

# Decision-theoretic foundations of simulation optimization

Peter I. Frazier

April 7, 2010

## Abstract

In simulation optimization, a task appearing frequently in applications, we wish to find a set of inputs to a simulator that cause its output to be maximal in some sense. Within any simulation optimization algorithm we must make a sequence of decisions about which input to test at each point in time. The problem of making this sequence of decisions so as to discover a near-optimal point as quickly as possible may be understood within a decision-theoretic framework. We describe this decision-theoretic framework in the context of two problems: ranking and selection and Bayesian global optimization. This decision-theoretic framework provides both a way of better understanding existing algorithms, and a path to developing new algorithms.

Simulation is an invaluable tool in operations research and management science. We may use it whenever we wish to manipulate a system to see what would happen, but actually performing the manipulation in reality would be too expensive, too risky, or simply impossible. Instead of manipulating the real system, we can build a simulation of the system and manipulate that instead.

In OR and MS, we often focus on stochastic simulation, and in particular on stochastic discrete event simulation (Law and Kelton, 2000). Important examples of systems simulated in this way include supply chains, queuing systems, manufacturing systems, and financial markets. The mathematical techniques that we will describe here may also be applied to simulations of systems of interest in engineering, medicine, and the natural sciences, e.g. simulations of the physical and chemical processes in a combustion engine.

Often, our goal in building a simulation is to optimize the real system. To do so, we may test a variety of possible actions using the simulation, find the one that performs best in simulation, and implement it in the real system. For example, we may simulate inventory policies in a supply chain to minimize expected costs, or we may simulate methods for repositioning ambulances in a city to maximize the number of calls answered in a timely fashion. Finding the inputs to a simulation that maximize some aspect of its output is called *simulation optimization*.

A related goal is *calibration* of a simulated system. In calibration, we search for parameter settings that cause the simulation to behave as similarly as possible to the real system. For example, we may search for a parametric model of credit risk in a financial market that causes predicted prices of liquidly traded securities to match market prices, or we may search for a representation of consumer behavior that cause simulated sales data to match what is observed in practice. Because we are searching for simulation inputs that cause its output to match a target as well as possible, calibration is also a simulation optimization problem.

In general, any simulation optimization problem may be written

$$\max_{x \in \mathcal{X}} \mathbb{E}[f(x, W)], \tag{1}$$

where  $f$  is a function implemented by the logic of the simulation (often highly complicated),  $W$  is random variable (often highly multivariate) generated by the random-number generator inside the simulation, and  $\mathcal{X}$  is the set of possible inputs to the simulation. In this article, we will consider two cases. In the first case,

called ranking and selection (R&S),  $\mathcal{X}$  a finite set. In the second case, called Bayesian Global Optimization (BGO)  $\mathcal{X}$  is a subset of  $\mathbb{R}^d$  and  $x \mapsto \mathbb{E}[f(x, W)]$  is a continuous function.

Simulation optimization differs from other types of optimization in that it does not rely on an analytic expression for  $\mathbb{E}[f(x, W)]$ . This makes simulation optimization more flexible than other types of optimization, but also more challenging. In most simulations,  $f$  is very complicated, and  $W$  is highly multivariate, and so an analytic expression is simply not available. Instead, the only way to estimate  $\mathbb{E}[f(x, W)]$  is to run the simulator many times with the same input  $x$  but different randomly generated values  $W$ , say  $W_1, \dots, W_n$ , and average the outputs  $f(x, W_1), \dots, f(x, W_n)$ . Accurately estimating  $\mathbb{E}[f(x, W)]$  for even a single value of  $x$  may require many simulations and a great deal of computer time.

The goal of the field of simulation optimization is to design algorithms that will solve (1) as quickly and accurately as possible. Simulations of highly complex systems often require significant amounts of time to generate samples of  $f(x, W)$ , and so an algorithm that solves the problem with fewer samples can save a significant amount of time. In situations with large search spaces  $\mathcal{X}$  or very time-consuming simulation, an efficient algorithm may make a solution to (1) available, where otherwise a solution would be impossible to find. For example, if a large discrete-event simulation takes 30 minutes of computer time to generate a single sample, then an algorithm that solves the problem to reasonable accuracy with 100 samples could be used in practice, while an algorithm that requires  $10^6$  samples would not (the first requires 50 hours of computer time, while the second requires more than 50 years).

At its heart, an algorithm is a sequence of decisions – decisions about which points  $x$  to sample with the simulation (“allocation or measurement decisions”), a decision of when to stop sampling (the “stopping decision”), and at the end, a decision of which point  $x^*$  to declare as an approximate solution (the “selection decision”). Thus, decision theory provides a framework in which to assess and compare the quality of different algorithms, and ultimately to guide us toward an optimal algorithm. In particular, since simulation optimization is concerned with the collection of information, it may be understood as a problem in the value of information (Howard, 1966, 1988).

