VALUE OF INFORMATION METHODS FOR PAIRWISE SAMPLING WITH CORRELATIONS

Peter I. Frazier
Jing Xie

Operations Research & Information Engineering
Cornell University
Rhodes Hall
Ithaca, NY, 14853, U.S.A.

Stephen E. Chick

Technology & Operations Management Area
INSEAD
Boulevard de Constance
77300 Fontainebleau, FRANCE

ABSTRACT

We consider optimization via simulation over a finite set of alternatives. We employ a Bayesian value-of-information approach in which we allow both correlated prior beliefs on the sampling means and correlated sampling. Correlation in the prior belief allow us to learn about an alternative's value from samples of similar alternatives. Correlation in sampling, achieved through common random numbers, allows us to reduce the variance in comparing one alternative to another. We allow for a more general combination of both types of correlation than has been offered previously in the Bayesian ranking and selection literature. We do so by giving an exact expression for the value of information for sampling the difference between a pair of alternatives, and derive new knowledge-gradient methods based on this valuation.

1 INTRODUCTION

In this article we consider optimization via simulation over a finite set of alternatives. We seek to allocate simulation effort efficiently across the alternatives to find the alternative with the largest sampling mean. We describe this effort in the context of ranking and selection (R&S) but our goal is ultimately to handle many alternatives as with discrete optimization via simulation (DOvS). The novelty of this paper, which reports initial progress toward this goal, is twofold. First, we extend the knowledge-gradient/value of information approaches to Bayesian R&S by allowing for a broader set of sequential sampling plans – those that sample from two alternatives in a given stage of sampling, whereas all prior work assumes only one alternative is sampled in a given stage of sampling. Second, we allow for both correlated prior beliefs about the unknown means of each alternative and correlation in samples, as might result from common random numbers (CRN). Prior work has focused on one or the other or has made stronger assumptions about the correlation or sampling structure than we believe may be desirable in applications.

We employ a Bayesian approach, in which we begin with a prior distribution on the values of the sampling means, updating this prior distribution based on sampling information, and use value of information computations to decide how to best allocate sampling effort. We allow a correlated prior distribution on the values of the unknown sampling means and correlated sampling for the simulation output. Either of these correlations can decrease the number of samples needed to find a good alternative.

Allowing a correlated prior distribution allows us to take advantage of relationships between alternatives. When performing optimization via simulation in a structured problem, for example in finding the best staffing level in a large hospital, when we learn about the value of one alternative (in the example, the quality of care at a particular staffing level), this suggests that similar alternatives (in the example, staffing levels with similar numbers of staff in each unit) are likely to have similar values. A correlated prior distribution allows this information to be used by an optimization algorithm and thereby improve efficiency.

Indeed, with a correlated prior distribution, one can perform well even when the number of samples taken is smaller than the number of alternatives, because we can learn about an unsampled alternative from

3979

samples of related alternatives. This opens the door to performing optimization via simulation over very large-scale problems, where the number of alternatives is much larger than the sampling budget. This is the approach, for example, in Bayesian global optimization for continuous (Brochu, Cora, and de Freitas 2009) and discrete problems (Negoescu, Frazier, and Powell 2011), in which correlated prior distributions are used to handle an infinite or combinatorially large number of alternatives. This use of correlated prior distributions contrasts the current approach with most algorithms for ranking and selection (R&S), which require at least one sample per alternative to perform well.

Although the emphasis on the use of correlated prior distributions differentiates the current work from much of the R&S literature, there is some previous work in Bayesian R&S employing correlated prior distributions (Chick and Inoue 2001, Frazier, Powell, and Dayanik 2009).

Strengths of Chick and Inoue (2001) include the ability to model both types of correlation (correlated prior distribution for unknown means and correlates samples with CRN) and that the covariance matrix can be inferred through sampling. Limitations of Chick and Inoue (2001) include a restriction to two-stages of sampling, and the requirement that the prior distribution of the unknown means and variances be a conjugate prior for the sampling distribution. Below, we model both types of correlation, allow for a broader class of prior distributions for the unknown mean that have been useful in the context of Bayesian global optimization (but which typically assumes independent sampling), but do not yet fully model inference of the covariance.

Frazier et al. (2009) allow a general correlated normal prior distribution on the sampling means, as we do, but assumes independent sampling. They call their approach the knowledge gradient (KG) approach because it assesses the expected value of information in a given increment of sampling. The value of information computation used by Frazier et al. (2009) has a scaling issue as the number of alternatives increases. This scaling issue is addressed by Scott, Frazier, and Powell (2011), which continues to assume independent sampling, and focuses on optimization of continuous functions.

Fu, Hu, Chen, and Xiong (2007) have also considered Bayesian R&S with CRN, but assume an independent prior distribution on the sampling means. Such independence causes the number of measurements required to find good alternatives to scale roughly as the number of alternatives, and requires the number of measurements to be at least as large as the number of alternatives. An independent prior distribution on the sampling means precludes potential benefits for sampling efficiency with a correlated prior distribution.

There are a number of non-Bayesian indifference-zone methods that allow CRN, including the two stage procedures of Clark and Yang (1986) and Nelson and Matejcik (1995). CRN have also been used for frequentist formulations of the sequential multiple comparisons problem (Yang and Nelson 1991, Nakayama 2000, Kim 2005), and CRN can be very effective at reducing sampling requirements.

In summary, this paper presents new techniques to handle both correlated means and correlated samples within the KG approach. It does so by examining the value of information of sampling plans that sample from more than one alternative at a time (unlike other KG work). This preliminary work indicates that the approach holds promise and identifies areas for further work to handle large-scale decision problems.

2 SAMPLING MODEL AND POSTERIOR INFERENCE

We suppose that we have a collection of k alternatives, and that if we sample from all k alternatives together using the same common random number seed, then we observe a normal random vector. Let the mean vector of this normal distribution be $\theta = (\theta_1, \dots, \theta_k)$, and let its covariance matrix be Λ. In practice neither θ nor Λ is known. In this paper we will assume that θ is unknown and that Λ is known fully for simplicity. Our interest here is in determining which alternative x has the largest sampling mean θ_x.

