Seminal Concepts for a New Approach to Continuous-Variable Optimization Under Uncertainty: Probabilistic Ordinal Optimization*

Vicente J. Romero,
Sandia National Laboratories†
Albuquerque, NM

Chun-Hung Chen‡
Dept of Systems Engineering and Operations Research
George Mason University, Fairfax, VA

Abstract

A very general and robust approach to solving optimization problems involving probabilistic uncertainty is through the use of Probabilistic Ordinal Optimization. At each step in the optimization problem, improvement is based only on a relative ranking of the probabilistic merits of local design alternatives, rather than on crisp quantification of the alternatives. Thus, we simply ask: "Is that alternative better or worse than this one?" to some level of quantified statistical confidence, not: "HOW MUCH better or worse is that alternative to this one?". The latter answer strictly implies asymptotically converged statistics (and corresponding expense of sampling/integration) associated with complete certainty (i.e., 100% confidence in the statistics). Looking at things from an ordinal optimization perspective instead, we can begin to quantify and utilize the tradeoff between computational expense and vagueness in the uncertainty characterization. This paper introduces fundamental ordinal optimization concepts using a low-dimensional probabilistic optimization problem as a vehicle. Advanced implementational possibilities are discussed, along with merits versus non-ordinal approaches to optimization under uncertainty.

1. Introduction

A very general and robust approach to solving optimization problems involving parameter uncertainty is through the use of Ordinal Optimization, where improvement at each step in the optimization problem is based only on a *relative ranking* of local alternatives, rather than on a crisp quantification of each alternative. This approach allows optimization under non-probabilistic and semi-quantitative descriptions of uncertainty. In fact, it can be seen as a formalism of the ordinal selection process that decision makers employ in the real world when deliberating over competing options whose outcomes are not precisely estimable.

For continuously variable design options ("continuous-variable" design problems), non-gradient-based local optimizers such as Simplex and Pattern Search methods ([7],[9]), and global-to-local methods such as Genetic Algorithms and DIRECT ([10]), can be used to perform ordinal optimization under uncertainty. In cases where *probabilistic* information on the uncertain alternatives is available, probabilistic ordinal optimization as discussed here may provide a "Gold Standard" reference against which the accuracy and efficiency of other OUU methods can be compared –analogous to the role that Monte Carlo Simulation plays in uncertainty propagation. This paper introduces fundamental probabilistic ordinal optimization concepts with reference to a low-dimensional probabilistic optimization problem. Advanced implementational possibilities are discussed, along with merits versus non-ordinal approaches to probabilistic optimization.

2. Engineering Optimization Problem motivating Probabilistic Ordinal Concepts

2.1 Deterministic Optimization Problem

The probabilistic optimization problem considered here is an outgrowth of a deterministic optimization problem ([19]) in which heating conditions were sought that put a hypothetical weapon subsystem design most at risk. The design problem is parameterized in terms of two key heating variables: 1) the radius *r* of a circular region of impinging fire on the top of the safing subsystem; and 2) a coor-

---

†Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the U.S. Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

‡Supported in part by NSF under Grants DMI-0002900 and DMI-0049062, by NASA Ames Research Center under Grant NAG-2-1565, by FAA under Grant 00-G-016,and by George Mason University Research Foundation.

dinate *x* that moves the center of the "blow torch" along the device's surface. A 3-D finite-element conduction/radiation thermal model (described in [19]) is used to calculate the transient temperature response of the device. Figure 1 shows the calculated temperature of the device (cut along its plane of symmetry) at some point in time for a particular radius *r* and location *x* of heating.

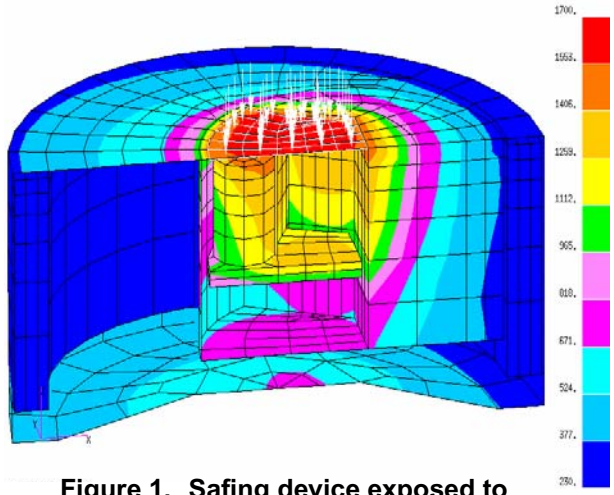

Figure 1. Safing device exposed to circular region of heating.

Figure 2 shows temperature histories for two safety-critical components within the safing device: a "strong link" (SL) and a "weak link" (WL). The components are so termed because the weak link must be thermally weaker and fail at a lower temperature than the thermally stronger

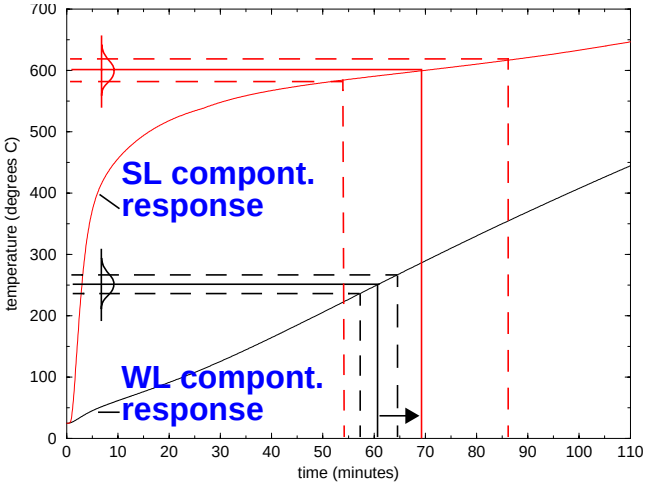

Figure 2 Solid lines represent a single realization of each component's failure temperature, reflected through its temperature response curve into a failure time. The difference in failure times, t_fail_{SL} - t_fail_{WL}, is the corresponding *safety margin*, shown by right-pointing arrow for positive margin above.

strong link in order to prevent inadvertent operation of the weapon.

For now focusing just on the depicted *nominal* (as though deterministic) failure temperatures of the components (nominal T_fail_{WL} =250° C and nominal T_fail_{SL} =600° C), we get an associated "safety margin" of time:

$$S = t\_fail_{SL} - t\_fail_{WL}. \qquad \text{EQ 1}$$

Here t_fail_{SL} and t_fail_{WL} are the elapsed times (from time zero at the beginning of the thermal simulation) required for the links to reach their nominal failure temperatures. In Figure 2, the safety margin associated with the nominal failure temperatures of the components is positive and has a value of about S=+9 minutes (right-pointing arrow in the figure). A negative or zero safety margin indicates that the strong link is not failing after the weak link as desired, so the safing device is experiencing a vulnerability for that set of heating conditions and component failure temperatures. The safety margin is a deterministic indicator of the safety of the device. A probabilistic indicator of the safety of the device will later be defined that takes into account the depicted uncertainty in the component failure temperatures.

