

# Efficient Sampling Technique for Optimization under Uncertainty

Urmila M. Diwekar and Jayant R. Kalagnanam

Center for Energy and Environmental Studies and Dept. of Engineering and Public Policy, Carnegie Mellon University, Pittsburgh, PA 15213

*The concept of robust design involves identification of design settings that make the product performance less sensitive to the effects of seasonal and environmental variations. This concept is discussed in this article in the context of batch distillation column design with feed stock variations, and internal and external uncertainties. Stochastic optimization methods provide a general approach to robust/parameter design as compared to conventional techniques. However, the computational burden of these approaches can be extreme and depends on the sample size used for characterizing the parametric variations and uncertainties. A novel sampling technique is presented that generates and inverts the Hammersley points (an optimal design for placing  $n$  points uniformly on a  $k$ -dimensional cube) to provide a representative sample for multivariate probability distributions. The example of robust batch-distillation column design illustrates that the new sampling technique offers significant computational savings and better accuracy.*

## Introduction

Robust/parameter design is an off-line quality control method popularized by the Japanese quality expert G. Taguchi, for designing products and manufacturing processes that are robust in the face of uncontrollable variations (Bendell et al., 1990). At the design stage, the goal of parameter design is to identify design settings that make the product performance less sensitive to the effects of manufacturing and environmental variations, and deterioration.

In parameter design, Taguchi's stated objective is to find settings of the product or process design parameters that minimize an average quadratic loss function, defined as the average standard deviation of the response from a target value. In order to select the settings of the design parameters, a set of measures called signal-to-noise ratio (SN) needs to be maximized. Taguchi uses orthogonal-array designs to arrive at the optimal settings. This approach is limited since it requires that the mean and variance of the output variable are not coupled (Hunter, 1985; Kacker, 1985; Nair, 1992). In order to evaluate whether this assumption holds, an *a priori* functional relationship between the input and output is required, but is seldom available for large and complex real-world problems. An alternative is to treat the design variables and the noise factors in a single matrix and develop the

response directly as a function of the control and noise factors and then optimize the loss functions via these approximations (Easterling, 1989; Welch et al., 1990; Yu et al., 1991). Clearly, the reliability of this optimization depends on the accuracy of the approximating models. Since many of the parameter design experiments are now run using computer models, the computational cost of running the simulator often dictates the size of the samples used for constructing the response surfaces. Latin Hypercube (McKay et al., 1979) designs are now commonly used for constructing response surface models since they require a smaller number of samples as compared to the orthogonal array designs (Welch and Sacks, 1991). However, if the factor ranges of the controls and noise are large, the input-output relationships are complex and larger samples are required to get reliable approximations (Currin et al., 1991; Sacks et al., 1989; Welch et al., 1992). Therefore the availability of an efficient sampling technique that provides reliable estimates of the performance statistics using a reasonably small sample size lies at the heart of these approaches.

The most general approach to the robust or parameter design problem is to couple an optimizer directly with the computer-simulation model using stochastic descriptions of the

noise factors (Boudriga, 1990; Diwekar and Rubin, 1994). Such an approach is more reliable than using a response surface model; however, it is also computationally more expensive. The use of such a direct method is warranted in process applications where input-output relations are highly nonlinear and rugged (Diwekar and Rubin, 1991). Furthermore, in process design, minimizing just the variance (in the form of a loss function) can lead to severe overdesigns, therefore, an economic function needs to be included along with the loss function in the robust design evaluations, and operating or control variables (such as reflux ratio) should be considered in the optimal decisions to compensate for the variations. The stochastic optimization problem involves the evaluation of an aggregate measure (used as a performance statistic) derived from a multivariate probability distribution. For nonlinear models, this is done numerically using a representative sample from the multivariate space and has to be repeated at each optimization iteration, as shown in Figure 1, where the inside loop is an iterative stochastic model. One can easily envision the computational intensity of the stochastic optimization problem presented in the Figure 1. Therefore, an efficient sampling scheme that reduces the number of samples required for each iteration can significantly improve the computational efficacy of the stochastic optimization procedure.

In this article, we present a new and efficient sampling technique using the shifted Hammersley points for uniformly sampling a  $k$ -dimensional unit hypercube. This new sampling technique requires far fewer samples, as compared to other techniques, to approximate the mean and variance of distributions derived by propagating a representative sample (for the inputs) over nonlinear functions. For the robust design problem posed in terms of stochastic optimization, the use of this efficient sampling technique can significantly alleviate the computational burden.

Chemical processes are subject to a high degree of uncertainties. Stochastic variabilities or uncertainties in batch chemical processes typically arise from variations in the initial conditions (e.g., feed-stock composition and temperature) and operating procedures, as well as from equipment failures and other unexpected reductions in resource availability, and

noise in measurements used for monitoring and control purposes. For units such as batch crystallizers (Mullin, 1993) the uncertainties are inherent in crystal size distributions, and related property predictions. Furthermore, most industrial batch processes are operated through open-loop applications of an off-line optimized input profile, such as feed or temperature, commonly known as optimal control profiles. These profiles even if reoptimized using on-line measurements often contain nonnegligible uncertainties due to the system state being inferred from indirect, so-called model-based measurements, which can be subject to both stochastic measurement noise and structural measurement-model mismatch (Terwiesch, 1995). The need to take these uncertainties into consideration in the design and planning stage is well recognized (Reklaitis et al., 1989; Cott and Machhietto, 1989; Watzdorf et al., 1993; Mignon et al., 1995; Tsai and Chang, 1996; Terapetritou and Pistikopoulos, 1996). This new sampling technique provides an efficient and generalized approach for handling uncertainties in batch process design, optimization, scheduling and planning. We illustrate this in the context of robust design of a batch distillation column that is subject to feedstock variations, modeling uncertainties, and measurement errors and report computational savings of up to a factor of 10.

