SAND2007-5345C

Effective Robust Design Strategy employing Ordinal Variance Minimization and Adaptive Mean Constraint Satisfaction

Francesca D. Reale-Levis* Vicente J. Romero† Laura P. Swiler‡

We present a practical and efficient methodology for design optimization where robustness to uncertain random variables is of paramount concern. In particular, the methodology addresses design of the system to meet a target nominal or mean desired behavior, while finding the point of minimum variability of system behavior in the feasible portion of the design space. The robust-design methodology first involves finding a hyperplane where the response is approximately at the mean value. Then, the method incorporates spatially-correlated ordinal optimization for exceedingly economical variance minimization in the design space. To estimate mean system behavior at any point in the design space, either of two methods are found to be most suitable, depending on how many random variables contribute to variability in system behavior. For less than four such random variables, a Moment Estimation method appears most efficient and effective, otherwise Latin Hypercube sampling with an identifiable number of samples to get appropriately small confidence intervals is most effective. We apply the optimization procedure to an automotive device design robustness problem. We consider a snap-fit composed of a moving part and a stationary part which are designed to the specification of three uncertain design variables. The three design variables and an additional noise variable determine the friction force between the two parts. Using our process, we find a point in the design space of minimum variance subject to a constraint that the mean of the friction force is a set level. We examine a two-dimensional example and a three-dimensional example. We consider an exhaustive search and reliability methods for comparison of results and cost.

I. Introduction

Optimization under uncertainty refers to performing an optimization procedure when there is uncertainty or noise in the variables to be optimized over. This algorithm has been designed to minimize the number of function evaluations necessary to perform optimization under uncertainty (OUU). We present a method that simultaneously minimizes the variance of a desired output response while maintaining the constraint that the mean is fixed at a certain level. We demonstrate the process using an automotive device design example. The optimization procedure incorporates robustness into the design. We find a set of values for the design variables that ensure repeatability throughout the manufacturing process and service life.

II. A Taguchi Snap-Fit Example

We demonstrate the process using an automotive device design robustness example. The optimization procedure incorporates repeatability into the target design leading to efficiency in the manufacturing process and improved satisfaction for the duration of customer use. The type of snap-fit we present is used to connect an encapsulated motor stator and motor housing.[1]

Consider a snap-fit design[12] that is composed of both a moving part and a stationary part that are designed to the specification of three uncertain design variables: spring constant K (N/mm), interference I (mm), and ramp angle θ (degrees). A fourth noise variable is the friction coefficient μ. The interference

*North Carolina State University, fdreale@ncsu.edu
†Sandia National Laboratories, AIAA Senior Member, vjromer@sandia.gov
‡Sandia National Laboratories, lpswile@sandia.gov

I represents the vertical distance between the tip on the fixed part and the bottom edge of the tip on the moving part. In order to snap the parts together, the moving part must move past the tip of the fixed part. Whether or not the parts permanently snap together is largely based on the ramp angle. The setup is demonstrated in Figure 1:

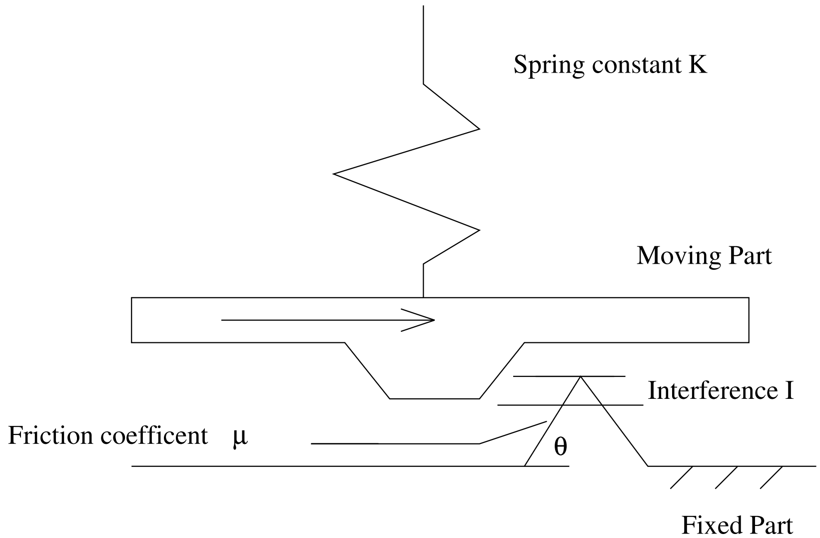


Figure 1. A snap-fit design example. Figure redrawn from Ref.[12]

All four variables help to determine the friction force Y (N) between the moving and stationary parts. This relationship is captured in (1),

$$Y = KI\frac{\mu + \tan\theta}{1 - \mu\tan\theta}. \tag{1}$$

The bounds on the nominal values of the design variables are provided below.

$$500 \text{ N/mm} \leq K \leq 600 \text{ N/mm}$$
$$0.1 \text{ mm} \leq I \leq 0.35 \text{ mm}$$
$$45 \text{ degrees} \leq \theta \leq 65 \text{ degrees}$$

We shall henceforth refer to the above intervals as the design ranges of each design variable. The average value of the friction coefficient μ is 0.17. As specified in,[12] all four variables are normally distributed with respective standard deviations $\sigma_K = 10$ N/mm, $\sigma_I = 0.017$ mm, $\sigma_\theta = 1$ degrees, and $\sigma_\mu = 0.017$. We assume that the above standard deviations hold for all nominal values of the four variables.

We aim to minimize variance and satisfy the constraint on the mean $\bar{Y} = 120$ N simultaneously.

III. Reduction to Two Design Variables

In this problem, there are three design variables K, I, and θ which specify how the moving and stationary parts are related. There are four variables K, I, θ, and μ which are uncertain or noise variables (tolerance variability) which lead to variability in the result Y.

Let us analyze the three design variables and their effect on Y. We vary the nominal value of each variable while the nominal value of each other design variable is fixed at the center of its design range. The value of μ is fixed at its mean. Figure 2 demonstrates that over the design range of K there is smaller change in Y than there is over the design ranges of I and θ. Furthermore, in Figure 3 we examine all four variables over a $\pm 2\sigma$ interval around the center of their design range or mean. In Figure 3, we also observe the friction force over its $\pm 2\sigma$ uncertainty.

We reduce to two design variables so that we can better demonstrate the following robust optimization algorithm visually. Based on the relative importance of the three design variables, we relegate K to be strictly a noise variable. Therefore, we fix the nominal value of the spring constant K and retain two design variables: the interference I and the ramp angle θ. Now there are two design variables I and θ for which

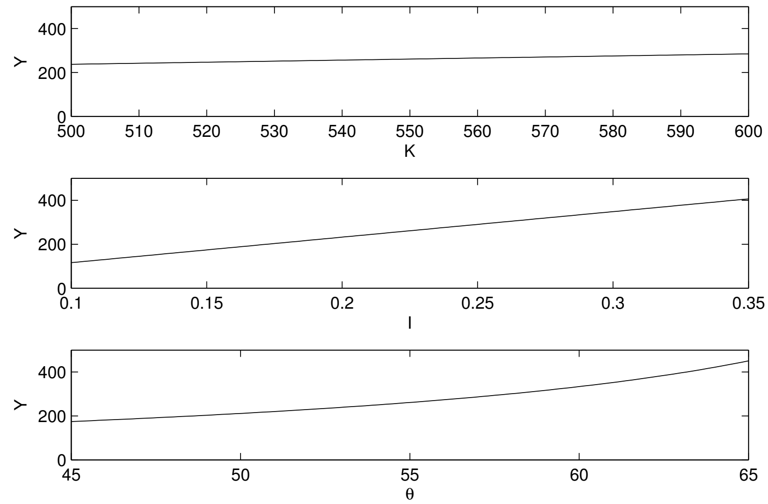


Figure 2. The friction force over the design range of each design variable.

we change the nominal values throughout the optimization procedure. There are still four uncertain (noise) variables K, I, θ, and μ. The nominal value of K if fixed at 550 N/mm and μ is fixed at 0.17 throughout the procedure.

IV. Analysis of the 2D Design Space

The problem statement has slightly changed. Now, we optimize over two variables I and θ. We still aim to minimize variance and satisfy the constraint on the mean $\bar{Y} = 120$ N. We study the behavior of the friction force, its mean, and its variance.

We allow the two design variables to vary over their design spaces while we fix the nominal values of the other variables at $K = 550$ N/mm and $\mu = 0.17$. We initially use a 25×25 linearly spaced grid of points in the two dimensional design space for visualization purposes.

