
Computational Optimization and Applications, 24, 335–371, 2003

A Novel Sampling Approach to Combinatorial Optimization Under Uncertainty

URMILA M. DIWEKAR* urmila@cmu.edu
CUSTOM (Center for Uncertain Systems: Tools for Optimization & Management),
Department of Civil & Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract. The generalized approach to stochastic optimization involves two computationally intensive recursive loops: (1) the outer optimization loop, (2) the inner sampling loop. Furthermore, inclusion of discrete decision variables adds to the complexity. The focus of the current endeavor is to reduce the computational intensity of the two recursive loops. The study achieves the goals through an improved understanding and description of the sampling phenomena based on the concepts of fractal geometry and incorporating the knowledge of the accuracy of the sampling (fractal model) in the stochastic optimization framework thereby, automating and improving the combinatorial optimization algorithm. The efficiency of the algorithm is presented in the context of a large scale real world problem, related to the nuclear waste at Hanford, involving discrete and continuous decision variables, and uncertainties. These new developments reduced the computational intensity for solving this problem from an estimated 20 days of CPU time on a dedicated Alpha workstation to 18 hours of CPU time on the same machine.

Keywords: stochastic annealing, HSS technique, efficient sampling, nuclear waste, stochastic optimization, combinatorial optimization

1. Introduction

Optimization can be described as a three step-decision-making process, where the first step is the knowledge of the system and hence is related to the modeling of the process. The second step involves finding a measure of system effectiveness. The third step is related to the theory of optimization, which involves application of a proper optimization algorithm to find the solution. Figure 1 shows the schematic of this three-step decision making process. The goal of the optimization problem shown in figure 1, is to determine the decision variables x that optimize some aspect of the deterministic model represented by the objective function Z, while ensuring that the model operates within established limits enforced by the equality constraints expressed in terms of function h and inequality constraints in terms of function g.

In general the models involved in optimization are deterministic in nature. However, many a time in realistic situation one cannot neglect the associated uncertainties and this results in a stochastic optimization or stochastic programming problem. One simple representation of a stochastic optimization problem is shown in figure 2. In the figure, the deterministic model in figure 1 is replaced by an iterative stochastic model with sampling loop. In the figure, u is the vector of uncertain parameters the $P_i(z(x, u))$ represents the probabilistic functional which for an expected value minimization reduces to Eq. (1), and where P represents the

*Currently at University of Illinois at Chicago, urmila@uic.edu

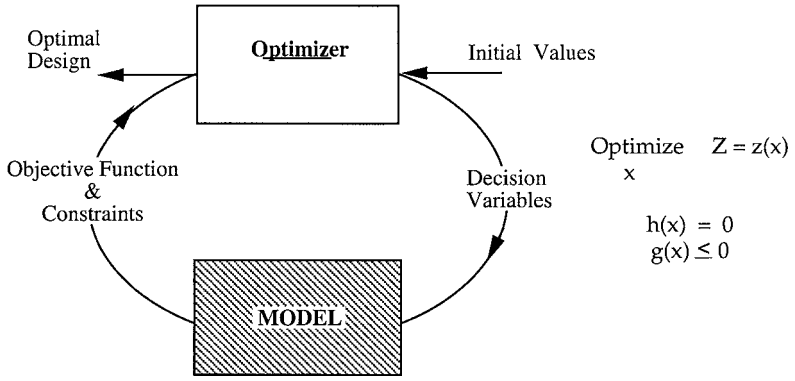

Figure 1. Pictorial representation of the optimization framework.

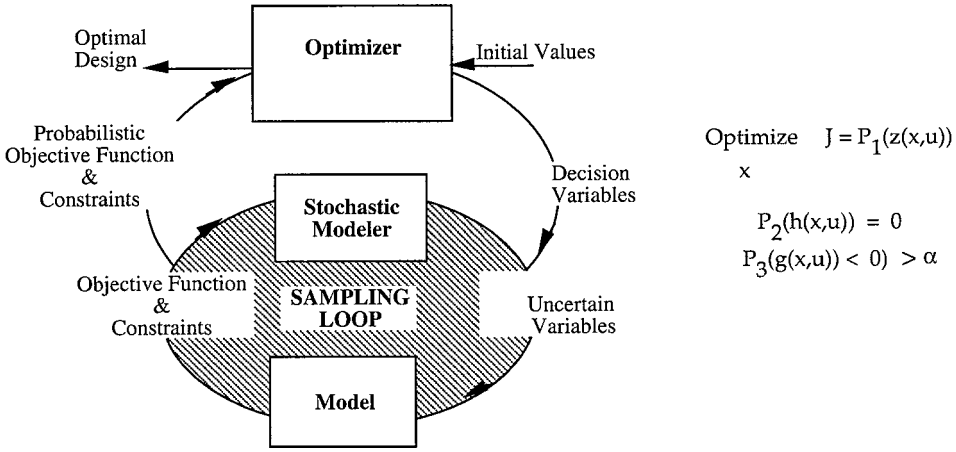

Figure 2. The stochastic optimization framework.

cumulative distribution function which can be calculated using a sampling technique.

$$\underset{x}{\text{Optimize}} \quad J = P_1(z(x,u)) = E(z(x,u)) = \sum_{k=1}^{N_{samp}} \frac{z(x,u_k)}{N_{samp}} \tag{1}$$

In recent years remarkable progress has been made in the development of algorithms for stochastic programming and stochastic optimization problems. However, most of these methods are based on exploiting a specific structure of the stochastic program and converting it into an equivalent deterministic problem. Moreover, some of these methods allow for analysis of random variations with specific distributions (e.g. Gaussian) and require quasi-concavity of the probability measures, or involve manipulation of probability distributions

(for more details, please see [35, 36, 42]). The strong need for generalized and efficient methods for solutions to stochastic programming/optimization problems [43] is the focus of the current endeavor.

The generalized stochastic optimization framework (figure 2) involves the two recursive loops; (1) the sampling loop, and (2) the optimization loop. The computational intensity of the sampling technique is the major bottleneck for these problems and poses one of the challenges in stochastic programming problems as pointed out by Wets [43]. This paper uses a new sampling technique [20] based on a minimum discrepancy Hammersley sequence to address this question.

The disadvantage of sampling approaches that solve the γ-th approximation completely is that some effort might be wasted on optimizing when approximation is not accurate [3]. For specific structures where the L-shaped method is applicable, two approaches avoid these problem by embedding sampling within another algorithm without complete optimization. These two approaches are the method of Danzing and Glynn [8] which uses importance sampling to reduce variance in each cut based on a large sample, and the Higle and Sen approach which utilizes a single stream to derive many cuts that eventually drop away as iteration numbers increase [13]. These methods require convexity conditions, dual-block angular structures, and are only applicable to continuous decision variables optimization. The central limit theorem is used to provide bounds for these methods. However, the bounds provided by the central limit theorem are too conservative for sampling methods other than the Monte Carlo technique. This paper presents a fractal dimension approach for defining the error bounds of the various sampling techniques. A new optimization algorithm based on simulated annealing which takes into account these error bounds to find the efficient interaction between the sampling and the optimization loop is described. This coupled approach provides a generalized and efficient approach to stochastic optimization in the presence of both continuous and discrete variables. Birge in the recent state-of-the-art review [4] on stochastic programming indicated the intractability of solving problems with discrete variables.

The paper illustrates these new developments in the context of a large scale, real world problem related to the environmental restoration of Hanford site. Buried beneath Hanford are 177 storage tanks, with capacities ranging from 50 thousand to one-million gallons, holding waste from nearly two-thirds of the weapons-grade nuclear fuel produced in the US following World War II. The short-term focus of the original fuel production did not consider the eventual disposal of its concomitant waste materials. The storage tanks, for instance, were originally built with an expected useful life of no more than a few decades; moreover, poor documentation was kept as chemicals were dumped into the tanks, removed for re-processing, evaporated to increase storage capacity, and juggled from one reservoir to the next [12]. The long-term consequences of this situation would not be quite as serious if the waste could be maintained underground indefinitely. Known and suspected leaks, the volatile nature of the reactions among the tank contents, and uncertainty about the future of the Hanford site, however, preclude this option. The current, through evolving, remediation strategy consists of a multistage process, beginning with characterization and preprocessing of the tank contents, separation of low- and high-level radioactive components, and conversion of both waste streams into glass “log” for permanent disposal in a national

repository [28]. The blending problem, presented in this paper focuses, applies to the glass formation process—known as "vitrification"—and seeks combinations of tank wastes that minimize the amount of glass produced. In order to convert these substances into glass, oxides collectively referred to as "frit" must be added to the waste as it melts. The objective is to maximize the amount of wastes per glass log, by keeping the amount of frit added to a minimum. Increasing the number of tanks (wastes), increases the combinatorial size of the problem, this coupled with uncertainties in waste composition and glass property models demand rigorous stochastic optimization procedures as outlined in this paper.

The paper is divided into six sections. Section 2 briefly describes the new sampling technique which enhances the computational efficiency of the sampling loop in the stochastic optimization procedure. Section 3 presents the fractal dimension approach as a measure for characterizing the accuracy and convergence of different sampling technique and investigates its applicability through a large test matrix. A new combinatorial optimization algorithm which uses this error characterization is described in Section 4. Section 5 presents a real world case study related to the nuclear waste problem at Hanford and illustrates the usefulness of these new algorithmic enhancements. Section 6 provides conclusions.

2. A novel sampling technique

A stochastic optimization problem involves integrals of any probabilistic function. Hence a sampling technique which provides a representative sample from the multi-variate probability distribution is crucial in obtaining true performance statistics for optimization. Here it is assumed that the uncertainty distributions are a-priori known. A common approach is to generate a sequence of points of size n on a k-dimensional unit hypercube, assuming a uniform distribution, U(0,1). The specific values for each input variable is selected by inverting the n samples over the cumulative distribution function. The convergence rate of the sampling technique depends in a critical way on how the sequence on the unit hypercube is chosen. In this section the commonly used sampling designs are discussed (on the unit hypercube) followed by the Hammersley sequence sampling technique [20] which provides a faster convergence rate as compared to commonly used techniques.

Perhaps one of the best known methods for sampling a probability distribution is the Monte Carlo Sampling (MCS) technique which is based on the use of a pseudo random number generator to approximate a uniform distribution, U(0, 1) with n samples on a k-dimensional unit hypercube. The specific values for each input variable are selected by inverting the N_{samp} samples over the cumulative distribution function. On average, the error ϵ of approximation is of the order $O(N_{samp}^{\frac{-1}{2}})$. The remarkable feature is that the bound is not dependent on the dimension k. However, one of the main disadvantages of the Monte Carlo method is that the bound is probabilistic, and there is no methodical way for constructing the sample points to achieve the probabilistic bound [27, 31].

Variance reduction techniques are statistical procedures designed to reduce the variance in the Monte Carlo estimate of the integrand. James [19] presented four categories of variance reduction techniques. Category 1 includes techniques like control variate technique that calls for model modification (reduction) and category 3 is specific to correlated uncertain variables. The most generalized and commonly used methods in these four categories are

stratified sampling techniques from category 2 and importance sampling techniques from category 3.

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 [21]. Once it is apparent that the uniformity properties are central to the design of sampling techniques, constrained or stratified sampling becomes appealing [25].

Stratified sampling techniques ensure that more samples are generated from high probability regions. On the other hand, importance sampling techniques guarantee full coverage of high consequence regions in the sample space, even if these regions are associated with low probabilities. This makes importance sampling techniques problem-dependent. Therefore for that reason importance sampling techniques are not considered further here.

Latin Hypercube Sampling (LHS, [17, 18, 24]) is one form of stratified sampling which can reduce the variance in the Monte Carlo estimate of the integrand significantly. The range of each input u_i is divided into non overlapping intervals of equal probability. One value from each interval is selected at random with respect to the probability density in the interval (In median Latin Hypercube (MLHS) this value is chosen as the mid-point of the interval. MLHS is similar to the descriptive sampling described by Saliby [38]). The n values thus obtained for u_1 are paired in a random manner with the n values of u_2 and these n pairs are combined with n values of u_3 and so on to form n k-tuplets. The random pairing is based on a pseudo random 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 (for k input variables). Moreover, this approach still only provides probabilistic bounds, however it reduces the variance as compared to the Monte Carlo approach.

The quasi-Monte Carlo methods seek to construct a sequence of points that perform significantly better than the average Monte Carlo bound. Some well known constructions for quasi-Monte Carlo sequences are the ones due to Halton, Hammersley, Sobol, Faure, Korobov, and Niederreiter [27]. The basic idea presented in my earlier work with Kalagnanam [20] is to replace a Monte Carlo integration by a quasi-Monte Carlo scheme in the stochastic optimization problem. We called this new technique—the Hammersley Sequence Sampling (HSS) technique. It uses the Hammersley points to uniformly sample a unit hypercube and inverts these points over the joint cumulative probability distribution to provide a sample set for the variables of interest. Appendix A summerizes the HSS designs. Although the original HSS technique designs start at the same initial point, it can be randomized by choosing the first prime number randomly.