In the optimization problem we wish to find the heating conditions that minimize the indicated safety of the device. This minimum corresponds to the heating conditions that most threaten the intended function of the safing device. If the vulnerability is deemed unacceptable, then the design can be modified to sufficiently harden the device to these worst-case heating conditions.

In the *deterministic* optimization problem the optimal values of these two variables *r* and *x* are sought that minimize the safety margin *S* of the device. As the {*r*, *x*} variables are changed, the SL and WL temperature responses change and (with fixed nominal failure temperatures for the components) the value of the safety margin changes accordingly. Thus, the deterministic safety-margin objective function is navigated (minimized) in the deterministic optimization problem.

The deterministic optimization problem is complicated by numerical noise resulting from discrete time and space approximation in the model (see [3], [19]). On a more global scale, the deterministic optimization problem contains mathematical difficulties associated with navigating to a global minimum at {r,x} = {1.62, 0.782} on a nonconvex design surface having a fold and several local minima as described in [3] and [19].

Toward solving the deterministic local optimization problem, Newton-based nonlinear programming (NLP) methods (sequential quadratic programming, BFGS quasi-

Newton, and quasi-Newton) from several different research and commercial optimization packages were tried on the problem with little success as explained in [3]. Nonconvexity and noise resulted respectively in ill-conditioning and inaccuracy of the Hessian matrix of finite differenced second-order derivatives.

First-order NLP conjugate gradient methods were more successful, but still exhibited some noise-induced nonrobustness. Choice of finite difference step size (FDSS) for computation of gradients proved to be important. Initial starting point also proved to be important because of the noise. While conjugate gradient (CG) was capable of being successful, it was certainly not robust algorithmically or numerically on this problem. The "right" line search algorithm had to be combined with suitable finite difference step sizes that were combatible with the discretization or numerical resolution level of the computational model. Investigating the algorithmic and numerical space in order to find successful and efficient combinations was an ad hoc, uncertain, labor and computer intensive process.

Not only this experimentation process, but also the relatively high degree of model resolution required for reliable CG navigation of the design space, added large cost to the problem. More noise-tolerant derivative-free optimization approaches were applied to this problem in [4]. A simple Coordinate Pattern Search (CPS) method ([7]) was shown to be successful with lower model resolution/cost requirements than the CG approach. CPS was more robust than CG and more cost effective than the least expensive CG run. This having been said, CPS was run with a step size selection that benefitted from previous experimentation with finite difference step sizes in the CG applications. If such prior information was not available, some experimentation with CPS step size would probably have to be done, as it is a free parameter in the method that directly affects its numerical performance –though probably not to the extent as in CG.

More recently, a global-to-local optimizer efficient for low design dimensions and tolerance of small scale noise was devised and applied to this problem in [13]. In terms of: number of function evaluations required; cost of function evaluations (model discretization level required for accuracy); and robustness to noise level, the method performed considerably better than all those previously tried. The method was later recognized to be a primative surrogate-based *trust-region* optimization approach with moving low-order polynomial local response surfaces. Unknown to the author, established approaches in this vein had already existed at the time ([1], [12]). The advantages of surrogate-based trust region methods for handling noisy optimization problems are now well recognized.

2.2 Probabilistic Optimization Problem

We now consider the associated probabilistic optimization problem. A glance at Figure 2 indicates that stronglink and weaklink failure temperatures trace out to relatively flat portions of the temperature response curves. This is more true for the stong link than for the weak link, but regardless, it can be seen that the component failure times are very sensitive to the failure temperature values. Thus, moderate uncertainties in the component failure temperatures T_fail_{WL} and T_fail_{SL} can have a substantial impact on the uncertainty of the safety margin. In fact, when uncertainty bands 5% above and below the nominal failure temperatures are considered in Figure 2, the corresponding bands in failure times indicate that the safety margin could vary from about 28 minutes at best to about -11 minutes at worst. Comparing this to the nominal safety margin of +9 minutes, it is apparent that the effects of uncertainty in component failure thresholds are very important in this problem.

2.2.1 Safety margin distribution due to uncertain component failure thresholds. For the purposes here the stronglink and weaklink failure temperatures T_fail_{WL} and T_fail_{SL} are each assumed to be described by truncated normal distributions with means μ equal to the respective mean failure values of 600° C and 250° C, and standard deviations σ equal to 3% of the means, e.g. 18° C and 7.5C respectively. (The distributions are truncated at 3σ above and below their mean values, and then renormalized to integrate to unit value.)

Multiple sets of weak link and strong link failure temperatures can be generated from standard Monte Carlo sampling, and then a safety margin can be computed for

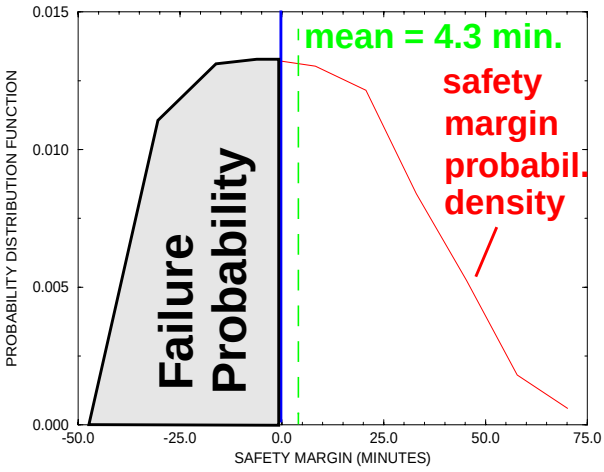

Figure 3. Safety Margin probability density function at a particular set (*r*,*x*) of heating conditions for uncertain SL,WL failure temperatures.

each set by using the time histories of the weak and strong links obtained from the thermal model run at the {*r,x*} heating conditions. The resulting population of safety margin realizations will be distributed with some probability density as illustrated in Figure 3. Associated statitistics such as mean, standard deviation, and probability of attaining a zero or negative safety margin can be calculated. For example, the mean safety margin denoted in the figure is +4.3 minutes for the deterministically worst case set of heating conditions ($\{r,x\}_{\text{opt-det}} = \{1.62, 0.782\}$).

The deterministic safety margin for this set of heating conditions is about 2.5 minutes or 40% less than the 4.3-minute mean. This problem is clearly nonlinear in the uncertain parameters T_fail_{WL} and T_fail_{SL}; otherwise the mean safety margin would equal the deterministic safety margin at the mean values of the uncertain parameters. This nonlinearity may result in different optimal (worst-case) heating conditions $\{r,x\}_{\text{opt}}$ if component failure uncertainty is being taken into account instead of being ignored as in the deterministic optimization problem.

Figure 4 shows a 3x3 grid of points over a small subset of the design space, centered about the deterministic optima $\{r,x\}_{\text{opt-det}} = \{1.62, 0.782\}$. Figure 5 shows magnitude bars for failure probabilities calculated at the nine points of the grid by Monte Carlo sampling over the SL and WL failure temperature uncertainties. Thus, safety margin distributions and corresponding failure probability magnitudes (like in Figure 3) are obtained at each design point. The optimization problem now is to *maximize* the *probabilistic* objective function in order to determine the worst-case heating values of *r* and *x*.