We use a Bayesian formulation, in which we begin with a multivariate normal prior on θ,

$$\theta \sim \mathcal{N}(\mu_0, \Sigma_0).$$

The ability to include correlation in the prior through the presence of a non-diagonal covariance matrix Σ_0 has important practical implications for the algorithm we propose, because it allows us to learn about

alternatives that we have not sampled from observations of other, similar, alternatives. We envisage the choice of Σ_0 in particular applications to allow for conjugate prior distributions (Chick and Inoue 2001) or more general kriging-type models (Brochu et al. 2009).

We assume that sampling proceeds as follows. At each point in time $n = 1, 2, \ldots$ we choose a subset of the alternatives $\{1, \ldots, k\}$ to sample, and sample each of the chosen alternatives once, using a stream of common random numbers. We then observe a vector Y_n with one entry for each alternative sampled. To describe the distribution of Y_n, we construct a matrix X_n that has one row for each alternative sampled, where each row has a value of 1 at the index corresponding to its alternative, and a value of 0 elsewhere. With X_n defined in this way, the distribution of Y_n is

$$Y_n \sim \mathcal{N}\left(X_n\theta, X_n\Lambda X_n'\right), \tag{1}$$

where $'$ indicates matrix transposition.

The sampling distribution in (1) also holds for more general matrices X_n. For example, if we observe only the *difference* between two alternatives, then we can take X_n to have a single row with a value of 1 at one of these alternatives and a value of -1 at the other alternative. We use this fact in Section 3 while computing the value of sampling information.

At each point in time n, we are assumed to choose a new CRN seed so that, conditioned on X_n and θ, Y_n is independent of the past. Although one can imagine more complex sampling schemes in which we may return to previous CRN seeds to sample alternatives not yet sampled with that seed, the sampling scheme we propose here has the advantage of simplicity, while still allowing considerable freedom. This simplicity allows more explicit value of information calculations, and reduces the implementation complexity and memory requirements of implementations of the algorithm we propose.

We will also consider the sampling distribution if we observe not just one but some fixed number of samples β_n from the distribution given by (1), each from a different independently chosen CRN seed. In this case, a sufficient statistic for θ is the average of the β_n observed random vectors, which we continue to call Y_n, and is distributed according to

$$Y_n \sim \mathcal{N}\left(X_n\theta, X_n\Lambda X_n'/\beta_n\right). \tag{2}$$

We allow the X_n and β_n to be chosen adaptively, so that X_n and β_n may depend upon all previous observations. We define $D_n = (X_1, \beta_1, Y_1, \ldots, X_n, \beta_n, Y_n)$, and this adaptivity requirement can be understood as requiring X_{n+1} and β_{n+1} to be functions of D_n.

2.1 Posterior Distribution

With this sampling scheme (and the assumption that the sampling covariance matrix Λ is known) we can compute a closed-form expression for the posterior distribution on θ. We let $\mathbb{E}_n$ and Var_n indicate the conditional expectation and variance respectively with respect to D_n. Define $\mu_n = \mathbb{E}_n\theta$ and $\Sigma_n = \mathrm{Var}_n\theta$. One can show using standard results from Bayesian linear regression (see, e.g., Gelman, Carlin, Stern, and Rubin 2004, Section 14.6) that the conditional distribution of θ given D_n is normal,

$$\theta \mid D_n \quad \sim \quad \mathcal{N}(\mu_n, \Sigma_n),$$

and that the mean μ_n and variance Σ_n of this multivariate normal distribution can be computed recursively,

$$\mu_n = \Sigma_n \left[\Sigma_{n-1}^{-1}\mu_{n-1} + X_n'\Lambda_n^{-1}Y_n\right]$$

$$\Sigma_n = \left[\Sigma_{n-1}^{-1} + X_n'\Lambda_n^{-1}X_n\right]^{-1},$$

where $\Lambda_n = X_n\Lambda X_n'/\beta_n$. Using the Sherman-Morrison-Woodbury formula (see, e.g., Golub and Van Loan 1996), we can re-express Σ_n and μ_n without inverting the large matrix Σ_{n-1} as

$$\Sigma_n = \Sigma_{n-1} - \Sigma_{n-1}X_n'\left(\Lambda_n + X_n\Sigma_{n-1}X_n'\right)^{-1}X_n\Sigma_{n-1},$$

$$\mu_n = \left[I_k - \Sigma_{n-1}X_n'\left(\Lambda_n + X_n\Sigma_{n-1}X_n'\right)^{-1}X_n\right]\mu_{n-1} + \Sigma_nX_n'\Lambda_n^{-1}Y_n, \tag{3}$$

where I_k is the $k \times k$ identity matrix. Although this expression does require inverting Λ_n and $\Lambda_n + X_n \Sigma_{n-1} X_n'$, these two matrices have a dimension equal to the number of rows of X_n, which is usually quite small. In our algorithm, they are either 1×1 or 2×2.

These and similar updates are used extensively for Bayesian linear regression and Kalman filtering, and many computational techniques exist for computing μ_n and Σ_n more rapidly and accurately than is allowed by naive application of these equations (Rasmussen and Williams 2006).

2.2 Predictive Distribution

Later, we will also need the predictive distribution for μ_{n+1} that results from a particular decision to sample X_{n+1}. Standard computations (Gelman et al. 2004) show that this predictive distribution is normal. The tower property of conditional expectation shows that its mean is μ_n, and the conditional variance formula shows that its variance is $\Sigma_n - \Sigma_{n+1}$ (note that Σ_{n+1} is known given D_n and X_{n+1}), which can be computed directly from (3). By this argument, the predictive distribution can be written,

$$\mu_{n+1} \mid D_n, X_{n+1} \quad \sim \quad \mathcal{N}\left(\mu_n, \Sigma_n X_{n+1}' \left(\Lambda_{n+1} + X_{n+1} \Sigma_n X_{n+1}'\right)^{-1} X_{n+1} \Sigma_n\right).$$

In Section 3 we will be particularly interested in the predictive distribution in the case when X_{n+1} is a $1 \times k$ matrix. In this case, both Λ_{n+1} and $X_{n+1} \Sigma_n X_{n+1}'$ are scalars, and the covariance matrix of the predictive distribution can be decomposed into the product $\tilde{\sigma}_n(X_{n+1}, \beta_{n+1}) \tilde{\sigma}_n(X_{n+1}, \beta_{n+1})'$, where we define the $k \times 1$ vector $\tilde{\sigma}_n(X, \beta)$ by

$$\tilde{\sigma}_n(X, \beta) = \Sigma_n X' / \sqrt{\beta^{-1} X \Lambda X' + X \Sigma_n X'}. \tag{4}$$

This will be used in Section 3.1 to compute the value of information explicitly.