The rate of convergence of Monte Carlo methods, based on central limit theorem, is $O(N_{samp}^{-1/2})$. Variance reduction methods such as importance sampling, control variates, or antithetic variates can substantially reduce the constant term but do not change the $O(N_{samp}^{-1/2})$ rate of convergence. Quasi-Monte Carlo methods like HSS provides faster convergence rate than $O(N_{samp}^{-1/2})$ [1]. This is because a sampling design based on Hammersley points has better uniformity properties as compared to MCS or LHS, since it uses optimal design scheme for placing N_{samp} points on a k-dimensional hypercube. A qualitative picture of the uniformity properties of the different sampling techniques on a unit square is presented in figure 3. It is

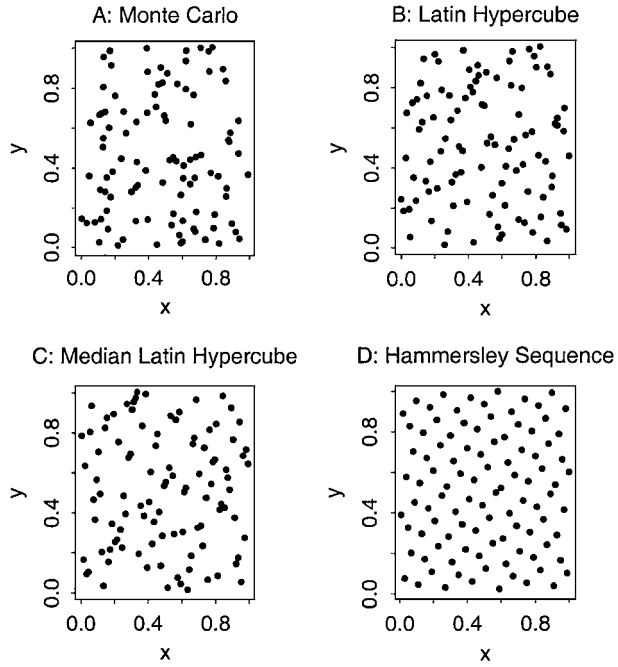

Figure 3. Sample points (100) placed on a unit square for different sampling techniques: (A) Monte Carlo, (B) Latin Hypercube Sampling, (C) Median-Latin Hypercube Sampling, and (D) Hammersley Sequence Sampling.

clearly observed that HSS shows greater uniformity than other stratified techniques such as LHS, which are uniform along a single dimension only and do not guarantee a homogeneous distribution of points over the multi-variate probability space. Similar behavior is observed for correlated samples also. The implementation of correlation structures is based on the use of rank correlations [16].

To compare the performance of the HSS technique with LHS and Monte Carlo (MCS), a large matrix of numerical tests were conducted, detailed results of which are available in [20]. 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 ($y = f(x_1, x_2, \ldots, x_n)$) and measuring the number of samples required to settle down to within an error of 1% of the "true" mean and variance of the derived distribution for y. These extensive numerical tests involved more than 12000 data sets, and problems ranging from small scale to large scale real world examples. The small scale problems involved functions of type 1 to 5 with 2 to 10 variables, various uncertainty distributions, and correlation structures. The large scale real world problem included in these data sets are: (1) a differential algebraic system of equations representing batch separation of industrial scale [9], (2) a quality control problem involving highly nonlinear system of chemical reactions [20], and (3) the nuclear waste problem presented in this paper. These tests were repeated using different random seeds to produce unbiased estimates. It was found that HSS is at least 3 to 100 times faster than LHS (and than MCS, as LHS is much faster than MCS). However, it should be noted

that for very high dimensional systems (e.g. the problem involving 360 uncertain variables in [1]), the quasi Monte carlo methods like HSS exhibit poor uniformity properties. Modified versions of these low discrepancy sequence produced by removing cycles can be used for higher dimensions. These methods are currently under investigation [22, 41].

The success of the outer loop in the figure 2 depends upon the accuracy of the stochastic model which is related to the number of samples used for generating the stochastic modeling information. Theoretically a true estimate of a probabilistic function requires infinite samples. However, in practice one can only take a finite number of samples. For any sampling technique, the number of samples required to obtain a given accuracy not only depends on types of uncertainty distributions, but also on the type of sampling technique. Further, for optimization problems, the number of samples required also depends on the location of the trial point solution in the optimization space [30]. Therefore, the selection of the number of samples for the stochastic optimization procedure is a crucial and challenging problem and depends on the error characterization of the sampling phenomena. The following section describes a novel approach to error characterization using a fractal dimension approach.

3. Characterizing accuracy of different sampling techniques

This section presents the error band-width associated with different sampling techniques, qualitatively. One strategy for determining the convergence properties of different sampling techniques is to ascertain the sampling accuracy, i.e., the number of samples required by any sampling scheme to converge for any probability function, e.g., the mean or the variance. Classical statistical methods provide good estimates for the bounds (and confidence interval) for truly random Monte Carlo sampling and are not applicable to other sampling techniques. For example, from the central limit theorem [25], it can be proved that the error-bandwidth (ϵ_μ) for the mean μ_y of the parameter $y = f(x_1, x_2 \ldots)$ calculated using N_{samp} samples is proportional to the ratio of standard deviation σ_y to $\sqrt{N_{samp}}$. However, this error-bandwidth expression presented above is only applicable to Monte Carlo simulations and not for other sampling technique. Furthermore, theoretical discussions such as the above are not possible for other techniques, as a result a systematic quantification of accuracy is lacking for non-Monte Carlo techniques for the various performance indices like the mean, the variance, or the fractiles. Figures 4 and 5 show the upper and lower bounds based on the true mean and the true standard deviation (both estimated at 10,000 samples), characterizing the error band-width of the estimated mean of a function of two uncertain parameters ($f(x_1, x_2) = x_1^2 + x_2; x_1$ is uniform with $U(0.1, 1.0)$, x_2 is uniform with $U(0.1, 1.0)$) for Monte Carlo and Hammersley sequence sampling at 95% confidence levels. It is clearly evident that classical methods to characterize the error band-width for any confidence level for the HSS technique overestimate the bounds. A similar trend is also observed for the bounds on the variance. Based on the assumption that the true probability distribution is normal, chi-square estimates for the error band-width for 95% confidence level provide true bounds for Monte Carlo sampling and overestimated bounds for Hammersley Sequence sampling. This again only provides an upper bound to error width and is not specific to any particular performance index. These figures show conclusively that there is a need to characterize precise estimates of the error band-widths for the improved sampling techniques, since it

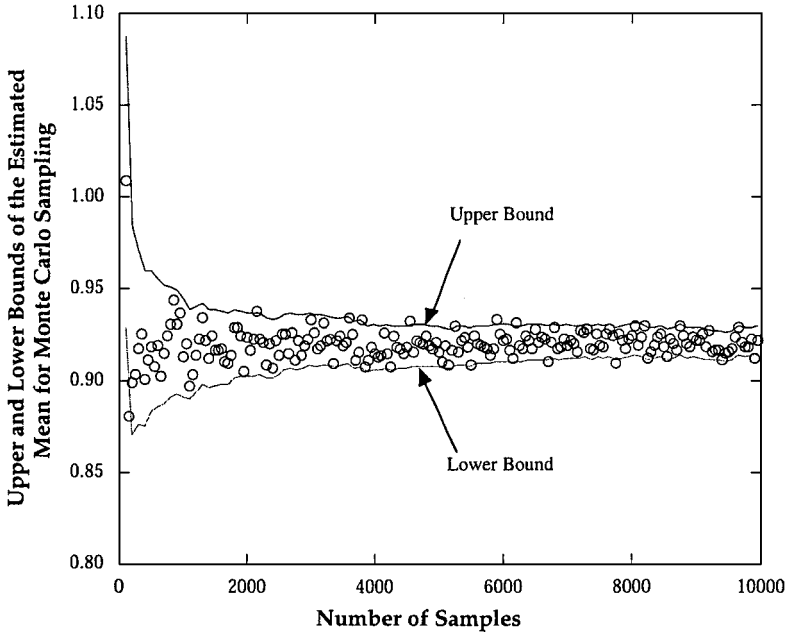

Figure 4. Upper and lower bounds of the estimated mean (95% confidence interval) for the Monte Carlo sampling based on classical statistics.

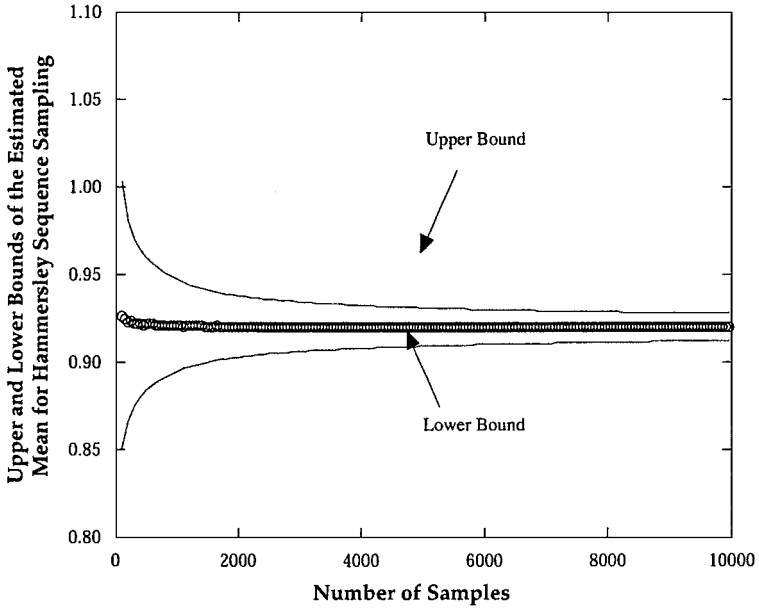

Figure 5. Upper and lower bounds of the estimated mean (95% confidence interval) for the Hammersley Sequence Sampling based on classical statistics.

has important implications on the sampling accuracy in stochastic modeling, optimization, and uncertainty analysis. Although theoretically it has been shown that for quasi-Monte Carlo techniques like the Hammersley Sequence Sampling, the number of sample points required to contain the approximation error within ϵ is proportional to $(\frac{1}{\epsilon})(\log(\frac{1}{\epsilon}))^{\frac{(k-1)}{2}}$ (for a k dimensional problem) [44], this again provides an upper bound to error-bandwidth that is not specific for the various performance indices like the mean, or the variance. The next section discusses how a self-affinity and scaling properties of fractal object can be used to characterize the error band-width of various performance indices calculated using the Hammersley Sequence Sampling.

3.1. A fractal dimension approach

Euclidean geometry describes regular objects like points, curves, surfaces, cubes using integer dimensions 0, 1, 2, and 3 respectively. However, many of these objects in nature, such as the surface of mountains, coastlines, microstructure of metals etc., are not Euclidean objects. Such objects are called fractals, and are described by a non-integral dimension called the *fractal dimension*.

For example, if one were to measure the length of a coastline, the length would depend on the size of measuring stick used. Decreasing the length of the measuring stick leads to a better resolution of the convolution of the coastline, and as the length of the stick approaches zero, the coastline's length tends to infinity. This is the fractal nature of the coastline. Since a coastline is so convoluted, it tends to fill space, and its fractal dimension lies somewhere between that of a line (which has a value 1) and a plane (which has a value of 2) [23].

Any measure of a fractal object such as the length, area or volume, is governed by a "scaling" relationship of the form:

$$M(L) \sim (L)^{d_f} \tag{2}$$

where '$\sim$' should be interpreted as 'scales as', and d_f is the fractal dimension of the object. Just like length, area, and volume have different Euclidean dimensions, fractal dimension of different measures of the fractal object are different.

The striking feature of fractals is their property of self-similarity i.e., subsets of a fractal object have primarily the same form as the whole object (like the convolutions of the coastline at different magnification). This concept of self-similarity is embedded in Eq. (2) which implies that the d_f calculated from the relationships in Eq. (2) is constant over a range of scale lengths. Mathematical or theoretical fractal objects are infinitesimally sub-divisible in this way, since each subset however 'small' contains no less detail than the complete set [7]. In other words fractals, both deterministic and random, are scale invariant i.e., if we take a part of the fractal and magnify it by the same factor in all directions, the magnified picture cannot be distinguished from the original one and provides no idea of the scale of magnification.

Unlike exact fractals like the Koch curve, Dürer Pentagon, Cantor set, *global* self-similarity is seldom observed and the fractal description is based on *statistical self-similarity* where the objects exhibits a fractal nature over a certain *local* range of the scale (L). This

concept of local, statistical self-similarity has been used in a wide variety of applications which range from the description of composite materials [34] to waveforms and signals [33].