2.2.2 Motivation for probabilistic ordinal optimization in this problem. A bi-quadratic response surface is also shown in Figure 5 that can actually be seen to not go exactly through the end points of the probability bars because it was created from probabilities obtained with 500 Monte Carlo samples, rather than the 1000 samples the bars are based on. The mismatch between bar height and surface height can be seen to vary somewhat over the design space. As depicted in Figure 6, the issue is related to confidence intervals (CI) on the Monte Carlo point estimates.

This raises the questions: "What is the effect of number of Monte Carlo samples on the accuracy with which a probabilistic optimum can be identified in the design space? If I want to identify the heating variables {*r,x*} that correspond to the highest probability of device failure, how can I control the impact of Monte Carlo sampling errors? How many samples do I need to take at each point in the design space to shrink the confidence intervals small enough so that it is unambiguous which design point corresponds to the highest failure probability?"

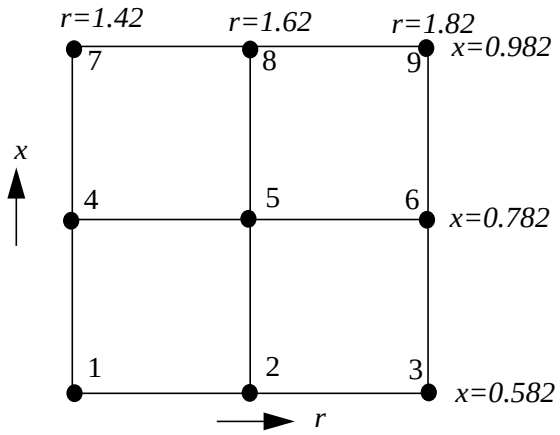

Figure 4. 9-point grid over important subset of 2-D optimization space, centered at point where deterministic optimum for worst-case heating occurs.

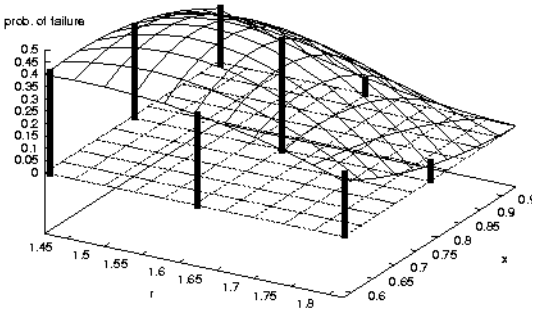

Figure 5. Probabilistic objective function (bi-quadratic response surface built from Monte Carlo point estimates of failure probabilities at the 9 grid points)

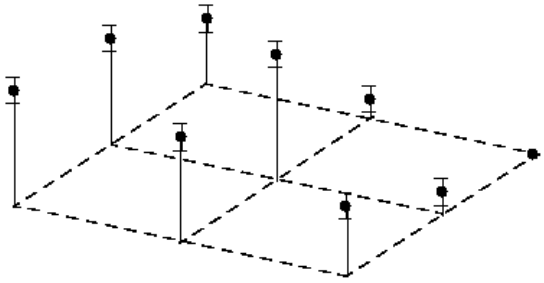

Figure 6. Monte Carlo point estimates of failure probabilities with associated Confidence Intervals shown

3. Fundamental Concepts of Probabilistic Ordinal Optimization

Probabilistic ordinal optimization has been studied for some time in the Operations Research field, in connection with evaluating discrete design alternatives (as opposed to continuous-variable problems like the one in this paper). In particular, the work has addressed apportioning Monte Carlo sampling amongst multiple uncertain or stochastic discrete systems in order to most efficiently resolve their statistical behavior for the purpose of selecting the best (or several best) option(s). The concepts and procedures apply whether it is desired to: A) minimize the total number of samples required to reach a desired probability of correct selection of the best (or several best) option(s); or B) maximize the probability of correct selection for a given budget of total samples N_T to be optimally apportioned among the various alternatives. In either case, the odds that the current identification of the best alternative(s) is correct can be estimated at every stage of the sampling.

Probabilistic ordinal optimization concepts can also be applied to continuous-variable optimization under uncertainty (OUU), though to the authors' knowledge they have not yet been applied with a systematic manner as discussed here. For continuous-variable (C-V) problems, improvement steps in the design space can be taken based on ordinal ranking of candidate alternatives according to relative merit, rather than by attempting to resolve the actual merit value of each alternative like non-ordinal approaches do (for a brief overview of non-ordinal OUU approaches see, *e.g.*, [5]). The relaxed conditions for progress in ordinal optimization allow it to be more efficient than surrogate-based methods for certain problems, as explained below. These latter approaches can introduce sources of noise and approximation error into the probabilistic objective function that can be avoided with probabilistic ordinal approaches ([15]).

3.1 Independent (Uncorrelated) Designs

Consider two different designs or processes that each have some variability in their output(s) or desired behavior(s). Focusing on one output or desired behavior to compare the relative merit of the two designs, we can compare statistics of their relative merit. Thus, we can ask whether design A or design B has the higher mean output; or the lower probability of meeting or not meeting some acceptable threshold of behavior; or the smallest variance in the product produced.

We can sample the behavior or output of these independent designs and generate statistics of their tendencies. For mean and probability we know from classical Confidence Interval (CI) theory that the values $\hat{S}_A$ and $\hat{S}_B$ of a statistic calculated from finite sampling will be realizations from normal distributions about the true statistical values S_A and S_B, with variances $\sigma^2_{\hat{S}_A}$ and $\sigma^2_{\hat{S}_B}$ that decrease as the number of samples increases. Figure 7 depicts normal distributions of the calculated statistics from two designs being compared.

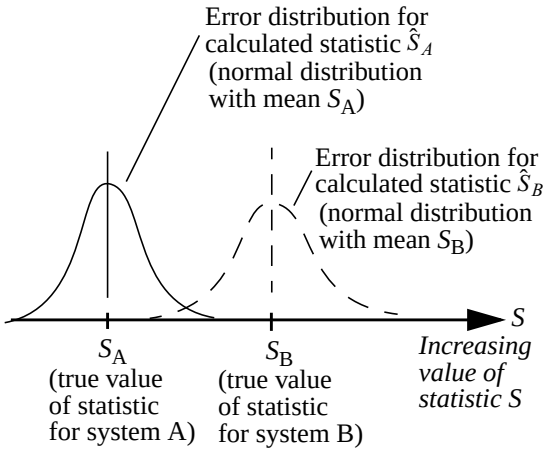

Figure 7. Normally distributed errors in Monte Carlo estimates of statistical behaviors of 2 different systems.

