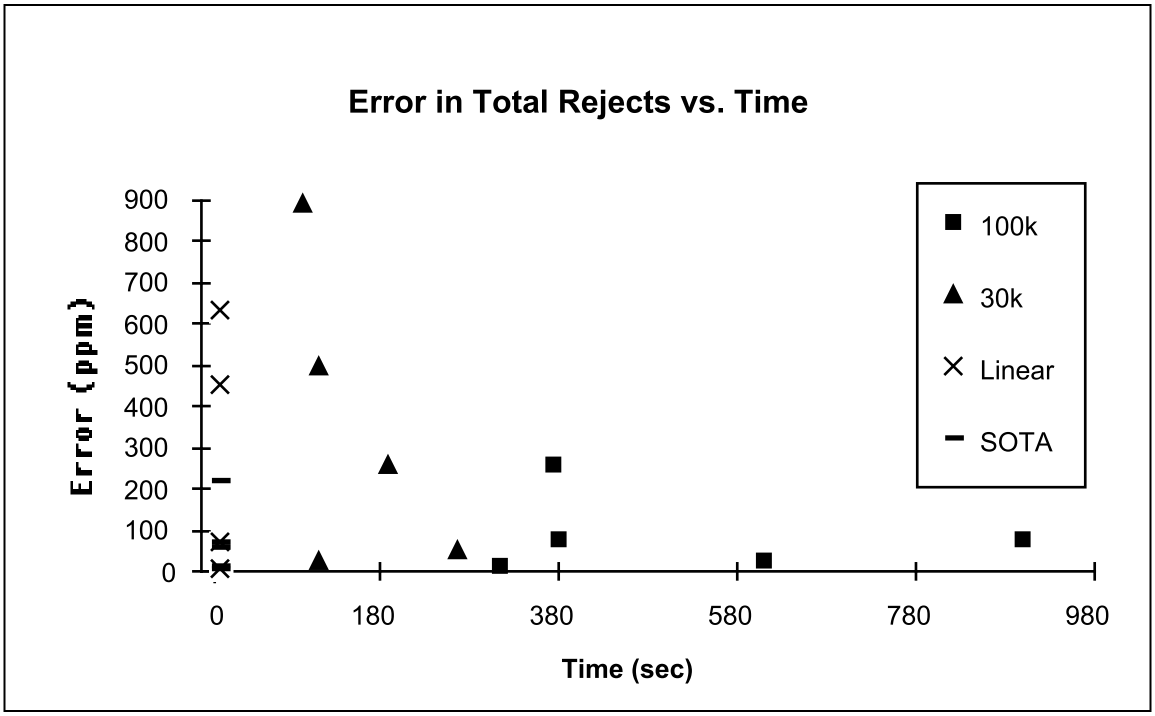

Figure 4.11: Error in Rejects vs. Time Chart

4.6.4 Relative Effort

While analysis time is useful for comparing the performance of the different tolerance analysis methods, it is not independent of the computer system. A more CPU independent measure is to estimate the computational effort of each method. For the three methods, Monte Carlo simulation, the SOTA method and the Linearized Method, a relative measure of effort is easily formulated. The common operation of these three tolerance analysis methods is that each must perform a linear solution of the loop equations. For example, the SOTA method and Monte Carlo simulation require a linear solution of the loop equations for each iteration of Newton's method. Of course the Linearized Method only requires a single linear solution. So, if the Linearized Method is given a effort value of 1, the relative effort of Monte Carlo simulation and the SOTA method may be evaluated by the following expressions:

$$\text{MCeffort} = (\text{sample size}) \times (\text{average Newton iterations}) \quad (4.1)$$

$$\text{SOTAeffort} = (2n^2 + 1) \times (\text{average Newton iterations}) \quad (4.2)$$

For Equation 4.2 the variable n is the number of component dimensions. The effort for traditional Monte Carlo systems would only be one explicit evaluation, or one iteration. However, since the case studies required the consideration of closed-loop constraints, an enhanced Monte Carlo system using an iterative solution had to be used.