3 VALUE OF INFORMATION

In this section we compute the value of information (VOI) resulting from two different types of sampling. These VOI calculations are then used to derive a sampling algorithm in Section 4.

VOI is a concept introduced by Howard (1966), in which information is valued according to the improvement it produces in some decision to be made later. In our problem, the general form of the VOI for a set of β samples collected by observing Y_{n+1} with matrix X according to (2) can be written,

$$V(X, \beta) = \mathbb{E}_n \left[\max_x \mu_{n+1,x} \mid X_{n+1} = X, \beta_{n+1} = \beta \right] - \max_x \mu_n. \tag{5}$$

For computational convenience we will employ an approximation in which we restrict the maximum over alternatives to only those alternatives in a set $A \subseteq \{1, \ldots, k\}$. This approximation to the full VOI can be written,

$$V(X, A, \beta) = \mathbb{E}_n \left[\max_{x \in A} \mu_{n+1,x} \mid X_{n+1} = X, \beta_{n+1} = \beta \right] - \max_{x \in A} \mu_{n,x}. \tag{6}$$

When $A = \{1, \ldots, k\}$ then $V(X, A, \beta) = V(X, \beta)$. This technique of using a subset A to approximate the VOI was also used by Scott et al. (2011) in a related context.

3.1 Exact Computation of the Value of Information

When X is a $1 \times k$ vector we can compute (6) in closed form by recalling the definition of the $k \times 1$ vector $\tilde{\sigma}_n(X, \beta)$ in (4). In this case, when D_n, X_{n+1} and β_{n+1} are given, we observe that μ_{n+1} is equal in distribution to $\mu_n + \tilde{\sigma}_n(X_{n+1}, \beta_{n+1}) Z$, where Z is an independent scalar standard normal random variable. Using this observation in (6) shows that

$$V(X, A, \beta) = \mathbb{E}_n \left[\max_{x \in A} \mu_{n,x} + \tilde{\sigma}_n(X, \beta) Z \right] - \max_{x \in A} \mu_{n,x}.$$

When A has exactly one element, one can show using the tower property of conditional expectation that $V(X,A,\beta) = 0$.

When A has exactly two elements, as would be the case for existing KG formulations when a single alternative is sampled in a given stage, computation of $V(X,A,\beta)$ is straightforward. In this case, let Δ be the absolute value of the difference of $\mu_{n,x}$ between the two different $x \in A$, and let s be the absolute value of the difference of the two components of $\tilde{\sigma}_n(X,\beta)$ corresponding to the $x \in A$. Then,

$$V(X,A,\beta) = sf(-\Delta/s),$$

where $f(-z) = \varphi(z) - z\Phi(-z)$, and φ and Φ are the normal density and cumulative distribution functions. When z is large, $f(-z)$ can be computed accurately using an asymptotic approximation to Mills' ratio from Gordon (1941), $\Phi(-z)/\varphi(z) \approx z/(z^2+1)$, so that $f(-z) \approx \varphi(z)/(z^2+1)$. In Matlab implementations, we use this approximation whenever $z > 10$.

When A contains more than two elements, as is the case for our new KG proposal when two alternatives are sampled in a given stage, computation of $V(X,A,\beta)$ is more involved, but still can be performed analytically. Given vectors a and b of equal size, let

$$h(a,b) = \mathbb{E}\left[\max_x a_x + b_x Z\right] - \max_x a_x.$$

Thus, $V(X,A,\beta) = h(\mu_{n,A}, \tilde{\sigma}_{n,A}(X,\beta))$, where $\mu_{n,A}$ and $\tilde{\sigma}_{n,A}(X,\beta)$ are $|A| \times 1$ vectors composed of elements of $\mu_{n,x}$ and $\tilde{\sigma}_{n,x}(X,\beta)$, respectively, for alternatives $x \in A$. Frazier et al. (2009) gives an algorithm for computing h and the matlabKG library provides a Matlab implementation of it (Frazier 2010).

3.2 Non-Myopic Value of Information

The VOI as calculated using (5) or (6) corresponds to a fixed batch of samples, and is exact only if this is the last set of samples collected before selecting the best. In our problem, sampling usually continues after the results of these samples are observed. To account for this, one can use dynamic programming to calculate a sequential VOI (Chick and Frazier 2009a), but this is computationally infeasible in many cases.

As a computationally tractable alternative to dynamic programming, we consider the average VOI earned per unit of computation time when taking β independent copies of X,

$$V(X,A,*) = \max_{\beta \geq 1} \frac{V(X,A,\beta)}{\beta c(X)}, \tag{7}$$

where $c(X)$ is the computational cost of taking the samples implied by X, measured by the computation time required. With this understanding, $V(X,A,*)$ is the best average rate at which VOI can be earned by taking a fixed number of samples according to X, assuming a selection must be made after those samples are observed. It thus extends the KG to allow for flexible-duration one-stage sampling with correlation.

More that one sample is considered because VOI is not concave in the number of samples (Frazier and Powell 2010), and $V(X,A,1)$, can be nearly zero even though $V(X,A,\beta)/\beta$ is much larger. A similar maximization of the VOI per sample over β was used successfully for Bayesian R&S with independent samples and prior distributions (Chick and Frazier 2009b, Frazier and Powell 2010), and for multi-armed bandit problems (Ryzhov, Frazier, and Powell 2010).