Under limited sampling the two normal distributions each have non-zero variance, which means that they overlap to some degree as illustrated in the figure. Hence, it is possible for a calculated statistic $\hat{S}_A$ of design A to have a value greater than the calculated statistic $\hat{S}_B$ of design B, even though this is not the case for the true values S_A and S_B. If the optimization goal here is to pick the design that increases the critical statistic (say, increases the mean output rate or probability of acceptance), then there is a non-zero probability that a misleading indication would be given that design A is better than design B. The more the error distributions overlap, the less certain it is that the chosen design is indeed the better design.

As the number of Monte Carlo samples regarding each design increases, the critical statistics are resolved better and better. That is, their confidence intervals decrease such that it begins to become more and more apparent which alternative has the favored statistical behavior. Hence, the probability of making a correct selection among the two alternatives (i.e., the "probability of correct selection", P{CS}) increases with more Monte Carlo sampling.

In [16] a methodology is described for calculating the probability of correct selection, P{CS}, when only two design options are considered at a time. We can then pose a requirement that we want to determine to a given level of statistical assurance P{CS} that the option picked has the better actual figure of merit. Answering this will require

sampling the alternatives' behaviors a sufficient number of times. The question then comes: "How many is adequate?"

A unique answer to the above question does not exist. In looking at Figure 7, a desired overlap (or non-overlap as the present requirement is stated) may be attainable by increased Monte Carlo sampling to manipulate the spread (variance) of alternative A's distribution, of alternative B's distribution, or some combination of both (only this last case is guaranteed to always work). If a combination of both is used, then there is also not a unique combination that will attain the desired result.

The way to approach the sufficiency issue, then, is to ask a second question: "What is the most *efficient* way I can achieve sampling sufficiency for a given P{CS} requirement?" This question *does* have a unique answer, and this answer minimizes the total number of samples N_T apportioned amongst the two alternatives that will achieve the desired P{CS} level. In comparing two alternatives, the optimal apportionment of samples between the two alternatives is relatively easy to obtain, as described in [17]). In comparing more than two alternatives, the optimal solution ("Optimal Computing Budget Allocation", OCBA[2]) is much more involved and difficult to obtain.[1] Furthermore, calculation of P{CS} itself is much more difficult than the well-known approach for two alternatives ([8]).

Though the possibilities of increasing the efficiency are myriad, a simple implementation of probabilistic ordinal optimization (comparing only two alternatives at a time[2]) was applied to the probabilistic optimization problem defined earlier. A requirement of 95% probability of correct selection was imposed in the ordinal comparisons, and simple Coordinate Pattern Search ([7]) was used to systematically progress toward the optimum in the continuous 2-D *r-x* design space. At the end of the process a high statistical confidence existed that probabilistic optimum obtained was the true optimum (within the resolution of the final CPS step size in the design space).

To the authors' knowledge, no other methods for probabilistic optimization currently offer this type of quantitative assessment of correctness, or even error estimates that are shown to be robust and reliable. Thus, probabilistic ordinal optimization may be usable to provide a "Gold Standard" reference against which the accuracy and efficiency of other competing methods can be compared –analogous to the role that Monte Carlo simulation plays in uncertainty propagation.

[1] In comparing more than two alternatives, a "simultaneous comparison" approach is more efficient than the sequential method of comparing two at a time and then dropping the "loser" and comparing the "winner" to the next alternative in the list, until all alternatives have been considered.

[2] A more sophisticated approach considering multiple alternatives simultaneously with OCBA is cited in section 3.2.3 of this paper.

3.2 Continuous-Variable Problems: Correlation of Uncertainty in the Design Space

The full-variance distributions of Figure 7 apply for completely uncorrelated designs, as is often the case in comparing discrete alternatives in the realm of Operations Research. For continuous design spaces, however, closely neighboring points in the design space can have closely correlated uncertainties. The efficiency enhancing prospects of spatial correlation in the design space for C-V problems is investigated here. Figure 8 illustrates the issue. The mapping of uncertainty distributions at two neighboring design points is shown. For convenience of illustration, the uncertainties in this particular figure derive from stochastic noise in the tolerances that can be held in the design variable. This uncertainty (as a function of location in the design space) maps through the deterministic input/output function of system behavior, as shown. Resulting output response uncertainties are depicted on the vertical axis.

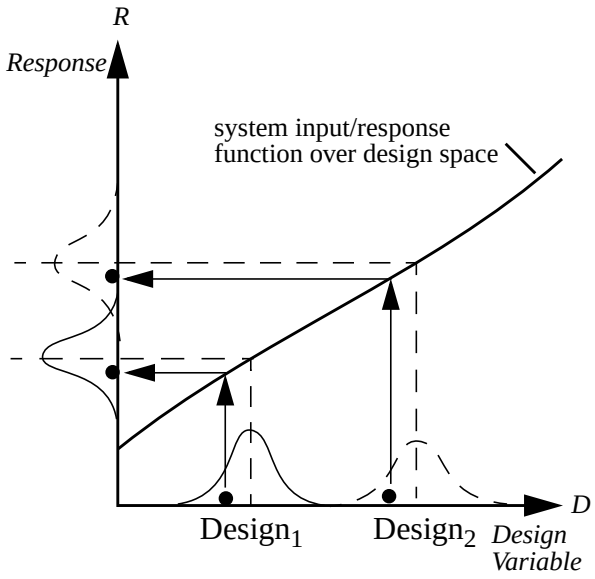

Figure 8. Correlated sampling of input/ response distributions at 2 different points in design space.

Though shown on the design-variable axis, uncertainties from other orthogonal (non-design-variable) inputs to the system may exist. These uncertainties would similarly map into the response distributions. Thus, in general the response distributions may have contributions of uncer-

tainty from design variables and/or orthogonal uncertainty variables.

In any case, as the compared design points in the space get closer and closer, the response distributions look more and more alike. Furthermore, the response function that maps points in the design and uncertainty spaces to points on the response axis becomes more and more alike. This means that if Monte Carlo realizations of the involved uncertainties are taken with the exact same random number generator (RNG) starting seed at the neighboring design points ("*spatially correlated sampling*"), the uncertainty realizations (samples) and the response realizations (samples) will be very strongly correlated as illustrated in the figure for the *n*th sample taken at each design point. That is, on the design-variable axis, uncertainty realizations at the neighboring design points will come at similar percentile locations in the input uncertainty distributions. Likewise, the mapped response values shown for these input uncertainty realizations will occur at correlated percentiles of the response distributions. This means that as the compared design points get closer and closer, the problem of comparing the associated response distributions approaches a *deterministic* comparison problem.

3.2.1 Probabilistic Comparison Efficiency Limit–
The Spatial Correlation Deterministic Efficiency Limit. In the probabilistic comparison efficiency limit, if input and response uncertainties do not change over the design space between design points being compared, then the OUU problem is *deterministic* –the only difference in the response distributions is their relative position on the response axis due to their mapping from the deterministic response function. Since the only difference in the response distributions is their relative position on the response axis, it is then trivial to identify which gives a higher mean response or probability of response exceeding some threshold value. Then, since only one correlated Monte Carlo sample of the uncertainty at each design point is needed to define the relative position of the response distributions, if we take the single sample at each design point to be that which occurs at the mean of the input uncertainties then this is recognized as the usual deterministic analogue of the probabilistic optimization problem. Note that we cannot say here what the mean or probability value is at each design point, but only which has the highest or lowest value –which is sufficient to select the best move for optimizing the probabilistic objective function.