We examine the design space for the contour $Y = 120$ N. Figure 4a. displays the contours of Y. We find that the friction force is mildly nonlinear. During the process, we depend on the $Y = 120$ N contour being "close" to the $\bar{Y} = 120$ N contour. Figure 4b. illustrates $\bar{Y} - Y$ where $\bar{Y}$ is found using 200 samples at each point. The absolute difference between Y and $\bar{Y}$ is less than one percent of 120 N throughout the design space.

We study the contours for both the mean and variance by creating a 25×25 grid of sampled points. Figure 5a. presents the mean contours. Figure 5b. shows an approximation to the standard deviation contours. At each point, the mean and standard deviation are found by sampling two hundred quadruplets over distributions of each variable. Based on the direction of decreasing standard deviation, we expect that the best point is at the lower, right end of the $Y = 120$ N contour.

By relegating K to be strictly a noise variable, we change the original problem. As we see from these plots, this is a mildly nonlinear that we can visually track as well. Some of the techniques used in the following iterative process can be applied to an array of problems, some cannot. We discuss the applicability of these methods to higher dimensions by applying them to a related example as well.

V. An iterative process: lowering variation while $Y = 120$ N

We employ an algorithm for finding an "optimal" point which has minimum variance while maintaining a friction force value of approximately 120 N. This process involves first finding the contour line $Y = 120$ N, then stepping along the contour to minimize variance.

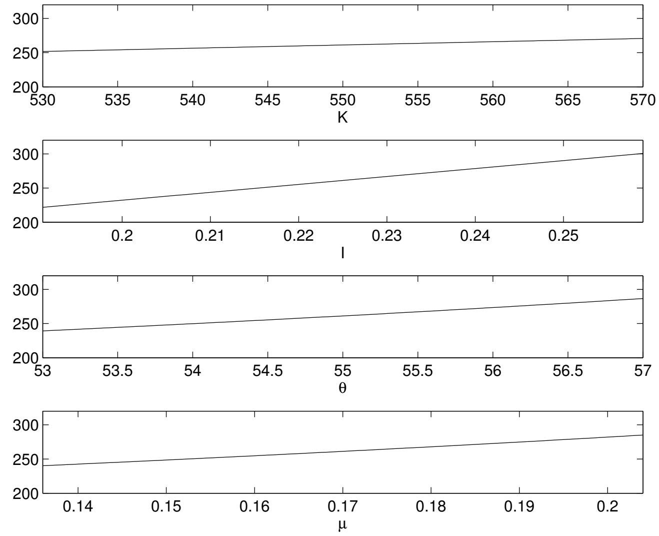


Figure 3. The friction force over a $\pm 2\sigma$ interval of each variable.

For a visual representation of the following heuristic process, see Figure 6 which provides the position of each iterate as well as the two tangent approximations to $Y = 120$ N.

A. Finding $Y = 120$ N

We maintain that two of the nominal values are fixed at $K = 550$ N/mm and $\mu = 0.17$. Consider a starting point at the center of the design space where $I = 0.225$ mm and $\theta = 55$ degrees.

1. An Analytic Line Search Coefficient Approach

We consider F a function of n variables where $F(\vec{x}) = F((x_1, x_2, ..., x_n))$. Suppose that we want to reach a target value of F which we call F_t.

The tangent hyperplane formula is,

$$(F(\vec{x}) - F(\vec{x}_0)) = \frac{\partial F}{\partial x_1}(x_1 - [x_1]_0) + \ldots + \frac{\partial F}{\partial \theta}(x_n - [x_n]_0). \tag{2}$$

We can write this as $\Delta F = \nabla F \cdot \Delta \vec{x}$.

Further, consider a single step of gradient descent $\vec{x}_1 = \vec{x}_0 - \lambda \nabla F$ where λ is the line search coefficient. We rewrite gradient descent in the form $\Delta \vec{x} = -\lambda \nabla F$.

Combining the tangent plane and gradient descent expressions yields

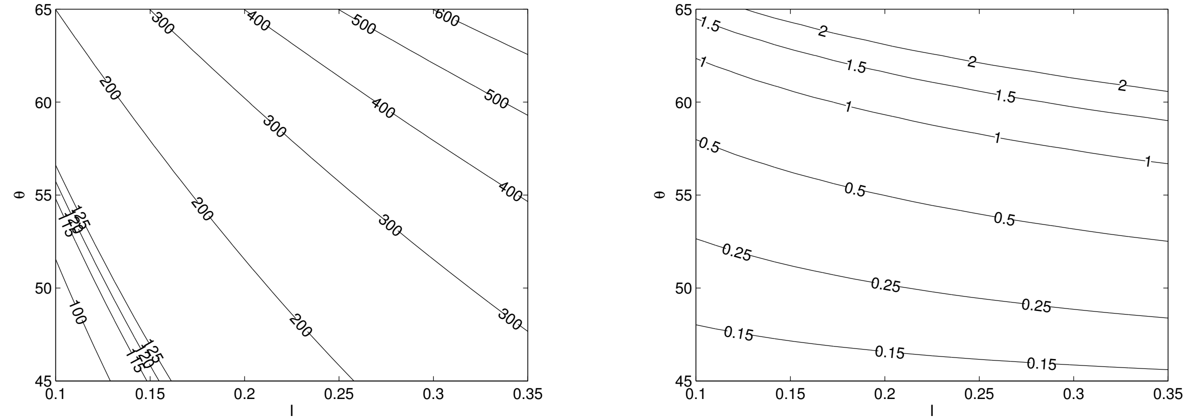
$$\Delta F = \nabla F \cdot \Delta \vec{x} = -\nabla F \cdot \lambda \nabla F = -\lambda \, ||\nabla F||_2^2 \, . \tag{3}$$

We solve for λ to obtain our analytic line search coefficient,

$$\lambda = -\frac{\Delta F}{||\nabla F||_2^2}. \tag{4}$$

This yields an iterative process. We perform gradient descent $\vec{x}_{n+1} = \vec{x}_n - \lambda_n \nabla F$ where we have analytic line search coefficient

$$\lambda_n = -\frac{F_t - F(\vec{x}_n)}{||\nabla F||_2^2} = \frac{F(\vec{x}_n) - F_t}{||\nabla F||_2^2}. \tag{5}$$



(a) Contour plot of the friction force Y.

(b) Contour plot of $\mid \bar{Y} - Y \mid$.

Figure 4. The friction force Y as well as the relationship between Y and the mean $\bar{Y}$.

Now, we apply the above process to our test problem. Although we can use analytic gradients for the test problem, to mimic a problem in which analytic derivatives are not available, we use forward difference perturbations of one percent of the design ranges. We retain this gradient for the remainder of the line search process. We want to find $Y = 120$ N or some reasonably close value to this target, e.g. within one percent. At each step, we perform gradient descent

$$\begin{bmatrix} I_{i+1} \\ T_{i+1} \end{bmatrix} = \begin{bmatrix} I_i \\ T_i \end{bmatrix} - \lambda_i \begin{bmatrix} \frac{\partial Y}{\partial I} \\ \frac{\partial Y}{\partial T} \end{bmatrix} \tag{6}$$

with our updated step coefficient,

$$\lambda_i = \frac{Y_i - 120}{\frac{\partial Y}{\partial I}^2 + \frac{\partial Y}{\partial T}^2} = \frac{Y_i - 120}{||\nabla Y||_2^2} \tag{7}$$

where $T = \frac{\pi}{180}\theta$ and $T_i = \frac{\pi}{180}\theta_i$ are in radians. After three line search steps, we obtain the point $I = 0.1208$ mm and $\theta = 51.5553$ degrees where $Y = 120.9170$ N. We next obtain a tangent line approximation to the $Y = 120$ N contour of Figure 4 a.,

$$L_1 = \{(I, \theta) | 1000.5857(I - 0.12085) + 284.6367(\theta \frac{\pi}{180} - 0.89981) = 0\}. \tag{8}$$

This line goes through the identified point where $Y = 120.9170$ N.