In this article the conventional sampling techniques used in the literature are compared with the new sampling technique based on Hammersley points using the results of a large set of numerical experiments. The concept of robust design of batch distillation column is presented, as well as the accuracy and computational efficiency gained by the use of this new sampling technique in solving the robust design problems.

#### Monte-Carlo and Latin-Hypercube Sampling Techniques

Perhaps one of the best known methods of sampling a probability distribution is the Monte-Carlo sampling technique, which is based on the use of a pseudorandom number generator to approximate a uniform distribution,  $U(0, 1)$  with  $n$  samples. The specific values for each input variables are selected by inverting the  $n$  samples over the cumulative distribution function. A Monte-Carlo sample has the property that successive points are independent. However, in most applications, the actual relationship between successive points in a sample has no physical significance, hence the independence/randomness of a sample for approximating a uniform distribution is not critical (Knuth, 1973). Moreover, the error of approximating a distribution by a finite sample depends on the equidistribution properties of the sample used for  $U(0, 1)$  rather than its randomness. Once it is apparent that the uniformity properties are central to the design of sampling techniques, constrained or stratified sampling becomes appealing (Morgan and Henrion, 1990).

Latin hypercube sampling (LHS) is one form of stratified sampling that can yield more precise estimates of the distribution function (Iman and Shortencarier, 1984). The range of each  $u_i$  is divided into nonoverlapping intervals of equal probability. One value from each interval is selected at random with respect to the probability density in the interval. The  $n$  values thus obtained for  $u_i$  are paired in a random

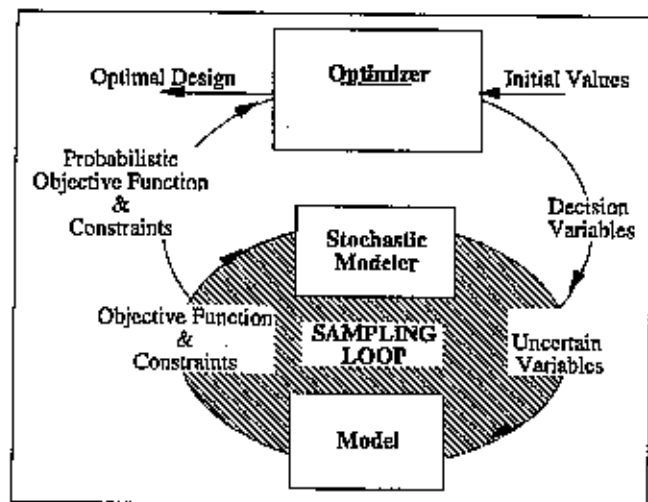


Figure 1. Stochastic optimization framework.

manner with the  $n$  values of  $X_2$ , and these  $n$  pairs are combined with  $n$  values of  $X_3$  and so on to form  $n$   $k$ -tuplets. The random pairing is based on a pseudorandom number generator. The main shortcoming with this stratification scheme is that it is one-dimensional and does not provide good uniformity properties on a  $k$ -dimensional unit hypercube. Until recently, the only known design for uniformity on an  $d$ -dimensional hypercube was a uniform grid (Papageorgiou and Wasilkowski, 1990). However, the uniform grid requires exponential sample points in the number of variables for good equidistribution properties. Recently, Wozniakowski (1991) showed that the shifted Hammersley points provide the location for  $n$   $k$ -dimensional sample points on a unit hypercube so as to minimize the discrepancy from a uniform grid. The next section describes the new sampling technique, which is based on generating the Hammersley points, a variant of the sequence used by Wozniakowski.

### New Sampling Technique

Since most of the stochastic optimization problems involve integrals of some probabilistic function, consider the approximation of an integral of a  $k$ -dimensional continuous function by sampling its values at a finite set of points. For the sake of simplicity let us assume that the integration is restricted to a  $k$ -dimensional unit cube. One straightforward approach is to place the points along equally spaced intervals on a  $d$ -dimensional grid. Although this is a good arrangement, the number of points needed to keep the average error less than  $\epsilon$  is roughly proportional to  $1/\epsilon^k$ . The traditional alternative is to use a Monte-Carlo method where the points are chosen completely randomly using a pseudorandom number generator. The approximation to the integral is then based on the function evaluation at these points. Although on the average the number of points required to keep to error within  $\epsilon$  is bound by  $1/\epsilon^2$  there is no methodical way for constructing the sample points to achieve the bound (Papageorgiou and Wasilkowski, 1990). Recently, Wozniakowski (1991) showed that the shifted Hammersley points provide a low-discrepancy design. The number of points required to contain the approximation error within  $\epsilon$  is proportional to  $(1/\epsilon)(\log(1/\epsilon))^{k-1/2}$ .

In this section we describe a new sampling technique designed using the Hammersley sequence. We call this new technique the Hammersley sequence sampling (HSS) technique. The basic idea behind this technique is to replace a Monte-Carlo integration with a quasi-Monte-Carlo scheme. This quasi-Monte-Carlo scheme uses a quasi-random number generator based on the Hammersley points to uniformly sample a unit hypercube, and inverts these points over joint cumulative probability distribution to provide a sample set for the variables of interest. In the following two subsections we describe an algorithm for generating the Hammersley points, and then describe the implementation of inversion and the imposition of a correlation structure on the sample.