Many fractals are made up of parts which are, in some way, similar to the whole. For example, the middle third Cantor set is the union of two similar copies of itself, and the Von Koch curve is made up of four similar copies. Furthermore, there are fractals which are defined by transformation. For example, self-affine sets form an important class of sets, which includes self-similar sets as a particular case. An affine transformation is a combination of translation, rotation, dilation, and perhaps, a reflection. Unlike similarities, affine transformations contract with different ratios in different directions [11]. It is possible to establish a relationship between the dimension of the self-affine sets and the fractal dimension. An example of the self-affine fractals is the one dimensional Brownian motion. Figure 6 shows the Brownian motion scaled with a factor 2 in the horizontal direction but with a vertical scaling factor of $2^{0.5}$. It can be easily seen that the curves look about the same. In fact, they are the same, statistically speaking. An analysis of mean, variance, moments and so forth would give the same statistical properties of the re-scaled curves. In general, a fractional Brownian motion exhibits statistical scale invariance for the random functions $X(t)$ and $\frac{X(rt)}{r^H}$, where x is the displacement (y axis) and t represents the time (x axis), and r is the scaling factor ($r = 2$ here). This exponent H is sometime called the *Hurst exponent*, after Hurst, a hydrologist who did some early work with Mandelbrot, on scaling properties of river fluctuations. There is a direct relationship between this exponent H and the fractal dimension for the fractional Brownian motion curves. For details, please refer to [32].

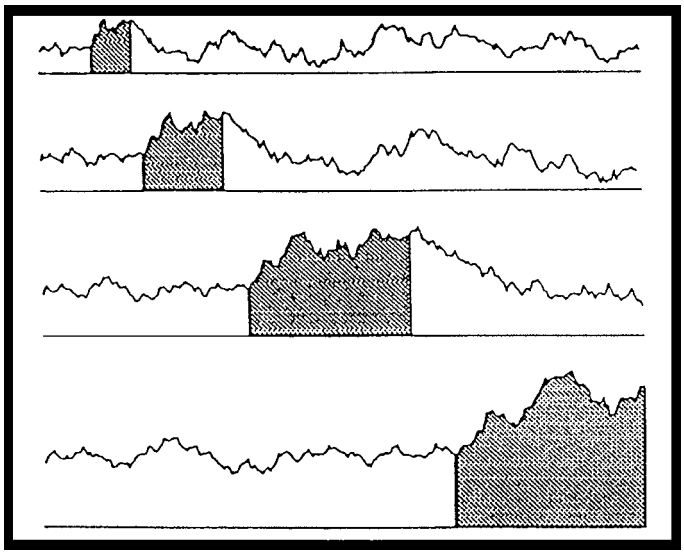

Figure 6. Scaling the fractal Brownian motion with horizontal scaling factor of 2 and the proper vertical scaling factor $2^{0.5}$. The curves are statistically equivalent, revealing the scaling law for Brownian motion. The shaded region shows the same shape properly re-scaled for the different stages (reproduced from [32]).

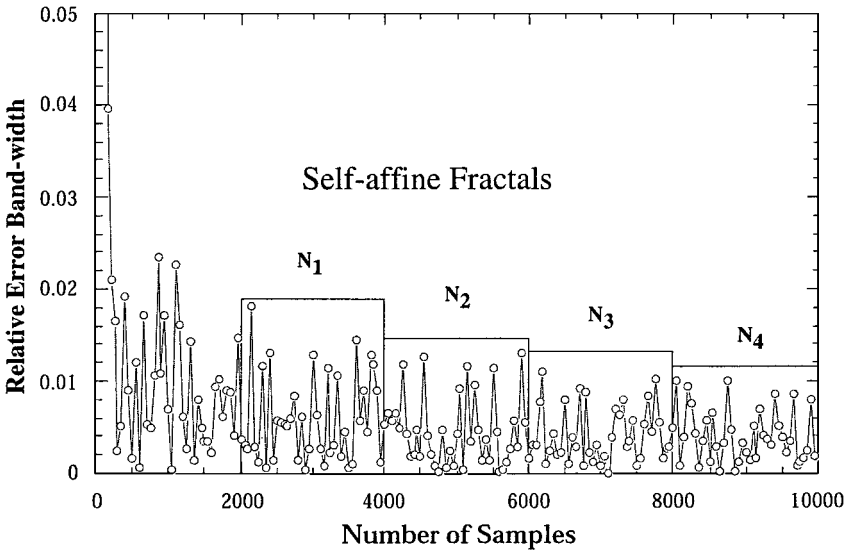

Figure 7. The relative error band-width shows self-affine properties and scales with the sample size.

3.2. Fractional Brownian motion and the error-bandwidth

What do fractional Brownian motion and the error-bandwidth of a probability function have in common? Figure 7 shows the relative error-bandwidth of a probability function around different sample sizes N_1, N_2, and so on. The characteristic probability function shown here is the mean for a function ($y = x_1^2 + x_2$) of two uncertain parameters x_1 (Uniform (0.1, 1.0)) and x_2 (Uniform (0.1, 1.0)). One can visualize a box whose height (representing the relative error) scales in a statistical self-similar fashion with the sample size. In other words, one can easily observe from a comparison of the number and nature of peaks in two adjacent boxes, that the adjacent boxes exhibit the same "irregular" scaling behavior of the error band-width with the sample size. The concept of statistical scale invariance is observable (similar to the characteristics of the fractional Brownian motion), where each of these boxes are similar (or self-affine) in an average sense and one can easily find a dimension H' (equivalent to the Hurst exponent in the fractional Brownian motion) such that the average error band-width around N_1, $\epsilon_{av}(N_1)$ and average error-bandwidth around N_n, $\epsilon_{av}(N_n)$ for $n = 1, 2, \ldots$ show the self-similarity property. This scale invariance property when transformed results in the following expression.

$$\epsilon_{av}(N_1) = \frac{\epsilon_{av}(N_n)}{(n)^{H'}} = \frac{\epsilon_{av}(N_n)}{(N_n/N_1)^{H'}} \tag{3}$$

In light of the discussion above, a scaling relationship is proposed between the error band-width and the sample size as:

$$\epsilon \sim N_{samp}^{H'} \tag{4}$$

where '$\sim$' is interpreted as 'scales as,' and H' is the exponent which is characteristic of different sampling techniques. Pickover and Khorosani [33] have also pointed out that a good exercise to test for self-similarity is to plot the scaling relationship in a log-log plot. If the phenomena has self-similar properties, then the line will be remarkably straight, suggesting probably a fractal nature.

3.3. Error characterization of Hammersley Sequence Sampling

This section illustrates the methodology to characterize the error band-width for the new sampling technique, Hammersley Sequence Sampling. The same approach was also applied to the Monte Carlo technique to compare the relationship of the error band-width with the sample size obtained using the proposed methodology and classical statistics.

The error band-width ϵ is defined, in this case, as the relative absolute error expressed as a percentage:

$$\epsilon = \frac{(x_i - x_{conv})}{x_{conv}} \times 100 \tag{5}$$

where x_i is any estimate of the probability function (e.g. mean, fractile, or variance), and x_{conv} is the estimate of the probability function after convergence at higher samples and reflects the actual value of the probability measure.

A large matrix of tests including the type of function, type of distribution, and the number of uncertain parameters was designed to study the fractal nature of the error band-width. The matrix structure involved the following:

- *Functions*: Five different types of functions were used as outlined below:
  - Type 1: Linear additive: $y = \sum_{i=1}^{nv} x_i \;\; nv = 2 \ldots 10$
  - Type 2: Multiplicative: $y = \pi_{i=1}^{nv} x_i \;\; nv = 2 \ldots 10$
  - Type 3: Quadratic: $y = \sum_{i=1}^{nv} x_i^2 \;\; nv = 2 \ldots 10$
  - Type 4: Exponential: $y = \sum_{i=1}^{nv} x_i \exp(x_i) \;\; nv = 2 \ldots 10$
  - Type 5: Logarithmic: $y = \sum_{i=1}^{nv} \log(x_i) \;\; nv = 2 \ldots 10$
- *Distributions*: Three types of distributions were selected to characterize the uncertainties in the input variables x_i. The first one is skewed (lognormal), while the last two are symmetric (normal and uniform).

The validity of the proposed relation for the error band-width and the sample size (Eq. (4)) can be shown through a log-log plot of the error band-width (characteristic of each of the self-affine adjacent boxes) and the sample size. If Eq. (4) holds true, then the log-log plot will yield a straight line. Tables 1–3 present the results for the error band-width for both the mean and variance with respect to different functions, and different distributions for the Monte Carlo and the Hammersley sequence sampling.

For the function ($f(x_1, x_2) = x_1^2 + x_2$) mentioned previously, log-log plots of the error band-width are plotted against the sample size for the mean (figure 8) for the Monte Carlo and

Table 1. Error characterization for the mean and variance for Monte Carlo sampling.

	Uncertainties					
	x_1 (log), x_2 (log)		x_1 (norm), x_2 (norm)		x_1 (unif), x_2 (unif)	
Function	H'	R^2	H'	R^2	H'	R^2
			Mean			
Type 1	−0.57	0.90	−0.63	0.93	−0.64	0.95
Type 2	−0.55	0.97	−0.51	0.99	−0.47	0.99
Type 3	−0.53	0.98	−0.53	0.95	−0.53	0.96
Type 4	−0.54	0.99	−0.54	0.99	−0.54	0.99
Type 5	−0.51	0.99	−0.60	0.99	−0.62	0.99
			Variance			
Type 1	−0.75	0.99	−0.75	0.99	−0.65	0.99
Type 2	−0.68	0.97	−0.68	0.97	−0.62	0.99
Type 3	−0.56	0.98	−0.56	0.98	−0.44	0.99
Type 4	−0.51	0.96	−0.51	0.96	−0.54	0.99
Type 5	−0.64	0.99	−0.64	0.99	−0.74	0.99

H': Exponent for defining error-band of a sampling technique.
R^2: Coefficient of determination.

Table 2. Error characterization for the mean and variance for Hammersley sequence sampling.

	Uncertainties					
	x_1 (log), x_2 (log)		x_1 (norm), x_2 (norm)		x_1 (unif), x_2 (unif)	
Function	H'	R^2	H'	R^2	H'	R^2
			Mean			
Type 1	−2.07	0.97	−2.12	0.96	−1.99	0.97
Type 2	−2.17	0.89	−2.26	0.89	−1.97	0.91
Type 3	−2.12	0.97	−2.15	0.96	−2.02	0.97
Type 4	−2.28	0.93	−2.35	0.90	−2.12	0.93
Type 5	−2.47	0.91	−1.76	0.85	−1.89	0.91
			Variance			
Type 1	−1.85	0.93	−1.28	0.95	−1.26	0.97
Type 2	−1.34	0.88	−1.63	0.91	−1.45	0.93
Type 3	−2.09	0.95	−1.19	0.99	−1.21	0.99
Type 4	−1.88	0.96	−1.24	0.99	−1.47	0.93
Type 5	−2.95	0.99	−1.64	0.90	−1.57	0.93

Table 3. Error characterization for the mean and variance for Hammersley sequence sampling for a range of uncertain parameters and for two different functional forms.

Function type 1			Function type 4		
No. of uncertain pars.	H'	R^2	No. of uncertain pars.	H'	R^2
		Mean			
2	−1.99	0.97	2	−2.12	0.93
3	−2.05	0.99	3	−2.10	0.95
4	−1.91	0.99	4	−1.89	0.98
5	−1.71	0.99	5	−1.79	0.99
6	−1.63	0.98	6	−1.58	0.99
7	−1.81	0.98	7	−1.62	0.99
8	−1.79	0.99	8	−1.77	0.98
9	−1.91	0.98	9	−1.87	0.98
10	−1.86	0.97	10	−1.78	0.98
		Variance			
2	−1.26	0.97	2	−1.47	0.93
3	−1.16	0.99	3	−1.29	0.96
4	−1.48	0.99	4	−1.32	0.96
5	−1.76	0.99	5	−1.32	0.99
6	−1.61	0.99	6	−1.32	0.99
7	−1.44	0.99	7	−1.39	0.99
8	−1.33	0.99	8	−1.19	0.99
9	−1.37	0.98	9	−1.01	0.97
10	−1.62	0.99	10	−1.30	0.99

the Hammersley sequence sampling based on two uncertain parameters, x_1 (x_1 is uniform, U(0.01, 1.0))) and x_2 (x_2 is uniform, U(0.01, 1.0)). The scaling relationship is obvious from the figure, which exhibit good coefficients of determination (R^2 values), suggesting the fractal nature. It is worth mentioning that a simple curve fit of the error band-width vs. the number of samples did not yield good R^2 values, proving conclusively that the relationship between the error band-width and the sample size is not a simple power-law relationship. Both self-affinity and the scaling properties must be taken into account to obtain good linear fits suggesting the fractal nature. The exponent H' is obtained from the slope of the linear relationship, which is different for the mean and variance and for Monte Carlo and Hammersley sequence sampling techniques. It is realized however, that this model may not be applicable for extremely low sample sizes (i.e., N_{samp} tending to 1). The higher dimensions obtained for the mean and the variance for the Hammersley sequence sampling is consistent with the fact that the mean and the variance converge at very few samples for the Hammersley sequence sampling as compared to the Monte Carlo sampling.