B. Moving in the Direction of Lower Variance

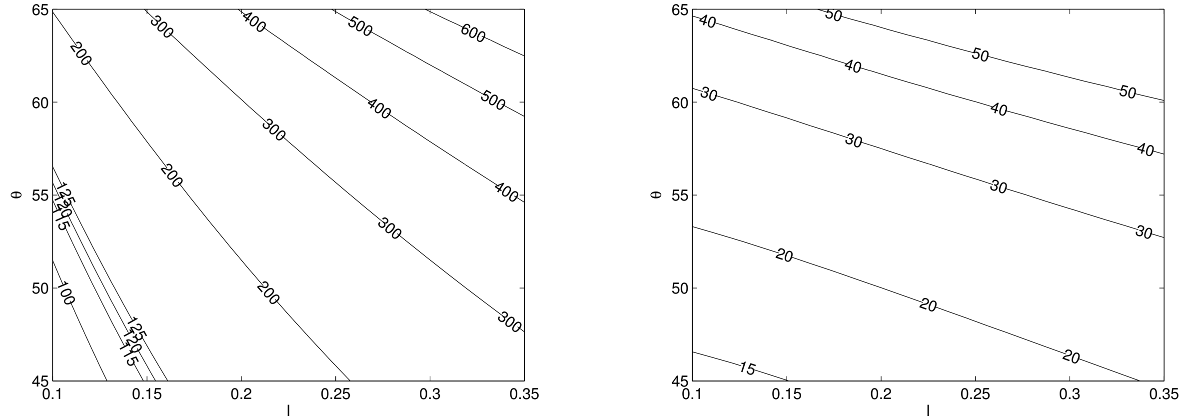
Now that we have found where $Y \approx 120$ N, we follow the tangent line approximation L_1 in the direction of lower variance. We test three methods to direct movement. We determine which direction to follow and how large of a step we should take. Referring to Figure 6, the stepsize is chosen to be relatively large, one-fifth of the length of L_1.

1. Linearized response function for calculating variance

The First-Order, Second Moment (FOSM) variance formula[5] is presented in 9,

$$\sigma_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \sigma_{i,j} \tag{9}$$

where $F = f(\vec{x}) = f((x_1, x_2, \ldots, x_n))$ and $\sigma_{i,j}$ is the covariance of x_i and x_j.



(a) A mean approximation using 200 random samples.

(b) A standard deviation approximation using 200 random samples.

Figure 5. The contours of the two-hundred sample mean and standard deviation.

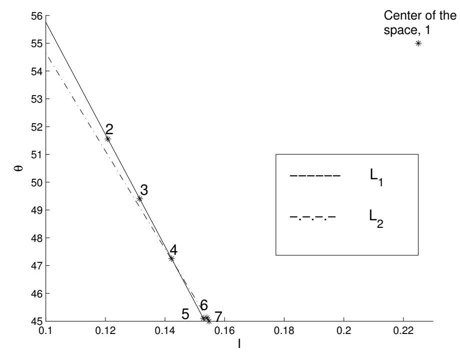


Figure 6. The iteration history and the approximations to $Y = 120$ N.

For the test problem, the variance is estimated by,

$$\sigma^2_{\text{FOSM}} = \frac{\partial Y}{\partial K}^2 \sigma^2_K + \frac{\partial Y}{\partial I}^2 \sigma^2_I + \frac{\partial Y}{\partial \theta}^2 \sigma^2_\theta + \frac{\partial Y}{\partial \mu}^2 \sigma^2_\mu. \tag{10}$$

Although we can determine the expressions for the analytic partial derivatives in this problem, we approximate the partial derivatives using forward difference calculations with one percent of the design range perturbations in I and θ. Correspondingly, we perturb K by 1 N/mm and μ by 0.0035. Each calculation of the variance requires five function evaluations. This is a large savings over the 200 sample calculations performed previously. We compute the variance at the second point using this method. We find a lower variance when we examine the third point on L_1 which is down and to the right as shown in Figure 6.

2. Two Sample Ordinal Variance Estimate

Figure 7 shows a crude approximation to the standard deviation contours using two samples at each point. The first sample is the quadruplet containing the nominal values at each point (i.e. Y at $[\bar{K}, I_i, \theta_i, \bar{\mu}]$ where the subscript i signifies the design space coordinates of point i in Figure 6) and the second sample is the quadruplet where we perturb each variable by one standard deviation simultaneously (i.e. Y at $[\bar{K} + \sigma_K, I_i + \sigma_I, \theta_i + \sigma_\theta, \bar{\mu} + \sigma_\mu]$). Each variance calculation requires only two function evaluations.

We compare the two sample ordinal standard deviation to the two hundred sample standard deviation. To make Figure 7, we perform $25 \times 25 \times 2 = 1,250$ function evaluations. We compare this number to the $25 \times 25 \times 200 = 125,000$ function evaluations used to create Figure 5b. Either set of contours correctly indicates that moving downward and to the right on L_1 reduces variance. Although two sample ordinal is not able to predict the same absolute behavior of the contours, it is able to predict the relative behavior which is sufficient for the comparisons between successive points.

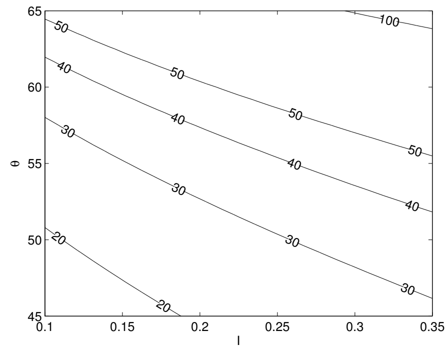


Figure 7. A standard deviation approximation using two previously chosen samples.

3. Sampling based on a Specified Confidence Level

The choice of the stepsize we take is based on a combination of hypothesis testing and seeking a low sample size. We try two different stepsizes and test at two different confidence levels. We attempt to determine the combination of stepsize and confidence level that allows us to have some evidence that we are in fact decreasing variance and also allows us to maintain a low cost.

We want to move along L_1. We do not use any more function evaluations to do this. We just need to find the end points of the line (I_s, θ_s) and (I_f, θ_f). As noted in section two, we know the bounds on both I and θ. In our expression for L_1, the initial point I_0 and θ_0 is known. We check our first bound where $I_s = 0.1$ mm and obtain $\theta_s = 55.7540$ degrees. Similarly, we check our second bound where $\theta_f = 45$ degrees to determine $I_f = 0.1534$ mm. We want to move some fraction s of the length of L_1. So the next point on along L_1 is $I = I_0 + s(I_f - I_s)$ and $\theta = \theta_0 + s(\theta_f - \theta_s)$ where $0 < s \leq 1$.

We begin with a one-tenth of the line step. Our first advance down the line takes us to a third point $I = 0.1262$ mm and $\theta = 50.4799$ degrees where $Y = 120.8303$ N. Using Latin hypercube sampling (LHS) that is implemented in DAKOTA (Sandia-developed software),[3] we perform an F-test. The null hypothesis is that the value of the variance at the third point is the same as the value of the variance at the second point. The alternative hypothesis is that the variance values at the two points are not equal. We require 41 samples to verify that the two variance values are different with 60 percent confidence. We require 995 to deem that they are different with 90 percent confidence.

Now, starting from the second point, we advance with a step of one-fifth of the line. Our first advance down the line takes us to another possible third point $I = 0.1315$ mm and $\theta = 49.4045$ degrees where $Y = 120.6423$ N. We require 12 samples to determine that the two variance values are different with 60 percent confidence. We require 256 to confirm that they are different with 90 percent confidence.

The cost and benefit considerations drive toward larger steps at the 60 percent confidence level. Therefore, the following steps are one-fifth of the line length apart and tested at the 60 percent confidence level. Consequently, the third point is $I = 0.1315$ mm and $\theta = 49.4045$ degrees as discussed above.

We take another one-fifth of the line step to the fourth point $I = 0.1422$ mm and $\theta = 47.2537$ degrees where $Y = 119.9849$ N. Observe the proximity of Y to 120 N. We do another F-test using LHS in DAKOTA. We find that 14 samples are required to confirm that the variance values at the third and fourth points are not equal with 60 percent confidence.

After another one-fifth of the line step, we obtain a fifth point $I = 0.1529$ mm and $\theta = 45.1029$ degrees. It takes 14 samples to confirm that the variance at the fourth point is not equal to the variance at the fifth point at the 60 percent confidence level. The value of $Y = 118.9819$ N at the fifth point.

We can take one more truncated step down the line to the end point on the lower 45 degree edge of the design space where $I = 0.1534$ mm. We call this point 6a. The value of the friction force $Y = 118.9258$ N. The friction force is still within one percent of 120 N and we cannot move in any direction on the line to further decrease the variance.

4. *Summary of Procedure that satisfies the original 1 % Criterion*

Under the one percent criterion we set for moving down the line, the last step from the fifth point to point 6a completes our process. We have made it to the final point following one single line L_1 without an additional side step. A summary of our results is included in Table 1. The table lists the point, its Y value, and its standard deviation calculated by FOSM, two-sample ordinal sampling, and sampling with the sample size chosen based on sixty percent confidence in the F-test.

Table 1. Iteration history of the heuristic with the original one percent criterion.