It would be expected that the number of iterations be greater for Monte Carlo simulation than for the SOTA method. For each nonlinear solution, Monte Carlo simulation changes the nominal value of all the component dimensions, whereas the SOTA method only changes one or two component dimensions for each solution. Furthermore, the step size used by the SOTA method will generally be very small compared to the variations required by Monte Carlo. The average number of Newton iterations for the five case studies is shown in Table 4.23.

Table 4.23: Average Iterations

	Tape Hub	Pawl	Blocks	Clutch	Bike (4°)
MC 1M	2.90	2.99	3.00	3.40	3.06
MC 100k	2.90	2.99	3.00	3.40	3.06
MC 30k	2.90	2.99	3.00	3.41	3.06
SOTA	2.00	2.00	2.00	2.16	2.21
Linear	1	1	1	1	1

Table 4.24 shows the relative effort of each method for the five case studies. The difference between the SOTA method and the Linearized Method is more evident with the relative effort than with analysis time. The advantage of the SOTA method over Monte Carlo simulation is clearly shown by the measure of relative effort.

Table 4.24: Relative Effort

	Tape Hub	Pawl	Blocks	Clutch	Bike (4°)
MC 1M	2,900,000	2,990,000	3,000,000	3,400,000	3,060,000
MC 100k	290,000	299,000	300,000	340,000	306,000
MC 30k	87,000	89,700	90,000	102,300	91,800
SOTA	326	198	146	41	285
Linear	1	1	1	1	1