### Hammersley sequences

The choice of an appropriate quasi-Monte-Carlo sequence is based on the concept of discrepancy. The deterministic upper and lower error bounds of any sequence for integration are expressed in terms of the discrepancy measure. Discrep-

ancy is a quantitative measure for the deviation of the sequence from uniform distribution. Therefore it is typically desirable to choose a low-discrepancy sequence. Some examples of low-discrepancy sequences are the Halton (1986) and Hammersley (1960) sequences. However, the constant terms on the error bounds for these sequences are a strong function of the dimension  $k$  of the unit hypercube and other sequences such as the Sobol sequences (Niederreiter, 1978) and Faure sequences (Fox, 1986) that have been developed to alleviate this problem. The other problem often encountered with the previously cited sequences is that the error bounds are not adequately sensitive to the form of the integrand. Several designs using "good lattice" points were introduced by Korobov (Niederreiter, 1978, 1988) in the literature to address these issues. Without embarking on a detailed discussion of these issues (the interested reader is referred to Niederreiter, 1992), it is apparent that we are faced with the issue of which sequence should one use for the design of a quasi-Monte-Carlo sampling technique.

In this article, we have chosen to use a variant of the Hammersley sequence. In the following paragraphs we provide a definition of the Hammersley points and explicate an algorithm for the Hammersley design.

Any integer  $n$  can be written in radix- $R$  notation ( $R$  is an integer) as follows:

$$n \equiv n_m n_{m-1} \dots n_2 n_1 n_0 = n_0 + n_1 R + n_2 R^2 + \dots + n_m R^m,$$

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

$$\phi_R(n) = 0.n_m n_{m-1} n_{m-2} \dots n_1 n_0 = n_0 R^{-1} + n_1 R^{-2} + \dots + n_m R^{-m-1}.$$

The Hammersley points on a  $k$ -dimensional cube is given by the following sequence:

$$z_k(n) = \left( \frac{n}{N}, \phi_{R_1}(n), \phi_{R_2}(n), \dots, \phi_{R_{k-1}}(n) \right) \quad n = 1, 2, \dots, N$$

where  $R_1, R_2, \dots, R_{k-1}$  are the first  $k-1$  prime numbers. The Hammersley points are  $x_k(n) = 1 - z_k(n)$ . Now we present an algorithm to generate  $N$  Hammersley points:

1.  $k \leftarrow$  dimension of the unit cube,  $N \leftarrow$  number of samples
2.  $i \leftarrow 1$
3.  $R_i \leftarrow \text{gen\_prime}(i)$ ,  $i \leftarrow i + 1$
4. If  $i \leq (k-1)$  go to 3
5.  $\text{count} \leftarrow 0$ ,  $n \leftarrow 1$
6.  $i \leftarrow 2$ ,  $z_1 \leftarrow \frac{n}{N}$
7.  $z_i \leftarrow \text{inv\_radix}(n)$ ,  $i \leftarrow i + 1$
8. if  $i \leq (k-1)$  go to 7
9.  $x_k(n) = 1 - z_k$
10.  $\text{count} \leftarrow \text{count} + 1$ ,  $n \leftarrow n + 1$ .
11. If  $\text{count} < N$  go to 5.

$\text{gen\_prime}(i)$  is a subroutine that generates the  $i$ th prime number;  $\text{inv\_radix}(n)$  is a subroutine that generates the in-

verse radix notation for  $n$ . We have implemented a version of this algorithm and tested it extensively using different distributions. A prime number generator from Knuth is used (1973, pp. 143-44).

### Implementation of correlation structures

The implementation of correlation structures is based on the use of rank correlations (Iman and Conover, 1982). The method is very similar to the one used for Latin hypercube samples with one difference: LHS uses a matrix of independent permutations of arbitrary scores for generating a correlation structure, whereas for HSS we use the Hammersley points for the same purpose. In this subsection we outline the method based on rank correlations used for generating a correlation structure in LHS and highlight the main difference in the implementation for the HSS technique.

Let  $X$  be a matrix of uncorrelated random vectors, and let  $C$  be the desired rank correlation matrix of  $X$ . Then since  $C$  is positive definite,  $C = P \times P'$  (Cholesky factorization), where  $P$  is a lower triangular matrix. Then for some matrix  $R$  of arbitrary scores the transformed matrix  $R^* = R \times P'$  has the desired rank correlation matrix  $C$ .  $R$  is chosen such that correlation matrix and the rank correlation matrix of  $R^*$  are the same. Now to introduce the desired rank correlation in  $X$ , the random vectors are arranged in the same rank order as  $R^*$ . For LHS, the matrix  $R$  is constructed from van der Warden scores (Iman and Conover, 1982), whereas for HSS the matrix is the set of Hammersley points.

The main impact of using rank correlations for HSS is that the uniform structure of the Hammersley points is somewhat distorted; however, its effect on the transformed sample is not easily characterized analytically. Whether the distortions are large enough to completely negate the advantages of the Hammersley points is an empirical question that is investigated in the following section by comparing the convergence properties of HSS with LHS for correlated samples.

### Uniformity properties of the Hammersley points

In our discussion of different sampling techniques, we explicate the importance of the uniformity properties of a sampling technique when the sample is used for approximating a distribution by finite samples. Figure 2 graphs the samples generated by different techniques on a unit square. This provides a qualitative picture of the uniformity properties of the different techniques. It is clear from Figure 2 that the Hammersley points have better uniformity properties compared to other techniques. The main reason for this is that the Hammersley points are an optimal design for placing  $n$  points on a  $k$ -dimensional hypercube. In contrast, other stratified techniques such as the Latin hypercube are designed for uniformity along a single dimension and then randomly paired for placement on a  $k$ -dimensional cube. Therefore the likelihood of such schemes providing good uniformity properties on high-dimensional cubes is extremely small. Figure 3 illustrates the effect of imposing a correlation structure on the