	I (mm)	θ (degrees)	Y (N)	FOSM	Ordinal	LHS
1	0.225	55	261.1819	25.9131	36.6749	
2	0.1208	51.5553	120.9170	18.4662(5)	22.3074(2)	20.8599(12)
3	0.1315	49.4045	120.6423	17.0504(5)	20.7772(2)	19.2423(12)
						18.0946(14)
4	0.1422	47.2537	119.9849	15.8064(5)	19.4389(2)	16.8051(14)
5	0.1529	45.1029	118.9819	14.7037(5)	18.2567(2)	15.6598(14)
6a	0.1534	45	118.9258	14.6540(5)	18.2029(2)	

C. Finding a new approximation to $Y = 120$ N

So that we can demonstrate a side step. We set a new side step criterion of 1 N. We do not see holding to this tight of a criterion, but for the purpose of illustration we return to the fifth point where Y is more than 1 N away from 120 N. Therefore, we find another approximation to $Y = 120$ N. We do this in one step of gradient descent with the analytic line search coefficient described in (7). This gives us a new point $I = 0.1541$ mm and $\theta = 45.1248$ degrees. At this sixth point, the value of $Y = 120.0013$ N. We obtain an updated tangent line approximation to $Y = 120$ N,

$$L_2 = \{(I, \theta) | 778.8915(I - 0.15407) + 253.4266(\theta \frac{\pi}{180} - 0.78758) = 0\}. \tag{11}$$

D. Proximity to $\bar{Y} = 120$ N

We continue in the direction of decreasing variance. We cannot complete a one-fifth of the line step without leaving the design space. Therefore we complete a truncated step to the lower edge of the design space. The seventh point we obtain is $I = 0.1548$ mm and $\theta = 45$ degrees where $Y = 119.9977$ N. Now we have found a point where the friction force is within 1 N of 120 N and we cannot move along the tangent line to further decrease variance. So, we are done with this part of the process.

Now, we must determine how close this point on the approximation to the $Y = 120$ N contour is to the $\bar{Y} = 120$ N contour. We want the true mean to be within two-fifths of the standard deviation of the target mean 120 N. We use two types of methods to study the mean at the final point. We begin with sampling for a specific confidence interval width. We also discuss point estimate methods.

1. *Sampled value of Mean used for Comparison*

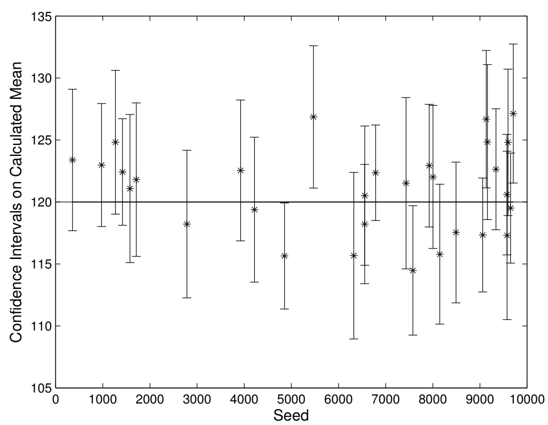
To compare the accuracy of the methods we use to study the mean, we determine a reference mean by taking 100,000 LHS samples. We obtain a sample mean value of $\bar{Y} = 120.1941$ N and a sample standard deviation of $s = 14.7477$ N. From these results, we obtain the 95 percent confidence interval on the mean $(120.1027, 120.2855)$ N. Now we know that the true mean is within two-fifths of the sample standard deviation ($\frac{2}{5}s = 5.8991$ N) away from our target mean with 95 percent confidence.

2. Sampling for specific confidence interval within $\frac{2}{5}\boldsymbol{\sigma}$ of $\bar{\mathbf{Y}}_{target}$

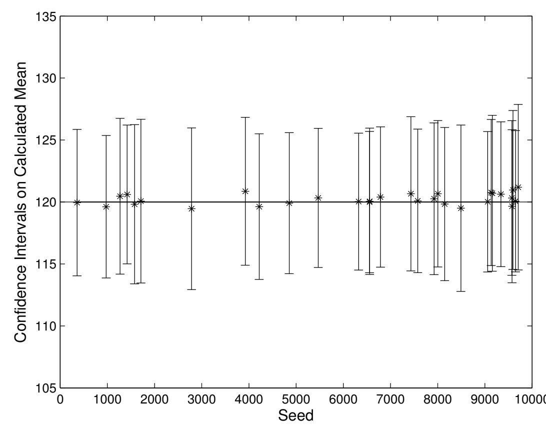
We use a confidence interval on the mean to determine if the seventh point is near the $\bar{Y} = 120$ N contour. We attempt to fix the width of the confidence interval. We determine the number of samples required to obtain a 90 percent two-sided confidence interval on the mean with half-width $\frac{2}{5}s$ where s is the sample standard deviation.

For the purpose of finding the required sample size n, we assume a standard normal distribution. Thus, we find n such that $2/5 \approx \frac{t(1-\alpha/2;n-1)}{\sqrt{n}}$ where $t(1-\alpha/2;n-1)$ refers to Student's-t distribution. Using $\alpha = .1$, we find that we require $n = 19$ for the half-width to be approximately $\frac{2}{5}s$. Notice that the sample size obtained is not dependent on the test problem nor the number of variables. If we wanted the half-width to be γs, we find n such that $\gamma \approx \frac{t(1-\alpha/2;n-1)}{\sqrt{n}}$.

We analyze the confidence intervals on the mean that are produced when $n = 19$ using simple random sampling (SRS) and LHS both implemented in DAKOTA. Among the sixty different samplings, we obtain five intervals which do not include 120 N. Table 2 provides the results from this sampling activity. Figure 8 provides a visual representation of the information from the table. The seed values are found using the MATLAB implementation of the Mersenne Twister pseudorandom number generator. Initially, the function is reset using the seed value 5489. We take the numbers produced by the routine and multiply them by 10,000. We obtain the current seed by taking the ceiling function of that product.



(a) Two-sided 90 percent confidence intervals using SRS.



(b) Two-sided 90 percent confidence intervals using LHS.

Figure 8. Examination of the two-sided 90 percent confidence intervals of the mean.

The mean estimates and associated confidence intervals determined by Monte Carlo sampling depend to some degree on the initial seed used in the sampling process. The above results demonstrate that SRS is more dependent on the choice of initial seed than LHS. Often, LHS is called a "low variance" estimator when compared to SRS.

We use 90 percent confidence intervals. Therefore, we expect that after many trials are done, the true mean will be outside of the confidence interval 10 percent of the time. Among our thirty SRS trials, the high variability of the SRS estimates leads us to find that more than 10 percent of the intervals did not contain the true mean. We did not perform a larger number of trials when we found the true mean is inside of the thirty SRS intervals only 83 percent of the time as opposed to 90 percent of the time. However, previous experience with a larger number of trials[9] indicates that you cannot expect $100(1-\alpha)$ percent SRS confidence intervals to contain the true mean the advertised $100(1-\alpha)$ percent of the time. This is an effect of using sample mean and standard deviation estimates with high variability to compute the confidence interval bounds. Nonetheless, the SRS confidence interval formula can give light to the number of samples needed to obtain a CI of a certain size as shown above.

We strongly recommend LHS over SRS. Notice from Figure 8 that the LHS estimates are much more conservative than the SRS estimates. In fact, it may be conservative to use the classical confidence interval $\bar{x} \pm \frac{t(1-\alpha/2;n-1)}{\sqrt{n}}s$ formula to determine the $100(1-\alpha)$ percent confidence intervals for LHS. Then the LHS

Table 2. SRS and LHS confidence intervals.