4.6.5 The Effect of a Nonlinear Problem

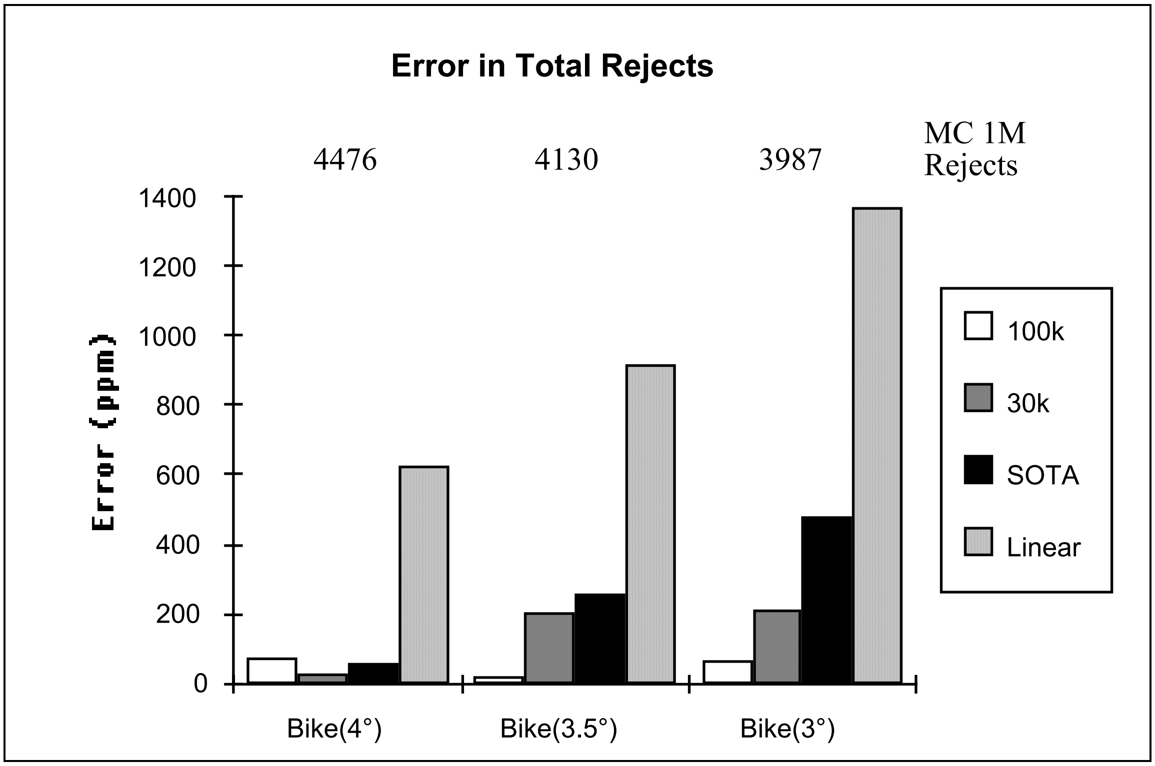

Figure 4.12: Error in Total Rejects for Bicycle Crank Cases

As explained above the bicycle crank problem was analyzed at three different wedge angle values, 4°, 3.5°, and 3°. The three sets of results clearly illustrated how a nonlinear problem can effect the accuracy of a tolerance analysis. In Figure 4.12 the effect is most pronounced for the Linearized Method. As the wedge angle decreased in size, the error in the prediction of total rejects increased. This same increase in error occurred for the SOTA method as well, although not as severe.

It should be noted that the variation in the wedge angle was ± 1.0°. So, the ratio of variation to nominal value ranged from 0.25% to 0.33%. The large magnitude of these ranges violates the basic premise of small variations about the nominal, making the problem highly nonlinear.

Also interesting to note are the ratios of variations to nominal values for the resultant assembly dimensions for all five case studies. Table 4.25 list these ratios as the percent change of the assembly dimensions. The percent change related to the nonlinearity of the assembly function for each case study. Notice that the two cases which displayed nonlinear assembly functions, the bike crank and the one-way clutch, have the highest percent change. The remaining three cases exhibited fairly linear assembly functions, and their percent changes are relatively small.

Table 4.25: Percent Change of Assembly Dimensions

CASE	Assembly Dimension	Standard Deviation	Percent Change
Bike (4°)	5.0852	0.625727	12.3%
Clutch	7.0184	0.219884	3.13%
Blocks	18.7182	0.099913	0.53%
Hub	1.79755	0.002277	0.13%
Pawl	99.0679	0.098336	0.10%

The percent change is not a fail-safe measure of nonlinearity. Certainly a simple linear stack of dimensions could exhibit a large percent change if the variations of each dimension were large. However, for these five two-dimensional assemblies, it does provide a useful correlation.

4.6.6 Non-normal Inputs

Two cases were analyzed for component dimensions with non-normal distributions. The two cases were the tape hub and the one-way clutch. The Linearized Method assumes all component distributions are normally distributed and was, therefore, not used in the analysis of these two cases. Figure 4.13 show the results of the two non-normal cases.