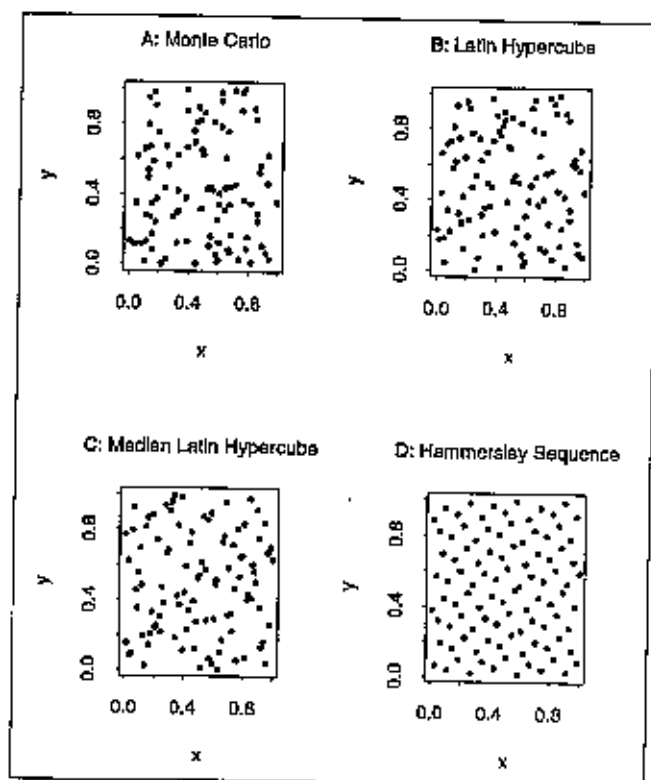


Figure 2. Sample points (100) on a unit square using (A) linear congruent generator, (B) random Latin hypercube, (C) median Latin hypercube, and (D) the Hammersley points.

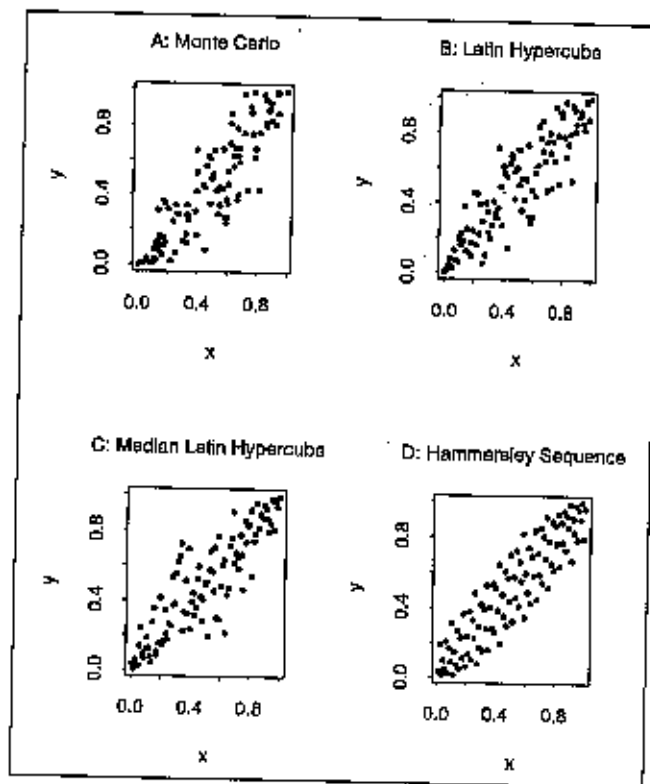


Figure 3. Sample points (100) on a unit square with correlation of 0.9 using (A) linear congruent generator, (B) random Latin hypercube, (C) median Latin hypercube, and (D) the Hammersley points.

sample sets. The approach used is described in the earlier section, which uses rank correlations to preserve the stratified design along each dimension. Although this approach preserves the uniformity properties of the stratified schemes, the optimal location of the Hammersley points is perturbed by imposing correlation structure. The effect of this on the uniformity properties is not apparent from Figures 2 and 3; however, we will examine this issue in detail in the following subsections.

### Convergence properties of samplers

In this subsection, we provide a comparison of the performance of the HSS technique to that of LHS and Monte-Carlo (MCS) techniques. The comparison is performed by propagating samples derived from each of the techniques for a set of  $n$ -input variables ( $X_i$ ), through various nonlinear functions ( $Y = f(X_1, X_2, \dots, X_n)$ ) and measuring the number of samples required to converge to the mean and variance of the derived distribution for  $Y$ . Since there are no analytic approaches (for stratified designs) to calculate the number of samples required for convergence, we have conducted a large matrix of numerical tests. The design of the test matrix included varying of the type of function, the number of input variables,  $X_i$ , type of input distribution, and the correlation structures between them. The details of the test matrix are described below:

**Sampling Techniques.** A total of four sampling techniques have been compared: Monte Carlo, random Latin hypercube, median Latin hypercube, and the Hammersley.

**Number of Variables.** The number of input variables used was varied between 2 and 10.

**Functions.** Five different kinds of functions were used and are outlined below:

1. Function 1: Linear additive function:  $Y = \sum_i X_i$
2. Function 2: Multiplicative function:  $Y = \prod_i X_i$
3. Function 3: Quadratic function:  $Y = \sum_i X_i^2$ ; for  $i=2$ ,  $Y = X_1^2 + X_2^2$
4. Function 4: Exponential function:  $Y = \sum_i X_i \times \exp X_{i+1}$
5. Function 5: Logarithmic function:  $Y = \sum_i X_i \times \log(X_{i+1})$

**Distributions.** Three types of distributions have been used for the input variables  $X_i$ . Two of them, uniform and normal, are symmetric and the third is a skewed distribution, lognormal.

**Correlations.** Three types of correlation structures have been used: the first is a zero correlation, and the other two sets use a correlation of 0.5 and 0.9 between the input variables.