A simple formulation of $c(X)$ is to assume homogeneous computation time for each sample, i.e., $c(X) = c \cdot \sum_{i,j} |X_{ij}|$, where c is a positive constant cost per sample. We adopt this assumption below.

4 SAMPLING ALGORITHM

Using the VOI calculations from Section 3, we first propose two new sampling algorithms that can be used when both the sampling covariance matrix and the sampling costs are known. The first, KG_β^2, is

described in Section 4.2. It extends the usual KG formulation by considering sampling stages where pairs of alternatives are sampled rather than only one. The second, KG_*^2, is described in Section 4.3 and extends the KG_β^2 policy by using the non-myopic VOI calculation proposed in Section 3.2. We continue to assume that the sampling covariance matrix and sampling costs are known.

4.1 Generic Sampling Algorithm

A generic sequential sampling algorithm assumes the choice of both an allocation rule that chooses which alternatives to sample, and a stopping rule that chooses when to stop sampling. Below in Section 4.2 - 4.3 we describe particular allocation and stopping rules (KG_β^2 and KG_*^2, which are new in this paper, and the previously proposed KG_β and KG_* rules). They are used in the following generic sequential sampling algorithm.

1. Choose allocation and stopping rules, as well as the mean vector μ_0 and covariance matrix Σ_0 of the prior distribution for the unknown means. Specify the sampling covariance Λ. Let $n = 0$.
2. If the stopping rule says to stop sampling, go to Step 6. If it does not, go to Step 3.
3. Use the allocation rule to choose a set of alternatives to sample next and a number of times β_{n+1} to sample them. Let X_{n+1} be the corresponding sampling matrix as described in (2).
4. Sample Y_{n+1} using CRN according to (2) with the chosen X_{n+1} and β_{n+1}.
5. Calculate the parameters μ_{n+1} and Σ_{n+1} of the new posterior distribution using (3) with Y_{n+1}, X_{n+1}, β_{n+1}, μ_n, and Σ_n. Increment n and go to Step 2.
6. Select as the best the alternative with the largest μ_{nx}.

4.2 KG_β^2 Allocation and Stopping Rules

To describe our new KG allocation and stopping rules, which consider the VOI in a single stage of sampling from two alternatives, we describe the notion of a KG factor for the VOI in a given increment of sampling. For a given fixed $\beta \geq 1$ (often $\beta = 1$) and for each $x \in \{1, \ldots, k\}$, we calculate the matrix X corresponding to a single sample from alternative x. The *single-alternative KG factor* is defined by

$$v^{KG}(x) = V(X, A(x), \beta)/(\beta c(X)), \tag{8}$$

where the set $A(x)$ contains x and the best other alternative. That is, $A(x) = \left\{ x, \text{argmax}_{x' \neq x} \mu_{nx'} \right\}$.

For each $x^{(1)}, x^{(2)}$, we calculate the matrix X corresponding to observing the difference in value between these two alternatives. We now define for the first time the *pairwise-difference* KG *factor*,

$$v^{KG}\left(x^{(1)}, x^{(2)}\right) = V\left(X, A\left(x^{(1)}, x^{(2)}\right), \beta\right) \Big/ (\beta c(X)), \tag{9}$$

where the set $A\left(x^{(1)}, x^{(2)}\right) = \left\{ x^{(1)}, x^{(2)}, \text{argmax}_{x' \neq x^{(1)}, x^{(2)}} \mu_{nx'} \right\}$ contains $x^{(1)}, x^{(2)}$, and the best other alternative.

The KG_β **allocation rule** allocates one sample (even if $\beta > 1$) to the alternative that maximizes the single-alternative KG factor. In the KG_β **stopping rule**, sampling stops if and only if all of the single-alternative KG factors drop below 1. When a KG factor corresponding to sampling design X drops below 1, $V(X, A, \beta)/(\beta c(X)) < 1$, implying that the (myopic) VOI $V(X, A, \beta)$ obtained is less than the cost $\beta c(X)$ required to obtain it. The KG_β allocation rule with $\beta = 1$ and $c = 1$ was introduced in Gupta and Miescke (1996), and later studied in greater detail in Frazier, Powell, and Dayanik (2008). The KG_β stopping rule with $\beta = 1$ and $c = 1$ was introduced in Frazier and Powell (2008). They were generalized to allow different values of β and c in Chick and Frazier (2009a).

The KG_β^2 **allocation rule**, which is new in this paper, looks over both the single-alternative KG factors and the pairwise-difference KG factors to find the largest one. If a single-alternative KG factor is maximal, it takes a single sample from this alternative (even if $\beta > 1$). If a pairwise-difference KG factor is maximal,

it takes one sample from each of these two alternatives. The KG_β^2 **stopping rule** is also new in this paper. In this stopping rule, sampling stops when all of the single-alternative and pairwise-difference KG factors drop below 1.

The KG_β^2 allocation and stopping rules are formally defined by the preceding paragraphs. Computing the many pairwise differences is an expensive computational step of KG_β^2 if k is large, and it may happen that a sampling correlation is negative for a given pair to be sampled. We now examine two adaptations of the KG_β^2 concept above that address these two issues.

First, we observe that there are $k(k-1)/2$ pairwise-difference KG factors and that computing each one requires calculating a VOI over a set A with 3 entries. Computing these KG factors is the most expensive step in the KG_β^2 allocation rule. To make this step faster, one may calculate the pairwise-difference KG factors for only some pairs of alternatives. In our implementation, we order the alternatives according to their single-alternative KG factors, and then take only the pairs of alternatives where the first alternative is among the top k_1 alternatives, and the second alternative is among the top k_2 alternatives.

Second, in situations where some entries in the sampling covariance Λ are negative, the KG_β^2 allocation and stopping rules can be improved by considering independent sampling rather than correlated sampling for these pairs of alternatives. This is shown by the following lemma. This proof of this lemma will appear in a later journal version of this paper.