A number of reviews have been written that consider decision theory applied to the decisions made in simulation, and many of these consider simulation optimization. Many of these reviews, e.g., Chick (2000, 2006), introduce decision theory to readers already familiar with simulation, while Merrick (2009) introduces problems in simulation to readers already familiar with decision theory. Also, we do not consider here the larger field of simulation optimization separated from its decision-theoretic aspects. For such broader discussion, see, e.g., Andradóttir (1998); Swisher et al. (2000); Fu (2002); Spall (2003).

In this article, we begin by describing the R&S problem in a Bayesian context. This problem is canonical, being simple enough to allow a thorough analysis, while at the same time exemplifying the issues at the core of all simulation optimization problems. We then discuss a worst-case analysis of R&S called the indifference-zone formulation, which is historically more prominent than the Bayesian formulation. We conclude by discussing the BGO problem, in which we optimize a continuous function over a continuous domain and must deal with an infinite number of alternatives.

## 1 Bayesian Ranking and Selection

The decision-theoretic understanding of simulation optimization is best introduced in the context of the (R&S) ranking and selection problem. In this problem, a collection of distinct “alternatives,” numbered 1 through  $k$ , comprises our set  $\mathcal{X}$  of possible inputs to the simulation. We might think of the alternatives as possible configurations of a manufacturing process, or as inventory policies that might be used in a supply chain.

Associated with each alternative  $x$  is a sampling distribution, which is the distribution of  $f(x, W)$ . In R&S, our goal is to maximize  $\mathbb{E}[f(x, W)]$ , i.e., find the alternative whose sampling distribution has the largest mean. We often assume that the sampling distributions are normal, which is justified by supposing that,

when samples are not normal, we average many together to obtain approximately normal batch averages. When the sampling distribution is normal, it is completely determined by its mean and variance. Let  $\theta_x$  and  $\lambda_x$  be the mean and variance respectively of the sampling distribution for alternative  $x$ . Let  $\theta = (\theta_1, \dots, \theta_k)$  and  $\lambda = (\lambda_1, \dots, \lambda_k)$ . These values are unknown a priori, and simulation is the tool through which we may learn about them.

Whenever we choose to simulate alternative  $x$ , we observe a sample from this sampling distribution. In R&S, we often assume that samples are independent over time and across the alternatives. Let  $n$  index the number of samples observed, and let  $x_n$  be the alternative that is observed at time  $n$ . We may choose  $x_n$  based on previous samples, and it is this choice or decision, along with the stopping decision, that is the main focus within R&S. For each  $n \geq 1$  we observe a value  $y_n$  that results from sampling alternative  $x_n$ . It has distribution

$$y_n \sim \text{Normal}(\theta_{x_n}, \lambda_{x_n}) \mid x_n, \theta, \lambda,$$

and is independent of all previous allocation decisions and samples, i.e., of  $x_m$  and  $y_m$ ,  $m < n$ .

By taking more samples, we learn more about the sampling distributions and we improve our ability to find the best one. However, taking more samples takes more time. In some formulations of the R&S problem, the amount of time that we may spend sampling is constrained. In others, we control how much time we spend sampling, but sampling incurs an explicit cost that we must balance against the value of the information obtained. In this section we will suppose an explicit cost  $c > 0$  for each sample taken.

Let  $\tau$  be the number of samples taken. Like the decision of which alternatives to sample, the decision to stop taking samples may be made adaptively, based on the samples observed so far. After we stop sampling, we select an alternative for implementation based on the accumulated evidence. Call the alternative selected  $x^*$ . Our selection decision may depend on  $x_1, \dots, x_\tau$  and  $y_1, \dots, y_\tau$ . As stated, we wish to select the alternative with the largest sampling mean,  $\arg \max_x \theta_x$ . When we make a mistake and we fail to select the one with the largest  $\theta_x$ , we suffer a loss  $L(x^*, \theta)$ , where  $L$  is our loss function. Common loss functions are the linear loss function  $L(x^*, \theta) = \max_x \theta_x - \theta_{x^*}$ , and the 0–1 loss function  $L(x^*, \theta) = \mathbf{1}_{\{x^* \notin \arg \max_x \theta_x\}}$ . One could also consider a loss function depending on both  $\theta$  and  $\lambda$ , although commonly this is not considered. In addition to this loss associated with an incorrect selection, we also pay the cost of sampling, which is  $c\tau$ . Thus, the total loss is  $L(x^*, \theta) + c\tau$ .