Seed	Sampling Technique	Mean $\bar{Y}$ (N)	s (N)	Confidence Interval (N)
8148	SRS	115.785	14.1854	(110.1416,121.4284)
	LHS	119.833	15.5234	(113.6573,126.0087)
9058	SRS	117.335	11.5491	(112.7404,121.9296)
	LHS	120.018	14.2497	(114.349,125.687)
1270	SRS	124.82	14.5887	(119.0162,130.6238)
	LHS	120.467	15.7959	(114.1829,126.7511)
9134	SRS	126.6791	13.9368	(121.1346,132.2236)
	LHS	120.76	14.8167	(114.8655,126.6545)
6324	SRS	115.667	16.8978	(108.9446,122.3894)
	LHS	120.032	13.881	(114.5097,125.5543)
976	SRS	122.976	12.4719	(118.0143,127.9377)
	LHS	119.615	14.4523	(113.8654,125.3646)
2785	SRS	118.216	14.9651	(112.2624,124.1696)
	LHS	119.454	16.4037	(112.9281,125.9799)
5469	SRS	126.864	14.4271	(121.1245,132.6035)
	LHS	120.329	14.1103	(114.7155,125.9425)
9576	SRS	117.298	17.0785	(110.5037,124.0923)
	LHS	119.655	15.5131	(113.4834,125.8266)
9649	SRS	119.506	11.1643	(115.0645,123.9475)
	LHS	120.059	14.3264	(114.3595,125.7585)
1577	SRS	121.082	15.0217	(115.1059,127.0581)
	LHS	119.819	16.144	(113.3964,126.2416)
9706	SRS	127.133	14.102	(121.5228,132.7432)
	LHS	121.192	16.7876	(114.5134,127.8706)
9572	SRS	120.597	12.2284	(115.7322,125.4618)
	LHS	120.321	15.6862	(114.0806,126.5614)
4854	SRS	115.641	10.7563	(111.3618,119.9202)
	LHS	119.907	14.3039	(114.2165,125.5975)
8003	SRS	122.019	14.5122	(116.2456,127.7924)
	LHS	120.661	14.8475	(114.7542,126.5678)
1419	SRS	122.417	10.8078	(118.1173,126.7167)
	LHS	120.605	14.0793	(115.0038,126.2062)
4218	SRS	119.381	14.6934	(113.5355,125.2265)
	LHS	119.624	14.7675	(113.749,125.499)
9158	SRS	124.835	15.7271	(118.5783,131.0917)
	LHS	120.703	15.8219	(114.4086,126.9974)
7923	SRS	122.932	12.4356	(117.9847,127.8793)
	LHS	120.262	15.3917	(114.1387,126.3853)
9595	SRS	124.81	14.8572	(118.8994,130.7206)
	LHS	120.97	16.1151	(114.5589,127.3811)
6558	SRS	120.511	14.1244	(114.8919,126.1301)
	LHS	120.055	14.8441	(114.1496,125.9604)
358	SRS	123.386	14.3511	(117.6767,129.0953)
	LHS	119.947	14.8508	(114.0389,125.8551)
8492	SRS	117.539	14.2748	(111.8601,123.2179)
	LHS	119.496	16.882	(112.7798,126.2122)
9340	SRS	122.638	12.2653	(117.7585,127.5175)
	LHS	120.627	14.6846	(114.785,126.469)
6788	SRS	122.358	9.6727	(118.5099,126.2061)
	LHS	120.4	14.2331	(114.7376,126.0624)
7578	SRS	114.476	13.1278	(109.2534,119.6986)
	LHS	120.09	14.5632	(114.2963,125.8837)
7432	SRS	121.515	17.3768	(114.602,128.428)
	LHS	120.661	15.6441	(114.4373,126.8847)
3923	SRS	122.546	14.2893	(116.8613,128.2307)
	LHS	120.862	14.9929	(114.8974,126.8266)
6555	SRS	118.218	12.1118	(113.3996,123.0364)
	LHS	119.994	14.3271	(114.2943,125.6937)
1712	SRS	121.802	15.5647	(115.6099,127.9941)
	LHS	120.068	16.6043	(113.4623,126.6737)

success rate of getting a true mean inside a confidence interval is greater than the advertised $100(1-\alpha)$ percent. Nevertheless, there is still a chance that a given LHS estimate of the mean and standard deviation could result in a confidence interval with a previously chosen magnitude that incorrectly does not contain the true mean when indeed the true mean is within that magnitude of the mean estimate. We call this a "false negative." For our test problem, this indicates that the current mean estimate for a given design point does not fall within $\frac{2}{5}\sigma$ of the target mean even though the true mean of the design point does fall within $\frac{2}{5}\sigma$ of the target mean. This prompts one to continue the optimization until a design point is located that has the minimum variance for a point with a mean within a fixed distance from the target mean.

Although the above circumstance gives a false indication that the optimization must be continued, when solving the problem, we cannot confirm that we have a "false negative." Thus we proceed as if the current point does not have a true mean within the chosen distance of the target mean. It is unwise to proceed with the same approach used for a "true negative" where the current point's confidence interval non-falsely suggests that the mean at the current point is not within the chosen distance from the target mean. At this point, we can no longer follow the $Y = \bar{Y}_{\text{target}}$ path. After all, that is what got us to this point in the first place. We switch to an objective function in the mean $\bar{Y}$. We perform gradient descent with an analytic line search coefficient. To compute the gradient approximation, we apply "spatially-correlated sampling" meaning we use the same seed to determine the mean at the design point and at the perturbed points. The use of a common seed at the perturbed design points is important because, as Figure 8 shows, different seeds have a random biasing effect on calculated means at a fixed point in the design space. This random biasing with different seeds also persists for closely perturbed design points in the space. Hence, random seed-induced errors in the calculated partial derivatives will exist unless the seed is held constant for the calculation of the gradient entries.

In the case of a "false negative," we must take special care with respect to the spatially-correlated sampling. If we reuse the seed that produced the false negative to compute the mean estimates at the perturbed points, then the mean estimates we obtain will be similarly biased. Furthermore, this extra work and cost should have never been expended in the first place because we got a false indication that this point did not meet the set criterion.

Therefore, we want to avoid false negatives. We always recalculate the mean using another seed at the terminal point of the optimization procedure with the objective function in Y to improve method efficiency and reliability. Given the high probability that the LHS confidence interval calculated with the classical SRS confidence interval formula will contain the true mean, the possibility of getting two false negatives in a row is unlikely. Therefore, if the second result is negative as well, there is evidence that the result is a true negative. Now, the optimization procedure using the objective function in $\bar{Y}$ can be pursued using spatially-correlated sampling with the first or second seed. If the second result is a positive indication that the target mean is contained within the LHS confidence interval, then the third calculation is required to resolve the contradiction. This third result is used as the true indication since the possibility of obtaining two false positives in a row is also unlikely.

Similarly, we avoid false positives by performing a second evaluation. A false positive would prematurely end the optimization procedure at the end of the first optimization phase in Y even though the second phase in $\bar{Y}$ should be entered to get to a design point that meets the objective of the target $\bar{Y}$ within the $\frac{2}{5}\sigma$ tolerance. A second positive essentially confirms that a true positive exists at the design point. A contradictory negative result following the initial positive one requires a third and deciding evaluation.

We next demonstrate the process change to an optimization in $\bar{Y}$ when an affirmed true negative arises at the terminal point in the optimization in Y. Now we believe that the target mean is not within $\frac{2}{5}\sigma$ of the value of the friction force at the final point. For the purpose of illustrating technique, we suppose that our target mean is 135 N. Notice that 135 N is not inside any one of the 90 percent confidence intervals. The next seed returned is 7061. The sample mean is 120.061 N and the standard deviation is 13.5418 N. This evaluation essentially confirms that a true negative exists at this point. Using this new seed, we perform spatially-correlated sampling at one standard deviation perturbations in the design variables I and θ. We use LHS implemented in DAKOTA with $n = 19$ samples to obtain the sample mean at the three points. This yields a gradient in $\bar{Y}$. We perform gradient descent with an analytic line coefficient to find the $\bar{Y} = 135$ N contour. After only one step of gradient descent with an analytic line search coefficient, we obtain the point $I = 0.1721$ mm and $\theta = 45.3282$ degrees where the sample mean is 135.1702 N and the sample standard deviation is 14.0585 N. Now, we minimize variance by following the tangent line approximation to the mean $\bar{Y} = 135$ N in the direction of decreasing variance. We find the point on the lower 45 degree edge of the design space where the interference is 0.1742 mm. At this point the sample mean is 135.1718 N and the sample standard deviation is 13.9177 N. We satisfy the criterion that the mean is within two-fifths of the standard deviation away from our target mean 135 N and we cannot move in a direction of lower variance on our tangent line thus we have completed the process.

3. Statistical-Moment Generation Approach (Response Mean by Optimal Placement and Weighting)

We use point estimate methods (PEMs) to determine the mean (first moment) by optimal placement and weighting of samples. These methods are also referred to as statistical-moment generating methods. Table 3 provides an overview of the results. A visual representation of this table is provided in Figure 9. Observe the proximity of each of these results to the mean obtained using 100,000 samples. We observe that these methods provide fairly high accuracy at low cost. Another benefit of using PEMs is that the same samples and weights used to compute the mean can also be used to compute the variance and standard deviation so that no further function evaluations are necessary.

Table 3. Increasing accuracy of mean (first moment) estimates.