Lemma 1 Consider an arbitrary pair of alternatives $(x^{(1)}, x^{(2)})$. Let $\mu_n, \Sigma_n, \Lambda_{x^{(1)}x^{(1)}}, \Lambda_{x^{(2)}x^{(2)}}$ be fixed, and let the matrix X correspond to observing the difference in value between $x^{(1)}$ and $x^{(2)}$. Then for any arbitrary A and β, $V(X, A, \beta)$ and $V(X, A, *)$ are increasing functions of the sampling correlation between $x^{(1)}$ and $x^{(2)}$, i.e., $\rho(x^{(1)}, x^{(2)})$.

To take advantage of the insight offered by this lemma, when calculating pairwise-difference KG factors we replace all of the strictly negative entries in the sampling covariance matrix Λ by 0. Then, if a pair of alternatives whose sampling covariance was replaced by 0 is selected for simulation by the KG_β^2 allocation rule, we use independent sampling rather than CRN to simulate these alternatives.

In numerical experiments below we use these speed-ups, and therefore require k_1 and k_2 to be specified. In summary, the KG_β^2 **allocation rule** is approximated by the following implementation.

1. For each alternative x, calculate the single-alternative KG factor using (8).
2. For each unique pair of alternatives $(x^{(1)}, x^{(2)})$ with $x^{(1)}$ among the top k_1 alternatives, and $x^{(2)}$ among the top k_2 alternatives according to their single-alternative KG factors, calculate the pairwise-difference KG factor using (9). If this pair has a negative sampling correlation, replace this negative correlation with 0.
3. Find the largest KG factor among those computed in Steps 1 and 2.
   (a) If the largest is a single-alternative KG factor, and x is the corresponding alternative, let X_{n+1} correspond to measuring x and $\beta_n = \beta$.
   (b) If the largest is a pairwise-difference KG factor, and $(x^{(1)}, x^{(2)})$ is the corresponding pair, let X_{n+1} correspond to measuring *both* $x^{(1)}$ and $x^{(2)}$ (not just their difference), and let $\beta_n = \beta$. If $x^{(1)}$ and $x^{(2)}$ originally had a negative sampling correlation, then direct the sampling algorithm to sample them independently (otherwise sample them using CRN).

The KG_β^2 **stopping rule** is approximated by the following algorithm in numerical tests and also uses the parameters k_1 and k_2 to reduce computation time by focusing on the pairwise-difference KG factors that are hoped to be of the greatest magnitude.

1. Calculate the single-factor and pairwise-difference KG factors as in Steps 1-2 of the KG_β^2 allocation.
2. If any of these KG factors is at least 1, continue sampling. Otherwise stop sampling.

When $k_1 = k_2 = k$, these approximations to the KG_β^2 allocation and stopping rules are exact.

When used by itself, without the KG_β^2 allocation rule, the KG_β^2 stopping rule can be streamlined by checking whether each KG factor exceeds 1 as it is computed. As soon as the first KG factor exceeding 1 is found, the stopping rule can cease computation and report the decision to continue sampling. It is only when every KG factor is strictly less than 1 that the KG_β^2 stopping rule requires that we compute all of the KG factors. If, however, the KG_β^2 stopping rule is used together with the KG_β^2 allocation rule (which computes all of the KG factors), then all of the KG factors will be available.

4.3 KG_*^2 Allocation and Stopping Rules

The KG_β and KG_β^2 allocation and stopping rules in Section 4.2 above assess the VOI relative to a given increment in sampling, β. Thus, the KG factors in (8) and (9) may depend upon β.

The KG_* **allocation and stopping rules** take advantage of this by evaluating the single-alternative KG factors in (8) with respect to the β that maximizes $v^{KG}(x)$ for each individual x (Chick and Frazier 2009a). More formally, the KG_* allocation rule assesses the VOI per unit effort with

$$v^{KG,*}(x) = \max_{\beta \geq 1} V(X, A(x), \beta) / (\beta c(X)). \tag{10}$$

instead of with $V(X, A, \beta)/(\beta c(X))$ for some fixed β, then allocates one sample to the alternative x that maximizes $v^{KG,*}(x)$. The KG_* stopping rule checks if any $V(X, A(\cdot), *)$ for single-alternative KG factors is at least one. This can result in a considerable improvement in selection efficiency, as compared to KG_β for a given fixed $\beta \geq 1$ (Chick and Frazier 2009b, Frazier and Powell 2010).

Here, we use the same principle of checking different values of β for a given pairwise VOI comparison in an attempt to obtain even greater efficiency. With this in mind, new KG_*^2 allocation and stopping rules are proposed that perform like the KG_β^2 allocation and stopping rules, except that their determination of KG factors uses the most favorable β for each individual VOI calculation. That is, the KG_*^2 **allocation rule** uses the single-alternative KG factors in (10) and the pairwise-difference KG factors it uses are determined by generalizing (9) to use the most favorable budget for each potential pair of alternatives to sample,

$$v^{KG,*}\left(x^{(1)}, x^{(2)}\right) = \max_{\beta \geq 1} V\left(X, A\left(x^{(1)}, x^{(2)}\right), \beta\right) / (\beta c(X)). \tag{11}$$

The sampling plan that maximizes $v^{KG,*}(x)$ or $v^{KG,*}\left(x^{(1)}, x^{(2)}\right)$ receives one sample for each alternative in the KG_*^2 allocation rule.

Similarly, the KG_*^2 **stopping rule** requires that sampling stop if and only if all $v^{KG,*}(x)$ and $v^{KG,*}\left(x^{(1)}, x^{(2)}\right)$ are less than one. This occurs if and only if there is no one-stage sampling plan of one alternative, or of two alternatives that receive the same number of samples, that results in a greater expected VOI than the cost to obtain the samples.

The above statements suffice to formally define the KG_* and KG_*^2 allocation and stopping rules. As suggested in Section 4.2, we will use approximations to speed up the computation of the allocations. In particular, the numerical experiments (i) uses the two speed-ups from Section 4.2 that consider only a subset of pairs of alternatives to simulate, and that sample independently from pairs if sampling with CRN would have a negative correlation, and (ii) only consider a finite subset $B()$ of budgets β rather than searching the continuum $\beta \geq 1$ in (10) and (11).