Let  $\pi$  be a generic set of rules for making the decisions  $\tau$ ,  $(x_n)_{n \leq \tau}$  and  $x^*$ . We call this set of rules a *policy*. It must obey the previously stated rules governing the information upon which decisions may be based. These rules require that the decision  $\mathbf{1}_{\{\tau=n\}}$  to stop and the decision  $x_{n+1}$  of which alternative to sample next (when not stopping) depend upon only  $x_{1:n} = (x_1, \dots, x_n)$  and  $y_{1:n} = (y_1, \dots, y_n)$ . They also require that the selection decision  $x^*$  depend upon only  $x_{1:\tau}$  and  $y_{1:\tau}$ . Given a policy  $\pi$  obeying these rules, and fixed values  $\theta$  and  $\lambda$ , the expected total loss suffered is

$$R^\pi(\theta, \lambda) = \mathbb{E}^\pi [L(x^*, \theta) + c\tau \mid \theta, \lambda],$$

where  $\mathbb{E}^\pi$  indicates that when the expectation is taken, the values  $\tau$ ,  $(x_n)_{n \leq \tau}$  and  $x^*$  are chosen according to the policy  $\pi$ .

If we knew the true  $\theta$  and  $\lambda$ , then we would prefer the policy  $\pi$  with the smallest value of  $R^\pi(\theta, \lambda)$ . However, we do not know these values at all. If we did, we could simply choose  $x^* \in \arg \max_x \theta_x$  directly without sampling at all, and achieve 0 loss every time. Since we do not know  $\theta$  and  $\lambda$ , we instead need a method for combining performance across multiple values of  $\theta$  and  $\lambda$  to establish a preference order over policies.

In the Bayesian formulation we place a prior probability distribution on the unknown quantities,  $\theta$  and  $\lambda$ . In some cases, this is justified because the simulation optimization problem at hand really is drawn at random from some larger set of simulation optimization problems with which we have a great deal of experience and whose distribution we can characterize. One such case might be the calibration of a financial model of credit risk based on current market prices for liquidly traded credit derivatives. The model must be calibrated each day to a new set of market prices, and because the calibration is performed repeatedly, we

may estimate the overall distribution of truths  $\theta$  and  $\lambda$  in the simulation optimization problems encountered, and use this as our prior. Much more frequently, however, the prior probability distribution encodes our subjective belief about  $\theta$  and  $\lambda$ . This is in keeping with Bayesian methods as used throughout decision theory (see, e.g., Berger (1985)). If one has very little information concerning  $\theta$  and  $\lambda$ , it is also possible to specify a non-informative prior distribution.

Given a prior distribution, the Bayes risk (expected total loss under the prior) of a policy  $\pi$  is

$$\mathbb{E}[R^\pi(\theta, \lambda)] = \mathbb{E}^\pi [L(x^*, \theta) + c\tau],$$

where the first expectation is over the truth  $\theta, \lambda$  as drawn from the prior, and the second expectation is over both  $\theta, \lambda$  and the intrinsic randomness introduced by the sampling noise. This first expectation lacks a  $\pi$  in the superscript because it is taken only over the prior on  $\theta, \lambda$ , which does not depend upon the policy (however, note that  $R^\pi(\theta, \lambda)$  does depend on  $\pi$ ), while the second expectation has a  $\pi$  in the superscript because the distribution of  $x^*$  and  $\tau$  depends upon the policy.

If a non-informative prior distribution is specified, then the expectation  $\mathbb{E}^\pi [L(x^*, \theta) + c\tau]$  may not exist. In such situations, one may perform a first stage of measurements in which each alternative is measured an equal number of times to produce a proper posterior belief. One may then take this resulting belief and use it as the prior under which we formulate the decision-theoretic problem.

Given this prior, whether it is the original prior or the belief resulting from an earlier noninformative prior and a first stage of measurements, we value a policy according to its Bayes risk. One policy is better than another if it has smaller Bayes risk, and an optimal policy  $\pi$  is one that attains minimal Bayes risk,

$$\inf_{\pi \in \Pi} \mathbb{E}^\pi [L(x^*, \theta) + c\tau]. \quad (2)$$

Here,  $\Pi$  is the set of policies considered, and may be the set of all policies, or may be a strict subset, e.g., the set of two-stage policies described in Section 1.2 below. Finding a policy  $\pi$  that attains this infimum is one of the ultimate goals in Bayesian R&S.

Beginning with a prior probability distribution, sampling results in a posterior distribution that combines the prior with evidence from sampling. This posterior may be calculated using Bayes' rule. Given a set of samples and the resulting posterior distribution, the conditional expected loss of a given selection decision  $x^*$  at time  $\tau$  is

$$\mathbb{E}[L(x^*, \theta) \mid x_{1:\tau}, y_{1:\tau}].$$

One can show that the best implementation decision  $x^*$  is any from the set  $\arg \min_x \mathbb{E}[L(x, \theta) \mid x_{1:\tau}, y_{1:\tau}]$ . For the linear loss function, this is the one with the largest posterior mean, but for other loss functions, including the 0 – 1 loss function, this is not necessarily true (Berger and Deely, 1988). Nevertheless, given a posterior distribution obtained by a given policy, choosing  $x^*$  is relatively straightforward.