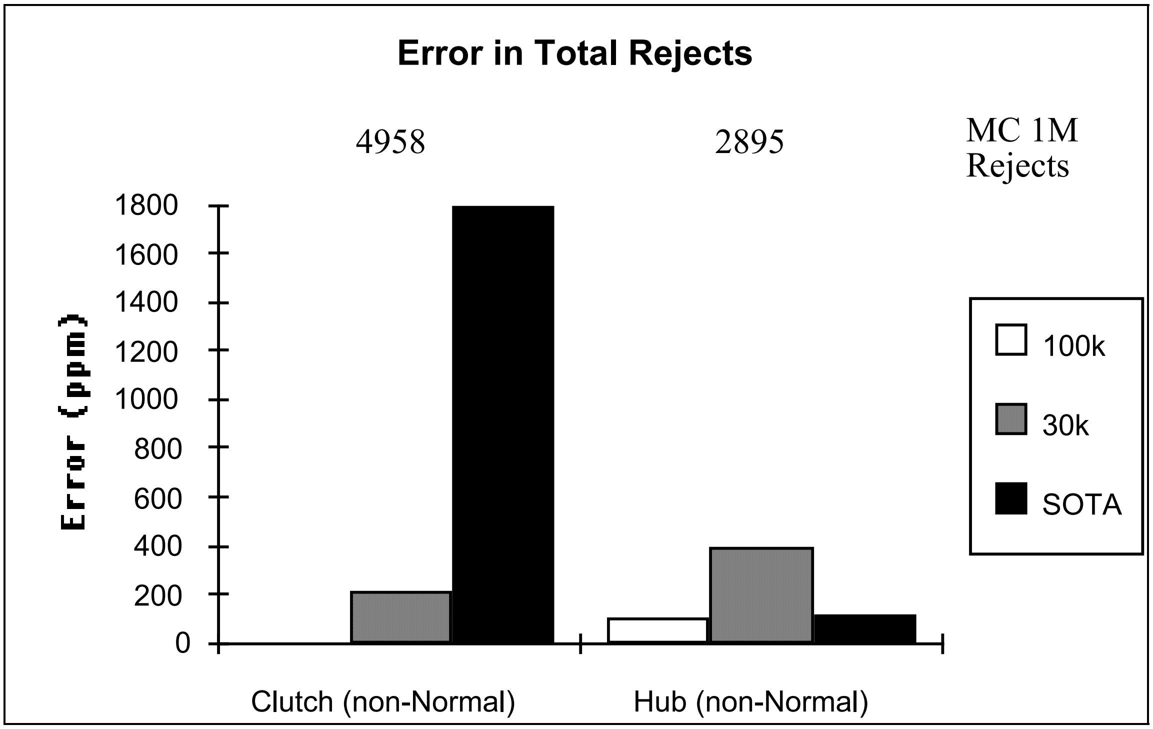

Figure 4.13: Error in Total Rejects for Non-normal Cases The tape hub had proved to be a fairly linear assembly function, while the one-way clutch had shown to be nonlinear. The error in prediction of rejects for the non-normal distributions in Figure 4.13 showed the SOTA method to perform much better for the tape hub, the linear case. The accuracy of the SOTA method for the tape hub case was comparable to the Monte Carlo 100k analysis. For the one-way clutch, the error of the SOTA method increased significantly, even though the skewness and kurtosis values, shown in Table 4.26, seemed to correspond well with the MC 1M analysis. The combination of nonlinear assembly function and non-normal inputs did effect the accuracy of the SOTA method.

Table 4.26: Skewness and Kurtosis for Non-normal Cases

Case		Components	Number of Variables	SOTA Assembly	MC 1M Assembly
Tape Hub	Skewness	0.5	8	0.05837	0.05902
	Kurtosis	3.5		3.1185	3.1109
Clutch	Skewness	0.5	3	-0.5206	-0.5258
	Kurtosis	3.5		3.3057	3.5822

In practice, when several component dimensions contribute to an assembly variation, they are not likely to be skewed in the same direction, as assumed in this problem. The skewness in such cases would tend to be self-canceling, resulting in an assembly distribution which is more nearly normal. A highly skewed assembly is more likely to occur when only one component is skewed and it is also the dominating source of variation, that is, an order of magnitude greater than any other component variation. Thus, highly skewed assembly distributions should be relatively rare.

The next chapter will list the contributions and conclusions of this thesis as well as some recommendations for future research.

Chapter 5

CONCLUSIONS AND RECOMMENDATIONS

The goal of this thesis was to develop a second-order tolerance analysis method for vector-loop models. This chapter contains a summary of the results of the thesis, including the conclusions of this research and suggestions for further study.

5.1 Contributions

A new tolerance analysis method, the Second-Order Tolerance Analysis (SOTA) method, has been developed for use with vector-loop-based assembly models. The SOTA method provides an alternative nonlinear analysis method to Monte Carlo simulation without a significant increase in computation time over the Linearized Method. The SOTA method is well suited for design iteration because of its speed. The SOTA method also includes the capability to handle non-normal distributions for the component dimensions as well as non-normal distributions for the assembly dimensions.

Monte Carlo systems are not well-suited for design work because of computation requirements and modeling restrictions. As described in Chapter 2, recent research [Gao 1993] has been able to increase the modeling capability of Monte Carlo simulation by adding closed-loop constraints, but it was demonstrated in Chapter 4 that the already excessive computation requirements were tripled. The strategy for the SOTA method was very different from adding capability to a Monte Carlo system. The strategy for the SOTA method was to take a very capable design tool, the Linearized Method, and increase the accuracy and flexibility of that tool even further. The result, the SOTA method, is a tolerance analysis method well-suited for design iteration, with added accuracy for nonlinear problems and increased flexibility for non-normal distributions.