The matrix represents a total of 180 data sets (4 sampling techniques  $\times$  3 types of distributions  $\times$  3 correlation structures  $\times$  5 functions) for each set of input variables,  $X_i$ . As the number of input variables is varied from 2 to 10, it adds another factor of 9, that is, 1,620 data sets. However, in the interests of space and clarity, we present only the results that highlight the main findings of this numerical experiment. Initially we present two figures (Figure 4 and 5) that illustrate the rate of convergence for the three sampling techniques: MCS, LHS, and HSS sampling techniques. Figure 4 plots the mean and the variance for Function 2 using two input variables that are uncorrelated. A uniform distribution  $U(0,1)$  is used for both the inputs. The results are unequal-

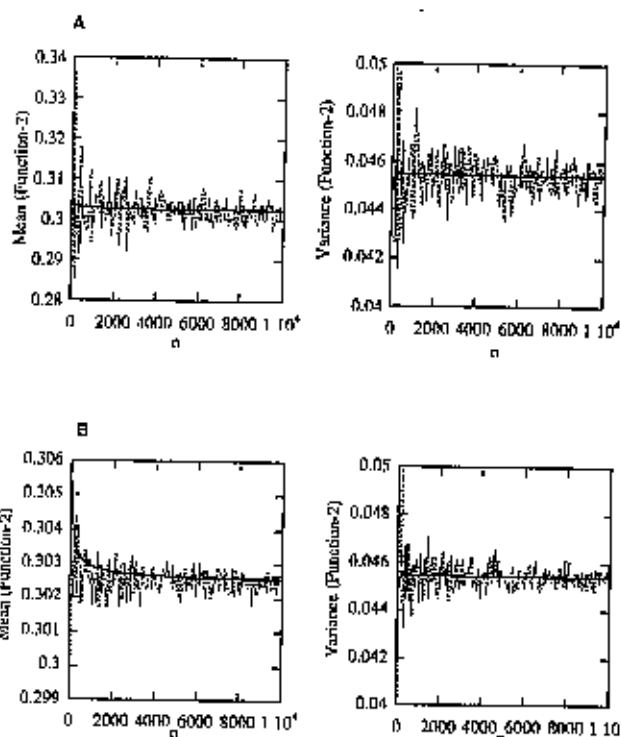


Figure 4. Mean and variance as a function of sample size for MCS and LHS (dotted line) and HSS (solid line) for 2 input variables without correlations: (A) MCS vs. HSS (B) LHS vs. HSS.

vocal—MCS requires a significantly large number of samples compared to LHS and HSS, and the HSS technique requires far fewer samples to converge to within 1% of the variance as compared to LHS. The results are similar for correlated samples (Figure 5).

Figure 6 presents a more comprehensive view of the comparisons conducted in the numerical experiment. This figure plots the ratio of the LHS-to-HSS sample size as a function of the design parameters (outlined in the matrix given earlier) of the numerical experiments. The sample size used for this comparison is the number of samples required to converge to within 1% of the actual value of variance. Each subgraph plots the ratio of the sample size against the number of

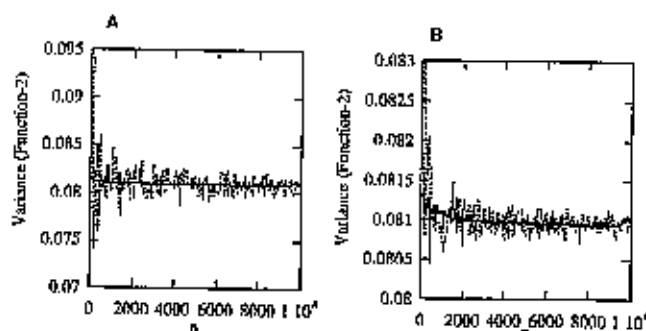
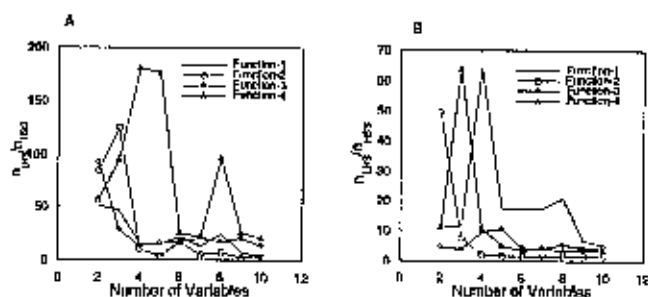


Figure 5. Variance as a function of sample size for MCS and LHS (dotted line) and HSS (solid line) for 2 input variables with correlation of 0.9: (A) MCS vs. HSS (B) LHS vs. HSS.



**Figure 6. Ratio of LHS-to-HSS sample size against different design parameters for 1% convergence of variance: (A) function-1 to function-4 for uniform input distributions, and (B) function-1 to function-4 for lognormal input distributions. Note that wherever LHS did not converge within 10,000 samples it was assumed that the convergence point is 12,500 samples (large number).**

input variables for the four functions (function 1 to function 4) and for two types of input distributions; uniform (symmetric) and lognormal (asymmetric) uniform input distributions. Once again the results are encouraging—the HSS sampling technique has a much faster convergence rate, anywhere from a factor of 1.5 to 100 and larger! The results presented here are qualitatively representative of the general trends observed for all the data sets that were analyzed.

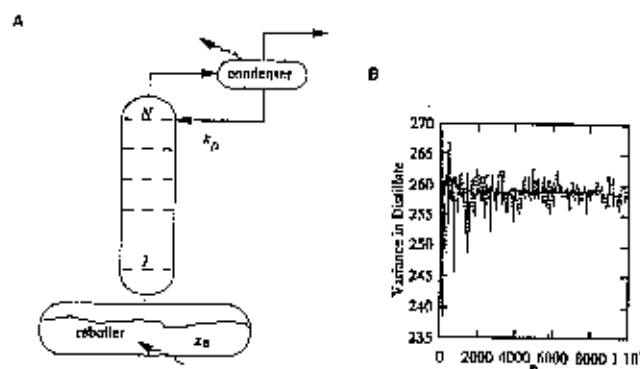
In the next section, we apply the new sampling scheme to the robust design of a batch distillation column and illustrate the computational savings as compared to using the conventional techniques such as Monte Carlo or Latin hypercube.