If we fix the rule for choosing  $x^*$  to the optimal method, we may rewrite (2) as

$$\inf_{\pi \in \Pi} \mathbb{E}^\pi \left[ \min_{x^*} \mathbb{E}[L(x^*, \theta) \mid x_{1:\tau}, y_{1:\tau}] + c\tau \right]. \quad (3)$$

Within this expression,  $\min_{x^*} \mathbb{E}[L(x^*, \theta) \mid x_{1:\tau}, y_{1:\tau}]$  is the best that we can do given the information we have collected by time  $\tau$ ,  $c\tau$  is the cost we pay for this information, and  $\mathbb{E}^\pi [\min_{x^*} \mathbb{E}[L(x^*, \theta) \mid x_{1:\tau}, y_{1:\tau}] + c\tau]$  is the overall expected net loss incurred by sampling according to the policy  $\pi$ . By subtracting this from the value of taking no samples at all, one obtains  $\min_{x'} \mathbb{E}[L(x', \theta)] - \mathbb{E}^\pi [\min_{x^*} \mathbb{E}[L(x^*, \theta) \mid x_{1:\tau}, y_{1:\tau}] + c\tau]$ , which is the net value of the information collected by  $\pi$ .

The challenge in designing algorithms for Bayesian R&S is choosing rules for making allocation and stopping decisions whose expected loss is as close to this minimal value (3) as possible. In theory, such rules may be found using dynamic programming (see Chick and Gans (2009); Frazier et al. (2008) for dynamic programming formulations of R&S), but the amount of computational power required to actually find them far exceeds what is currently available. Instead, a number of heuristic rules have been suggested in the literature. We discuss these after introducing the most commonly used prior distributions.

## 1.1 Choice of Prior Probability Distribution

Bayesian R&S is most convenient if the prior resides within a conjugate family for the sampling distribution. The most commonly considered case is

$$\theta_x \mid \lambda_x \sim \text{Normal}(\mu_{0x}, \lambda_x/t_{0x}), \quad \lambda_x^{-1} \sim \text{Gamma}(a_{0x}, b_{0x}),$$

which is parameterized by  $\mu_{0x}$ ,  $t_{0x}$ ,  $a_{0x}$ ,  $b_{0x}$ ,  $x = 1, \dots, k$ . With this prior and a sequence of samples, the posterior is

$$\theta_x \mid \lambda_x, x_{1:n}, y_{1:n} \sim \text{Normal}(\mu_{nx}, \lambda_x/t_{nx}), \quad \lambda_x^{-1} \mid x_{1:n}, y_{1:n} \sim \text{Gamma}(a_{nx}, b_{nx}),$$

where the parameters for the unmeasured alternatives remain unchanged from  $n$  to  $n+1$ , and the parameters for the measured alternative,  $x = x_{n+1}$ , can be calculated recursively as

$$\begin{aligned} \mu_{n+1,x} &= (t_{nx}\mu_{nx} + y_{n+1}) / (t_{nx} + 1), & t_{n+1,x} &= t_{nx} + 1, \\ b_{n+1,x} &= (y_{n+1} - \mu_{nx})^2 t_{nx} / (2t_{nx} + 2), & a_{n+1,x} &= a_{nx} + \frac{1}{2}. \end{aligned}$$

Details, derivations, and the limiting non-informative case may be found in DeGroot (1970), Sections 9.6 and 10.2.

If  $\lambda$  is known, then the prior-posterior analysis simplifies. The priors and posteriors, for  $n \geq 0$ , are  $\theta_x \sim \text{Normal}(\mu_{nx}, \lambda_x/t_{nx})$ , and the parameters of the posterior are updated recursively for  $x = x_{n+1}$  as  $\mu_{n+1,x} = (t_{nx}\mu_{nx} + y_{n+1}) / (t_{nx} + 1)$ , and  $t_{n+1,x} = t_{nx} + 1$ .

In both cases, as we continue to sample an alternative  $x$ , our posterior belief about its sampling mean concentrates around its true sampling mean  $\theta_x$ . The mean of our belief,  $\mu_{nx}$ , converges to  $\theta_x$ , and the variance of our belief shrinks to 0.

## 1.2 Allocation and Stopping Rules

In the framework discussed so far, the decisions made at time  $n+1$  (the measurement decisions  $x_{n+1}$  and the decision of whether to stop or not i.e., whether to set  $\tau = n+1$ ) may be made based upon *all* of the information  $x_{1:n}, y_{1:n}$  available.

In some situations, it may be useful to restrict the set of information upon which a decision may be based. This is done for two reasons. The first is to run simulations in parallel – if we have multiple processors, we would like to start several simulations at once without waiting for the results of the last simulation to decide which alternative to simulate next. The second reason is simplicity of analysis. Complex adaptive policies are usually more difficult to analyze and develop than those with more restrictive information dependencies.

