Zhige Zhou
Graduate Student Research Assistant
e-mail: z.g.zhou@263.net

Wenzhen Huang
Professor
e-mail: wzhuang@mail1.sjtu.edu.cn

Li Zhang
Graduate Student Research Assistant
e-mail: angel-henry@263.net

The State Key Laboratory for Vibration Shock & Noise,
Mechanical Engineering Institute,
Shanghai Jiao Tong University,
No. 1954, Huashan Road, Shanghai,
P. R. China, 200030

Sequential Algorithm Based on Number Theoretic Method for Statistical Tolerance Analysis and Synthesis

Tolerancing is one of the most important tasks in product and manufacturing process design. In the literature, both Monte Carlo simulation and numerical optimization method have been widely used in the process of statistical tolerance analysis and synthesis, but the computational effort is huge. This paper presents two techniques, quasi random numbers based on the Number Theoretic Method and sequential algorithm based on the Number Theoretic net, to calculate yield and to perform tolerance allocation. An example demonstrates the optimal tolerance allocation design and is employed to investigate the efficiency and accuracy of this solution. This algorithm can efficiently obtain the global optimum, and the amount of calculation is considerably reduced.
[DOI: 10.1115/1.1378795]

1 Introduction

The allocation of tolerances is closely tied to the overall quality and cost of a product. If the tolerance is too loose, the probability of an assembly function acceptably (yield) is low. On the other hand, if the tolerance is too tight, the manufacturing cost becomes high. Thus the tolerance allocation becomes an optimization problem to determine the optimal allotment of the tolerance under the constraints of the functional requirements and acceptance probability (spec yield).

A process diagram for a typical statistical tolerance synthesis process is shown in Fig. 1 [1]. The manufacturing process model takes the values of the tolerances and uses them to determine the manufacturing cost and the statistics of the manufacturing variations. The tolerance analysis process uses the manufacturing variation statistics to generate instances of the part or assembly that are then analyzed. The analysis calculates a design function value for this instance, which is a measure of the product's performance. Design function statistics are then calculated for the instances that have been analyzed. Typically, the statistics calculation is matched to the algorithm for generating the variation instances [2]. The optimization algorithm to determine the next set of tolerances to test uses the design function statistics and the manufacturing cost. The process continues until the optimum tolerances have been found.

The yield (Y) is computed as the probability. Let x_{il} and x_{iu} represent the lower and the upper limits of an individual dimension x_i in an assembly. Then the yield is represented as

$$Y=\int_{x_{1l}}^{x_{1u}}\cdots\int_{x_{nl}}^{x_{nu}}q(x_1,\cdots,x_n)\phi(x_1,\cdots,x_n)dx_1\cdots x_n\,, \qquad (1)$$

where $\phi(x_1,\cdots,x_n)$ is the multivariate normal probability density function, and $q(x_1,\cdots,x_n)$ is the test function which checks whether a stochastically selected point is in the reliable region. If $F_i(x_1,\cdots,x_n)>0$ for all design functions, where F_i is an assembly constraint function, $q(x_1,\cdots,x_n)=1$, otherwise $q(x_1,\cdots,x_n)=0$.

The Monte Carlo simulation, or random sampling, is being used extensively for statistical tolerance analysis and synthesis, but computing effort is large. On the other hand, the number optimization algorithms such as genetic algorithm and simulated annealing [3] is often used to optimize the tolerance, but several million cases must be analyzed to find the optimum using these algorithms [4].

In the method presented here, techniques based on the Number Theoretic Method are used in tolerance analysis and syntheses to estimate spec yield and to optimize the tolerance respectively. Instead of random numbers based on Monte Carlo Method, pseudo-random numbers based on the Number Theoretic Method are utilized to approximate by using a very small number of sampling points. These greatly decrease the computing effort required for the Monte Carlo simulation. At the same time, a sequential algorithm based on the Number Theoretic Method is effectively applied to solve questions with nonlinear constraints in the process of tolerance synthesis.

Previous research on tolerance analysis and synthesis is reviewed in the followed section; the technique for reducing calculation effort in tolerance analysis is discussed in section 3; and the technique for optimizing the tolerance is described in section 4; an example is presented in section 5.