### Robust Design of a Batch-Distillation Column

The sudden increase in the production of high value-added, low-volume specialty chemicals and biochemicals in recent years has generated renewed interest in batch-distillation design. However, the current design procedures are still based on deterministic framework. We first present a description of the robust design problem in the context of batch-distillation column design and then outline the sample size required to characterize the variance of the output from the process that needs to be controlled for quality. Finally, we present the computational burden of solving the robust design problem for both the sampling techniques.

#### Problem definition

Figure 7a shows a conventional batch-distillation column with a reboiler at the bottom and a condenser at the top, which essentially performs the rectifying operation. For simplicity, consider a binary mixture with the feed composition  $x_F$  of a more volatile component (mole fraction) and total feed  $F$  (moles) to be separated in a batch distillation column. There are variabilities and fluctuations in feed composition and feed amount over different batches, which amounts to saying that the feed composition ( $x_F$ ), and feed ( $F$ ) are uncertain quantities. There are measurement errors in quantities like vapor boilup rate  $V$  (function of heat input to the reboiler), and reflux ratio  $R$ . The thermodynamic errors lead



**Figure 7. Robust design of a batch column: (A) a batch distillation column, (B) comparison of LHS (dotted) vs. HSS (solid line) for variance calculations for the batch column.**

to uncertainties in relative volatility ( $\alpha$ ) predictions. Given these variabilities and uncertainties, we want to design a batch column that will maintain the amount of product with the given purity. These variations are assumed to be at  $E_i$ % error levels in the inputs,  $E_i = \sigma_i/\mu_i \times 100$ , normally distributed. Numerical values of the design and noise variables (nominal values,  $\mu_i$  and % error levels  $E_i$ ) are given in Table 1. So the problem of robust design translates into finding the number of theoretical plates and reflux ratio required to minimize the variance in the distillate amount (expressed in terms of a loss function described in the next subsection) of specific purity  $x_D^*$ .

Since the objective from a robust design perspective involves minimization of variance in the amount of distillate (loss function), we characterize the number of samples required to estimate the variance to within 1% of its value using both Latin hypercube and Hammersley points. It should also be noted that calculation of the mean requires a significantly smaller number of samples than variance, and variance calculation is the major efficiency bottleneck for the robust design problem. The Hammersley points requires about 650 points to converge as compared to 6,550 points required by the Latin hypercube sample. Figure 7b plots the variance as a function of number of samples. However, minimizing just the variance (or a loss function based on variance) can lead to severe overdesign, therefore, we are considering an economic criterion with the associated loss due to variations as the objective function.

#### Objective function

Here we formulate the economic objective function based on the mean values of the parameters. The objective function

**Table 1 Parameters and Their Values Used in the Study**

Parameter	Values	Units	$E_i$ %
$\alpha$	2.0		5
$x_F$	0.5	mol fraction	10
$R$	2.5		5
$V$	100	mol/h	5
$F$	100	mol	10
$x_D^*$	0.95	mol fraction	
$D_{spec}$	36	mol	
$N$	10.95		

is derived from the profit function described in Diwekar (1995). The objective was to maximize profit. Assuming  $D$  = average number of moles of product distilled,  $P_r$  = sales value of the product relative to the cost of feed \$0.1/mol, the profit function can be written as

$$P = \frac{24(365)DP_r}{T + t_s} - \frac{c_1VN}{G_a} - \frac{c_2V}{G_b} - \frac{24(365)c_3VT}{T + t_s} \quad (1)$$

where

- $c_1$  = amortized incremental investment cost [\$15.0/m<sup>2</sup>/plate/yr]
- $c_2$  = amortized incremental cost of the equipment [\$1.65/m<sup>2</sup>/yr]
- $c_3$  = cost of the steam and coolant to vaporize or condense, respectively, 1 kg of distillate [\$0.00935/kg]
- $G_a$  = allowable vapor velocity [15.0 kg/h/m<sup>2</sup>]
- $G_b$  = vapor-handling capacity of the equipment [0.1028 kg/h/m<sup>2</sup>]
- $N$  = number of plates
- $T$  = batch time (average) [h]
- $t_s$  = setup time for each batch [0.1 h]
- $V$  = vapor boilup rate [kg/h]

To incorporate the robust design concepts in the objective function we need to add the measurement of quality during design. Following Phadke (1989) a quadratic loss function is included in the profit function described earlier. Here we are using the asymmetric loss function shown below.

$$P = \frac{24(365)DP_r}{T + t_s} - \frac{c_1VN}{G_a} - \frac{c_2V}{G_b} - \frac{24(365)c_3VT}{T + t_s} - L(y_i) \quad (2)$$

$$L(y_i) = K_1(y_i - \mu)^2, \quad y_i > \mu \quad (3)$$

$$= K_2(y_i - \mu)^2, \quad y_i \leq \mu \quad (4)$$

The quality variable in this exercise is the amount of distillate. If the total distillate is less than the specified value, then there is a definite loss, therefore, we are using  $K_1 = (24(365)P_r)/(T + t_s)/30.0$ . In this exercise we are assuming that if you produce more product  $y_i \leq \mu$ , then there is no loss that is equivalent to  $K_2 = 0$ .