Policies are classified according to the complexity of their information dependence. The simplest policies are those in which the alternatives to sample,  $x_{1:\tau}$  are chosen up front. In this case, the  $x_n$  and  $\tau$  are deterministic quantities. These are called “one-shot”, “single-stage” or “deterministic” policies. One can often obtain better performance than with one-shot policies by taking a fixed number of samples from each alternative in what is called a “first stage”, and then allocating a second stage of measurements based on the results. These are called “two stage policies.” Three-stage and multi-stage policies are defined similarly. Finally, a policy that uses the results from all previous samples to make each measurement decision is called *fully sequential*. As we increase the number of stages and decrease the number of measurements per stage down to the limit established by the number of parallel processors, we allow greater efficiency but must contend with more complicated policies.

In the special case of  $k = 2$  alternatives, the optimal one-shot policy was found in Gupta and Miescke (1996). This one-shot policy was shown to be optimal among the larger class of fully sequential policies with fixed measurement horizons in Frazier et al. (2008). If the prior variances and measurement variances are

equal, this optimal policy is the one that allocates an equal number of measurements to each alternative. For more than 2 alternatives, however, finding optimal single-stage allocation policies is non-trivial, and finding optimal multi-stage and fully sequential policies is even more difficult.

We briefly review heuristic policies developed within the Bayesian decision-theoretic framework. Many earlier papers considered allocation rules in isolation, without considering the stopping rule. Using a decision-theoretic analysis and an asymptotic approximation, Chick and Inoue (2001) introduce the LL and 0-1 allocation rules (for the linear loss and 0 – 1 loss functions respectively), in both one-shot, two-stage, and multi-stage forms. In a different line of research also motivated by Bayesian methods, a number of allocation rules have been developed under the collective name of optimal computing budget allocation (OCBA) in Chen (1996); Chen et al. (2000); He et al. (2007) and other works. A fully sequential allocation rule derived using a one-step heuristic in a decision-theoretic dynamic-programming framework with known sampling variance was introduced in Gupta and Miescke (1996) and analyzed more fully in Frazier et al. (2008), where it was noted that the policy was one a larger collection of myopic methods called knowledge-gradient (KG) methods. This allocation rule was extended to handle unknown sampling-variance in Chick et al. (2009) and given the name LL(1). A variety of adaptive stopping rules derived within a Bayesian framework were introduced in Branke et al. (2005); Frazier and Powell (2008); Chick and Frazier (2009), and Branke et al. (2005) showed that good adaptive stopping rules perform much better than does stopping at a fixed time. Chick and Gans (2009) considered both allocation and stopping rules for a natural extension of the R&S problem that includes discounting in time of the implementation payoff. A large empirical study of many of these policies was conducted in Branke et al. (2007).

## 2 Other Formulations

We have focused on R&S for pedagogical reasons — it is simple enough to be easily and thoroughly explained while still exhibiting the core characteristics shared by all simulation optimization problems. We have further focused on Bayesian R&S because modern decision-theory tends to be more interested in Bayesian methods. Despite this focus on Bayesian R&S, there are other rigorous and useful formulations of the simulation optimization problem. Here we briefly describe two of them, the indifference-zone formulation of the R&S problem, and Bayesian global optimization.

### 2.1 Indifference-Zone Ranking and Selection

The historical weight of the R&S literature employs a worst-case analysis known as the indifference zone (IZ) formulation. Although worst-case analyses have well-known drawbacks that have been highlighted within the decision-theoretic literature (e.g., Berger 1985 Section 5.5), the IZ formulation represents an important approach within R&S.

In the IZ formulation of R&S, the decision-maker’s goal is to implement the best alternative with a probability that exceeds a lower bound whenever the true configuration falls within a particular set of configurations. In detail, the IZ formulation begins by defining our loss function as

$$L(x^*, \theta) = \mathbf{1}_{\{x^* \in \arg \max_x \theta\}},$$

and the conditional risk of a policy is

$$R^\pi(\theta, \lambda) = \mathbb{P}^\pi \left\{ x^* \in \arg \max_x \theta_x \mid \theta, \lambda \right\}.$$

This is also known as the *probability of incorrect selection*. The quantity  $\text{PCS}^\pi(\theta, \lambda) = 1 - R^\pi(\theta, \lambda)$  is known as the *probability of correct selection*.