The finite set of budgets is taken to be of the form $B(b, \bar{\beta}) = \{\bar{\beta}^{a/b} : a = 0, 1, 2, \ldots, b\}$ to assess single-alternative KG factors and of the form $B(b, \bar{\beta}/2)$ to assess pairwise-difference KG factors (so that the total number of samples is the same in each assessment). Here, b is a parameter that controls the number of values of β to check, and $\bar{\beta}$ is the maximum value of β to test.

5   NUMERICAL RESULTS

This section explores the performance of the new KG^2_β and KG^2_* allocation and stopping rules relative to the existing KG_β and KG_* allocation and stopping rules in illustrative example problems. We report on two main experiments.

The first main experiment compares the new KG^2_β and KG^2_* rules with their predecessors, KG_β and KG_*, to assess the benefit of checking pairs of alternatives as opposed to the more-usual single alternative checks of KG factors, and the role of correlations. In this experiment, each procedure matched the allocation with its corresponding stopping rule so they use the same KG factor criterion. For example, the KG^2_β allocation is used only with the KG^2_β stopping rule in this experiment. For KG^2_β and KG_β we set $\beta = 1$.

We measure performance by the expected total penalty $\mathbb{E}[\text{OC} + c \cdot \tau]$ of not knowing the true means, where $\text{OC} = \max_x \theta_x - \max_x \mu_{\tau,x}$ is the opportunity cost when sampling stops and τ is the stopping time. This expected total penalty is computed as an expectation over random problem instances, where the true means θ are drawn from the same prior distribution used by the sampling algorithm.

Figure 1 shows the results with $k = 2$ alternatives in which we plot $\mathbb{E}[\text{OC} + c\tau]$ as the sampling correlation ρ ranges from 0 to 1. In this experiment, we used an independent prior $\Sigma_0 = 10^8 \times [1,0;0,1]$, with prior mean $\mu_0 = [0,0]'$, sampling covariance $\Lambda = 10^{10} \times [1,\rho;\rho,1]$, and sampling cost $c = 10$. We set $k_1 = k_2 = k$, $b = 30$, and the total number of samples initially available $\bar{\beta}$ to be 1000. We performed 5×10^4 replications, and the resulting standard error was less than 32.52 in all cases.

The figure shows that both KG^2_β and KG^2_* improve as the sampling correlation increases, as these policies take advantage of the correlation between the two alternatives by sampling the pair together to obtain low-variance observations of their difference. In contrast, the KG_β and KG_* methods are not able to sample pairs, and so cannot take advantage of the sampling correlation.

Across all values of the correlation, KG^2_* is better than KG^2_1, and KG_* is better than KG_1. We believe that this is due largely to the stopping rule. Because they consider many possible budgets instead of just one, the KG^2_* and KG_* values of sampling information are larger than the corresponding KG^2_1 and KG_1 values ($V(X,A,*) \geq V(X,A,1)/c(X)$ by (7)), and are closer to the true non-myopic VOI of sampling. This causes the KG^2_* and KG_* stopping rules to sample longer, substantially reducing the opportunity cost while only marginally increasing sampling costs. As evidence, we describe the stopping times of $\mathbb{E}[\tau]$ of the four algorithms. For KG_1, KG_*, and KG^2_1, $\mathbb{E}[\tau]$ is relatively flat across the difference values of sampling correlation, and falls in the ranges $[6.39, 6.48]$, $[34.02, 34.47]$ and $[11.45, 12.73]$ respectively. For KG^2_*, $\mathbb{E}[\tau]$ is uniformly larger, and decreases from 46.33 to 27.42 as ρ increases.

When $\rho = 0$, i.e., when samples from the pair of alternatives are independent, KG^2_* still performs better than KG_*, and KG^2_1 still performs better than KG_1. This may seem surprising, because when $\rho = 0$ it would seem that there is no advantage to sampling a pair of alternatives together rather than sampling them separately. This surprising advantage is present again because of the stopping rule: KG^2_* and KG^2_1 consider both single-alternative and pairwise-difference KG or KG_* factors, instead of just single-alternative factors. This causes them to give a higher value to sampling information that is closer to its true non-myopic value, which allows them to sample more and reduce opportunity cost.

These three tendencies (pairwise-KG methods improve as correlation increases while single-alternative KG methods do not, KG_* stopping rules are better than KG_1 stopping rules, and pairwise-KG stopping rules are better than single-alternative KG stopping rules with independent sampling) explain the overall ordering of the policies observed in Figure 1: When ρ is large ($\rho > 0.6$), the advantage of pairs outweighs the advantage of KG_* and KG^2_1 is better than KG_*. When ρ is small ($\rho < 0.6$), the advantage of pairs is less important and KG_* is better than KG^2_1. Across all values of the sampling correlation, KG^2_* possess both advantages and is best, while KG_1 possesses neither advantage and its performance is the least desirable.