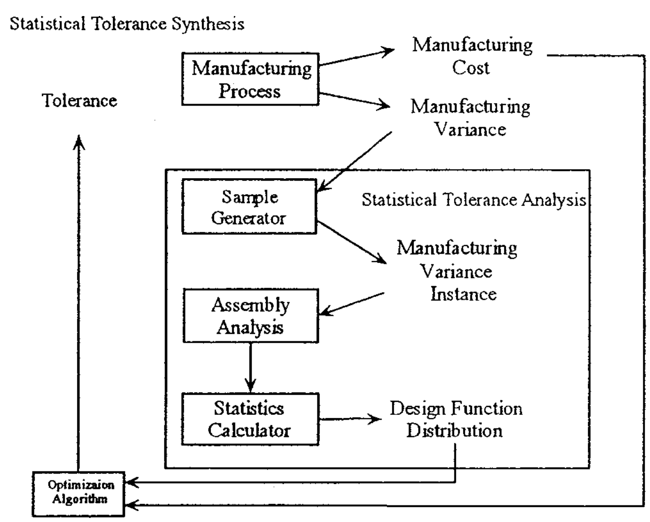

Fig. 1 Typical statistical tolerance synthesis process

Contributed by the Manufacturing Engineering Division for publication in the JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING. Manuscript received June 1999; revised July 2000. Associate Editor: E. DeMeter.

2 Background

Initial research on statistical tolerance synthesis made two assumptions: (1) individual variations are normally distributed; and (2) the overall design function of the part or assembly is linear with respect to the individual variations. These assumptions are based on two simplifications: (1) the distributions of the individual variations can be expressed as functions of the tolerance; and (2) all values of the overall design function are normally distributed with a variance, which is a linear combination of the individual variations. Finally, using standard optimization techniques can solve the statistical tolerance synthesis problem.

However, the assembly analysis shown in Fig. 1 usually produces a nonlinear design function. For nonlinear design functions, the distribution of values is not normal, and the design function variance is no longer a linear combination of the variance of the individual variations. In general, the variances and other parameters of the design function distribution cannot be expressed as a simple function of the tolerances.

Gadallah and ElMaraghy [5] proposed a regression analysis technique. It is similar to finite differences but the points are more widely spaced. The result was a quadratic formula approximated by the actual design function. This technique does provide sensitivity information, so that the efficient optimization algorithms based on gradient could be used, but the requirements for Monte Carlo simulations for each point of the orthogonal array means that a large number of individual analyses were performed.

Skowronski and Turner [1] applied two techniques (correlation and approximation function, important sampling) in Monte Carlo simulation to increase efficiency. The important sampling technique is especially useful when there are some regions of the probability space that are more important than others. In tolerance synthesis, this region would be the vicinity of the boundary of the acceptance zone. However, constructing an important function for such a region may be difficult, since finding the location of this boundary in manufacturing variation space is part of the problem of tolerance analysis.

Because the computed yield is virtually the same as the actual yield if accurate stochastic analysis methods are used, Monte Carlo simulation is used widely to calculate the design function values [3,6,7]. Unfortunately, Monte Carlo simulation (and other stochastic methods) requires many sampling points to assure high accuracy. In fact, the accuracy of the basic Monte Carlo technique is proportional to the square root of the number samples used. It may be unsuitable for the inner iteration of many classical optimization algorithms because of the enormous computational efforts.

On the other hand, some optimization algorithms use gradient or derivative information to optimize tolerance [8,9]. They are not capable of some complex cases, and cannot obtain the global optimization results sometimes. Current research on statistical tolerance synthesis has focused on a class of optimization algorithms called direct search techniques. These algorithms do not use gradient or derivative information in determining the optimistic tolerance since gradient information is difficult to compute. Although the direct search algorithm can obtain global results, the convergence rate is slow.

An example of a direct search algorithm that has been used in tolerance synthesis is the genetic algorithm [10]. In a genetic algorithm, sets of tolerance values, call their fitness, are analyzed and ordered according to their fitness (objective function and value). Those that are most fit are used to create the next generation. Lee determined that a very inaccurate measure of fitness would still allow them to calculate an optimum if they extended the algorithm through enough generations. However, a large number of assembly analyses (30 samples per case, 100 cases per generation, and 300 generations for 900,000 analyses) were required and the exact optimum was not found.