Method	Number of Function Evaluations	Mean $\bar{Y}$ N
Mean value	1	119.9977
Hong's "$2n$"	8	120.1934
Hong's "$2n+1$"	9	120.1933
Two-point Rosenblueth"2^n"	16	120.1935
Three-point Seo-Kwak "3^n"	81	120.1938
LHS with 100,000 samples	100,000	120.1941

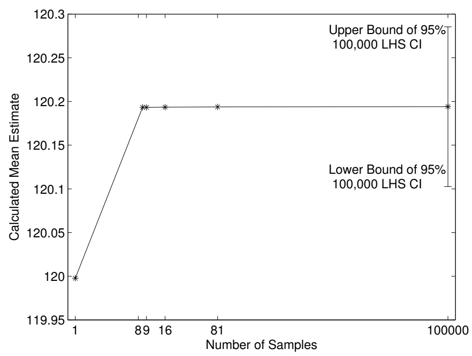


Figure 9. Increasing accuracy of mean (first moment) estimates.

Using the final point as our only sample, we find the mean value estimate[2] of 119.9977 N.

Now, we use a $2n$ PEM[7] where n is the number of variables. For our example, $n = 4$. This case corresponds to $m = 2$. Let $k = 1, \ldots, n$ and $i = 1, \ldots, m$. We sample the points

$$x_{k,i} = \bar{x}_k + \xi_{k,i} s_k \tag{12}$$

where $\bar{x}_k$ is the nominal value of the k^{th} variable and s_k is the standard deviation of the k^{th} variable. Under the assumption of normality, we have

$$\xi_{k,i} = (-1)^{3-i}\sqrt{n} \tag{13}$$

and equal weighting is used so that the weights of each point are defined by $p_{k,i} = \frac{1}{2n}$. This method yields the mean 120.1934 N.

A $2n+1$ method[7] uses the center point as well. This is the $m = 3$ case. To clarify the formula, the kurtosis of a normal distribution is considered to be three. We use the same point-generation formula (12) as before, but now, under the assumption of normality, we have

$$\begin{cases} \xi_{k,i} = (-1)^{3-i}\sqrt{3},\ i = 1, 2 \\ \xi_{k,3} = 0. \end{cases} \tag{14}$$

The weights also change. Now, $p_{k,1} = p_{k,2} = \frac{1}{6}$ and $p_{k,3} = \frac{1}{n} - p_{k,1} - p_{k,2} = \frac{1}{n} - \frac{1}{3}$. When $n = 4$, $p_{k,3} = -\frac{1}{12}$. We see that $i = 3$ four times thus the same point at the center of the space is reached four times. Since we evaluate the value of Y at this point only once, the middle point only counts once and ends up getting a weight of $4 \times -\frac{1}{12} = -\frac{1}{3}$. From this exercise, we obtain the mean estimate 120.1933 N.

The previous two methods required sampling at the center of each face of the hypercube or at the center of the hypercube where each variable is fixed at its nominal value. The next two methods involve sampling at the corners of the hypercube. The final method uses samples at the corners, the center of each face, the center of each edge, and the center of the hypercube.

We try a 2^n[10] method. Let $i = 1, 2$ and $k = 1, \ldots, n$. We use (12) just as before, but now

$$\xi_{k,i} = (-1)^{3-i}. \tag{15}$$

We have equal weighting so $p_{k,i} = \frac{1}{2^n}$. We find the mean estimate 120.1935 N.

The final estimation method we try is a three-point Seo-Kwak.[11] Now $i = 1, 2, 3$ and $k = 1, \ldots, n$. We use (12) with

$$\begin{cases} \xi_{k,i} = (-1)^{3-i}\sqrt{3},\ i = 1, 2 \\ \xi_{k,3} = 0. \end{cases} \tag{16}$$

The weighting in this case is a bit trickier. The Gauss-Hermite three-point weights are used. Therefore $p_{k,1} = p_{k,2} = \frac{1}{6}$ and $p_{k,3} = \frac{2}{3}$. A product weighting system is used. That means that we take our weight w at a point to be

$$w = \prod_{k=1}^{n} p(k, i_k) \tag{17}$$

where i_k indicates the value of i for the k^{th} variable at the point. The first-moment result from this method is 120.1938 N.

4. Assessment of various other approaches for calculating mean

No other methods that the authors are aware of for calculating mean are cost/performance competitive, including AMV, AMV+/FORM/SORM, and response-surface meta-modeling approaches.

E. Iteration History

Table 4 lists the iteration history starting from the middle of the design space. The variance and standard deviation provided are found using FOSM. The partial derivatives are computed by a forward difference approximation. For the design variables I and θ, the perturbations are one percent of their design ranges. For K, the perturbation is 1 N/mm. For the friction coefficient μ, the perturbation is 0.0035. For a visual history of the algorithm, refer back to Figure 6.

F. Cost

We first determine the cost of moving along the $Y = 120$ N contours. We then discuss the variance comparison cost along the way. We end with the cost of the studies on the mean. Observe that this is the cost associated with the 1.2 N side step criterion.

We begin with gradient descent with an analytic line search to get our first point on $Y = 120$ N. This costs six function evaluations. Determining L_1 costs us three more. We follow the line down to the fifth point while we check to make sure we are within some tolerance of 120 N. This costs four function evaluations. Altogether we spend 13 function evaluations to just move down L_1 .

Now, we consider the cost of the variance comparisons. We have already computed the function value at all five points on the tangent line so we use the previously computed friction force values in the variance estimates. Using linear approximations costs us $5 \times 4 = 20$ more function evaluations. Using the ordinal

Table 4. Iteration history the heuristic.

Iteration	I (mm)	θ (degrees)	Y (N)	s^2 (N^2)	s (N)
1	0.225	55	261.1819	671.4899	25.9131
2	0.1208	51.5553	120.9170	341.0013	18.4662
3	0.1315	49.4045	120.6423	290.7151	17.0504
4	0.1422	47.2537	119.9849	249.8437	15.8064
5	0.1529	45.1029	118.9819	216.1979	14.7037
6a	0.1534	45	118.9258	214.7400	14.6540
6	0.1541	45.1248	120.0013	217.2102	14.7380
7	0.1548	45	119.9977	215.4828	14.6793

sampling requires $5 \times 1 = 5$ additional function evaluations. Alternatively, we do the F-tests to compare the points that are one-fifth of the line apart using 66 function evaluations. The values used in the F-tests were generated by LHS which is not incremental. If we had used SRS, we could take advantage of incremental sampling and the cost would be 54. However, we do not know if the number of samples required to deem that the variances differed would be the same for SRS. Although not specifically part of the process, the cost of determining that we need to travel one-fifth of the line is 2,608 function evaluations.

Finally, we tally the number of function evaluations needed to perform the analysis of the mean at the final point. We attempt to fix the confidence interval magnitude. To do this and do a check for false positives/negatives, we require $19 \times 2 = 38$ function evaluations. However, one confidence interval computed with 19 LHS results may be sufficient as it was in the example we provided. We save significantly by using the PEMs. Since we have already computed the friction force Y at the final point, mean value is free. For the same reason, using $2n+1$ requires only 8 more function evaluations. Since $2n$ and $2n+1$ are not sampled at the same points, we use 8 more to get the $2n$ result. Sixteen samples are averaged to obtain the two-point Rosenblueth result. Since the 3^n uses the center sample and the same samples at the center of each face used for $2n+1$, we use only 72 more function evaluations to get the Seo-Kwak estimate.

VI. Comparison of results to other methods

We compare the results of a few other methods to the result of the algorithm when a 1.2 N side step criterion is used. We use 1,000 LHS samples to determine the mean and standard deviation at the returned point for each method. The sample mean at the final point is 119.1069 N and the sample standard deviation is 14.7263 N.

Using DAKOTA we perform an exhaustive search to find the point of best variance where the mean value of the friction force is approximately 120 N. We implement the nested optimization routines in DAKOTA. The outer loop is an optimization algorithm while the mean and variance used in the constraints and objective function are determined in the inner loop.

The outer loop is a Coliny division of rectangles (DIRECT) optimization routine. The global search balancing parameter is set to zero while the local search balancing parameter is set to 1×10^{-8}. We test DIRECT using two inner loops.

First, the mean and variance are determined by 10 LHS samples in the inner loop. The objective of DIRECT is to minimize variance subject to the constraint that the mean be within 1 N of the target mean. We set a solution threshold of 15. This nested routine performs 1,250 function evaluation and returns the point $I = 0.1519$ mm and $\theta = 45.3705$ degrees.