The second main experiment explores the allocation rules in more detail. In this experiment, sampling stops after a given fixed number of samples are taken. Only the allocation rule changes. We assume that samples are independent, but that the prior distribution allows for correlation in the means of each
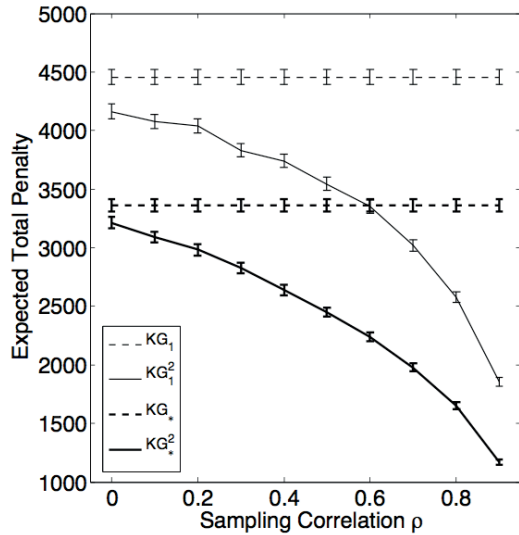
Figure 1: Expected total penalty, $\mathbb{E}[OC + c\tau]$, for not knowing means as a function of sampling correlation ρ, with $k = 2$ alternatives whose means are *a priori* independent, and when the allocation and stopping rules use the same criterion.
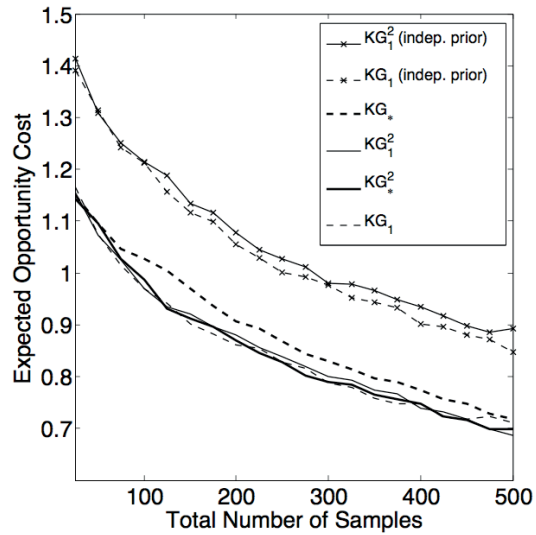
Figure 2: Expected opportunity cost, $\mathbb{E}[OC]$, for a given allocation rule as a function of the total number of samples, with $k = 100$ alternatives, independent sampling and unknown means that are correlated via a Gaussian process prior.

alternative. A comparison is made with respect to average-case performance by generating random problem instances, applying the procedures, and averaging the results.

In this experiment, we generated stylized problem instances by assuming that $k = 100$ alternatives sit on a 10×10 lattice and that the means are correlated via a discrete Gaussian process (Brochu, Cora, and de Freitas 2009). This type of lattice is typical of a decision problem with two discrete decision variables. The covariance in the unknown means of any two alternatives is a decreasing function of their distance from each other on the grid, corresponding to our belief that similar alternatives tend to behave similarly. Specifically, the covariance between an alternative x with grid coordinates x_1 and x_2, and an alternative x' with grid coordinates x'_1 and x'_2 is $\Sigma_0(x, x') = \alpha_0 \exp\left(-\sum_{i=1,2} \alpha_i |x_i - x'_i|^2\right)$, with $\alpha_0 = 1$, $\alpha_1 = \alpha_2 = 0.01$. Our prior mean μ_0 is uniformly 0.

We tested the KG_1, KG_*, KG_1^2, and KG_*^2 allocation rules, under the assumption that the prior distribution of the modeler matches the distribution that generates the problem instances. We also tested KG_1^2 and KG_1 with an independent prior distribution for the means (μ_0 was the 0 vector and Σ_0 was a diagonal matrix with diagonal entries equal to α_0). Thus, we test the benefit of accounting for correlated means rather than incorrectly assuming that they are independent (which is often done in ranking and selection practice).

We use independent sampling with sampling variance 100 for each alternative. We set $b = 10$ for our KG_* and KG_*^2 rules, and use $k_1 = 1, k_2 = 50$ when implementing the KG_1^2 and KG_*^2 rules. We average over 1000 sample paths for KG_* and KG_*^2, and 2000 sample paths for KG_1 and KG_1^2. Each of the replications chooses a new set of sampling means θ from the prior. The standard error was below 0.03 in all cases.

Figure 2 shows expected opportunity cost of a potentially incorrect selection, $\mathbb{E}[OC]$, as a function of the total number of samples. No adaptive stopping rule is used. The figure shows that methods using a correlated prior significantly outperform methods using use an independent prior. This is because the correlated beliefs allow us to learn about all of the alternatives near the sampled alternative, not just the one that was sampled. Through only a few samples, we obtain a rough idea about which areas of the grid have large values, on which we should focus our later effort. In contrast, those policies using an independent prior must sample each point once before it can make such judgments.

The plot shows little difference between any of the four policies KG_1, KG_*, KG_1^2, KG_*^2 using correlated beliefs, and none of the observed differences are statistically significant. Although it may appear that KG_* has consistently higher expected opportunity cost that the other three policies, this could be caused, for example, by KG_* being randomly assigned especially difficult θ to optimize on a few replications. Similarly, there is no statistically significant difference between the two policies KG_1 and KG_1^2 using independent beliefs. In future work we will perform more experiments with more replications to better distinguish between these policies on similar problems using correlated prior beliefs.

The lack of a strong difference in performance between the four policies using correlated prior beliefs may be because we sample independently in this experiment, and use only an allocation rule without an adaptive stopping rule. In contrast, the previously observed differences between these four policies in Figure 1 were due largely to the ability to take advantage of correlated sampling, and to improvements to the stopping rule. Thus, the ability to sample pairs may be less important for problems that do not allow adaptive stopping or correlated sampling. Although sampling pairs may be less important in such restricted situations, introducing correlated sampling and adaptive stopping can improve performance considerably. In these more flexible situations, sampling pairs offers an important benefit beyond sampling just one alternative at a time.

6   CONCLUSIONS AND FUTURE WORK

We have developed new KG allocation and stopping rules that are able to take advantage of both correlated prior beliefs and correlated sampling distributions. These new techniques offer advantages in efficiency over existing KG methods, requiring fewer samples than KG methods that do not use correlations. While KG methods that use correlated beliefs were previously developed, this is the first set of KG methods to take advantage of both correlated beliefs and correlated sampling.

The new stopping rules represent a theoretical improvement over prior KG-type procedures even in the absence of correlation. They can sample from pairs of alternatives in a given stage of sampling if the VOI justifies it. Prior KG-type work does not, and only examines the VOI of sampling from one alternative in a given stage of sampling.

While the results presented here are promising, two important challenges remain before these methods can be applied successfully to real-world problems in optimization via simulation.

First, while the KG_β^2 and KG_*^2 allocation and stopping rules demonstrate an improved ability to select the best alternative with fewer simulation samples, they require more computation than KG_β and KG_* to compute their allocation and stopping decisions. This is especially true for the large-scale applications for which correlated beliefs offer the most promise, in which k could be 10^6 or larger. For such large-scale applications, more research needs to be done to develop computational techniques that deliver the sampling efficiency of KG_β^2 and KG_*^2, but with less computational overhead.