The above property is termed the "*spatial correlation efficiency limit*" ([18]) in C-V probabilistic ordinal optimization. It enables probabilistic OUU problems to be recognized as deterministic optimization problems from an ordinal perspective –with the attendant simplification and cost savings that this brings. For example, the relatively expensive and complex machinery employed to solve the OUU model problems in [11] and [20] is not necessary, as the input and response uncertainties do not change over the design space. These are cases where an ordinal approach to OUU is more efficient than surrogate-based approaches.

3.2.2 Efficiency from Domination of "Uncertainty Variation" by Underlying Deterministic Variation

Another aspect of the nature of probabilistic optimization problems frequently allows successful initial treatment of general probabilistic optimization problems as deterministic optimization problems to inexpensively find a point (deterministic optimum) in the design space that is relatively close to the probabilistic optimum. From this starting point, OUU methods are then applied to find the probabilistic optimum.

Movement from the deterministic optimum to the probabilistic optimum is driven by combined influences from local trends in the deterministic and uncertainty behaviors. Figure 9 illustrates the issues. The middle histogram in the figure corresponds to the safety-margin uncertainty distribution (Figure 3) at the deterministic optimum $\{r,x\}_{\text{opt-det}} = \{1.62, 0.782\}$. The histograms on either side correspond to the neighboring points 4 and 6 of the 9-Point Grid (Figure 4) over the optimum region of the design space. In Figure 9 the deterministic response curve is a quadratic fit through the 50th percentile locations on the safety margin distributions. The mathematical optimum of this curve (not drawn exactly in the figure) is at $r_{\text{opt-det}} = 1.62$. There is an obvious "uncertainty gradient" there; safety margin variance increases noticably as the r coordinate is traversed from right to left (the histograms of response and their associated 25th, 50th, and 75th percentile curves increasingly spread out in going from right to left). The nonzero uncertainty gradient at the deterministic optimum will push the probabilistic optimum off of the deterministic optimum.

Given that the deterministic trend curve is horizontal at the deterministic optimum, the direction in which the variance of the safety margin increases is the direction in which more of the distribution will fall below the S=0 threshold (horizontal axis in the figure), and therefore the direction in which failure probability will increase and the direction in which the probabilistic optimum (failure probability maximum) will lie. Simultaneously, the deterministic safety margin (deterministic behavior function) rises to the left of its minimum at $r_{\text{opt-det}} = 1.62$. As the deterministic response curve rises, the variance of the safety margin distribution has to increase more and more in order that the amount of the distribution beneath the S=0 axis keep increasing. At some point, even though response variance keeps increasing in this direction, the effect of the rising

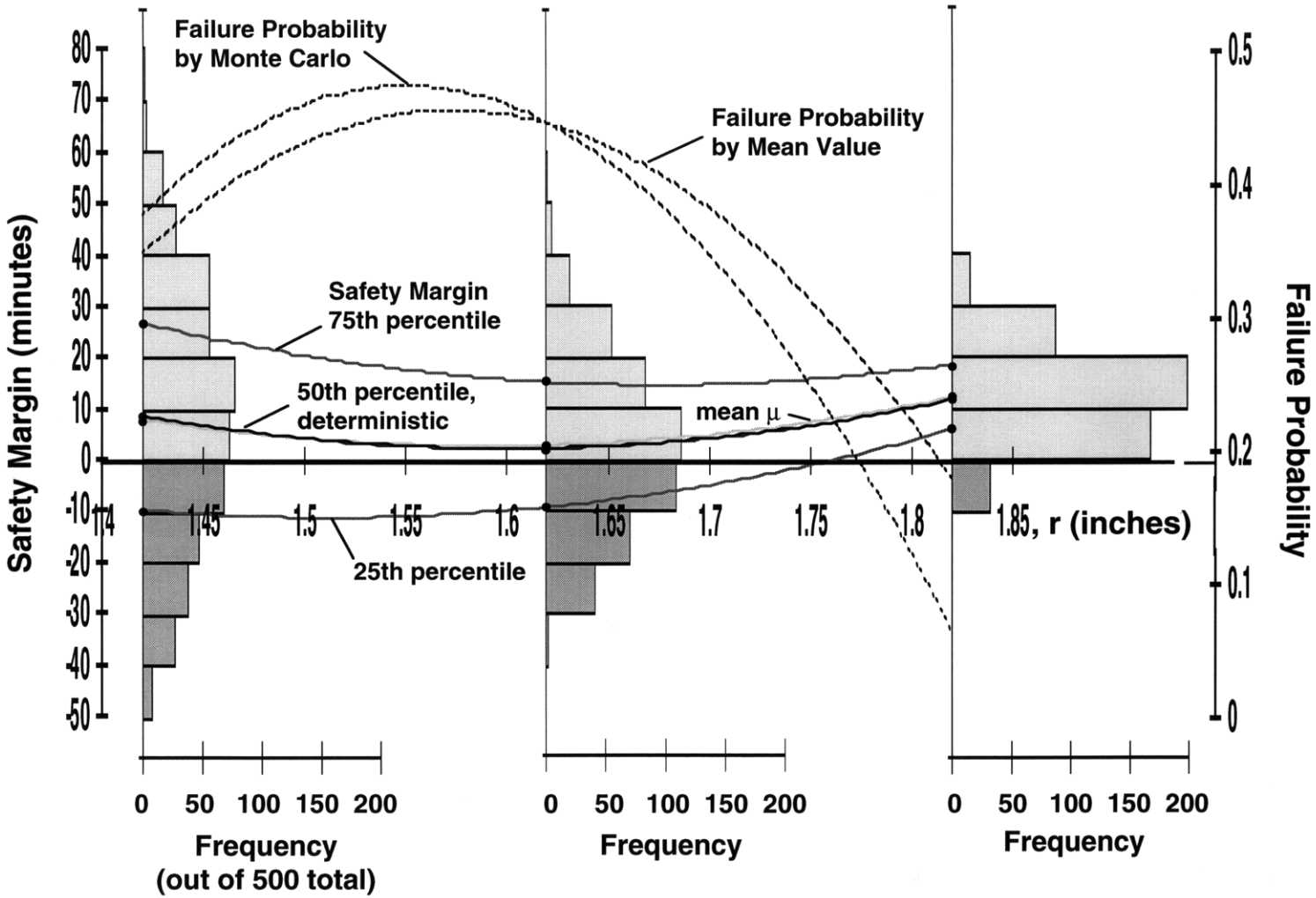

Figure 9. Histograms of response (safety margin) distributions at three points (# 4,5,6 of the 9-Point Grid in Figure 4) along an x=0.782 cut of the optimum region of the design space.

underlying deterministic trend overwhelms the effect of the increasing variance and that the failure probability begins to decline. This is reflected in the figure by the failure probability curves' maxima somewhat to the left of the deterministic optimum. (These curves are quadratic fits to failure probabilities at Grid Points 4, 5, and 6 calculated by: 1) 500 Monte Carlo samples; and 2) a 5-sample central difference first-order second-moment (FOSM) "Mean Value" method –see [14]).