Another kind of direct search algorithm is the simulated annealing algorithm [3,11]. Simulating annealing algorithms are traditionally applied to discrete optimization problems. The algorithm ensures that all constraints are met, and an optimal solution may be found in the meantime. Zhang and Wang [11] present a methodology to maximize a product's robustness by appropriately allocating assembly and machining tolerances. The robust tolerance design problem is formulated as a mixed nonlinear optimization model. A simulated annealing algorithm is employed to solve the model and an example is presented to illustrate the methodology.

3 Numerical Integration Based on the Number Theoretic Method

The essence of the Number Theory Method is to find uniformly distributed points in a domain of I^s, where $I=[0,1]$. This sequence can be used to replace the random numbers based on Monte Carlo simulation. There are a number of methods to produce sets of points $\{b_k, k=1,\cdots,n\}$ in I^s [12]. The algorithm to produce representative points used in this paper is as follows: Let $(n;h_1,\ldots,h_s)$ be an integral vector, where $h_1=1$, $0<h_i<n$ and greatest common divisor $(n,h_i)=1$, $i=1,\cdots,s$. Let

$$p_n(k)=(kh_1,\cdots,kh_s)=(q_{k1},\cdots,q_{ks})(\bmod n), \quad k=1,\cdots,n, \tag{2}$$

where $0<q_{ki}\leqslant n$. Set

$$b_{ki}=(2q_{ki}-1)/2n, \quad i=1,\cdots,s, \quad k=1,\cdots,n$$

Then $\{b_k\}$ is a set of points in I^s with lower discrepancy if $(n;h_1,\ldots,h_s)$ are carefully selected.

It has been proved that the sequence discrepancy based on the above method for any $\varepsilon>0$ is as follows [13]

$$DP(P,D)=O(n^{-1/s+\varepsilon}) \tag{3}$$

Because the Number Theoretic net is uniformly distributed, it can be proven that the average integral rate of convergence based on the Number Theoretic Method is $O(n^{-1+\varepsilon})$, and it may be $O(n^{-1}(\log n)^s)$ in some cases in the process of integration. On the other hand, the efficiency of Monte Carlo simulation is low because the random numbers are not uniformly distributed. The average rate of convergence of Monte Carlo simulation is $O(1\sqrt{n})$, and it is not less than $O(\sqrt{\ln(\ln(n))/n})$ in any case [13].

4 Sequential Algorithm Based on the Number Theoretic Method

There are many gradient methods for optimization problems such as the iteration method, Newton's method, Brown's method, quasi Newton's method, etc. Unfortunately, there appear only few cases that the global maximum can be reached, and we can obtain usually a local maximum if the function is not unimodal.

A sequential algorithm for optimization based on the Number Theoretic Method is contained in detail in a report of Fang and Wang [14], and it is denoted by SNTO. The continuities are required only for the function f_i' in SNTO so that the convergence of the approximate minimum M^* and the maximum point x^* to the respective M and x_0 are ensured. Besides, it is easy to work out a program in SNTO.

There are many advantages to apply the SNTO optimization method. First, it is assumed that function is continuous, not unimodal and derivative. Second, this method only calculates the function values at representative points, and it need not calculate differential quotient value. Third, SNTO does not depend on the initial value of optimization, and this is not the same as the traditional optimization method.

Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two vectors of $\boldsymbol{R}^s$, where $a_i<b_i$, $i=1,\cdots,s$. We use $[\boldsymbol{a},\boldsymbol{b}]$ to denote the rectangle $[a_1,b_1]\times\cdots\times[a_s,b_s]$. In this paper, the initial area is defined as (0.1, 1, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1, 1, 0.1, 0.1, 0.1) considering the coefficient of $x2$ and $x9$ in objective function.

The sequential algorithm for solving this example as follows:

Step 1. Take n_1 points $P_i=\{\boldsymbol{y}_k^{(1)}=(y_{k1}^{(1)},\cdots,y_{ks}^{(1)}),\ k=1,\cdots,n_1\}$, which are uniformly scattered on $D_1=[\boldsymbol{a},\boldsymbol{b}]$ by the methods in section 3. Find out the minimum M_1 and a minimum point $\boldsymbol{x}^{(1)}$ of $L(\boldsymbol{x})$ on P_1.