Next, we use the DAKOTA reliability package[2] to further confirm our results. We aim to minimize $E[| Y - 120 |] + C\sigma$ with a solution threshold of 55. The mean and standard deviation are computed using mean value with a gradient stepsize of 1×10^{-4}. We set the convergence tolerance to 1×10^{-4} and the threshold delta to 1×10^{-8}. This optimization is sensitive to the formulation of the objective function. The C reflects the $C\sigma$ reliability.[4] When we test $C = 3$, we find the point $I = 0.1365$ mm and $\theta = 48.3338$ degrees using 105 function evaluations. Here the mean is 120.3002 N and the standard deviation is 16.5004 N.

We test another optimization routine that is implemented by the Coliny pattern search in DAKOTA.

The objective is to minimize variance subject to the constraint that the mean be within 1 N of the target mean 120 N. We set a solution threshold of 15. The solution accuracy is chosen to be 1×10^{-8}. We pick an initial delta of .5. The threshold delta is set to 1×10^{-13}. We choose a contraction factor of .85. Using 1,360 function evaluations, the routine yields the point $I = 0.1535$ mm and $\theta = 45.4438$ degrees where the mean is 121.1675 N and the standard deviation is 14.9873 N.

VII. Cost Comparison

The cost is the number of function evaluations required. In other words, we compare how many times we need to compute the friction force Y for each method. It is clear that the cost of our algorithm is dependent on which methods are used for each task and the problem. We report the cost of performing the algorithm when the original 1.2 N side step criterion is used.

We see that all of the methods yield similar results. The main reason to use the algorithm we have provided is the ultimate cost savings.

The cost of the algorithm is dependent on which methods are used. The entire algorithm can be implemented in 26 function evaluations using two sample ordinal variance comparisons and Hong's "$2n$" estimate of the mean and standard deviation at the final point. The maximum cost of the algorithm is 151 function evaluations when F-test variance comparisons are used and the final mean and standard deviation estimate are determined by "3^n" Seo-Kwak. For the types of applications we consider, the three-point Seo-Kwak is not cost-performance competitive to the two-point Rosenblueth. We suggest using two sample ordinal variance comparisons and the "2^n" two-point Rosenblueth method to estimate the mean and standard deviation at the final point. The number of samples required for the PEMs is dependent on the number of variables n. If $n \geq 5$, we suggest using the fixed width CI. A higher dimension case is discussed in detail later.

Overall, we see that there are considerable savings over the stand alone methods discussed in the previous two sections. Table 5 displays the costs associated with the various methods.

Table 5. Costs of the optimization.

Method	Task	Options	Cost
Heuristic			
	Find $Y = Y_{\text{target}}$		13
	Lower Variance		
		FOSM	20
		Ordinal	5
		Hypothesis Testing	66
	Determine Mean		
		Fixed-width CI	19
		Mean value	0
		Hong's "2n"	8
		Hong's "2n+1"	8
		Two-point Rosenblueth "2^n"	16
		Three-point Seo-Kwak "3^n"	72
DIRECT		using LHS	1,250
		with reliability	113
Pattern search			1,360

VIII. Three dimensional Design Space Optimization

Now that we understand the algorithm, we return to the original problem. The above optimization process reflects a two design variable scenario. We reincorporate K as a design variable. Now there are three design variables, K, I, and θ, and four variables that lead to variability in the friction force, K, I, θ, and μ.

We employ the algorithm to find an "optimal" point which has minimum variance subject to the constraint that the friction force mean $\bar{Y} = 120$ N. Figure 10 presents the iteration history of the three-dimensional design space problem.

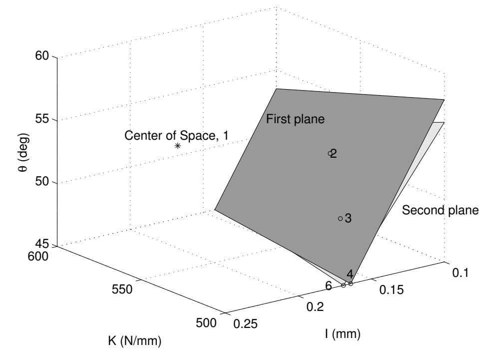


Figure 10. The iteration history and the approximations to $Y = 120$ N.

IX. Finding $Y = 120$ N

We maintain that the nominal value of the friction coefficient is fixed at $\mu = 0.17$. We consider a starting point at the center of the design space where $K = 550$ N/mm, $I = 0.225$ mm, and $\theta = 55$ degrees.

After seven line search steps, we find a point $I = 0.1207$ mm and $\theta = 51.5493$ degrees where the friction force is 120.7044 N which is within one percent of 120 N. We find the first tangent plane approximation,

$$P_1 = \{(K, I, \theta)|0.2195(K - 550) + 1000.3361(I - 0.1207) + 284.099(\theta\frac{\pi}{180} - 0.8997) = 0\}. \quad (18)$$

X. Moving in the Direction of Lower Variance

Now that we have found where $Y \approx 120$ N, we follow the tangent plane approximation P_1 in the direction of lower variance. Since we are on a plane, the direction to go is not as obvious as it was in the two-dimensional case.

A. Optimization towards the Corners of the Hyperplane

In the 2D design space optimization procedure, we moved towards the lower right hand end of the tangent line. We generalize this procedure to N dimensions. In our case, the hyperplane satisfies that constraint that the $\bar{f} = 120$ and the function g to minimize on the plane is the variance.

Suppose $\vec{x} \in R^N$ where $N \geq 2$. For all i, x_i is bounded. Let $f, g \in R$ where $f = f(\vec{x}; \vec{y})$ and $g = g(\vec{x}; \vec{z})$. The entries of $\vec{y}$ and $\vec{z}$ are parameters that are fixed throughout the optimization. The objective is to minimize g on the hyperplane $f(\vec{x}) = \bar{f}$. Let $f_k = f(\vec{x}_k)$ and $g_k = g(\vec{x}_k)$.

1. Foundations

After gradient descent with an analytic line search coefficient, we obtain a point $\vec{x}_0 \in R^N$ where $f(\vec{x}_0) = \bar{f}$. We find the tangent plane

$$\sum_{i=1}^{N} \frac{\partial f}{\partial x_i}|_{\vec{x}=\vec{x}_0}(x_i - [x_0]_i) = d = 0. \quad (19)$$

Now a hypersphere that represents all points a fixed distance ℓ from $\vec{x}_0$ is expressed by

$$\sum_{i=1}^{N} \Delta x_i^2 = \ell^2 \quad (20)$$

where $\vec{\Delta x} = \vec{x} - \vec{x}_0$ and $\ell > 0$ is the problem and point specific line step.

Since $\ell > d = 0$, there are two or infinitely-many points in the intersection of the hyperplane and the hypersphere. See Table 6. Therefore, we know the points we are looking for in the next section exist.

Table 6. Intersections of the hyperplane and hypersphere.

N	Intersection
2	2 points
3	(great) circle
4	sphere

2. *Finding p_k*

Since the entries x_i are bounded, the corners of the tangent hyperplane are the points of intersection of the hyperplane with the hypercube formed by the bounds on the x_i. The hyperplane has 2^{N-1} corners. Let L_k be the line passing through both $\vec{x}_0$ and corner C_k where $k = 1, \ldots, 2^{N-1}$. We want to move along L_k. We want p_k to be some fraction s of the length of L_k away from $\vec{x}_0$. Subsequently, the i^{th} component of the point p_k is expressed by

$$[p_k]_i = [\vec{x}_0]_i + s([C_k]_i - [\vec{x}_0]_i) \tag{21}$$

where $0 < s \leq 1$.

3. *Moving on the plane*

Let K be the index k where $\min\limits_{1 \leq k \leq N} g$ is minimized. Let $\vec{x}_1 = p_K$.

If $g_1 \leq g_0$, keep moving along L_K. If $g_1 > g_0$, either decrease ℓ and move on L_K starting from $\vec{x}_0$ or start over using P_K as your initial starting point and moving a smaller fraction of the L_k. Recall that if we move more than a chosen tolerance away from $f = \bar{f}$, we must compute a new tangent plane and find the corners of that plane.

B. Implementation of the Corners Method

For our problem, we compute g at the p_k with the 2 sample ordinal variance. So, the overall cost of checking all of the corners is 2^N or 8 function evaluations.

We set $s = \frac{1}{2}$. We choose a step threshold of one-fifth of the shortest edge of the plane so that if we cannot move at least one-fifth of the plane, we move directly to the corner of lowest variance. We find the lowest variance by stepping toward the corner of the hyperplane on the 500 N/mm - 45 degree edge. At the first step toward the corner, the friction force is 119.2661 N which is within one percent of the target 120 N. However, at our next step which brings us to the corner, the friction force is 115.6586 N. So, we must find a new tangent plane approximation.