### Solution procedure

The iterative nature of the robust design calls for use of simplified models. The shortcut method presented by Diwekar and Madhavan (1991) provides an efficient alternative for robust design. Apart from computational efficiency, lower memory requirements, and the algebraic equation oriented form of the shortcut method, it is also useful in identifying the feasible region of operation crucial for design problems. Furthermore, the dimensionality of the problem does not increase with increasing number of plates, and the design variable number of theoretical plates is not an integer. These attributes make the shortcut method desirable for robust design procedures. For details of the shortcut method, please refer to Diwekar (1995).

As stated earlier, the robust design problem involves solution of a stochastic optimization problem. Please note that it is easier to manipulate operating variables to adjust for the variations, therefore, along with the key design variable,  $N$ , we are including the operating variable,  $R$ , in the optimal

decisions. The following procedure outlines the steps for solving the robust design problem stated earlier.

1. In the stochastic optimization the optimizer in the outer loop finds the decision variable  $N$  and reflux ratio  $R$ , and the inner loop is the stochastic model.

2. Outer loop—At first we need to find the feasible region for the decision variable  $N$  for the convergence of the optimizer. The minimum number of plates  $N_{min}$  provides the lower bound for  $N$ , the upper bound can be specified from the economic constraint, which is specified to be 25 in this case.

$$N_{min} = \frac{\ln \left[ \frac{x_D^*}{1 - x_D^*} \frac{1 - x_F}{x_F} \right]}{\ln[\alpha]}$$

Using the preceding relationship for the minimum number of plates, the lower bound for the problem given in Table 1 is found to be 4.25 (including reboiler as one plate). Therefore the following bounds are used for the preceding problem:

$$3.25 \leq N \leq 25,$$

where  $N$  does not include the reboiler.

3. For the given value of  $N$  and  $R$ , perform the inner-loop calculations.

(a) Inner loop— $R_{MIN}$  provides the lower bound to  $R$ . For the given value of  $N$ , the lower bound for  $R$  is calculated.

(b) Sample the noise variables  $R$ ,  $V$ ,  $x_F$ ,  $F$ , and  $\alpha$  using the HSS sampling technique. (We have also used LHS for evaluation of the performance of the new sampling technique.)

(c) If  $R$  for the sample is less than or equal to  $R_{MIN}$  (where  $R_{MIN}$  is defined as the value of  $R$  required to obtain the distillate composition of the key component ( $x_D^{(1)}$ ) equal to the specified average distillate composition ( $x_D^*$ ) at the initial conditions for the given  $N$ ), then equate the total distillate  $D = 0$  for that sample. Otherwise, run the shortcut model, stop when the average distillate composition is equal to the  $x_D^*$  value for each sample, and find the distillate amount  $D$  for each sample.

(d) After all the samples are passed through the shortcut model. Find the objective function for the outer loop. If the profit is maximum, then the number of plates and the reflux profile represent the robust design, else, repeat the inner loop calculations. Figure 8 shows the variation in the product

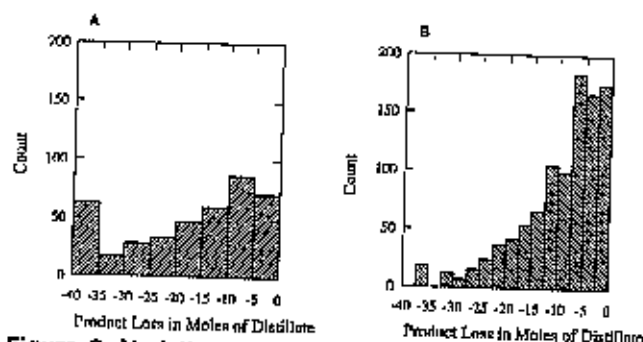


Figure 8. Variations in the amount of distillate, before (A) and after (B) design for quality.

amount before and after the robust design, where the variance in the loss function (when the total distillate is less than the specified value) is reduced from 256 to 77. The new optimal design requires 17.5 theoretical plates and a time-dependent reflux profile of  $R_t = 2.5t$  as compared to the original design of  $N = 7.5$  and  $R_t = 3.5t + 0.25$ . The optimal design resulted in a threefold increase in the annual profit.

## Conclusions

This article presented a new sampling technique based on Hammersley points. This new sampling technique is shown to have better uniformity properties, which reduces the computational intensity of the stochastic optimization problem considerably. A robust design concept was introduced in the context of a batch-distillation column design operating under internal and external uncertainties. Since the robust design concept essentially involves solution of the stochastic optimization problem, it was found that this sampling technique is always preferred for robust/parameter design problems. This is because of its high precision and consistent behavior coupled with great computational efficiency. The computational efficiency of this new sampling technique shows great promise for its use in other application areas such as stochastic modeling, Monte-Carlo simulations, and experimental design.