It is seen that the characterization of the error band-width for the mean based on the self-affine fractal characterization in the case of Monte Carlo technique is almost identical to

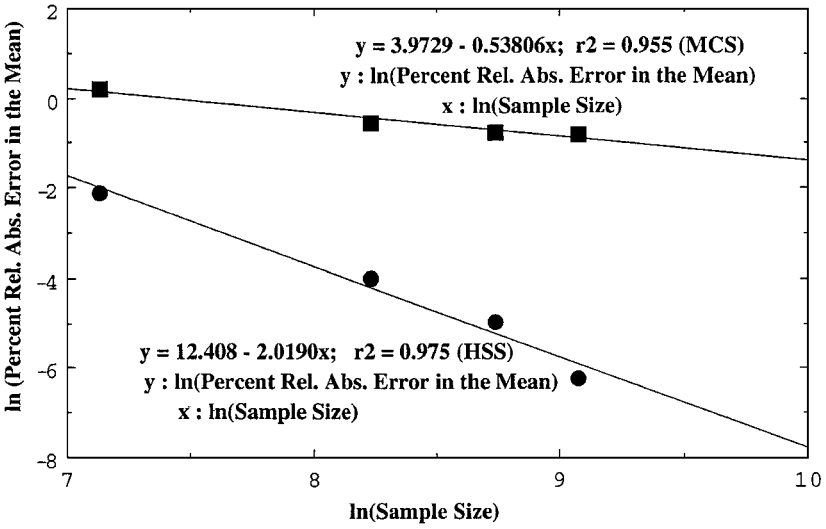

Figure 8. Characterization of the error band-width associated with the mean for the Monte Carlo and the Hammersley sequence sampling.

the relationship predicted from classical statistics. It may be recalled that classical statistics predict that the error band-width for Monte Carlo techniques should scale as $N_{samp}^{-0.5}$, where N_{samp} is the sample size. The experimentally determined exponent for the relationship between the error band-width and the number of samples is almost -0.5, with good linear fits, indicating that the proposed model characterizes the error band-width for the Monte Carlo sampling technique fairly well.

It is also observed, based on experiments with several uncertain parameters for the Hammersley sequence sampling and for the Monte Carlo sampling that the exponent H' is independent of the number of uncertain parameters, or the output relationship, or the type of uncertainty distributions. This is important, since it shows that this method of characterizing the error band-width is robust, and quantifies the behavior of the error band-width with the sample size globally. It was observed that the Latin Hypercube sampling technique did not consistently predict the same exponent when the number of variables, functional relationship, or distributions are changed. It should be noted that the results of the fractal dimension model presented here are based on empirical studies. Theoretical investigation of this phenomena is the focus of ongoing work.

4. Stochastic annealing: A combinatorial optimization algorithm for handling uncertainties

> For broad applicability, stochastic programming methods must include discrete variables in a more comprehensive manner. –Birge, 1997 [4].

Stochastic annealing is an algorithm recently proposed for handling discrete optimization problems under uncertainty. In this section, the efficiency of this novel algorithm will be

enhanced further. The algorithm is derived from the simulated annealing algorithm a brief description of which is presented below.

Simulated annealing. Simulated Annealing (SA) is a probabilistic method for discrete optimization based on ideas from statistical mechanics [40]. The analogy in simulated annealing is to the behavior of physical systems in the presence of a heat bath: in physical annealing, all atomic particles arrange themselves in a lattice formation that minimizes the amount of energy in the substance, provided the initial temperature (T_{init}) is sufficiently high and the cooling is carried out slowly. At each temperature T, the system is allowed to reach thermal equilibrium, which is characterized by the probability of being in a state with energy E given by the Boltzmann distribution as a function of E and T. In simulated annealing, the objective function (usually cost) becomes the energy of the system. The goal is to minimize the cost/energy. Simulating the behavior of the system then becomes a question of generating a random perturbation that displaces a "particle" (moving the system to another configuration). If the configuration S representing the set of the decision variables that results from the move has a lower energy state, the move is accepted. However, if the move is to a higher energy state, the move is accepted according to the Metropolis criteria (accepted with probability $= \frac{1}{Z_t} e^{\frac{-E}{K_b T}}$) [40]. This implies that at high temperatures, a large percentage of uphill moves are accepted. However, as the temperature gets colder, a small percentage of uphill moves are accepted. After the system has evolved to thermal equilibrium at a given temperature, the temperature is lowered and the annealing process continues until the system reaches a temperature corresponding to a certain "freezing" temperature defined a priori. Thus simulated annealing combines both iterative improvement in local areas and random jumping to help ensure the system does not get stuck in a local optimum. Although, the original SA algorithm was not designed to handle constraints or infeasibilities, there are various ways of dealing with this problem. For example, one can use explicit penalties for constraint violation [29], or use infeasible path optimization, or use a coupled Simulated Annealing-NonLinear Programming (SA-NLP) approach where the problem is divided into two levels, similar to the Mixed Integer NonLinear Programming (MINLP) algorithms [10]. The outer level is SA which decides the discrete variables. The inner level is NLP for continuous variables and can be used to obtain feasible solution to the outer SA.

It must be realized that simulated annealing is not a panacea for all combinatorial optimization problems. The convergence of the annealing algorithm is guaranteed asymptotically, if the move sequences are Markov chains. Computationally, this does not provide a practical guarantee of convergence in a real problem. Therefore, for the real world problem presented in this work, different algorithms are used to guarantee convergence of the framework.

As SA is a probabilistic method, several random probability functions are involved in this procedure. It is known that the efficiency of the annealing algorithm is little affected by the use of different acceptance probability distributions [40]. However, G_{ij}, probability for generating configuration j from i at each temperature can significantly affect the overall efficiency of annealing process. Generally G_{ij} is a random probability given by the uniform distribution within the neighborhood. Thus recent research efforts for SA improvement have

been focused on modifying or changing G_{ij}. These new simulated annealing algorithms differ mainly in the choice of G_{ij} and the cooling schedule [37]. Among the proposed simulated annealing variants, Fast Simulated Annealing (FSA) [39] and Hybrid Simulated Annealing (HSA) [37] are worth mentioning. G_{ij} in FSA has a Gaussian-like peak and Lorentzian long tails which can make an occasional long jump from the current configuration to increase the speed of annealing. But this G_{ij} cannot guarantee uniform coverage of the moves over the configuration surface. HSA applies the Hybrid Monte Carlo method to obtain G_{ij} in which the design variable $\mathbf{x}$ and a Gaussian-like auxiliary momenta $\mathbf{p}$ are mapped using the Hamilton's equations of motion. The acceptance probability is similar to the Metropolis criterion, but the energy difference ΔH is replaced by the Hamiltonian function difference $\Delta H(\mathbf{x}, \mathbf{p})$. Although this algorithm is found to be very fast, HSA requires an evaluation of derivative of the objective function, $-\frac{\partial f(\mathbf{x})}{\partial x_i}$ for mapping and hence this destroys one of the advantages of simulated annealing that SA does not require derivative information.

Although SA can be viewed as an ordinary hill-descent method with artificial noise added, most of these algorithms are used for deterministic objective function. Recent variants of SA consider noise in the objective function that resembles expected value minimization problems in stochastic programming [2]. These variants generate appropriate transition probabilities, G_{ij}, based on the probability distributions of uncertain variables, so as to reach global optimum. However, the convergence of these algorithms requires large number of samples. Furthermore, the stochastic problems that can be solved using these variants are restricted to expected value minimization formulation.

Stochastic annealing. The stochastic annealing algorithm, presented in the earlier work [6, 30] is an algorithm designed to efficiently optimize any probabilistic objective function. In the stochastic annealing algorithm, the optimizer (figure 2) not only obtains the decision variables but also the number of samples required for the stochastic model. Furthermore, it provides the trade-off between accuracy and efficiency by selecting an increased number of samples as one approaches the optimum.

A substitute objective function in stochastic annealing therefore, consists of a probabilistic objective value P_1 and the penalty function, which is represented as follows:

$$\min \quad Z(cost) = P_1(x, u) + b(t)\epsilon_p \tag{6}$$

In the above equations, the first term represents the real objective function which is a probabilistic function in terms of the decision variables x and uncertain variables u and all other terms following the first term signifies the penalty function consisting of the weighting function $b(t)$ and the error band-width ϵ_p (as a function of number samples). The stochastic annealing algorithm minimizes the CPU time by balancing the trade-off between computational efficiency and solution accuracy by the introduction of a penalty function in the objective function. This is necessary, since at high temperature the algorithm is mainly exploring the solution space and does not require precise estimates of any probabilistic function. The algorithm must select a greater number of samples as the solution approaches. The weight of the penalty term, as mentioned before, is governed by $b(t)$, and is based on the annealing temperature. Thus, $b(t)$ increases as the temperature decreases. The choice of

a penalty term, on the other hand, must depend on the error-bandwidth ϵ_p of the function that is optimized, and must incorporate the effect of the number of samples. Therefore, the success of this algorithm depends on the accurate characterization of the error-bandwidth.

4.1. The new variant of stochastic annealing

Based on the results presented in the earlier tables, the error band-width for the mean in the case of Hammersley sequence sampling scales with N_{samp}, the sample size, as $N_{samp}^{-1.8}$. Consequently, the penalty term in the stochastic annealing algorithm for the mean in the case of Hammersley sequence sampling is defined as:

$$b(t)\epsilon_p = \left(\frac{b_0}{k^t}\right)_{HSS} N_{samp}^{-1.8}$$

where, b_0 and k are empirical constants, t is the corresponding temperature level, which increases as the annealing proceeds, and N_{samp} is the number of samples. Usually b_0 in the weighting term is small (e.g. 0.001), and k lies between 0.8 and 0.95. These constants are determined a-priori so that the penalty is less than 5% of the real objective function. It may be recalled that, the penalty term in the stochastic annealing algorithm for the mean in the case of Monte Carlo sampling is given by:

$$b(t)\epsilon_p = \left(\frac{b_0}{k^t}\right)_{MCS} N_{samp}^{-0.5}$$

The main steps in the stochastic annealing algorithm (for expected value minimization) are given below.

1. Initialize variables: $T_{initial}$, T_{freeze}, accept and reject limits, initial configuration S.
2. If ($T > T_{freeze}$) then
   (a) Perform the following loop (an optimization step) (i = (i)..(viii)) N (number of moves at a given temperature) times.
      i. Generate a move S' from the current configuration S as follows:
         A. Select the number of samples, N_{samp} by a random move.
            if $rand(0, 1) \leq 0.5$ then

            $$N_{samp} = N_{samp} + 5 \times rand(0, 1)$$

            else

            $$N_{samp} = N_{samp} - 5 \times rand(0, 1)$$

         B. Select the decision variables (zero-one integer, discrete variables). The decision variables are selected so that they follow the Markov process similar to the number of samples selection above.

ii. Generate N_{samp} samples of the uncertain parameters.
iii. Perform the following loop (iii(A)..iii(B)) N_{samp} times.
   A. Run the model.
   B. Calculate the objective function cost(S').
iv. Evaluate the expected value $E(cost(S'))$ of the cost function.
v. Generate the weighting function $b(t) = \frac{b_o}{k'}$
vi. Calculate the modified objective function given below.
   For Monte Carlo Sampling:

$$Obj(S') = E(Cost(S')) + b(t)\frac{1}{N_{samp}^{-0.5}}$$

   For Hammersley Sequence Sampling:

$$Obj(S') = E(Cost(S')) + b(t)\frac{1}{N_{samp}^{-1.8}}$$

vii. let $\Delta = Obj(S') - Obj(S)$
viii. If $\Delta \leq 0$ then accept the move. Set $S = S'$ else if ($\Delta \geq 0$) then accept with a probability $\exp(\frac{-\Delta}{T})$.

(b) Return to 2(a)

3. If $T > T_{freeze}$, set $T = \alpha T$ and return to 2(a)
4. Stop

The salient features of this algorithm are demonstrated using a simple example in Appendix B. The next section illustrates the computational efficiency of the stochastic annealing algorithm implemented with the Hammersley sequence sampling, for which the error band-width was characterized based on the self-affinity and scaling properties mentioned previously.