Central to the achievement of the development of the SOTA method was the application of the Method of System Moments (MSM) to the implicit variables of a system of nonlinear equations. For vector-loop tolerance models the assembly dimensions are contained implicitly in the loop equations. The challenge of this research involved finding a general method of determining the first and second-order derivatives of the assembly dimensions with respect to the component dimensions. This challenge was overcome using a numerical method to approximate these derivatives.

5.2 Conclusions

This thesis included a comparison of the SOTA method, the Linearized Method, and Monte Carlo simulation using sample sizes of one million, one hundred thousand and thirty thousand.

5.2.1 Comparison of the three methods

The SOTA method's performance for the case studies was impressive. For the five case studies, the accuracy of the SOTA method was comparable to Monte Carlo simulation with 100,000 samples. The accuracy of the Linearized Method compared to Monte Carlo simulation with 30,000 samples. The analysis speed of the SOTA method was comparable to the Linearized Method. The SOTA method is better suited to design iteration than Monte Carlo simulation.

The SOTA method was developed to meet six specific tolerance analysis features. Conclusions relevant to these six features are discussed below.

5.2.2 Speed

The analysis speed of the SOTA method was dramatically less than Monte Carlo simulation and, surprisingly, was very comparable to the Linearized Method. For all the case studies, the analysis time using the SOTA method was under one second.

5.2.3 Tolerance allocation

While tolerance allocation has not yet been incorporated into the SOTA computer program, the application of the same allocation algorithms used by the Linearized Method should be straight forward. The information needed for such algorithms is already available from the SOTA analysis. The existing allocation methods simply require the linear partial derivatives as well as the tolerance values for each controlled dimension.

5.2.4 Closed-loop constraints

Closed-loop constraints, or kinematic constraints, were included through the use of vector-loop models of the five case studies.

5.2.5 Nonlinear effects

The Method of System Moments provided a second-order model of the nonlinear loop equations. The SOTA method showed significant improvement in accuracy over the Linearized Method for nonlinear problems. The accuracy of the SOTA method was comparable to Monte Carlo simulation with 100,000 samples.

5.2.6 Non-normal inputs

The Method of System Moments required the distributions of the controlled dimensions be characterized by the first eight standardized moments. Two case studies were analyzed with non-normal inputs. The accuracy was very good for the linear case. However, for the nonlinear case the error increased. The prediction of rejects for the non-normal, nonlinear case was the least accurate, however, the predicted value was of the same order of magnitude as the more accurate answer.

5.2.7 Non-normal outputs

The output of the Method of System Moments was the first four distribution moments of the assembly dimensions. The four moments were then fitted with the Generalized Lambda Distribution in order to provide an estimate of the fraction of unacceptable assemblies. The first four moments calculated by the SOTA method compared very well with the Monte Carlo simulation of one million. The SOTA method's estimate of skewness was exceptionally good.

5.2.8 Differentiation Scheme

The differentiation scheme employed by the SOTA method consisted of three difference equations and a nonlinear system solver, Newton's method for systems. This scheme functioned well. To obtain all the necessary partial derivatives for the Method of System Moments required $2n^2 + 1$ function evaluations. On average, each function evaluation needed only 2.0 iterations of Newton's method before reaching the convergence tolerance of 1.0e-6.

5.3 Recommendations

1. *Add tolerance allocation to the SOTA method*. To complete the six features of the SOTA method tolerance allocation algorithms should be added. Including this capability in the SOTA method does not require any additional research. Existing allocation methodology uses the linear partial derivatives and component tolerances. This information is already available from the SOTA analysis. Additional research into expanding current allocation methodology to include second-order terms may also be profitable.

2. *Add open loop capability to the SOTA method*. Most design specifications, including gap, orientation and position, require the capability to analyze open vector loops. The addition of this capability is essential to the build the functionality of the SOTA method.