## Literature Cited

- Bendell, A., J. Disney, and W. A. Pridmore, *Taguchi Methods: Applications in World Industry*, IFS Publication, Springer-Verlag, UK (1989).
- Boudriga, S., *Evaluation of Parameter Design Methodologies for the Design of Chemical Processing Units*, M. S. Thesis, Univ. of Ottawa, Ottawa, Ont. Canada (1990).
- Cott, B. L., and S. Maechelitto, "Minimizing the Effects of Batch Process Variability Using On-line Schedule Modification," *Comput. Chem. Eng.*, **13**, 105 (1989).
- Currin, C., T. Mitchell, M. Morris, and D. Ylvisaker, "Bayesian Prediction of Deterministic Functions, with Applications to the Design and Analysis of Computer Experiments," *J. Amer. Stat. Assoc.*, **86**, 953 (1991).
- Dwykar, U. M., *Batch Distillation: Simulation, Optimal Design and Control*, Taylor & Francis, Washington, DC (1995).
- Dwykar, U. M., and K. P. Madhavan, "Multicomponent Batch Distillation Column Design," *Ind. Eng. Chem. Res.*, **30**, 713 (1991).
- Dwykar, U. M., and E. S. Rubin, "Stochastic Modeling of Chemical Processes," *Comput. Chem. Eng.*, **15**, 105 (1991).
- Dwykar, U. M., and E. S. Rubin, "Parameter Design Method Using Stochastic Optimization with ASPEN," *Ind. Eng. Chem. Res.*, **33**, 292 (1994).
- Easterling, R. G., "Discussion of Kacker's Paper," *J. Qual. Technol.*, **17**, 191 (1989).
- Fox, B. L., "Algorithm 647: Implementation and Relative Efficiency of Quasi-Random Sequence Generators," *ACM Trans. Math. Software*, **12**, 362 (1986).
- Halton, J. H., "On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals," *Numer. Math.*, **2**, 84 (1986).
- Hammersley, J. M., "Monte Carlo Methods for Solving Multivariate Problems," *Ann. N.Y. Acad. Sci.*, **86**, 844 (1960).
- Hunter, J. S., "Statistical Design Applied to Product Design," *J. Qual. Technol.*, **17**, 215 (1985).
- Ierapetritou, M. G., and E. N. Pistikopoulos, "Batch Plant Design and Operations under Uncertainty," *Ind. Eng. Comput. Res.*, **35**, 772 (1996).
- Iman, R. J., and M. J. Shortencarier, A FORTRAN77 Program and User's Guide for Generation of Latin Hypercube and Random Samples for Use with Computer Models, NUREG/CR-3624, SAND83-2365, Sandia National Lab., Albuquerque, NM (1984).
- Iman, R. L., and W. J. Conover, "Small Sample Sensitivity Analysis Techniques for Computer Models, with an Application to Risk Assessment," *Commun. Stat.*, **A17**, 1749 (1982).
- Kacker, R. S., "Off-line Quality Control, Parameter Design, and the Taguchi Method," *J. Qual. Technol.*, **17**, 176 (1985).
- Knuth, D. E., *The Art of Computer Programming*, Vol. 1, *Fundamental Algorithms*, Addison-Wesley, Reading, MA (1973).
- McKay, M. D., W. J. Conover, and R. J. Bockman, "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code," *Technometrics*, **21**, 239 (1979).
- Mignon, D. J., S. J. Honkomp, and G. V. Reklaitis, "A Framework for Investigating Schedule Robustness under Uncertainty," *Comput. Chem. Eng. Suppl.*, Vol. 19, S357 (1995).
- Morgan, G., and M. Henrion, *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, New York (1990).
- Mullin, J. W., *Crystallization*, Butterworth-Heinemann, Oxford (1993).
- Nair, V. N., ed., "Taguchi's Parameter Design: A Panel Discussion," *Technometrics*, **34**(2), 127 (1992).
- Niederreiter, H., "Quasi-Monte Carlo Methods and Pseudo-Random Numbers," *Bull. Amer. Math. Soc.*, **84**, 957 (1978).
- Niederreiter, H., *Multidimensional Numerical Integration Using Pseudorandom Numbers*, *Stochastic Programming*, **84**, Part I, A. Prekopa and R. J.-B. Wets, eds., *Programming Study*, Vol. 27, North-Holland, Amsterdam, p. 17 (1988).
- Niederreiter, H., *Random Number Generation and Quasi-Monte Carlo Methods*, SIAM, Philadelphia, PA (1992).
- Papageorgiou, A., and G. W. Wasilkowski, "On Average Case Complexity of Multivariate Problems," *J. Complexity*, **6**, 1 (1990).
- Phadke, M. S., *Quality Engineering Using Robust Design*, Prentice Hall, Englewood Cliffs, NJ (1989).
- Reklaitis, G. V., "Process and Issues in Computer-Aided Batch Process Design," *Foundations of Computer-Aided Process Design*, J. J. Sirola, I. E. Grossmann, and G. Stephanopoulos, eds., CACHE, Elsevier, Amsterdam, p. 241 (1989).
- Sacks, J., S. B. Schiller, and W. J. Welch, "Designs for Computer Experiments," *Technometrics*, **31**, 41 (1989).
- Terwiesch, P., "Cautious On-Line Correction of Batch Process Operation," *AIChE J.*, **41**, 1337 (1995).
- Tsai, C. S., and C. T. Cheng, "Statistical Operating Strategies for Charging Batch Reactors," *AIChE J.*, **42**, 1304 (1996).
- Watzdorf, R., U. G. Naef, P. I. Barton, and C. C. Pantelides, "Deterministic and Stochastic Simulation of Batch/Semicontinuous Processes," *Comput. Chem. Eng. Suppl.*, Vol. 17, S343 (1993).
- Welch, W. J., R. J. Buck, J. Sacks, H. P. Wynn, H. J. Mitchell, and M. D. Morris, "Screening, Predicting, and Computer Experiments," *Technometrics*, **34**, 15 (1992).
- Welch, W. J., and J. Sacks, "A System for Quality Improvement via Computer Experiments," *Commun. Stat.—Theory Methods*, **20**, 477 (1991).
- Welch, W. J., T. K. Yu, S. M. Kang, and J. Sacks, "Computer Experiments for Quality Control by Parameter Design," *J. Qual. Technol.*, **22**, 15 (1990).
- Wozniakowski, H., "Average Case Complexity of Multivariate Integration," *Bull. Amer. Math. Soc.*, **24**(1), 185 (1991).
- Yu, T. K., S. M. Kang, J. Sacks, and W. J. Welch, "Parametric Yield Optimization of CMOS Analog Circuits by Quadratic Statistical Circuit Performance Models," *Int. J. Circuit Theory, Appl.*, Vol. 19, p. 579 (1991).

Manuscript received Apr. 20, 1995, and revision received July 18, 1996.