5. Synthesizing optimal waste blends under uncertainty

The Hanford site in southeastern Washington produced nuclear materials using various processes for nearly fifty years. The Hanford site has 177 tanks (capacity ranging from 50,000 to 1 million gallons) containing radioactive waste. At present, it is desired that the high-level liquid waste be converted to borosilicate glass for storage in a geologic repository since radioactivity does not leak easily through glass. This process, termed as vitrification, requires certain conditions related to 'processibility' and 'durability' be satisfied, so that the conversion is achievable. The processibility conditions ensure that during the processing stage, the glass melt has properties such as viscosity, electrical conductivity, and liquidus temperature, that lie within ranges known to be acceptable for the vitrification process. The durability considerations ensure that the resultant glass meets the quantitative criteria for storage in a repository.

Selective blending of the tank wastes prior to vitrification reduces the amount of glass produced only because the tank contents differ [26]. A non-homogeneous mixture of gases, liquids, slurries, and solids, the tank contents include salts, complex organic compounds, various metals, low- and high-level nuclear waste, and water [12]. In order to convert these substances into glass, oxides collectively referred to as "frit" (e.g., SiO_2, B_2O_3, Na_2O, Li_2O, CaO, and MgO) must be added to the waste as it melts. Blending takes advantage of the fact that the frit constituents are present to varying degrees in each tank; a selective combination of wastes therefore reduces the need to add frit during vitrification by matching tanks with complementary frit requirements [26]. In addition, blending decreases the proportion of so-called "limiting" components in the combined waste streams. The presence of these components adversely affects the vitrification process; sodium, for instance, reduces glass durability, while aluminum increases both the melt temperature of the material and its viscosity. Hence, selective blending of the tank contents increases the probability that vitrification will succeed, reduces frit requirements (i.e., achieves a greater waste-to-frit mass ratio), and minimizes the volume of glass produced. Blending involves a two-stage decision process: assignment of individual tanks to a given blend and determination of frit requirements. The latter decision depends on the contents of each tank, and is governed by both analytical and empirical glass property models derived for the Hanford tank wastes. The resulting constraints pertain to the vitrification process—rather than characteristics of the subsequent glass—and include: bounds on the waste component mass fractions; crystallinity requirements; solubility limits; and attributes of the molten glass, including its liquidus temperature, viscosity and electrical conductivity. Appendix C presents the glass property models in their guise as deterministic constraints.

This nested structure is typical of problems in combinatorial optimization, where a number of discrete decisions (e.g., the tank-blend assignments) must be made prior to the optimization of some function of continuous characteristics (e.g., the frit mass of each blend). This problem of determining the optimal waste blend configuration was studied previously by [26] based on data pertaining to a subset of the total number of tanks at the Hanford site, using various techniques including a coupled simulated annealing and nonlinear programming technique. The study assumed that there are no uncertainties related to the waste composition in the tanks, and in the glass physical property models.

As shown in the present study, uncertainties play a critical role in the optimal configuration of waste blends. The uncertainties are associated with (i) the range of applicability of the glass property models, and (ii) waste composition in the tanks. The waste composition uncertainties primarily arise due to the variability in the composition of the wastes, non-homogeneity of the tank contents, and poor documentation of the contents of each tank.

This leads to a challenging problem of determining the optimal waste blend configuration subject to the inherent uncertainties in the waste compositions and in the glass physical property models. The formulation is based on the stochastic optimization framework, and addresses the question as to how uncertainties affect the calculated waste loading and the synthesis of optimal waste blends. It was observed that this exercise of synthesizing optimal waste blends under uncertainty is a computational intensive problem, requiring several

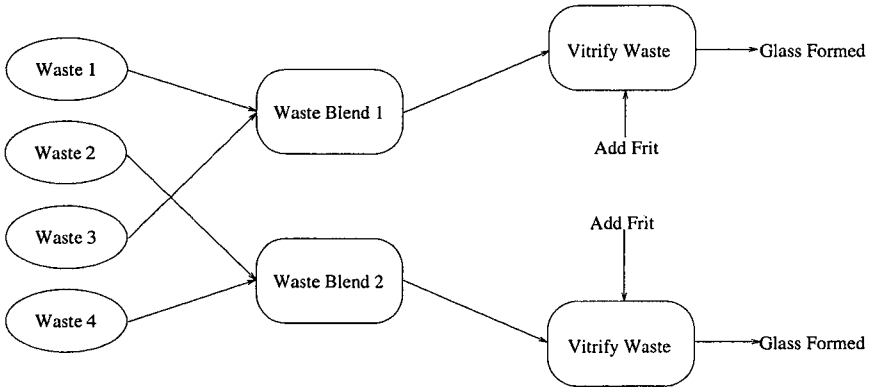

Figure 9. Conversion of waste to glass.

days of CPU time. The new approach based on stochastic annealing coupled with nonlinear programming (STA-NLP), using the efficient sampling technique yields an optimal solution in 18 hours of CPU time, without any significant loss of solution accuracy.

5.1. Problem description

The blending problem addressed in this section consists of mixing the N different waste sources or tanks to form a discrete number of blends B. It is required that *all* the waste from any given tank must combine with other wastes to form a single blend, each blend containing wastes from N/B sources. The blends must therefore be of equal size (i.e., same number of wastes per blend) or alternatively, blends could be formulated to have approximately the same waste masses. These constraints ensure that all the wastes did not go into a single blend [15]. Figure 9 shows a set of four wastes, which needs to be partitioned into two parts to form two blends.

The following section presents the formulation of the deterministic blending problem, which is also the basis for the stochastic optimization formulation for determining the optimal blend configuration in the presence of uncertainties.

5.2. The mathematical formulation of the blending problem

Given a set of tanks (N) and the number of blends (B) that needs to be formed, the number of tanks (TNK) that forms a blend is therefore given by:

$$TNK = \frac{N}{B} \tag{7}$$

If $w^{(i)}$ is the mass of the ith component in the waste, $f^{(i)}$ the mass of the ith component in the frit, and $g^{(i)}$ the mass of the ith component in glass, the following equality constraints

result:

$$g^{(i)} = w^{(i)} + f^{(i)} \tag{8}$$

$$G = \sum_{i=1}^{n} g^{(i)} \tag{9}$$

$$fg^{(i)} = \frac{g^{(i)}}{G} \tag{10}$$

where, G is the total mass of the glass formed, n the total number of components, and $fg^{(i)}$ denotes the fraction of the ith component in the glass. The formation of glass from the blend is governed by several constraints which are briefly described below. A detailed description of the constraints and the data from each tank can be found in [26].

- *Individual component bounds*: These constraints limit the range of the mass fractions each component can have in the calculated glass composition. For each of the components in glass, lower ($fg_{LL}^{(i)}$) and upper ($fg_{UL}^{(i)}$) limits are specified for each mass fraction:

$$fg_{LL}^{(i)} \leq fg^{(i)} \leq fg_{UL}^{(i)} \tag{11}$$

- *Crystallinity constraints*: These are the multiple component constraints which limit the combined fractions of different components. There are five such constraints.
- *Solubility constraints*: These constraints limit the maximum value for the mass fraction of one or a combination of components (e.g., oxides of noble metals: $Rh_2O_3 + PdO + Ru_2O_3$). They are intended to represent the solubility limits for the specified components and are of the form:

$$fg^{(k)} \leq fg_{UL}^{(k)} \tag{12}$$

  where k is the solubility component and $fg_{UL}^{(k)}$ denotes the mass fractions of the particular solubility component.
- *Glass property constraints*: These constraints govern the glass properties such as viscosity, electrical conductivity, and durability. They are of the form:

$$ln(minpropval) \leq \sum_{i=1}^{n} b_i fg^{(i)} + \sum_{i=1}^{n} \sum_{j \geq i} b_{ij} fg^{(i)} fg^{(j)} \leq ln(maxpropval) \tag{13}$$

  where b_i and b_{ij} are the first-order and second-order coefficients in the model, respectively; $fg^{(i)}$ is the mass fraction of component i, and n is the number of components. The *minpropval* and *maxpropval* are the lower and upper bounds of the glass property values, respectively.

The objective is to select the combination of the wastes in the tanks that would form each blend, so that the total amount of frit is minimized. The combinatorial size of this

mixed-discrete optimization problem increases with the number of wastes and the number of blends. The possible number of combinations for a total of N tanks forming B blends, with each blend consisting of TNK tanks is given by:

$$\text{Possible Combinations} = \frac{N!}{B!(TNK!)^B} \tag{14}$$

For a small subset (21) of the total number of tanks at the Hanford site, to be partitioned into 3 blends, there are 66,512,160 possible ways to form the three blends. The number of possible configurations of the blends thus poses an overwhelming problem for common search algorithms to determine the optimal blend configuration. The next section illustrates the procedure for solving the deterministic problem of finding the optimal blend configuration proposed by [26].

The presence of uncertainties not only increases the computational intensity of the problem, but it also affects the optimal blend configuration. The uncertainties associated with the waste blending problem are described in [14] and in the following paragraphs. For details please see [5].

5.2.1. Characterization of uncertainties in the model. This section outlines the methodology adopted to characterize the uncertainties in the waste composition and the glass property models. Since this is a preliminary study, several assumptions have been made to keep the problem manageable and to focus on the key objective; namely, to develop an efficient method for solving this large-scale problem in computationally affordable time, and to illustrate how uncertainties affect the optimal blend configuration. The uncertainties currently being addressed in the optimal waste loading problem fall into two categories.

5.2.1.1. Uncertainties in waste composition. The wastes in the tanks were formed as by-products in different processes used to produce radioactive materials. Consequently, there is associated with each of these tanks a certain degree of variability. Furthermore, over a period of 40–50 years, physical and chemical transformations within a tank has resulted in a non-uniform, non-homogenous mixture. Any experimental sample of the waste withdrawn from the tank is not representative of the tank as a whole, which contributes significantly to the uncertainty associated with the waste composition. This is supplemented, to a lesser extent, by the uncertainties associated with the analytical measurements in determining the waste compositions.

- For this study, 'waste composition uncertainty' is a general term, covering all possible uncertainties in waste feed composition. These sources include batch-to-batch (within a waste type), sampling within a batch, and analytical uncertainties.
- The only estimate of this 'lumped' uncertainty in the composition of the waste feed for high-level vitrification was based on the information available (i.e., analytical tank composition data).
- There is no bias in the composition estimates, the sample values are distributed about the true mean. The derived component mass fractions were assumed to follow normal distributions. The uncertainties of the species in the waste were assumed to be relatively independent of each other (i.e., uncorrelated).

- The relative standard deviation for each component in a particular waste tank was taken to be representative of all the tanks in the study. This assumption needs to be refined as subsequent data becomes available.

The procedure employed in characterizing the waste composition uncertainties is as follows:

- Based on the mean and the relative standard deviation (RSD) for each component in the tank, normal probability distributions were developed for the individual mass fractions, $m_j^{(i)}$, for component i in tank j.
- The above distributions were sampled to develop N_{samp} waste composition input sets (mass fractions, $m_{jk}^{(i)}, k = 1, \ldots, N_{samp}$). A stratified sampling technique (Latin Hypercube sampling or LHS), and the Hammersley sequence sampling (HSS) were both used to generate the samples, and to observe the implication of different sampling techniques on the optimum blend configuration and the computational time.
- Given the mass fractions and the total mass of the wastes (w_j^T) in each tank (j), the mass fractions were normalized to 1.0 to obtain the expected value of waste composition for each blend, as given below.

$$w_k^{(i)} = \sum_{j=1}^{TNK} \frac{m_{jk}^{(i)}}{\sum_{i=1} nm_{jk}^{(i)}} \times w_j^T \quad k = 1, 2, \ldots, N_{samp} \tag{15}$$

$$E\left[w^{(i)}\right] = \sum_{k=1}^{N_{samp}} w_k^{(i)} \tag{16}$$

where $E[w^{(i)}]$ signifies the expected value of the waste mass of the ith component in the waste.

- The mean of the input waste mass for each component, based on N_{samp} samples of the component mass fractions, was then used in the model run.

5.2.1.2. Uncertainties in glass property models. The glass property models are empirical equations fitted to the data (i.e., glass property values against glass compositions). Predictions made with a fitted property model are subjected to uncertainty in the fitted model coefficients. The uncertainties result from the random errors in property values introduced during the testing and measurements, as well as the minor lack-of-fit of the empirical model relative to the actual one. Uncertainties in glass property models reduce the feasible range of the application of the glass property models, thereby affecting the optimal waste loading. The uncertainty in a predicted property value for a given glass composition is defined as [14]:

$$Uncert_{prop} = M[\mathbf{x^T S x}]^{0.5} \tag{17}$$

where, M = multiplier, which is usually the upper 95th percentile of a t-distribution $[t_{0.95}(n - p)]$, n is the number of data points used to fit the model and p is the number of fitted parameters (coefficients) in the model, $\mathbf{x}$ = glass composition vector expanded in the form of the model, $\mathbf{S}$ = covariance matrix of the estimated parameters (coefficients) i.e., b_is and b_{ij}s.