3. *Include geometric tolerances*. Geometric tolerances include important variation sources caused by irregularities in form, shape and size. These variation source have been shown to make significant contributions in a tolerance analysis. In fact, the non-normal distribution capability of the SOTA method will enhance the accuracy of modeling geometric tolerance, since many geometric tolerances are known to be skewed or one-tail distributions.

4. *Expand SOTA method capability for three dimensional assemblies*. The enhancements required to include three dimension capability for a vector-loop model are not challenging. However, the combination of three dimensions and geometric tolerances has proven to be difficult. The use of non-contributing rotations inserted into the vector loops may help to orient each geometric tolerance properly.

5. *Implement a line search in addition to Newton's method for overdetermined assemblies*. With the addition of three dimension capability comes the need to handle the case when the vector-loop tolerance model supplies more equations than unknown assembly dimensions. This situation is referred to as an overdetermined system. A numerical approach such as a line search is useful in this case to prevent Newton's method from diverging.

6. *Add Pearson distribution empirical fit.* While the Generalized Lambda Distribution covers most of the skewness and kurtosis plane, a Johnson or Pearson fit would be able to cover the entire plane. Initial investigation has shown that matching of moments would be very difficult to implement for the Johnson bounded region. The

SOTA method estimates the first four moments of the assembly dimension, so the matching of moments must be used. It is suggested that the Pearson family be investigated for possible use.

7. *Test accuracy of empirical fits, especially out in the tails at high quality levels*. More information needs to be gathered on the accuracy of various empirical fits including, normal, Generalized Lambda, Johnson and Pearson.

8. *Benchmark results between the SOTA method, Linearized Method and Monte Carlo simulation at different quality levels*. Monte Carlo simulation requires larger sample sizes for specification limits at higher quality levels. It would be useful to understand the sample sizes need by Monte Carlo to achieve the same accuracy as the SOTA method and the Linearized Method at quality levels such as 4.5 sigma and 6 sigma.

9. *Calculate optimal step size for accurate derivatives numerically*. To ensure accurate derivatives numerical methods may be implemented to estimate the optimal step size based on the specific vector-loop tolerance model. Such numerical methods should be investigated for application in the differentiation scheme of the SOTA method.

APPENDIX A

Optimal Step Sizes For Equations 3.3, 3.4 and 3.5

Finite difference equations are subject to both truncation error and round-off error. Care must be used to select a step size, Δx , to minimize both these error sources. In general, as the step size increases, the truncation error will increase, and as the step size decreases, the round-off error will increase. An optimal step size can be found which balances the effect of truncation error and round-off error and minimizes the total error.

For example, the total error for the central difference equation, Equation 3.3, is

$$E(\Delta x) = \frac{e}{\Delta x} + \frac{\Delta x^2}{6} u'''(\xi) \tag{A.1}$$

The first term of Equation 3.4 is the round-off error. The variable e in the round-off term represents the precision of the numerical values. This value is effected by the convergence tolerance of Newton's method as well as the precision of the nominal values of the controlled dimensions. The second term of Equation A.1 is the truncation error. The notation $u'''(\xi)$ represents the third derivative of the resultant dimension function at some value ξ, where $x - \Delta x < \xi < x + \Delta x$.

To find the optimal step size, the derivative of Equation A.1 with respect to the step size is set equal to zero and then solved for the step size, Δx. Following these steps the optimal step size for the central difference Equation 3.3 is

$$\Delta x = \left[\frac{3e}{u'''(\xi)} \right]^{\frac{1}{3}} \tag{A.2}$$

The same procedure can be performed for difference Equations 3.4 and 3.5. The total error equation for approximation to the quadratic partial derivatives, difference Equation 3.4, is

$$E(\Delta x) = \frac{4e}{\Delta x^2} + \frac{\Delta x^2}{12} u^{(4)}(\xi) \tag{A.3}$$