Second, we have assumed that the sampling variances and covariances are known, while in practice they are unknown. One approach would be to use a full Bayesian prior on both the unknown means and the unknown covariance matrix. To compute values of information in this richer Bayesian framework, one could employ approximations where the covariances are estimated by their most likely value based on the most recent data, for the purpose of deciding which point to sample next.

In spite of these challenges, the work presented here represents an initial step to efficient VOI-based optimization via simulation for large-scale problems that takes advantage of correlated sampling.

ACKNOWLEDGMENTS

REFERENCES

Brochu, E., V. M. Cora, and N. de Freitas. 2009, November. "A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning". Technical Report TR-2009-23, Department of Computer Science, University of British Columbia.

Chick, S., and P. Frazier. 2009a, December. "The Conjunction of the Knowledge Gradient and Economic Approach to Simulation Selection". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 528–539. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chick, S., and P. Frazier. 2009b. "Sequential Sampling for Selection with ESP". in review.

Chick, S., and K. Inoue. 2001. "New Procedures to Select the Best Simulated System Using Common Random Numbers". *Management Science* 47 (8): 1133–1149.

Clark, G., and W. Yang. 1986, December. "A Bonferroni selection procedure when using common random numbers with unknown variances". In *Proceedings of the 1986 Winter Simulation Conference*, edited by J. Wilson, J. Henrikson, and S. Robert, 313–315. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Frazier, P. 2009–2010. http://people.orie.cornell.edu/pfrazier/src.html.

Frazier, P., and W. Powell. 2008, December. "The Knowledge-Gradient Stopping Rule for Ranking and Selection". In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Moench, O. Rose, T. Jefferson, and J. W. Fowler, 305–312. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Frazier, P., and W. Powell. 2010. "Paradoxes in Learning and the Marginal Value of Information". *Decision Analysis* 7 (4): 378–403.

Frazier, P., W. B. Powell, and S. Dayanik. 2008. "A Knowledge Gradient Policy for Sequential Information Collection". *SIAM Journal on Control and Optimization* 47 (5): 2410–2439.

Frazier, P., W. B. Powell, and S. Dayanik. 2009. "The Knowledge Gradient Policy for Correlated Normal Beliefs". *INFORMS Journal on Computing* 21 (4): 599–613.

Fu, M. C., J.-Q. Hu, C.-H. Chen, and X. Xiong. 2007. "Simulation Allocation for Determining the Best Design in the Presence of Correlated Sampling". *INFORMS J. on Computing* 19 (1): 101–111.

Gelman, A., J. Carlin, H. Stern, and D. Rubin. 2004. *Bayesian data analysis*. second ed. Boca Raton, FL: CRC Press.

Golub, G., and C. Van Loan. 1996. *Matrix Computations*. Baltimore, MD: John Hopkins University Press.

Gordon, R. 1941. "Values of Mills ratio of area to bounding ordinate and of the normal probability integral for large values of the argument". *Ann. Math. Statist* 12:364–366.

Gupta, S., and K. Miescke. 1996. "Bayesian look ahead one-stage sampling allocations for selection of the best population". *Journal of statistical planning and inference* 54 (2): 229–244.

Howard, R. 1966. "Information Value Theory". *IEEE Trans. Systems Science and Cybernetics* 2 (1): 22–26.

Kim, S. 2005. "Comparison with a standard via fully sequential procedures". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 15 (2): 155–174.

Nakayama, M. 2000. "Multiple comparisons with the best using common random numbers for steady-state simulations". *Journal of Statistical Planning and Inference* 85 (1-2): 37–48.

Negoescu, D., P. Frazier, and W. Powell. 2011. "The Knowledge Gradient Algorithm for Sequencing Experiments in Drug Discovery". *INFORMS Journal on Computing* 23 (3): 346–363.

Nelson, B., and F. Matejcik. 1995. "Using Common Random Numbers for Indifference-Zone Selection and Multiple Comparisons in Simulation". *Management Science* 41 (12): 1935–1945.

Rasmussen, C., and C. Williams. 2006. *Gaussian Processes for Machine Learning*. Cambridge, MA: MIT Press.

Ryzhov, I., P. Frazier, and W. Powell. 2010. "On the robustness of a one-period look-ahead policy in multi-armed bandit problems". *Procedia Computer Science* 1 (1): 1635–1644.

Scott, W., P. I. Frazier, and W. B. Powell. 2011. "The Correlated Knowledge Gradient for Simulation Optimization of Continuous Parameters Using Gaussian Process Regression". *SIAM Journal on Optimization* 21:996–1026.

Yang, W., and B. Nelson. 1991. "Using Common Random Numbers and Control Variates in Multiple-Comparison Procedures". *Operations Research* 39 (4): 583–591.

AUTHOR BIOGRAPHIES

PETER I. FRAZIER is an assistant professor in the School of Operations Research and Information Engineering at Cornell University. He received a Ph.D. in Operations Research and Financial Engineering from Princeton University in 2009. He is the recipient of an AFOSR Young Investigator Award in 2010. His research interest is in dynamic programming and Bayesian statistics, focusing on the optimal acquisition of information. He works on applications in simulation, optimization, operations management, and medicine. His email address is pf98@cornell.edu and his web page can be found via www.orie.cornell.edu.

JING XIE is a PhD student in the School of Operations Research and Information Engineering at Cornell University. She joined the Chinese National Honor Program for Fundamental Sciences in 2006, and received a B.S. in mathematics from Fudan University at Shanghai, China in 2009. She has a research interest in simulation optimization and Bayesian statistics. Her e-mail is jx66@cornell.edu.

STEPHEN E. CHICK is a professor of Technology and Operations Management at INSEAD, and the Novartis Chair of Healthcare Management. He has worked in the automotive and software sectors prior to joining academia, and now teaches operations with applications in manufacturing and services, particularly the health care sector. He enjoys Bayesian statistics, stochastic models, simulation, and the music of the Grateful Dead. His email address is stephen.chick@insead.edu.