Step 2. The domain D_1 is contracted to $D_1=[\boldsymbol{a}^{(2)},\boldsymbol{b}^{(2)}]$, where $\boldsymbol{a}^{(2)}=(a_1^{(2)},\cdots,a_s^{(2)})$, $\boldsymbol{b}^{(2)}=(b_1^{(2)},\cdots,b_s^{(2)})$, $a_i^{(2)}=\max(x_i^{(1)}-c_i^{(1)}/2,a_i)$, $b_i^{(2)}=\max(x_i^{(1)}+c_i^{(1)}/2,b_i)$, $i=1,\cdots,s$, and $c^{(1)}=(\boldsymbol{b}-\boldsymbol{a})/2=(c_1^{(1)},\cdots,c_s^{(1)})$. Then, take n_2 the points $P_2=\{\boldsymbol{y}_k^{(2)},k=1,\cdots,n_2\}$, which are uniformly scattered on D_2, and find out the minimum.

Step 3. Suppose that in the jth step we have found the minimum of the function and the corresponding point x_j^*. By a similar method we can reduce the domain D_j to D_{j+1}, and make a set of points on D_{j+1}, by which we can find another minimum M_{j+1} of the function and the corresponding point x_{j+1}^*.

Repeat Step 3 until the search domain is smaller. The last minimum M_{j+1} is expected to be closed to the global minimum M of the function.

5 Tolerance Synthesis Example

Next, a nonlinear-constraint problem posed by Lee was tried. The sample of the assembly is shown in Fig. 2.

This example involves the assembly of two mating parts. Two of the mating surfaces are at an angle to horizontal. The surfaces have complementary angles and discrepancy between these two angles is a design function.

The corresponding design functions derived from Fig. 2 are as follows [15]:

$$f1=(x6-x5)-(x8-x7)$$

$$f2=(x3-x4)-(x11-x10)$$

$$f3=(x8-x7)(x2-x3)-(x6-x5)(x10-x9)+\tan(pi/180)(x10-x9)(x2-x3)$$

$$f4=(x6-x5)(x10-x9)-(x8-x7)(x2-x3)+\tan(pi/180)(x10-x9)(x2-x3)$$

$$f5=-x1+x12+0.01$$

$$f6=x1-x12+0.01$$

The first two design functions are the vertical and the horizontal clearance conditions of the two parts. The third and the fourth design functions restrict the difference between the angle θ_1 and θ_2 to be within $\pm\pi/180$ rad for successful assembly. The last two conditions require the size difference between two parts to be within ± 0.01. The nominal dimensions are given as $X^T=(50.0, 40.00125, 20.05, 9.9985, 9.9985, 30.0, 10.0, 30.0, 10.05, 30.0, 40.0, 50.0)$. The dimensions are assumed to vary as normal distribution random variables. Their variances are assumed to be equal to $(t_i/6)$, where t_i is the tolerance of the ith dimension.

The modified reciprocal cost-function model is used to define the total manufacturing cost as [15]:

$$C_i(\sigma_i)=a_i\times 10^{-3}/(6\sigma_i)^{b_i} \tag{4}$$

The coefficients in the above equation were set by Lee and Woo [15] as $a_1=0.2$, $a_2=1.0$, $a_3=a_4=0.015$, $a_5=0.008$, $a_6=0.009$, $a_7=0.008$, $a_8=0.006$, $a_9=1.0$, $a_{10}=0.01$, $a_{11}=0.015$ and $a_{12}=0.2$, and $b_1=b_2=\cdots=b_{12}=2.0$.

In the process of tolerance analysis and synthesis, the spec yield is 95 percent. At the last calculation, Monte Carlo simulation was performed using 20000 sampling points for the calculation result to estimate the yields precisely.

In this example, the numbers of the Number Theoretic net are 101 and 521 respectively to calculate the yield and optimize tolerance. The calculation vector for obtaining the Number Theoretic net is derived from the primitive root of the prime number. These vectors are shown in Table 1.