Then, rather than valuing a policy  $\pi$  according to its average-case performance under a prior, we specify a fixed parameter  $\delta > 0$  and consider the policy’s worst-case performance over the set of configurations  $\text{PZ}(\delta)$

defined by

$$\text{PZ}(\delta) = \{\theta \in \mathbb{R}^k : \theta_{[1]} - \theta_{[2]} \geq \delta\},$$

where  $[1], \dots, [k]$  are a permutation of the indices  $1, \dots, k$  so that  $\theta_{[1]} \geq \dots \geq \theta_{[k]}$ . This set of configurations is called the *preference zone*. In addition to  $\delta$ , we also specify a level  $\alpha \in (0, 1 - 1/k)$ , often .01 or .05, and require that  $\pi$  satisfies

$$\inf_{\theta \in \text{PZ}(\delta), \lambda \in \mathbb{R}_+^k} \text{PCS}^\pi(\theta, \lambda) \geq 1 - \alpha.$$

If a policy satisfies this, then we say that it satisfies the indifference-zone guarantee for  $\alpha$  and  $\delta$ .

Whenever a configuration falls within the preference zone  $\text{PZ}(\delta)$ , the difference between best and all other alternatives is at least  $\delta$ , and the decision-maker is said to prefer this best alternative. Whenever the configuration falls outside the preference zone, the difference between the best and second best alternatives is small enough that the decision-maker is said to be indifferent to selecting the best. The indifference-zone guarantee then guarantees that the decision-maker selects the best with probability at least  $1 - \alpha$ , unless the configuration is such that he is indifferent. A more recent and stringent version of this formulation proposed in Nelson and Banerjee (2001) additionally requires that the probability of selecting a “good” alternative (one whose sampling mean is within  $\delta$  of the best) exceeds  $1 - \alpha$  over *all* configurations  $\theta \in \mathbb{R}^k$ .

A large literature exists on the IZ formulation of the R&S problem. This literature develops policies that satisfy this indifference-zone guarantee while at the same time taking as few samples as possible. This literature dates to the 1950s and the seminal works of Bechhofer (Bechhofer, 1954; Bechhofer et al., 1968). For a more recent monograph see Bechhofer et al. (1995), and for a current survey of the literature, see Kim and Nelson (2006, 2007).

Relative to Bayesian methods, the theoretical guarantee that IZ procedures offer is often quite appealing. This guarantee comes at a price however, because IZ methods tend to be quite conservative, often taking many more samples than is strictly required to meet the guarantee for most configurations. This conservatism has two sources. First is the need to protect against extremely unfavorable configurations, even if these are seldom met in practice. This type of conservatism is intrinsic to the worst-case analysis used by the IZ formulation. Second, existing procedures often over-sample, even under the most unfavorable configurations. This is because current methods for proving the IZ guarantee rely on inequalities that are often quite loose, especially for large numbers of alternatives. Recent efforts including Nelson et al. (2001); Kim and Nelson (2001); Goldsman et al. (2002) have greatly reduced but not eliminated this second type of conservatism. Thus, if one desires a well-understood statistical guarantee on the performance of the R&S procedure used, and is willing to spend some extra time sampling, then IZ methods are appropriate. If instead one desires good performance on average, and is willing to use a procedure whose performance is not as well characterized theoretically, then Bayesian methods are appropriate. For further discussion and extensive numerical comparisons between Bayesian and IZ procedures, see Branke et al. (2007).

## 2.2 Bayesian Global Optimization

R&S makes no assumptions about the relationship between different alternatives. In that problem, learning the value of one alternative tells us nothing about the value of other alternatives. This is a consequence of the independence of the alternatives’ true sampling means under the prior. In many applications of simulation optimization, however, the alternatives *are* related to each other. For example, in the classic problem of finding the best  $(s, S)$  inventory policy (Law and Kelton, 2000), two inventory policies with similar values for  $s$  and  $S$  usually perform similarly.

One can incorporate such information about the relationship between alternatives into the prior. One may replace the independent prior used in R&S with a prior under which the values of the alternatives are correlated. Conceptually, this is an easy extension of the Bayesian formulation of the R&S problem, but it presents many numerical and algorithmic challenges.

When the number of alternatives is finite, as it would be with the inventory policy example, one can use a multivariate normal prior. This case is considered in Frazier et al. (2009). When the space of alternatives is a multi-dimensional continuous space, as it would be when optimizing several continuous parameters, and the true sampling mean is a continuous function of these parameters, one can place a Gaussian process prior (see, e.g., Rasmussen and Williams (2006)) on the unknown sampling mean. The resulting problem is known as Bayesian Global Optimization problem (BGO).

In BGO, we have an infinite number of alternatives, but we have a characterization of their relationships to each other that allows efficient search for an optimum. Much of the literature on BGO assumes that the simulation is deterministic, but some of the literature considers stochastic simulation. Deterministic simulation is considered in Žilinskas (1975); Mockus (1989); Žilinskas (1992); Schonlau (1997); Jones et al. (1998); Schonlau et al. (1998); Sasena (2002), and stochastic simulation in Williams et al. (2000); Huang et al. (2006). For broader issues in the use of Bayesian methods for inference on continuous functions produced as output from a simulator, see Kennedy and O’Hagan (2001); Santner et al. (2003); OHagan (2006).