5.3. The solution procedure

The blending process addressed in this paper is similar to some of the processing applications in the coal and petroleum industry. A survey of the techniques used to determine the optimal blend configuration in such cases revealed that most of the methods were based on successive linear, integer, and nonlinear programming. The highly non-convex nature of the constraints and the large combinatorial size of the discrete blending problem at hand, makes it necessary to apply innovative combinatorial optimization techniques to determine the optimal blend configuration.

5.3.1. The deterministic optimization problem. Narayan et al. [26] studied the deterministic problem of determining the optimal blend configuration based on 21 tanks and 3 blends, each blend consisting of $\frac{21}{3}$ or 7 tanks. The composition of the 21 tanks are presented in the reference [26]. The approaches adopted to study the blending problem were as follows:

- *Heuristic approach*: In this case the limiting constraint was identified taking into consideration all (21) tanks. The blends were then formulated such that each blend had the same limiting constraint. If this was achievable, then the frit required would have been the same for the total blend. This approach was found to be very difficult to implement in practice, rather, the blends were formulated so that all blends were near the limiting value of the limiting constraint. The minimum frit requirement found using this approach was 11,736 Kgs. of frit.
- *GAMS-based MINLP approach*: The GAMS-based MINLP approach was very dependent on the initial conditions for the calculation. The best solution found using this approach was 12,341 Kgs. of frit. The GAMS-based MINLP model failed to find the global optimal solution due to the non-convexities in the problem structure.
- *Coupled simulated annealing-nonlinear programming approach (SA-NLP)*: They [26] proposed a two-stage approach based on simulated annealing and nonlinear programming to determine the optimal blend configuration for this problem. The schematic diagram of the solution procedure for the deterministic problem is shown in figure 10.

  The objective was to select a combination of blends for which the total amount of frit added was minimum. The discrete decision was related to the distribution of the tanks among the 3 blends, and was generated by the outer loop of the two-stages SA-NLP algorithm. The objective function for the outer loop (SA) was formulated as the minimization of the total mass of the frit over a combination of blends:

$$\text{Min} \sum_{j=1}^{B} \sum_{i=1}^{n} f_j^{(i)} \qquad (SA) \tag{18}$$

  where $f_j^{(i)}$ is the mass of the ith component in the frit for the jth waste blend, and n and B denote the total number of components, and the given number (3) of blends that needs to be formed, respectively.

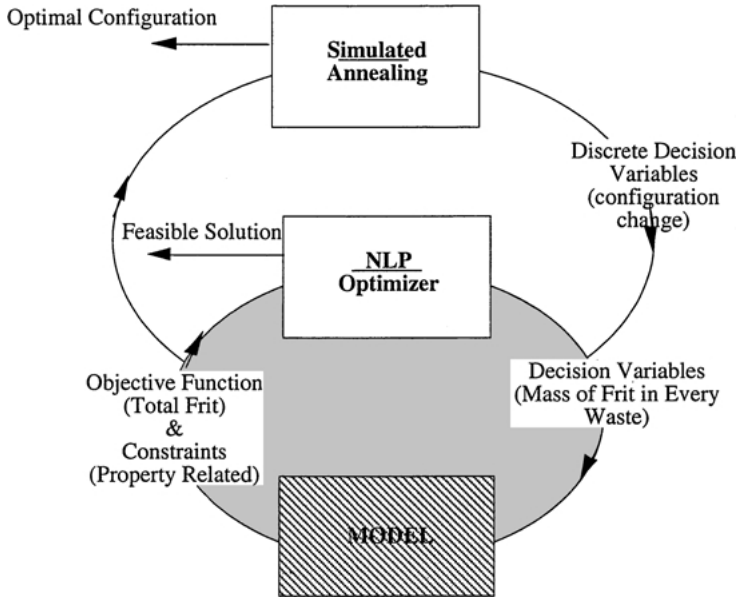

Figure 10. The solution procedure for the SA-NLP algorithm proposed by Narayan et al. for the waste blending problem (deterministic case).

Once the blend (j) was fixed, the resultant NLP sub-problem which consisted of a linear objective function and a mixture of linear and nonlinear constraints, was formulated as:

$$\text{Min} \sum_{i=1}^{n} f^{(i)} \qquad (NLP) \tag{19}$$

subject to: *equality constraints (Eqs. (7)–(10))*
individual component bounds (Eq. (11))
crystallinity constraints
solubility constraints (Eq. (11))
glass property constraints (Eq. (12))

The combined SA-NLP approach was able to identify an optimal solution (11,028 Kgs.) which was lower than the solution predicted by the heuristic and the GAMS-based MINLP approaches. The optimal solution is presented in Table 4. The optimum using the simulated annealing-NLP approach was verified to be the global optimum using a branch and bound strategy. It was observed that branch and bound was computationally intensive compared to the proposed SA-NLP algorithm (e.g., branch and bound took 3 days of CPU time, as opposed to 45 minutes of computational time using SA-NLP approach on a DEC-ALPHA 400 machine), indicating clearly that the SA-NLP approach can be a promising tool for solving large-scale, computationally intensive, mixed-discrete nonlinear optimization problems.

Table 4. Optimal waste blend configuration in the absence and presence of uncertainties in the waste composition and glass physical property models (deterministic and stochastic case).

	Tank distribution		
Blends	Deterministic	LHS	HSS
Blend-1	3, 4, 5, 6, 8, 9, 20	7, 13, 14, 17, 18, 19, 21	7, 13, 14, 17, 18, 19, 21
Blend-2	1, 10, 11, 12, 16, 19, 21	4, 5, 6, 8, 9, 16, 20	4, 5, 6, 8, 9, 16, 20
Blend-3	2, 7, 13, 14, 15, 17, 18	1, 2, 3, 10, 11, 12, 15	1, 2, 3, 10, 11, 12, 15
Objective function	det (expected)		
Total Frit (Kgs)	11,028 (12,022)	11,037	11,037

	Mass in Frit $f_e^{(i)}$ (Kgs.)								
	Blend-1			Blend-2			Blend-3		
Component	Det	LHS	HSS	Det	LHS	HSS	Det	LHS	HSS
SiO_2	293.78	356.49	356.81	680.950	5489.1	5489.3	4550.6	923.19	947.63
B_2O_3	31.350	37.997	38.000	2.186	826.70	828.07	1212.4	0.6956	1.0557
Na_2O	38.683	51.624	51.741	375.06	826.74	825.30	1130.3	427.28	427.37
Li_2O	43.890	51.784	51.817	64.709	756.86	756.83	302.97	46.428	55.064
CaO	0.000	0.000	0.000	11.466	25.355	25.279	344.20	5.7003	2.1108
MgO	0.000	0.000	0.000	66.866	0.000	0.000	485.78	43.944	14.208
Fe_2O_3	0.000	0.000	0.000	0.000	395.51	394.64	502.11	0.000	0.000
Al_2O_3	0.000	0.000	0.000	0.000	1020.0	1020.6	640.96	0.000	0.000
ZrO_2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Other	0.000	0.000	0.000	0.000	21.784	21.590	250.07	0.000	0.000

The sampling exercise was performed using LHS and HSS techniques.
Det denotes deterministics results.

5.4. The stochastic optimization problem

The problem of determining the optimal blend configuration in the presence of uncertainties in the waste composition as well as in the physical property models is posed as a stochastic optimization problem. In the previous papers, it has been shown that stochastic annealing provides an automated, efficient framework for addressing such problems. The stochastic optimization problem requires that the quantities for the waste composition must be represented in terms of their expected values and is represented as follows.

$$g_e^{(i)} = E\left[w^{(i)}\right] + f_e^{(i)} \tag{20}$$

$$G_e = \sum_{i=1}^{n} g_e^{(i)} \tag{21}$$

$$fg_e^{(i)} = \frac{g_e^{(i)}}{G_e} \tag{22}$$

where the subscript 'e' signifies that the quantities are based on the expected value, and $E[w^{(i)}]$ signifies the expected value of the waste mass of the ith component in the waste and $f_e^{(i)}$ is the composition of ith-component in the frit, a decision variable for the stochastic optimization problem. In this case expected formulation is chosen rather than chance constrained formulation as the information about probability of allowable violations for the different constraints in the problem was not available. Later studies will use different formulation and a multi-objective framework to deal with this issue.

Similarly, the individual component bounds, crystallinity constraints, solubility constraints, and the glass property constraints are formulated as:

$$fg_{LL}^{(i)} \le fg_e^{(i)} \le fg_{UL}^{(i)} \tag{23}$$

$$fg_e^{(k)} \le fg_{UL}^{(k)} \tag{24}$$

$$\begin{aligned} ln(minpropval) + Uncert_{prop} &\le \sum_{i=1}^{n} b_i fg_e^{(i)} + \sum_{i=1}^{n} \sum_{j \ge i} b_{ij} fg_e^{(i)} fg_e^{(j)} \\ &\le ln(maxpropval) - Uncert_{prop} \end{aligned} \tag{25}$$

The approach adopted for this waste blending problem is based on a coupled stochastic annealing-nonlinear programming (STA-NLP) technique, which is illustrated in figure 11. The solution procedure incorporates a sequence of three loops nested within one another. The inner loop corresponds to the sampling loop, which generates the samples for the normal distributions of mass fractions (or masses) of the different components in the waste, evaluates the mean of the waste mass for each tank, which is then propagated through the model that determines the glass property constraints. It must be noted that since uncertainties

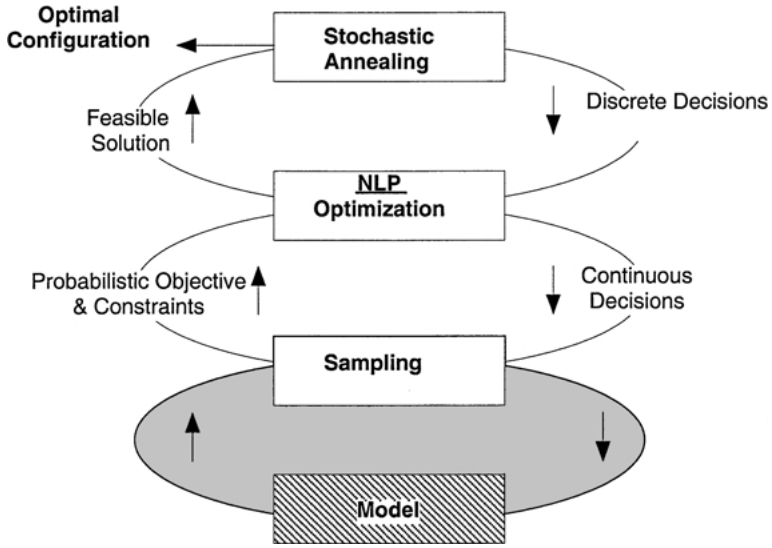

Figure 11. Schematic diagram of the three-stage stochastic-annealing (STA-NLP) algorithm.

in the glass property models were incorporated by reducing the feasible region, as mentioned previously, a sampling exercise to account for uncertainties in the property models is not necessary. The loop above the sampling loop is the NLP optimization loop based on Successive Quadratic Programming, a widely used technique for solving large-scale, nonlinear optimization problems. The objective function for the NLP optimizer identifies the minimum amount of frit for a *given* blend configuration based on the expected value of the masses of the components in the waste blend.

$$\text{Min} \sum_{i=1}^{n} f_e^{(i)} \qquad (NLP) \tag{26}$$

subject to: *equality constraints (Eqs. (7), (20)–(22))*
individual component bounds (Eq. (23))
crystallinity constraints
solubility constraints (Eq. (24))
glass property constraints (Eq. (25))

where, $f_e^{(i)}$ is the composition of ith-component in the frit, a decision variable for the stochastic optimization problem. These decision variables are obtained by optimizing the model containing expected value of the waste composition and uncertain physical property bounds.

$$\text{Min} \sum_{j=1}^{B} \sum_{i=1}^{n} f_{j_e}^{(i)} \qquad (STA) \tag{27}$$

where, $f_{j_e}^{(i)}$ is the mass of the ith component in the frit, a decision variable based on the expected values for the waste composition and the uncertainties in the physical property models for the jth waste blend, n and B denote the total number of components and the given number (3) of blends that needs to be formed, respectively.

Finally, the outer loop in the sequence consists of the stochastic annealing algorithm which predicts the sample size for the recursive sampling loop, and generates the blend configuration such that the total amount of frit is minimum over all the blends.

5.5. Results and discussion

In order to study the effect of the uncertainties in waste composition and in the glass property models, the stochastic optimization problem of determining the optimal blend configuration was solved using two sampling techniques; namely, Latin Hypercube with central limit theorem like the one used in [6, 30] and Hammersley sequence sampling. As mentioned previously, the presence of uncertainties in the waste composition, makes this problem highly computationally intensive. In fact, a fixed sample framework for stochastic optimization using 200 samples and Hammersley sequence sampling was unable to converge on an optimal solution in 5 days (total run time was expected to be approximately 20 days), on

a DEC-ALPHA 400 machine! This demanded the use of the coupled stochastic annealing-nonlinear programming (STA-NLP) approach to identify an optimal solution in a reasonable computational time.