Note that if we have sufficiently robust tools to determine the direction of the uncertainty gradient, or use probabilistic ordinal optimization, we can progress to the probabilistic optimum from the determinstic one if the conditions in the problem are like those just described. In some cases, like when the uncertainty gradient is also zero at the deterministic optimum, we may not be guaranteed that the probabilistic optimum (even if there is only one in the design space) can be found by starting from the deterministic optimum.

3.2.3 "Point of First Separation" Ordinal Selection Efficiency Mechanism. There are several situations in which determinstic-like efficiency in the probabilistic optimization problem can be had when the influence of the uncertainty gradient is overwhelmed by the influence of the underlying deterministic behavior. This is generally the case when the distance between compared design points is substantial, like in the global and early local stages of searching a design space. In this case a deterministic ordinal comparison will suffice to make effective optimization decisions in the probabilistic OUU problem. This is exemplified in Figure 9 where an ordinal comparison of deterministic safety margins at points 4, 5, and 6 in the design space yields a selection of the middle point, #5, which is pleasantly close to the probabilistic optimum. In fact, it is fairly common practice to initially treat a probabilistic optimization problem as a deterministic optimization problem to efficiently get an initial starting point for a hopefully short local search in the probabilistic optimization problem.

The influence of the uncertainty gradient being dominated by the underlying deterministic behavior is identically true when the uncertainty gradient is zero, *i.e.*, when the spatial correlation efficiency limit applies. In this case the OUU problem can be treated identically as a deterministic optimization problem.

When neither of the above situations apply, we must turn from a deterministic treatment of the OUU problem to one in which more than one correlated sample per design point may need to be taken for successful progress in the OUU problem. The nearest-to-deterministic treatment that can be applied is here coined the "*Point of First Separation*" ordinal selection method. It is an efficient implementation of the regular ordinal-optimization method that treats designs as completely uncorrelated, with truncation of the sampling at the earliest possible break point. As elaborated later, empirical experience on the present OUU problem was that the method is surprisingly efficient compared to the regular uncorrelated approach.

The method is explained with the help of Figure 10. Three design alternatives are being compared in the figure to determine which has the lowest probability of failure. This illustration concerns failure probability minimization instead of maximization as in the vulnerability searching problem, but the same principles apply whether seeking to maximize or minimize probability.

To start, a single correlated Monte Carlo sample of the response uncertainty of each design is evaluated. The sampled response of design (c) has a response value greater than the upper operating threshold level for acceptable system response. The first sample for the other two designs does not produce a threshold exceedence. On the assumption of closely spatially correlated sample realizations (i.e., the *n*th correlated Monte Carlo sample of each response distribution lies approximately at the same percentile on each distribution), this first sample already implies that alternative (c) will have a larger integrated failure probability than designs (a) and (b). Since here we want to identify the alternative with lowest failure probability, we can immediately eliminate design (c) from further consideration.

A second correlated sample is then taken from the designs still in contention. Neither of the remaining designs (a) and (b) indicate a failure (threshold exceedence) with this second sample. Of the two, alternative (a) has lower response values to this point, so would appear to be the better candidate from purely deterministic considerations. However, here we have a case where the influence of the uncertainty gradient is stronger than that of the underlying deterministic trend and dominates the determination of local failure probability magnitude. Hence, selection of the best alternative from a purely deterministic comparison would lead to an incorrect selection here.

The least expensive strategy that can often be successful under dominant influence of the uncertainty gradient is to continue the correlated sampling of the remaining designs until one or several become distinguished from the rest by a difference in response status relative to the applicable threshold. In the figure this occurs on the third correlated sample. The response value of Alternative (a) lies above the threshold, while that of Alternative (b) lies below the threshold. On the assumption of spatially correlated percentile realizations, this third sample implies that similar percentiles of response for Alternatives (a) and (b) lie respectively above and below the threshold. Therefore,

it is implied that Design (b) will have a lower integrated failure probability than Design (a).

In this contrived problem, it takes a total of 7 Monte Carlo samples apportioned among the three initial candidate designs to identify the correct design. With a different sample placement and ordering that would accompany a different RNG initial seed, a different total number of samples would be required. If, for instance, the sampling order in the figure is reversed, then with the first sample of each alternative (sample #3 in the figure), both designs (a) and (c) would be discarded immediately. Alternative (b) would be correctly selected at a total cost of only 3 samples. So, the total number of samples employed in this truncated version of the full probabilistic ordinal selection process will vary depending on the specific RNG starting seed. (This is true for full probabilistic ordinal optimization as well.)

Figure 11 shows some results of an empirical investigation of this "Point of First Separation" (PFS) approach. Calculated failure probabilities at points 4 and 5 of the 9-Point Grid (Figure 4) are shown versus number of correlated Monte Carlo samples of each design. After the third sample of each design, a separation in the calculated failure probabilities occurs. This separation wavers but generally grows as the sampling increases. Once the separation occurs, the calculated probability values never cross each other, so a correct ordinal selection can be made here just after the point of first separation. Any added sampling simply serves to decrease the confidence intervals about the calculated probability estimates, thereby increasing the P{CS} probability that the initial ordinal separation is correct. The correlated sampling shown in Figure 11 takes 101 samples each (202 total) to achieve a P{CS} of 95% under equal sampling of the two alternatives. Only 134 samples are required under optimal (non-equal) sample allocation among the two designs (see [16]).[3] In contrast, the point of first indication occurred after only four samples of each alternative (eight samples total), which is a reduction of almost 17X from the formal certainty of 95% P{CS} under the extreme assumption of completely uncorrelated designs.

The non-crossing of probability curves (once initially separated) that is exhibited in Figure 11 also occurs for all of 15 other two-alternative comparison tests that have been performed within the subspace defined by the 9-Point Grid. Though necessary and sufficient conditions for non-crossing after initial separation have not yet been rigorously formulated by the authors, the conditions do not appear to be very strict upon initial consideration. The conditions are met in the engineering OUU problem here, and may be met in many other OUU problems of engineering interest. For any problem, it appears that the more local the comparisons –as when the probabilistic optimum is being converged to in the final stage(s) of local optimization– the more similar the sampled response percentiles should be under correlated MC sampling, and therefore the better the chances of correct selection by the Point of First Separation mechanism.

As an empirical test of the PFS approach to the OUU problem in this paper, the procedure in [16] was essentially repeated in [18], but with sampling truncated to the PFS ordinal selection. For a total cost of 31 samples, the same final probabilistic optimum (maximum failure probability) was found as with formal 95% confidence in pairwise ordinal selections in [16]. The latter cost a total of 2721 samples. Thus, there was a cost reduction of 90X with PFS ordinal selection. This is not a fair comparison, however, because the 31 samples identified the correct final optimum, but only with a formal P{CS} of 9.7%. Very recent work (not yet published) used full OCBA in a second stage of sampling (as opposed to the less efficient pairwise OCBA used in [16]) to bring the P{CS} to 90%. This operation added 977 samples for a total of 1008 all together. Extrapolating the P{CS} convergence curve to a value of 95% yields a total about 1500 samples, as opposed to the 997 for a 90% P{CS}. Our current best estimate is therefore a total of 1531 samples to get an equivalent quality solution to that in [16] which didn't utilize spatial correlation or full OCBA. The cost savings with the more advanced ordinal optimization procedures is therefore estimated to be 44%. Actual final values and methodology will be presented in a future paper.