A broad literature exists on global optimization beyond that which is based in decision-theory. Much of this literature assumes that function evaluations are free from noise – see the recent survey Floudas and Gounaris (2008) and also the recent book Zhigljavsky and Žilinskas (2008), both of which discuss noise-free BGO methods in this broader context. For a discussion of methods designed to handle noise in function evaluations see the surveys Andradóttir (1998); Swisher et al. (2000); Fu (2002); Spall (2003) mentioned in the introduction. Relative to other methods of global optimization, BGO requires more computational effort per measurement to decide which points to sample, but tends to find a good solution with fewer measurements. This is discussed and demonstrated numerically on a few example problems in Huang et al. (2006). Thus, BGO tends to be more useful for optimizing complex long-running simulations, while other techniques are preferable for short-running simulations.

### 3 Conclusion

Simulation optimization is a difficult and important problem that is ubiquitous in OR, MS, and other fields. Decision theory offers us a way to understand how decisions should be made within simulation optimization algorithms. Better algorithms and deeper understanding of existing algorithms have resulted from the application of decision theory to these problems.

At the same time, many exciting challenges remain. Computing the optimal policy for all but the simplest problems, especially in sequential settings, remains practically impossible. In addition, existing work on simulation optimization has only begun to consider many aspects of the way that simulation optimization problems present themselves in practice, including aspects like risk aversion, the relationship between alternatives, steady-state simulations, common random numbers, and the difference between the simulation’s output and the behavior of the real system. By understanding the decision-theoretic foundations of simulation optimization, we may more easily understand the proper response to these novel aspects of the problem, and more easily design the algorithms of the future.

### References

- Andradóttir, S. (1998). Simulation optimization. In *Handbook of simulation: Principles, methodology, advances, applications, and practice*, pages 307–333. Wiley-Interscience. 2, 8
- Bechhofer, R. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *The Annals of Mathematical Statistics*, 25(1):16–39. 7
- Bechhofer, R., Kiefer, J., and Sobel, M. (1968). *Sequential Identification and Ranking Procedures*. University of Chicago Press, Chicago. 7

- Bechhofer, R., Santner, T., and Goldsman, D. (1995). *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons*. J.Wiley & Sons, New York. 7
- Berger, J. and Deely, J. (1988). A Bayesian Approach to Ranking and Selection of Related Means With Alternatives to Analysis-of-Variance Methodology. *Journal of the American Statistical Association*, 83(402):364–373. 4
- Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis*. Springer-Verlag, New York, second edition. 4, 6
- Branke, J., Chick, S., and Schmidt, C. (2005). New developments in ranking and selection: an empirical comparison of the three main approaches. In Kuhl, M., Steiger, N., Argstrong, F., and Joines, J., editors, *Proc. 2005 Winter Simulation Conference*, pages 708–717, Piscataway, NJ. IEEE, Inc. 6
- Branke, J., Chick, S., and Schmidt, C. (2007). Selecting a selection procedure. *Management Sci.*, 53(12):1916–1932. 6, 7
- Chen, C. (1996). A lower bound for the correct subset-selection probability and its application to discrete-event system simulations. *IEEE Transactions on Automatic Control*, 41:1227–1231. 6
- Chen, C., Lin, J., Yücesan, E., and Chick, S. (2000). Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems*, 10(3):251–270. 6
- Chick, S. (2000). Bayesian methods: bayesian methods for simulation. *Proceedings of the 32nd conference on Winter simulation*, pages 109–118. 2
- Chick, S. (2006). Bayesian ideas and discrete event simulation: why, what and how. In *Proceedings of the 37th conference on Winter simulation*, pages 96–105. Winter Simulation Conference. 2
- Chick, S., Branke, J., and Schmidt, C. (2009). Sequential sampling to myopically maximize the expected value of information. *INFORMS J. on Computing*. to appear. 6
- Chick, S. and Frazier, P. (2009). The conjunction of the knowledge gradient and economic approach to simulation selection. *Winter Simul. Conf. Proc., 2009*. 6
- Chick, S. and Gans, N. (2009). Economic analysis of simulation selection problems. *Management Sci.*, to appear. 4, 6
- Chick, S. and Inoue, K. (2001). New two-stage and sequential procedures for selecting the best simulated system. *Operations Research*, 49(5):732–743. 6
- DeGroot, M. H. (1970). *Optimal Statistical Decisions*. John Wiley and Sons. 5
- Floudas, C. and Gounaris, C. (2008). A review of recent advances in global optimization. *Journal of Global Optimization*, pages 1–36. 8
- Frazier, P. and Powell, W. (2008). The knowledge-gradient stopping rule for ranking and selection. *Winter Simul. Conf. Proc., 2008*. 6
- Frazier, P., Powell, W. B., and Dayanik, S. (2008). A knowledge gradient policy for sequential information collection. *SIAM J. on Control and Optimization*, 47(5):2410–2439. 4, 5, 6
- Frazier, P., Powell, W. B., and Dayanik, S. (2009). The knowledge gradient policy for correlated normal beliefs. *INFORMS Journal on Computing*, 21(4):599–613. 8
- Fu, M. (2002). Optimization for simulation: Theory vs. practice. *INFORMS Journal on Computing*, 14(3):192–215. 2, 8
- Goldsman, D., Kim, S., Marshall, W., and Nelson, B. (2002). Ranking and selection for steady-state simulation: Procedures and perspectives. *INFORMS Journal on Computing*, 14(1):2–19. 7