The optimal design configuration identified by the coupled STA-NLP approach using Latin Hypercube sampling and Hammersley sequence sampling are presented in Table 4. The minimum quantity of frit required using both Latin Hypercube and Hammersley sequence sampling is 11,307 Kgs. Although both approaches provided the same optimum objective function, the values of the continuous decision variables appear to be different (e.g. blend 3, Table 4) for the two cases. This is due to the non-convex nature of the problem. Nevertheless, the STA-NLP approach involving Hammersley sequence sampling, for which the error band-width was characterized based on a scaling relationship, was found to be computationally less intensive. For example, the STA-NLP technique using Hammersley sequence sampling, and improved formulation of the penalty term in the stochastic annealing algorithm, through accurate error band-width characterizations based on the scaling relationship took 18 hours, as opposed to 4 days using Latin Hypercube sampling with central limit theorem.

It is observed that the presence of uncertainties affect the optimal blend configuration, compared to a deterministic analysis, significantly. The difference between taking the average value of the uncertain variable as the solution as compared to using stochastic analysis (propagating the uncertainties through the model and finding the effect on the objective function) is defined as the *Value of Stochastic Solution, VSS*. In fact, given the uncertainties in the waste composition and the physical property models, the optimal design configuration obtained by Narayan et al. [26] for the deterministic case, estimates the total frit requirement with uncertainty to be 12,022 Kgs resulting in the value of stochastic solution, VSS to be 715 Kgs. This study re-emphasizes the need for characterizing and incorporating uncertainties in process of determining the optimal design configuration and shows the efficiency of the new algorithms.

6. Conclusions

This paper presented a generalized and efficient approach to stochastic optimization problems involving discrete and continuous decision variables. Sampling across a multi-variate probability distribution is an integral part of stochastic modeling and synthesis. An important aspect of probabilistic modeling is the characterization of the sampling error manifested through the error band-width of any output probability function in a simulation experiment. Classical statistics can estimate the error band-width associated with the mean and variance for Monte Carlo sampling, and is not applicable to any other sampling techniques. On the other hand, efficient and more uniform sampling techniques exist, for which the error band-widths have not been characterized in the past. A methodology based on the self-affinity and scaling properties of the error with the sample size has been proposed, and was shown to estimate the error band-width more accurately for more uniform sampling schemes. The methodology is robust, and is reasonably independent of the functional forms, probability distributions for the uncertain parameters, and the number of uncertain parameters. This new approach using Hammersley sequence sampling was implemented in the stochastic

annealing framework resulting in increased computational savings. This approach made it possible to solve a real world problem of obtaining optimal design configuration of radioactive waste blends to be transformed into glass for long-term storage in a repository in a reasonable amount of computational time. The new capability for stochastic optimization shows great promise for solving large-scale optimization problems under uncertainty.

Appendix A: The Hammersley points

This paragraph provides a definition of the Hammersley points and explicate a procedure for it's design. Any integer n can be written in radix-R notation (R is an integer) as follows:

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

where $m = [\log_R n] = [\frac{lnn}{lnR}]$, 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:

$$\phi_R(n) = 0 \cdot n_0 n_1 n_2 \cdots n_m = n_0 R^{-1} + n_1 R^{-2} + \cdots n_m R^{-m-1}$$

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

$$\vec{z}_k(n) = \left(\frac{n}{N}, \phi_{R_1}(n), \phi_{R_2}(n), \ldots, \phi_{R_{k-1}}(n) \right) \quad n = 1, 2, \ldots, N$$

where $R_1, R_2, \ldots, R_{k-1}$ are the $k-1$ prime numbers where R_1 is the first prime number selected randomly. The Hammersley points are $\vec{x}_k(n) = 1 - \vec{z}_k(n)$.

Appendix B: Stochastic annealing: An illustrative example

The salient features of the stochastic annealing algorithm is best understood through a simple example involving integer and continuous variables. Consider a simple function:

$$Cost = (u_1 y_1 - 3)^2 + (u_2 y_2 - 3)^2 + 2.0\left(x_1^2 - x_2\right)^2 + (x_1 - 1)^2 \tag{28}$$

where

u_1, u_2: uncertain parameters
y_1, y_2: integer variables
x_1, x_2: continuous variables

The uncertain parameter u_1 was obtained from a uniform distribution (Uniform (0.9, 1.1)) while u_2 was obtained from a normal distribution (Normal (0.8, 1.2)) using Monte Carlo and Latin Hypercube sampling methods. The bounds for the integer and continuous

variables are as follows:

$$1 \leq y_1 \leq 4$$
$$1 \leq y_2 \leq 5$$
$$0 \leq x_1 \leq 6$$
$$0 \leq x_2 \leq 5$$

The problem can be stated as:

$$\begin{aligned} &\text{Minimize } E(Cost) \\ &\text{subject to } \quad 1 \leq y_1 \leq 4 \\ &\qquad\qquad\quad 1 \leq y_2 \leq 5 \\ &\qquad\qquad\quad 0 \leq x_1 \leq 6 \\ &\qquad\qquad\quad 0 \leq x_2 \leq 5 \end{aligned}$$

From observations, the minimum of the cost function is zero and occurs when $y_1 = y_2$ approximately equals 3, $x_1 = x_2 = 1$ and the average value of the uncertain parameters u_1 and u_2 are set to unity.

In figure 12, the objective function, the expected value of the cost function $E(Cost)$ is plotted against the annealing temperature. The objective function, $E(Cost)$, was computed for each move using a fixed number (100) of samples and no penalty was imposed for the sample size. On the other hand, figure 13 shows the results of the run using the stochastic annealing algorithm, which incorporates the penalty function approach for Monte Carlo

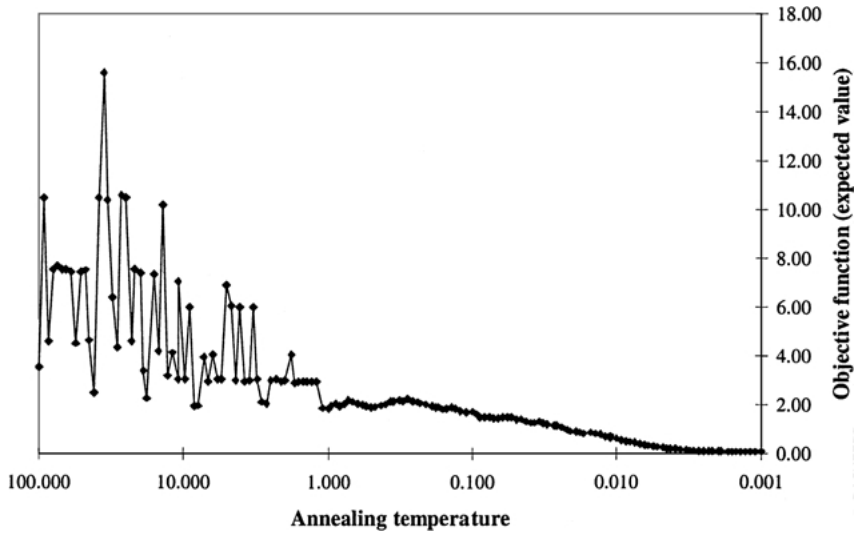

Figure 12. The objective function, $E(Cost)$ vs. annealing temperature for fixed number (100) of samples.

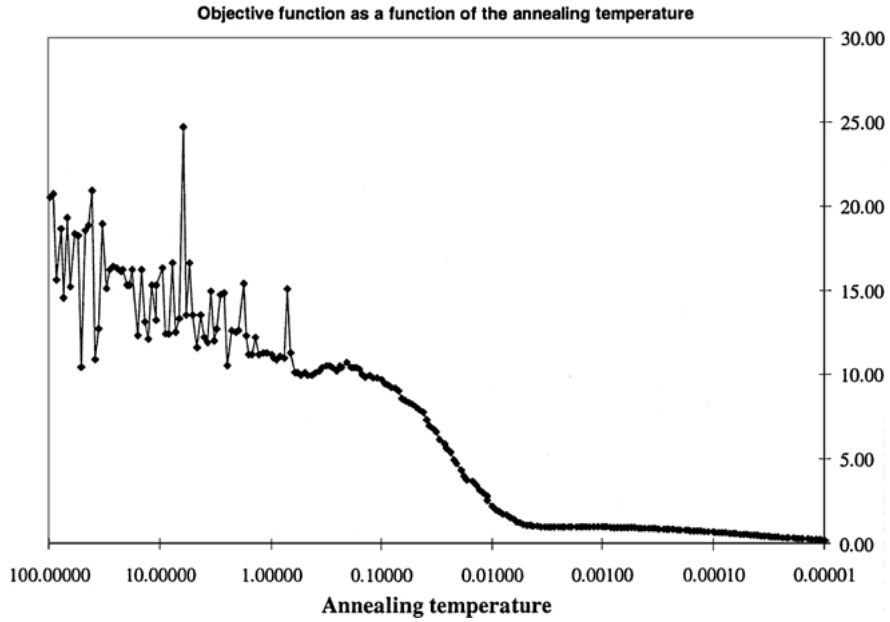

Figure 13. The objective function, $E(Cost)$ vs. annealing temperature for stochastic annealing which automates the sample size selection.

Method. The optimum reached in both cases were essentially the same, but the stochastic annealing algorithm took on an average 32 samples per temperature level to arrive at the minimum, compared to the fixed sample approach based on conventional simulated annealing. These results are summarized in Table 5.

It is easily observed that the stochastic annealing algorithm uses less CPU time and finds the optimum using less number of samples on an average. In fact, there is a trade-off between the computational efficiency and the solution precision, and in most cases the deviation of the solution based on stochastic annealing from the fixed sample approach using conventional simulated annealing was found to be less than 2%.

Figure 14 shows the penalty function (as a percentage of the objective function) and the number of sample size against the annealing temperature. Since the penalty function is related to the sample size, the sample size exhibits random variation with the annealing temperature as the configuration of the system changes with each neighborhood move.

Table 5. Summary of the results for the example problem.

Function	Algorithm	Samp./temp level	y_1	y_2	x_1	x_2	Optimum	CPU time(s)
Exp. val.	Sim. ann.	100	3	3	1.0001	1.0002	0.067	93.26
	Stoch. ann.	32	3	3	1.0007	1.0008	0.069	74.54

Note: The CPU time does not include the sampling time. However, the total sampling time observed was much less compared to the execution time of the algorithms. The computations were performed on VAX-4000.

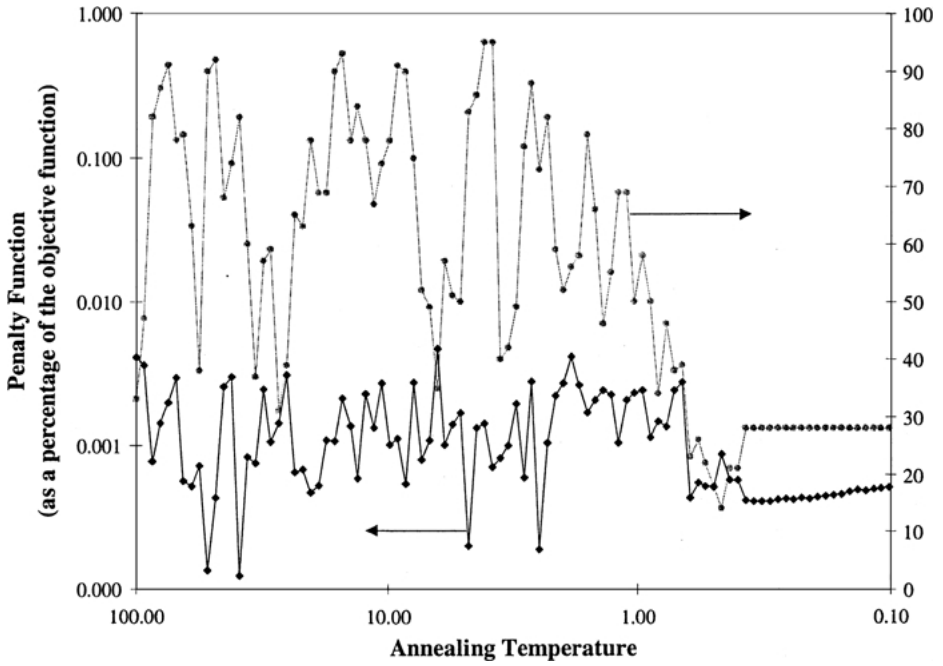

Figure 14. Effect of the number of samples on the penalty term for the expected value, $E(Cost)$ for stochastic annealing.