XI. Finding a new approximation to $Y = 120$ N

We find the fifth point where the friction force is within one percent of 120 N after two line search steps. Now, we obtain a second tangent plane approximation P_2,

$$P_2 = \{(K, I, \theta) | 0.2383(K - 500) + 707.114(I - 0.1685) + 251.5537(\theta\frac{\pi}{180} - 0.78693) = 0\}. \tag{22}$$

We apply the corners method to decide to head toward the corner on the 500 N/mm-45 degree edge. We cannot move one-fifth of the shortest edge of the plane thus we take a truncated step to the sixth point. At this point, the friction force is within one percent of 120 N and we cannot move in a lower variance direction on this plane. We do an optimality check at the end by comparing the variance at the final point to the two sample ordinal variance at two other points on the plane. The final point is a corner of the plane. We sample at points on the two edges that intersect at the final point. We find the variance at two points which are one percent of the edge length from our final point. Among these three points, the final point has the lowest variance. Therefore, we have completed this part of the process.

XII. Studies on the Mean at the Final Point

We study the mean at the final point using point estimate methods. Table 7 presents the mean and standard deviation estimates as well as reference values generated from 1,000 LHS samples. Overall, we conclude that the mean at the final point is within two-fifths of the standard deviation from the target mean 120 N.

Table 7. Increasing accuracy of mean (first moment) and standard deviation estimates.

Method	Function Evaluations	Mean $\bar{Y}$ N	s (N)
Mean value	1	119.1658	
Hong's "$2n$"	8	119.3602	13.6487
Hong's "$2n+1$"	9	119.3601	13.6464
Two-point Rosenblueth"2^n"	16	119.3603	13.6836
Three-point Seo-Kwak "3^n"	81	119.3605	13.6882
LHS with 1,000 samples	1,000	119.368	13.6908

XIII. Iteration History of the 3D Design Space Optimization

We present the iteration history in Table 8. For a visual representation of this table, refer to Figure 10. We note that a similar point (500 N/mm, 0.17 mm, 45 degrees) was obtained by Tsai.[12]

Table 8. Iteration history the heuristic.

	K (N/mm)	I (mm)	θ (degrees)	Y (N)	s^2 (N^2)	s (N)
1	550.0000	0.2250	55.0000	261.1819	671.4901	25.9131
2	550.0000	0.1207	51.5493	120.7044	340.6632	18.4571
3	525.0000	0.1424	48.2746	119.2661	248.0990	15.7512
4	500.0000	0.1641	45.0000	115.6586	183.2965	13.5387
5	500.0000	0.1685	45.0878	119.1681	186.7583	13.6660
6	500.0000	0.1691	45.0000	119.1658	185.7426	13.6287

XIV. Cost of the 3D design space Optimization Process

To find the first plane P_1 we use 14 function evaluations. We check the corners of that plane to find the direction of decreasing variance using 8 function evaluations. We take two steps on the plane and do variance comparisons using 3 more function evaluations. Now, we must find a second tangent plane approximation P_2. This requires 8 function evaluations. We determine the direction of decreasing variance using another eight function evaluations. We move along that line and check the variance with a single function evaluation. At that corner, we do an optimality check using four function evaluations. Finally, we determine whether the mean at the final point is within two-fifths of the standard deviation using Hong's "$2n$" which requires 8 function evaluations. Therefore, the overall cost is 54 function evaluations.

XV. Comparison of Cost and Results to Other Methods

Just as we did the 2D problem, we compare the results and cost of the algorithm to those of other methods. We use 1,000 LHS samples to determine the mean and standard deviation at the returned point for each method. The sample mean at the final point is 119.368 N and the sample standard deviation is 13.6908 N.

Using DAKOTA we perform an exhaustive search to find the point of best variance where the mean value of the friction force is approximately 120 N. We implement the nested optimization routines in DAKOTA. The outer loop is an optimization algorithm while the mean and variance used in the constraints and objective function are determined in the inner loop.

The outer loop is a Coliny division of rectangles (DIRECT) optimization routine. The global search balancing parameter is set to zero while the local search balancing parameter is set to 1×10^{-8}. We test DIRECT using two inner loops.

First, the mean and variance are determined by 10 LHS samples in the inner loop. The objective of DIRECT is to minimize variance subject to the constraint that the mean be within 1 N of the target mean. We set a solution threshold of 15. This nested routine performs 1,390 function evaluation and returns the point $K = 531.07$ N/mm, $I = 0.1417$ mm, and $\theta = 48.3347$ degrees. There the mean is 120.537 N and the standard deviation is 16.0440 N.

Next, we use mean value in the inner loop. We aim to minimize $E[\| Y - 120 \|] + 3\sigma$ with a solution threshold of 48. The standard deviation is computed with a gradient stepsize of 1×10^{-4}. We set the convergence tolerance to 1×10^{-4} and the threshold delta to 1×10^{-8}. We find the point $K = 529.0123$ N/mm, $I = 0.1417$ mm, and $\theta = 48.3338$ degrees using 515 function evaluations. We use Hong's "$2n$" to confirm that the mean is within two-fifths of the standard deviation from 120 N. Here the mean is 120.0699 N and the standard deviation is 15.9831 N.

Now, we use the Coliny pattern search in the outer loop. The objective is to minimize variance subject to the constraint that the mean be within 1 N of the target mean 120 N. We set a solution threshold of 15. The solution accuracy is chosen to be 1×10^{-8}. We pick an initial delta of .5. The threshold delta is set to 1×10^{-13}. We choose a contraction factor of .85. Using 2,940 function evaluations, the routine yields the point $K = 505.2176$ N/mm, $I = 0.1415$ mm and $\theta = 49.3775$ degrees where the mean is 120.2938 N and the standard deviation is 15.9609 N.

XVI. Conclusions

We have described an algorithm designed to reduce the number of function evaluations required to perform optimization under uncertainty. This method minimizes variance while ensuring that a prescribed mean target level of system output is achieved. We applied the optimization procedure to an automotive device design robustness problem. We studied a two-dimensional example and a three-dimensional example. We compared the results and cost of our algorithm to other algorithms. We found that the suggested algorithm finds similar or better results at a lower cost.

References

[1] *Fundamentals of annular snap-fit joints*, Machine Design.com, Penton Media, Inc (2005).

[2] M. S. Eldred, H. Agarwal, V. M. Perez, Wojtkiewicz Jr., and J. E. Renaud, *Investigation of reliability method formulations in DAKOTA/UQ*.

[3] M. S. Eldred, S. L. Brown, B. M. Adams, D. M. Dunlavy, D. M. Gay, L. P. Swiler, A. A. Giunta, W. E. Hart, J. P. Watson, J. P. Eddy, J. D. Griffin, P. D. Hough, T. G. Kolda, M. L. Martinez-Canales, and P. J. Williams, *DAKOTA, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 4.0 user's manual*, SAND2006-6337, (2006).

[4] M. S. Eldred, A. A. Giunta, S. F. Wojtkiewicz Jr., and T. J. Truncano, *Formulations for surrogate-based optimization under uncertainty*, American Institute of Aeronautics and Astronautics, (2002).

[5] A. Haldar and S. Mahadevan, *Probability, Reliability, and Statistical Methods in Engineering Design*, John Wiley and Sons, New York, 2000.

[6] M. E. Harr, *Probabilistic estimates for multivariate analyses*, Applied Mathematical Modeling, 13 (1989).

[7] H. Hong, *An efficient point estimate method for probabilistic analysis*, Reliability Engineering and System Safety, 59 (1998), pp. 261–267.

[8] J. Neter, W. Wasserman, and M. Kutner, *Applied Linear Statistical Models*, Irwin, Homewood, IL, 1985.

[9] V. J. Romero and C.-H. Chen, *Refinements in a new adaptive ordinal approach to continuous-variable probabilistic optimization*, American Institute of Aeronautics and Astronautics, (2007).

[10] E. Rosenblueth, *Two-point estimates in probability*, Applied Mathematical Modeling, (1981), pp. 329–335.

[11] H. S. Seo and B. M. Kwak, *Efficient statistical tolerance analysis for general distributions using three-point information*, International Journal of Production Research, 40 (4), pp. 931–944.

[12] S.-C. Tsai, *Taguchi S/N ratios and direct robustness measurement for computational robust design*, SAE2006-01-0738, 2006 SAE World Congress, Detroit, MI, April 3-6, 2006.

[13] L. Wang, D. Beeson, and G. Wiggs, *Efficient moment and probability distribution estimation using the point estimate method for high-dimensional engineering problems*, American Institute of Aeronautics and Astronautics, (2007).