- Gupta, S. and Miescke, K. (1996). Bayesian look ahead one-stage sampling allocations for selection of the best population. *Journal of statistical planning and inference*, 54(2):229–244. 5, 6
- He, D., Chick, S., and Chen, C. (2007). Opportunity cost and ocba selection procedures in ordinal optimization for a fixed number of alternative systems. *IEEE Transactions on Systems Man and Cybernetics Part C-Applications and Reviews*, 37(5):951–961. 6
- Howard, R. (1966). Information Value Theory. *Systems Science and Cybernetics, IEEE Transactions on*, 2(1):22–26. 2
- Howard, R. (1988). Decision analysis: practice and promise. *Management Science*, 34(6):679–695. 2
- Huang, D., Allen, T., Notz, W., and Zeng, N. (2006). Global Optimization of Stochastic Black-Box Systems via Sequential Kriging Meta-Models. *Journal of Global Optimization*, 34(3):441–466. 8
- Jones, D., Schonlau, M., and Welch, W. (1998). Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13(4):455–492. 8
- Kennedy, M. and O’Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464. 8
- Kim, S. and Nelson, B. (2006). *Handbook in Operations Research and Management Science: Simulation*, chapter Selecting the best system. Elsevier. 7
- Kim, S. and Nelson, B. (2007). Recent advances in ranking and selection. In *Proceedings of the 39th conference on Winter simulation: 40 years! The best is yet to come*, pages 162–172. IEEE Press. 7
- Kim, S.-H. and Nelson, B. L. (2001). A fully sequential procedure for indifference-zone selection in simulation. *ACM Trans. Model. Comput. Simul.*, 11(3):251–273. 7
- Law, A. M. and Kelton, W. D. (2000). *Simulation Modeling and Analysis*. McGraw-Hill, New York, third edition. 1, 7
- Merrick, J. (2009). Bayesian Simulation and Decision Analysis: An Expository Survey. *Decision Analysis*. 2
- Mockus, J. (1989). *Bayesian approach to global optimization: theory and applications*. Kluwer Academic, Dordrecht. 8
- Nelson, B. and Banerjee, S. (2001). Selecting a good system: Procedures and inference. *IIE Transactions*, 33(3):149–166. 7
- Nelson, B., Swann, J., Goldsman, D., and Song, W. (2001). Simple procedures for selecting the best simulated system when the number of alternatives is large. *Operations Research*, 49(6):950–963. 7
- O’Hagan, A. (2006). Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety*, 91(10-11):1290–1300. 8
- Rasmussen, C. and Williams, C. (2006). *Gaussian Processes for Machine Learning*. MIT Press. 8
- Santner, T., Williams, B. W., and Notz, W. (2003). *The Design and Analysis of Computer Experiments*. Springer. 8
- Sasena, M. (2002). *Flexibility and Efficiency Enhancements for Constrained Global Design Optimization with Kriging Approximations*. PhD thesis, University of Michigan. 8
- Schonlau, M. (1997). *Computer experiments and global optimization*. PhD thesis, University of Waterloo. 8
- Schonlau, M., Welch, W., and Jones, D. (1998). *New Developments and Applications in Experimental Design*, volume 34, chapter Global versus local search in constrained optimization of computer models, pages 11–25. Institute of Mathematical Statistics. 8

- Spall, J. (2003). *Introduction to stochastic search and optimization: estimation, simulation, and control*. John Wiley and Sons. 2, 8
- Swisher, J., Hyden, P., Jacobson, S., Schruben, L., Hosp, M., and Fredericksburg, V. (2000). A survey of simulation optimization techniques and procedures. *Simulation Conference Proceedings, 2000. Winter*, 1, 2, 8
- Žilinskas, A. (1975). Single-step Bayesian search method for an extremum of functions of a single variable. *Cybernetics and Systems Analysis*, 11(1):160–166. 8
- Williams, B., Santner, T., and Notz, W. (2000). Sequential design of computer experiments to minimize integrated response functions. *Statistica Sinica*, 10:1133–1152. 8
- Zhigljavsky, A. and Žilinskas, A. (2008). *Stochastic global optimization*. Springer. 8
- Žilinskas, A. (1992). A review of statistical models for global optimization. *Journal of Global Optimization*, 2(2):145–153. 8