The stochastic annealing therefore selects a large sample size when the penalty is large, and automatically manipulates the sample size as the annealing proceeds. As the configuration of the system gets closer to the optimum configuration, the weighted term $b(t)$ becomes larger, which prevents the sample size from increasing indefinitely. This is the significant aspect of the algorithm and allows the selection of optimum number of samples, without significant loss of precision in the computation of the probabilistic objective function.

Appendix C: The glass-property models as deterministic constraint functions

Notation

C_1	Bound for Crystal1—3.0
C_2	Bound for Crystal2—0.08
C_3	Bound for Crystal3—0.225
C_4	Bound for Crystal4—0.18
C_5	Bound for Crystal5—0.18
$k_{\min}$	Lower limit for conductivity—18
$k_{\max}$	Upper limit for conductivity—50
$\mu_{\min}$	Lower limit for viscosity (PaS)—2.0

$\mu_{\max}$ Upper limit for viscosity (PaS)—10.0
$D^{PCT}_{\max}$ Max release rate (product consistency test) (g per m^2)—10.0
$D^{MCC}_{\max}$ Max release rate (materials characterization center) (g per m^2)—28.0
μ^i_a Linear coefficients of viscosity model
μ^{ij}_b Cross term coefficients of viscosity model
k^i_a Linear coefficients of electrical conductivity model
k^{ij}_b Cross term coefficients of electrical conductivity model
Dp^i_a Linear coefficients of Durability (PCT) model (for Boron)
Dp^{ij}_b Cross term coefficients of Durability (PCT) model for Boron
Dm^i_a Linear coefficients of Durability (MCC) model (for Boron)
Dm^{ij}_b Cross term coefficients of Durability (MCC) model (for Boron);

Bounds

1. Component bounds:
   (a) $0.42 \le fg^{(\mathrm{SiO_2})} \le 0.57$
   (b) $0.05 \le fg^{(\mathrm{B_2O_3})} \le 0.20$
   (c) $0.05 \le fg^{(\mathrm{Na_2O})} \le 0.20$
   (d) $0.01 \le fg^{(\mathrm{Li_2O})} \le 0.07$
   (e) $0.0 \le fg^{(\mathrm{CaO})} \le 0.10$
   (f) $0.0 \le fg^{(\mathrm{MgO})} \le 0.08$
   (g) $0.02 \le fg^{(\mathrm{Fe_2O_3})} \le 0.15$
   (h) $0.0 \le fg^{(\mathrm{Al_2O_3})} \le 0.15$
   (i) $0.0 \le fg^{(\mathrm{ZrO_2})} \le 0.13$
   (j) $0.01 \le fg^{(\mathrm{other})} \le 0.10$
2. Five glass crystallinity constraints:
   (a) $fg^{(\mathrm{SiO_2})} > fg^{(\mathrm{Al_2O_3})} * C_1$
   (b) $fg^{(\mathrm{MgO})} + fg^{(\mathrm{CaO})} < C_2$
   (c) $fg^{(\mathrm{Fe_2O_3})} + fg^{(\mathrm{Al_2O_3})} + fg^{(\mathrm{ZrO_2})} + fg^{(\mathrm{'Other'})} < C_3$
   (d) $fg^{(\mathrm{Al_2O_3})} + fg^{(\mathrm{ZrO_2})} < C_4$
   (e) $fg^{(\mathrm{MgO})} + fg^{(\mathrm{CaO})} + fg^{(\mathrm{ZrO_2})} < C_5$
3. Solubility constraints:
   (a) $fg^{(\mathrm{Cr_2O_3})} < 0.005$
   (b) $fg^{(F)} < 0.017$
   (c) $fg^{(\mathrm{P_2O_5})} < 0.01$
   (d) $fg^{(\mathrm{SO_3})} < 0.005$
   (e) $fg^{(\mathrm{Rh_2O_3+PdO+Ru_2O_3})} < 0.025$
4. Viscosity constraints:
   (a) $\sum_{i=1}^{n} \mu^i_a * fg^{(i)} + \sum_{j=1}^{n} \sum_{i=1}^{n} \mu^{ij}_b * fg^{(i)} * fg^{(j)} < \log(\mu_m in)$
   (b) $\sum_{i=1}^{n} \mu^i_a * fg^{(i)} + \sum_{j=1}^{n} \sum_{i=1}^{n} \mu^{ij}_b * fg^{(i)} * fg^{(j)} < \log(\mu_m ax)$

5. Conductivity constraints:

(a) $\sum_{i=1}^{n} k_a^i * fg^{(i)} + \sum_{j=1}^{n} \sum_{i=1}^{n} k_b^{ij} * fg^{(i)} * fg^{(j)} > \log(k_m in)$

(b) $\sum_{i=1}^{n} k_a^i * fg^{(i)} + \sum_{j=1}^{n} \sum_{i=1}^{n} k_b^{ij} * fg^{(i)} * fg^{(j)} < \log(k_m ax)$

6. Dissolution rate for boron by PCT test (DissPCTbor):

$$\sum_{i=1}^{n} Dp_a^i * fg^i + \sum_{j=1}^{n} \sum_{i=1}^{n} Dp_b^{ij} * fg^{(i)} * fg^{(j)} < \log(D_{max}^{PCT})$$

7. Dissolution rate for boron by MCC test (DissMCCbor):

$$\sum_{i=1}^{n} Dm_a^i * fg^i + \sum_{j=1}^{n} \sum_{i=1}^{n} Dm_b^{ij} * fg^{(i)} * fg^{(j)} < \log(D_{max}^{MCC})$$

Acknowledgments

Author gratefully acknowledge the extensive discussions with Professor Rangaranjan Pitchumani, Department of Mechanical Engineering, University of Connecticut, Storrs, CT. This work is funded by National Science Foundation grant no. 0119430.

References

1. F. Akesson and J.P. Lehoczky, "Path generation for quasi-Monte Carlo simulation of mortgage-backed securities," Mangement Science, vol. 46, pp. 1171–1187, 2000.
2. M.H. Alrefaei and S. Andradottir, "A simulated annealing algorithm with constant temperature for discrete stochastic optimization," Management Science, vol. 45, pp. 748–764, 1999.
3. J.R. Birge and F. Louveaux, Introduction to Stochastic Programming, Springer Series in Operations Research, Springer: Berlin, 1997.
4. J.R. Birge, "Stochastic programming computation and applications," INFORMS Journal on Computing, vol. 9, no. 2, 1997.
5. P. Chaudhuri, "Process synthesis under uncertainty," Ph.D. Thesis, Department of Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, 1996.
6. P. Chaudhuri and U.M. Diwekar, "Synthesis under uncertainty: A penalty function approach," AIChE Journal, vol. 42, pp. 742–752, 1996.
7. A.J. Crilly, R.A. Earnshow, and J. Jones, Fractals and Chaos, Springer-Verlag: Berlin, 1991.
8. G.B. Dantzig and P. Glynn, "Parallel processors for planning under uncertainty," Annals of Operations Research, vol. 22, pp. 1–21, 1990.
9. U.M. Diwekar and J.R. Kalagnanam, "An efficient sampling technique for optimization under uncertainty," AIChE Journal, vol. 43, pp. 440–449, 1997.
10. M.A. Duran and I.E. Grossmann, "An outer-approximation algorithm for a class of mixed integer nonlinear programs," Math. Prog., vol. 36, pp. 307–339, 1988.
11. Falconer, Fractal Geometry: Mathematical Foundations and Applications, John Wiley & Sons: New York, 1990.
12. R.E. Gephart and R.E. Lundgren, "Hanford tank clean up: A guide to understanding the technical issues," Report BNWL-645, Richland, WA: Pacific Northwest Laboratory, 1995.
13. J. Higle and S. Sen, "Stochastic decomposition: An algorithm for two stage linear programs with recourse," Mathematics of Operations Research, vol. 16, pp. 650–669, 1991.
14. D.F. Hopkins, M. Hoza, and C.A. Lo Presti, "FY94 optimal waste loading models development," Report prepared for U.S. Department of Energy under contract DE-AC06-76RLO 1830, 1994.
15. M. Hoza, "Optimal waste loading models for vitrification of Hanford high-level waste," Report prepared for U.S. Department of Energy under contract DE-AC06-76RLO 1830, 1993.
16. R.L. Iman and W.J. Conover, "Small sample sensitivity analysis techniques for computer models, with an application to risk assessment," Communications in Statistics, vol. A17, pp. 1749–1842, 1982.

17. R.L. Iman and J.C. Helton, "An investigation of uncertainty and sensitivity analysis techniques for computer models," Risk Analysis, vol. 8, no. 1, pp. 71–90, 1988.
18. R.J. Iman 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 Laboratories, Albuquerque, N.M., 1984.
19. B.A.P. James, "Variance reduction techniques," J. Operations Research Society, vol. 36, no. 6, p. 525, 1985.
20. J.R. Kalgnanam and U.M. Diwekar, "An efficient sampling technique for off-line quality control," Technometrics, vol. 39, no. 3, pp. 308–319, 1997.
21. D.E. Knuth, The Art of Computer Programming, Vol. 1: Fundamental Algorithms, Reading, MA: Addison-Wesley, 1973.
22. L. Kocis and W.J. Whiten, "Computational investigation of low-discrepancy sequences," ACM Transactions of Mathematical Software, vol. 23, pp. 266–294, 1997.
23. B.B. Mandelbrot, The Fractal Geometry of Nature, W.H. Freeman: New York, 1983.
24. M.D. Mckay, R.J. Beckman, and W.J. Conover, "A comparison of three methods of selecting values of input variables in the analysis of output from a computer code," Technometrics, vol. 21, no. 2, pp. 239–245, 1979.
25. G. Morgan and M. Henrion, Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis, Cambridge: Cambridge University Press, 1990.
26. V. Narayan, U.M. Diwekar, and M. Hoza, "Synthesizing optimal waste blends," Industrial & Engineering Chemistry Research, vol. 35, pp. 3519–3527, 1996.
27. H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAM: Philadelphia, 1992.
28. Nuclear News, "DOE selects Hanford tank waste cleanup plan," Nuclear News, vol. 40, p. 49, 1997.
29. L.A. Painton and U.M. Diwekar, "Synthesizing optimal design configurations for a Brayton cycle power plant," Computers and Chemical Engineering, vol. 5, pp. 369–381, 1994.
30. L.A. Painton and U.M. Diwekar, "Stochastic annealing under uncertainty," European Journal of Operations Research, vol. 83, pp. 489–502, 1995.
31. A. Papageorgiou and G.W. Wasilkowski, "On average case complexity of multivariate problems," Journal of Complexity, vol. 6, pp. 1–6, 1990.
32. H. Peitgen, H. Jurgens, and D. Saupe, Fractal for the Classroom Part One: Introduction to Fractals and Chaos, Springer-Verlag: Berlin, 1991.
33. C.A. Pickover and A. Khorasani, "On the fractal structure of speech waveforms and other sampled data," Research Report No. 11305, Computer Science Dept., IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, 1985.
34. R. Pitchumani and S.C. Yao, "Correlation of thermal conductivities of unidirectional fibrous composites using local fractal techniques," ASME Journal of Heat Transfer, vol. 113, no. 4, pp. 788–796, 1991.
35. A. Prèkopa, "Logarithmic concave measures and related topics," in Stochastic Programming, M.A.H. Dempster (Ed.), Academic Press: New York, NY, 1980.
36. A. Prèkopa, Stochastic Programming, Kluwer Academic Publishers: Dordrecht, Netherlands, 1995.
37. R. Salazar and R. Toral, "Simulated annealing using hybrid Monte Carlo," Journal of Statistical Physics, vol. 89, pp. 1047–1060, 1997.
38. E. Saliby, "Descriptive sampling: A better approach to Monte Carlo simulations," J. Operations Research Society, vol. 41, no. 12, pp. 1133–1142, 1990.
39. H. Szu and R. Hartley, "Fast simulated annealing," Physics Letter A, vol. 3, pp. 157–162, 1987.
40. P.J.M. vanLaarhoven and E.H.L. Aarts, Simulated Annealing: Theory and Applications, Reidel Publishing Co., 1987.
41. R. Wang and U. Diwekar, Latin hypercube Hammersley sequence sampling and leaped Hammersley sequence sampling, in preparation.
42. R.-J.-B. Wets, "Stochastic programming," in Optimization (Handbooks in Operations Research and Management Science, Vol. 1, G.L. Nemhauser, A.H.G. Rinooy Kan, and M.J. Todd (Eds.), North-Holland: Amsterdam, 1990.
43. R.J.B. Wets, "Challenges in stochastic programming," Math. Progr., vol. 75, pp. 115–135, 1996.
44. H. Wozniakowski, "Average case complexity of multivariate integration," Bulletin of the American Mathematical Society, vol. 24, pp. 185–194, 1991.