Significantly, the least-expensive non-ordinal OUU approach possible (one-sided-difference FOSM Mean-Value) would have cost a minimum of 30 samples to get to the final optimum, and the more trustworthy central-difference FOSM would have cost 50 samples total. These costs are applicable only if the FOSM OUU was as efficient as possible (no missteps in the design space due approximation error in the failure probability estimate), like the PFS ordinal selection method was. Hence, at best, the lowest-cost non-ordinal approaches could only be as inexpensive as the PFS ordinal method on this problem is, and still not yield an error or confidence estimate like the ~10% confidence indicator for the final PFS result. This is a good first indicator of the relative efficiency of the PFS approach, given that the application problem is an actual one and not contrived to be favorable to the method.

[3] Recall that these numbers are dependent on the RNG initial seed. A trial with a different seed required 386 (equal allocation) and 364 (optimal allocations) correlated MC samples to attain 95% P{CS}.

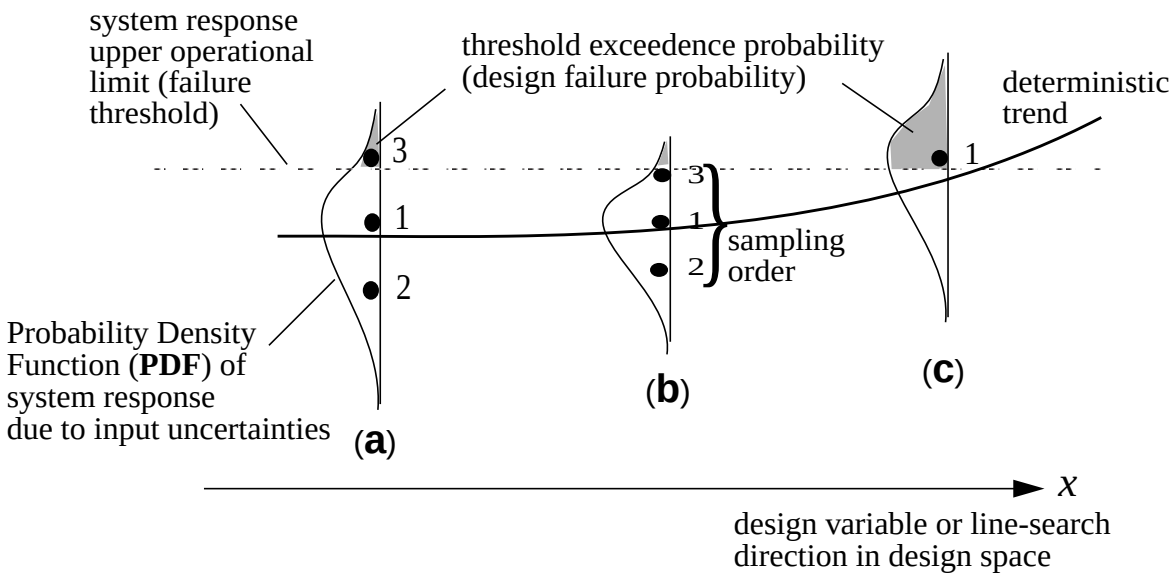

Figure 10. Correlated sampling of response behavior of three design alternatives at points (a), (b), and (c) in the design space.

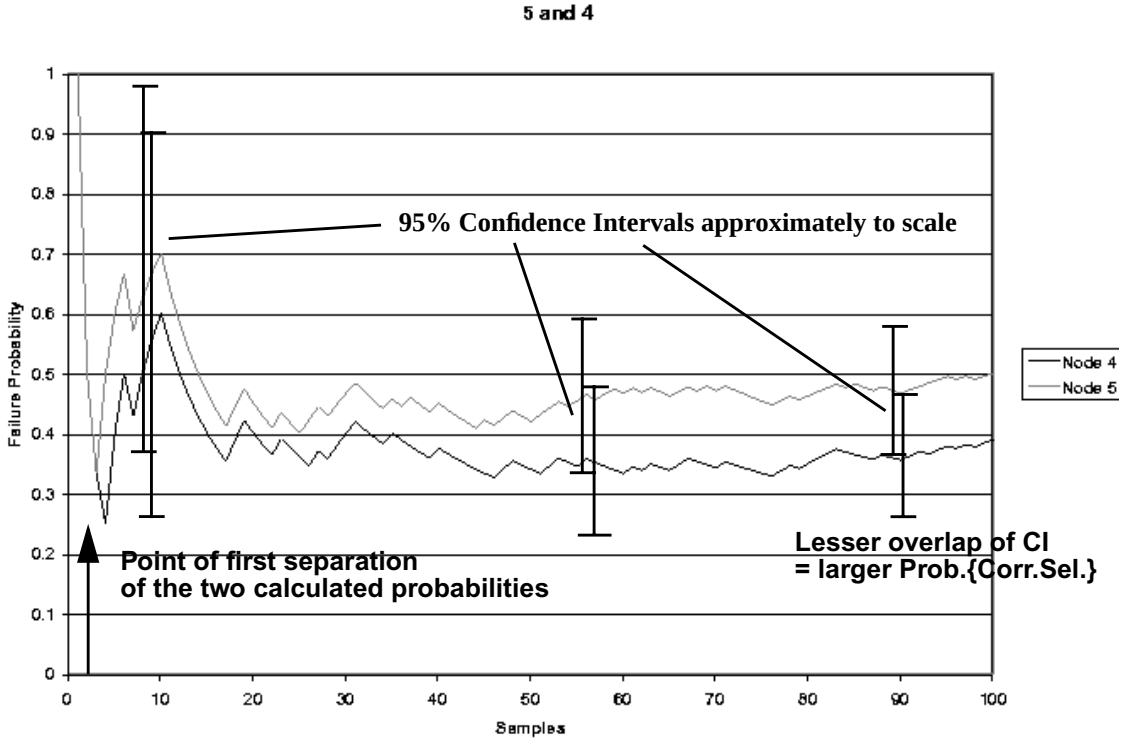

Figure 11. Failure Probabilities by correlated sampling of response behavior at design points 4 and 5 of the 9-Point Grid subspace (Figure 4).

4. Concluding Remarks

Several fundamental concepts of continuous-variable Ordinal optimization under uncertainty have been introduced here. These are: 1) *ordinal* selection or *ranking* of design alternatives based on statistical figures of merit of the probabilistic behaviors of the various designs; 2) the corresponding probability P{CS} of correctly selecting the best (or several best) design(s) among the available alternatives; 3) Optimal Computing Budget Allocation for optimally efficient sample allocation for maximizing P{CS} given a fixed number of total samples to be distributed among the alternatives, or for minimizing the total number of samples required to attain a stipulated P{CS} level; and 4) efficiencies to be gained from exploiting local spatial correlation of uncertainty in continuous-variable design problems.