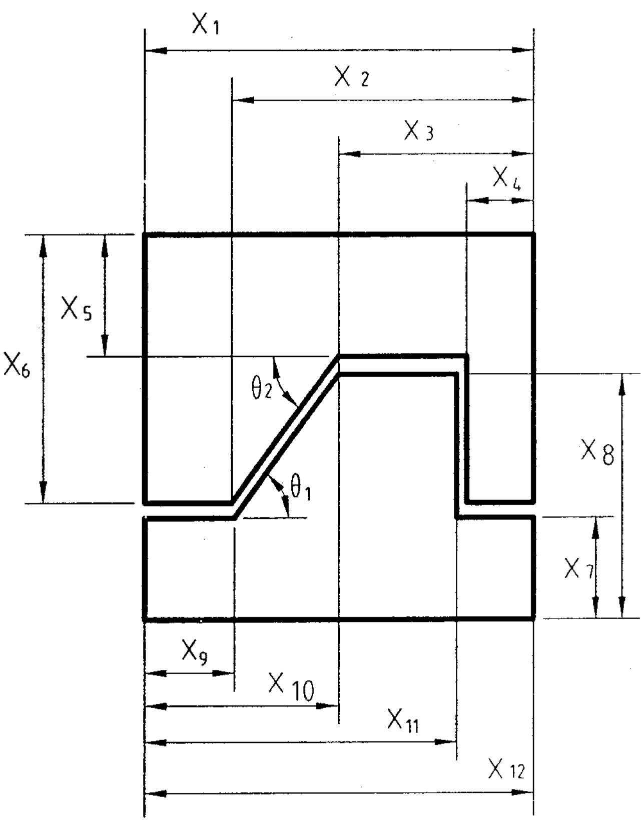

Fig. 2 Assembly for the non-linear constraint

Applying the SNTO to this problem, we set $n_0=n_1=\cdots=521$ for each step, and the results are given in Table 2.

The results of a test run are shown in Table 2. For the test run, the number of representative points was 521, and the number of sampling points was 101. Although there is little difference from Lee and Woo's [15] results for Lee and Johnson's [10] results, the cost and yield had been improved. Compared with the cost obtained by Lee and Woo's [15] algorithm or Lee and Johnson's [10] algorithm, the result shows that about 34 percent or 5 percent of the cost was reduced respectively on the basis of higher yield.

The computational complexity of an algorithm can be represented by various measurements, such as the number of function arithmetic operations, the number of function evaluations, the number of iterations, and the computation time. Two measures, the number of assembly analyses used in the synthesis and the number of floating point operations (flops), are used to compare the computational effort required for the synthesis described above.

The number of assembly analysis is used as a measure of efficiency because the assembly analysis step is the most computationally expensive. Comparing assembly analyses also compen-

Table 1 Calculation Vector for Obtaining Number Theoretic net

n	h1	h2	h3	h4	h5	h6	h7	h8	h9	h10	h11	h12
101	1	3	9	27	81	41	22	66	97	89	65	94
521	1	3	9	27	81	243	208	103	309	406	176	7

Table 2 Performance of optimization tolerance

Tolerances	Lee and Woo's (1990) results	Lee and Johnson's (1993) results			Calculated results
		result1	result2	result3	
x1	0.0187	0.01647	0.01647	0.01294	0.01300
x2	1.231	0.18824	0.20707	0.86591	0.1935
x3	0.0579	0.05897	0.05189	0.03385	0.005543
x4	0.0705	0.05646	0.05646	0.06587	0.005973
x5	0.0019	0.00235	0.00235	0.00235	0.002804
x6	0.0022	0.00212	0.00212	0.00212	0.002763
x7	0.0019	0.00306	0.00306	0.00306	0.002299
x8	0.0015	0.00212	0.00212	0.00212	0.002445
x9	1.2385	0.82765	0.63847	0.82765	0.18238
x10	0.0714	0.05788	0.06212	0.05647	0.03578
x11	0.0579	0.02353	0.01224	0.02259	0.011585
x12	0.0168	0.1176	0.01176	0.01294	0.01617
Cost	10.38	7.90	7.97	8.08	7.74
Yield	94.5%	94.8%	95.4%	95.5%	95.6%

sates for the difference between the manually generated design function used in the preceding example and design functions generated by actual CAD packages.

Lee and Johnson [10] use 30 assembly analyses per data point and evaluate 30,000 data points total in their tolerance synthesis, and the computational effort is 900,000 individual analyses (30 samples/case, 100 cases/generation, 300 generations for 900,000 analyses). One point required about 18,000 flops to evaluate one case with 30 individual analyses, and 540 Mflops was required for the entire tolerance synthesis.