Setting the first derivative equal to zero and solving for the step size gives

$$\Delta x = \left[\frac{48e}{u^{(4)}(\xi)} \right]^{\frac{1}{4}} \tag{A.4}$$

For the approximation to the cross-partial derivatives, difference Equation 3.5, the total error equation is

$$E(\Delta x) = \frac{e}{\Delta x^2} + \frac{\Delta x}{6} u'''(\xi) \tag{A.5}$$

The optimal step size for difference Equation 3.5 is then

$$\Delta x = \left[\frac{12e}{u'''(\xi)} \right]^{\frac{1}{3}} \tag{A.6}$$

The optimal step size equations may not seem to be valuable, since third and fourth derivatives are needed when first and second derivatives are being approximated. However, there are some tolerance models for which closed form solutions exist, and the bounds on the third and fourth derivative may be calculated. Such closed form solutions could help to establish a good guess for the perturbation step size for a wide range of problems.

APPENDIX B

Derivatives for the Case Studies

B.1 Tape Hub Derivatives

First-order derivatives of length H:

A	B	θ	D	E	F	G	I
-0.2679	1.0000	-0.3783	1.0352	1.0000	1.0000	0.2679	0.2679

Second-order derivatives of length H:

	A	B	θ	D	E	F	G	I
A	0.0000	0.0000	1.0717	0.0000	0.0000	0.0000	0.0000	0.0000
B	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
θ	1.0717	0.0000	0.2648	-0.2774	0.0000	0.0000	-1.0717	-1.0717
D	0.0000	0.0000	-0.2774	0.0000	0.0000	0.0000	0.0000	0.0000
E	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
F	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
G	0.0000	0.0000	-1.0717	0.0000	0.0000	0.0000	0.0000	0.0000
I	0.0000	0.0000	-1.0717	0.0000	0.0000	0.0000	0.0000	0.0000

B.2 Stacked Blocks Derivatives

First-order derivatives of length A:

F	G	I	K	J	B,C
-0.2731	0.7418	1.0366	0.2581	-0.0705	1.3097

Second-order derivatives of length A:

	F	G	I	K	J	B,C
F	0.0000	0.0412	0.0000	-0.0412	0.0112	0.0000
G	0.0412	0.0205	-0.0108	-0.0205	0.0162	-0.0521
I	0.0000	-0.0108	0.0000	0.0108	-0.0029	0.0000
K	-0.0412	-0.0205	0.0108	0.0205	-0.0162	0.0521
J	0.0112	0.0162	-0.0029	-0.0162	0.0073	-0.0142
B,C	0.0000	-0.0521	0.0000	0.0521	-0.0142	0.0000

B.3 Pawl and Ratchet Derivatives

First-order derivatives of angle $\phi 4$:

A	B	C	D	H	θ	E
2.4053	-0.6959	-2.5040	-2.3211	0.3081	0.0290	0.3490

Second-order derivatives of angle $\phi 4$:

	A	B	C	D	H	θ	E
A	0.6668	2.2400	-0.0179	-5.4569	-8.2944	-0.7817	-9.3940
B	2.2400	7.6934	-0.0134	-4.0927	-6.2208	-0.5862	-7.0455
C	-0.0179	-0.0134	0.0135	4.1044	6.2386	0.5879	7.0657
D	-5.4569	-4.0927	4.1044	11.1322	10.6696	1.4008	12.0841
H	-8.2944	-6.2208	6.2386	10.6696	6.6845	0.6299	7.5707
θ	-0.7817	-0.5862	0.5879	1.4008	0.6299	0.1077	0.8990
E	-9.3940	-7.0455	7.0657	12.0841	7.5707	0.8990	8.5743

B.4 One-way Clutch Derivatives

First-order derivatives of angle $\phi 1$:

A	C,D	E
-0.2078	-0.4141	0.206319

Second-order derivatives of angle $\phi 1$:

	A	C,D	E
A	-0.3510	-0.7046	0.353659
C,D	-0.7046	-1.4145	0.709908
E	0.3536	0.7099	-0.356250

B.5 Bicycle Crank Derivatives

First-order derivatives of length U2:

A	B	C	D	E	F	θ	G
-14.300	-14.300	1.000	-14.335	14.335	-14.335	-124.962	14.300

Second-order derivatives of length U2:

	A	B	C	D	E	F	θ	G
A	0.000	0.000	0.000	0.000	0.000	0.000	-205.509	0.000
B	0.000	0.000	0.000	0.000	0.000	0.000	-205.509	0.000
C	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
D	0.000	0.000	0.000	0.000	0.000	0.000	-205.008	0.000
E	0.000	0.000	0.000	0.000	0.000	0.000	205.008	0.000
F	-0.001	0.000	0.000	0.000	0.000	0.000	-205.008	0.000
θ	-205.509	-205.509	0.000	-205.008	205.008	-205.008	-3526.649	205.509
G	0.000	0.000	0.000	0.000	0.000	0.000	205.509	-0.000

BIBLIOGRAPHY

Burden, Richard L., Faires, J. Douglas, *Numerical Analysis Fifth Edition*, PWS Publishing Co., 1993

Chase, Kenneth, *CATS 2-D User's Guide*, ADCATS Report No. 92-2, July 29, 1992

Cox, N. D., “Tolerance Analysis by Computer”, *Journal of Quality Technology*, Vol. 11, No. 2, April 1979

Cox, N. D., “Volume 11: How to Perform Statistical Tolerance Analysis”, American Society for Quality Control, Statistical Division.

Dudewicz, E. J., Ramberg, J. S., Tadikamalla, P. R., "A Distribution for Data Fitting and Simulation", An. Tech. Conf. Trans. A.S.Q.C. 28, pp. 407--418, 1974

Evans, D. H., “Statistical Tolerancing: State of the Art, Part 1. Background”, *Journal of Quality Technology*, v 6, n 4, pp. 188-195, Oct. 1974

Evans, D. H., “Statistical Tolerancing: State of the Art, Part 2. Methods of Estimating Moments”, *Journal of Quality Technology*, v 7, n 1, pp. 1-12, Jan. 1975

Evans, D. H., “Statistical Tolerancing: State of the Art, Part 3. Shifts and Drifts”, *Journal of Quality Technology*, v 7, n 2, pp. 71-76, Apr. 1975

Gao, J., “Nonlinear Tolerance Analysis of Mechanical Assemblies”, Dissertation, Mechanical Engineering Department, Brigham Young University, 1993

Greenwood, W. H., “A New Tolerance Analysis Method for Designers and Manufacturers”, Dissertation, Mechanical Engineering Department, Brigham Young University, 1987

Hahn, G. J., Shapiro, S. S., *Statistical Modeling in Engineering*, John Wiley & Sons, 1967

Hastings, C., Mosteller, F., Tukey, J., and Winsor, C. "Low moments for small samples: A comparative study of order statistics", *Ann. Math. Stat. 18*, pp. 413-426, 1947

Johnson, N. L., "Tables to facilitate fitting S_U frequency curves", *Biometrika 52*, pp. 547-558, 1965

Johnson, N. L., "Systems of frequency curves generated by methods of translation", *Biometrika 36*, pp. 274-282, 1949

Larsen, D. V., "An efficient Method for Iterative Tolerance Design Using Monte Carlo Simulation", M.S. Thesis, Mechanical Engineering Department, Brigham Young University, 1989

Marler, Jaren D., "Nonlinear Tolerance Analysis Using the Direct Linearizating Method", M.S. Thesis, Mechanical Engineering Department, Brigham Young University, 1988

Press, W. H., Vetterling, W. T., Teukolsky, S. A., Flannery, B. P., *Numerical Recipes in C, The Art of Scientific Computing, Second Edition*, Cambridge University Press, 1992

Ramberg, J. S., Tadikamalla, P. R., Dudewicz, E. J., Mykytha, E. F., "A probability distribution and its uses in fitting data", *Technometrics 21*, pp. 201-214, 1979

Sandor, G. N., Erdman, A. G., *Advanced Mechanism Design: Analysis and Synthesis Volume 2*, Prentice-Hall, Inc., 1984

Shapiro, Samuel S., Alan J. Gross, *Statistical Modeling Techniques,* Marcel Dekker, Inc., New York, 1981