Though only very simple and elementary implementations of ordinal optimization concepts have been discussed, much more sophisticated and efficient possibilities are foreseen. Certainly, there are many areas to be researched in the future involving: *i*) optimization constraints; *ii*) advanced ordinal optimizers for searching and progressing in the design space; and *iii*) extensions to specialized uncertainty sampling schemes such as Importance Sampling MC methods for small-probability values (*e.g.*, [6]), and advanced sampling schemes to account for epistemic uncertainties. The many possibilities have only just begun to be identified.

Ultimately, efficient implementations of probabilistic ordinal optimization may prove to be the reference "Gold Standard" method by which the efficiency and accuracy of other probabilistic optimization approaches can be compared and evaluated. This is analogous to the position that Monte Carlo sampling holds among the various uncertainty propagation approaches.

REFERENCES

[1] Alexandrov, N., Dennis, J.E. Jr., Lewis, R.M., and Torczon, V., "A trust region framework for managing the use of approximation models in optimization," *Structural Optimization*, Vol. 15, No. 1, February 1998, pp. 16-23.

[2] Chen, C-H, J. Lin, E. Yucesan, S.E. Chick, "Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization," *Journal of DEDS*, Vol.10, No. 3, pp. 251-270, July 2000.

[3] Eldred, M.S., Outka, D.E., Bohnhoff, W.J., Witkowski, W.P., Romero, V.J., Ponslet, E.J., and Chen, K.S., "Optimization of Complex Mechanics Simulations with Object-Oriented Software Design," in *Computer Modeling and Simulation in Engineering*, Vol. 1 No. 3, August, 1996, pp. 323-352.

[4] Eldred, M.S., Hart, W.E., Bohnhoff, W.J., Romero, V.J., Hutchinson, S.A., and Salinger, A.G., "Utilizing Object-Oriented Design to Build Advanced Optimization Strategies with Generic Implementation," Proceedings of the Sixth Multi-Disciplinary Optimization Symposium, Belleview, WA, Sept. 4-6 1996.

[5] Eldred, M.S., Giunta, A.A., Wojtkiewicz, S.F., Trucano, T.G., "Formulations for Surrogate-Based Optimization Under Uncertainty," paper AIAA-2002-5585, 9th AIAA/ISSMO Symposium of Multidisciplinary Analysis and Optimization, Atlanta, GA, Sept. 4-6, 2002.

[6] Harbitz, A., "Efficient and Accurate Probability of Failure Calculation by Use of the Importance Sampling Technique," Proc. of the 4th Int'l. Conf. on Application of Statistics and Probability in Soils and Structural Engineering, ICASP-4, Pitagora Editrice, Bologna, Italy, 1983.

[7] Hart, W. E., "Evolutionary Pattern Search Algorithms," Sandia National Laboratories report SAND95-2293, printed September 1995.

[8] Haugen, E.B., *Probabilistic Mechanical Design*, Wiley & Sons, 1980.

[9] Hough, P.D., T.G. Kolda, V.J. Torczon, "Asynchronous Parallel Pattern Search for Nonlinear Optimization," *SIAM J. Scientific Computing*, 23(1):134-156, June, 2001.

[10] Jones, D.R., C.D. Perttunen, and B.E. Stuckman, "Lipschitzian Optimization Without the Lipschitz Constant," Journal of Optimization Theory and Applications, Vol. 79, No. 1, October 1993.

[11] Levy, M.N., M.W. Trossett, and R.R. Kincaid, "Quasi-Newton Methods for Stochastic Optimization," 4th International Symposium on Uncertainty Modeling and Analysis (ISUMA'03), Sept. 21-24, 2003, University of Maryland, College Park, MD.

[12] Rodriguez, J.F., Renaud, J.D., and Watson, L.T., "Trust Region Augmented Lagrangian Methods for Sequential Response Surface Approximation and Optimization," ASME paper #97-DETC/DAC-3773, Proceedings of the ASME Design Engineering Technical Conference, ISBN 0-7918-1243-X, Sacramento, CA, Sept. 14-17, 1997.

[13] Romero, V. J., "Efficient Global Optimization Under Conditions of Noise and Uncertainty – A Multi-Model Multi-Grid Windowing Approach," in the Proceedings of the 3rd WCSMO (World Congress of Structural and Multidisciplinary Optimization) Conference, Amhearst, NY, May 17-21, 1999.

[14] Romero, V.J., "Characterization, Costing, and Selection of Uncertainty Probagation Methods for Use with Large Computational Physics Models," paper AIAA-2001-1653, 42nd Structures, Structural Dynamics, and Materials Conference, April 16-19, 2001, Seattle, WA. Updated revision available from the author.

[15] Romero, V.J., "Probabilistic Ordinal Optimization versus Surrogate-Based Methods for Industial Scale Optimization Under Uncertainty," presented at the SIAM Conference on

Optimization, May 20-22, 2002, Westin Harbour Castle Hotel, Toronto, Canada.

[16] Romero, V.J., D. Ayon, and C.-H. Chen, "Application of Probabilistic Ordinal Optimization Concepts to a Continuous-Variable Probabilistic Optimization Problem," Sandia National Laboratories library archive SAND2003-3671C, extended version of same-titled paper presented at the 4th International Symposium on Uncertainty Modeling and Analysis (ISUMA'03), Sept. 21-24, 2003, University of Maryland, College Park, MD.

[17] Romero, V.J., D. Ayon, and C.-H. Chen, "Demonstration of Probabilistic Ordinal Optimization Concepts on a Continuous-Variable Probabilistic Optimization Problem," submitted for *Optimization and Engineering* special issue of extended OUU papers from ISUMA'03 Conference (4th International Symposium on Uncertainty Modeling and Analysis, Sept. 21-24, 2003, University of Maryland, College Park, MD).

[18] Romero, V.J., "Efficiencies from Spatial Correlation of Uncertainty and Correlated Sampling in Continuous-Variable Ordinal Optimization Under Uncertainty," submitted to *AIAA Journal*.

[19] Romero, V.J., Eldred, M.S., Bohnhoff, W.J., and Outka, D.E., "Application of Optimization to the Inverse Problem of Finding the Worst-Case Heating Configuration in a Fire," Numerical Methods in Thermal Problems, Vol. IX, Part 2, pp. 1022-1033, Pineridge Press, R.W. Lewis and P. Durbetaki, eds. (Proceedings of the 9th Int'l. Conf. on Numerical Methods in Thermal Problems, Atlanta, GA., July 17-21, 1995.)

[20] Wan, Z., and Igusa, T., "Integration of a Trust Region Method with Response Surfaces," submitted for *Optimization and Engineering* special issue of extended OUU papers from ISUMA'03 Conference (4th International Symposium on Uncertainty Modeling and Analysis, Sept. 21-24, 2003, University of Maryland, College Park, MD).