In this paper, the number of tolerance analysis is 157863 (101 representative points/analysis, 521 analyses/iteration, 3 iterations). The computational effort is about 95 Mflops, and this represents a considerable reduction in the number of analyses comparing with Lee and Johnson's [10] results.

6 Summary

This paper presents a novel method for tolerance analysis and synthesis by distributing tolerances so as to satisfy the stack-up conditions. The strategy developed in this paper is unique and practical. As a global criterion, cost minimization is used. Two techniques, quasi random numbers based on the Number Theoretic Method and a sequential algorithm based on the Number Theoretic net, are introduced for tolerance analysis and synthesis. For the nonlinear constraint problem considered here, the algorithm successfully identified a good solution with small numbers of quasi random numbers, and the algorithm satisfied the spec yield with sufficient precision.

The optimal costs are significantly reduced to compare with the costs of the domain-approximation scheme [15] for the cases we have investigated, and the computational effort in this paper is much less than that of the genetic algorithm [10].

The basic idea presented here (quasi random numbers coupled with a sequential algorithm) could be used to solve other stochastic-optimization problems. This is a worthwhile subject for future research.

Acknowledgments

The authors gratefully acknowledge the financial support provided by the National Educational Department Foundation No. 97024832.

References

[1] Skowronski, V. J., and Turner, J. U., 1997, ''Using Monte-Carlo Variance Reduction in Statistical Tolerance Synthesis,'' Computer Aided Design, **29**, pp. 63–69.

[2] Nigam, S., and Turner, J., 1995, ''A Review of Statistical Approach to Tolerance Analysis,'' Computer Aided Design, **27**, pp. 6–15.

[3] Ashiagbor, A., Liu, H.-C., and Nnaji, B. O., 1998, ''Tolerance Control and Propagation for the Product Assembly Modeller,'' Int. J. Prod. Res., **36**, pp. 75–93.

[4] Iannuzzi, M. P., and Sandgren, E., 1995, ''Tolerance Optimization Using Genetic Algorithms,'' *Benchmarking with Manual Analysis, CIRP Conference on Computer-Aided Tolerancing*, CIRP Secretariat, Paris.

[5] Gadallah, M., and ElMaraghy, H., 1994, ''The Tolerance Optimization Problem Using a System of Experimental Design,'' *Advances in Design Automation*, Vol. 69, ASME Publishers, New York, pp. 251–265.

[6] Lin, C. Y., Huang, W. H., and Jeng, M. C., et al., 1997, ''Study of An Assembly Tolerance Allocation Model Based on Monte Carlo Simulation,'' J. Mater. Process Technol., **70**, pp. 9–16.

[7] Arai, T., and Takeuchi, K., 1992, ''Simulation System on Assembly Accuracy,'' CIRP Ann., **41**, pp, 37–40.

[8] Salomous, O. W., Haalboom, F. J., and Poerink, H. J. J., et al., 1996, ''Computer Aided Tolerance Tool II: Tolerance Analysis,'' Computers in Industry, **31**, pp. 175–186.

[9] Chen, M.-S., 1996, ''Optimizing Tolerance Allocation for Mechanical Components Correlated by Selective Assembly,'' International Journal of Advanced Manufacturing Technology, **12**, pp. 349–355.

[10] Lee, J., and Johnson, G. E., 1993, ''Optimal Tolerance Allotment Using a Genetic Algorithm and Truncated Monte Carlo Simulation,'' Computer Aided Design, **25**, pp. 601–611.

[11] Zhang, C., and Wang, H. P., 1998, ''Robust Design of Assembly and Machining Tolerance Allocations,'' IIE Transactions (Institute of Industrial Engineers), **30**, pp. 17–29.

[12] Hua, L. K., and Wang, Y., 1981, *Applications of Number Theory to Numerical Analysis*, Springer-Verlag and Science Press, Berlin and Beijing.

[13] Fang, Kaitai, and Wang, Yuan, 1996, *Application of Number-Theoretic Methods in Statistics*, Scientific Press, Beijing.

[14] Fang, K. T., and Wang, Y., 1989, ''A Sequential Algorithm for Optimization and Its Applications to Regression Analysis,'' Technical Report, 002, Institute of Applied Mathematics, Academic Sinica, Beijing.

[15] Lee, W. J., and Woo, T. C., 1990, ''Tolerance: Their Analysis and Synthesis,'' ASME J. Eng. Ind., **112/113**, pp. 113–121.