The Knowledge-Gradient Algorithm for Sequencing Experiments in Drug Discovery

Diana M. Negoescu

Department of Management Science and Engineering, Stanford University, Stanford, California 94305,
negoescu@stanford.edu

Peter I. Frazier

School of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14853,
pf98@cornell.edu

Warren B. Powell

Department of Operations Research and Financial Engineering, Princeton University, Princeton, New Jersey 08544,
powell@princeton.edu

We present a new technique for adaptively choosing the sequence of molecular compounds to test in drug discovery. Beginning with a base compound, we consider the problem of searching for a chemical derivative of the molecule that best treats a given disease. The problem of choosing molecules to test to maximize the expected quality of the best compound discovered may be formulated mathematically as a ranking-and-selection problem in which each molecule is an alternative. We apply a recently developed algorithm, known as the knowledge-gradient algorithm, that uses correlations in our Bayesian prior distribution between the performance of different alternatives (molecules) to dramatically reduce the number of molecular tests required, but it has heavy computational requirements that limit the number of possible alternatives to a few thousand. We develop computational improvements that allow the knowledge-gradient method to consider much larger sets of alternatives, and we demonstrate the method on a problem with 87,120 alternatives.

Key words: simulation: design of experiments; decision analysis: sequential; statistics, Bayesian
History: Accepted by Allen Holder, Area Editor for Applications in Biology, Medicine, and Health Care; received August 2009; revised May 2010; accepted August 2010. Published online in *Articles in Advance* December 29, 2010.

1. Introduction

In drug discovery, medical researchers often begin with a single molecule that shows some promise for treating a given disease and then test many variations of that molecule to find one that produces the best results. These variations are obtained by substituting atoms or groups of atoms at certain locations (sites) on the original molecule with other atoms or groups of atoms (substituents). The number of variations increases exponentially in the number of sites and substituents; therefore, the number of candidate compounds is usually extremely large. Synthesizing and testing a compound may require several days' work and a significant investment of lab materials, which strictly limits the number of tests that a research team can perform. A critical problem is therefore deciding which compounds should be tested to most accurately and quickly find compounds with good disease-treating ability.

The problem of deciding which compounds to evaluate can be modeled mathematically as a ranking-and-selection problem. In this problem, we have a budget of measurements that we allocate sequentially to alternatives (molecular compounds in our case) so that when we finish our experiments, we have collected the information needed to maximize our ability to find the best alternative. In deciding which compound to measure, we may use the fact that compounds with similar structures often have similar properties. Although a vast literature exists for ranking-and-selection problems in which the alternatives are treated independently, the use of structure within the set of alternatives has received much less attention. A recent ranking-and-selection algorithm that uses correlated Bayesian beliefs to take advantage of structure within the set of alternatives is the knowledge-gradient algorithm for correlated beliefs (KGCB) (Frazier et al. 2009). However, the standard implementation of this algorithm requires storing and manipulating a covariance matrix whose dimension is the number of alternatives. In drug discovery, where one commonly has tens or hundreds of thousands of alternatives, this strategy is computationally expensive at best and often computationally infeasible.

In this paper, we represent beliefs about molecules using a linear, additive model. We take advantage of the structure of this model to develop two computational improvements to the knowledge-gradient algorithm that substantially streamline the procedure. When the number of molecules considered is large, as it almost always is during drug discovery, our new procedures improve computation time by several orders of magnitude. These improvements allow the KGCB algorithm to be used in applied settings for which the standard implementation is simply too slow to be useful. After introducing these new computation-saving procedures, we evaluate the performance of the KGCB algorithm using a previously published data set.

In this paper we restrict our discussion to drug discovery; however, it is worth noting that these improvements have applications not just in drug discovery, but in any problem in which beliefs obey a linear additive model. One example is that of choosing the shortest path through a network when the length along a path is exactly or approximately equal to the sum of the lengths along each link. Some of the computational improvements we develop can also be applied to problems whose beliefs are only approximated by a linear additive model, such as identifying which ads produce the most ad clicks based on a linear model using features from each ad.

There are two main approaches to ranking and selection: the frequentist approach, which is based entirely on observed data, and the Bayesian approach, which uses subjective a priori beliefs on the values of the compounds. We briefly review the Bayesian approach here. For a more thorough review of both approaches, see Swisher et al. (2003) and Kim and Nelson (2006).

Within the Bayesian approach, there are two main directions of research. The first is the optimal computing budget allocation (OCBA) (see, e.g., Chen et al. 1996), in which the probability of correct selection is maximized and the posterior variance is reduced under the assumption that sampling will not change the posterior means. The second direction consists of value of information procedures (VIP) (see, e.g., Chick and Inoue 2001b) and is similar to OCBA in that it maximizes the improvement in a single stage's allocation, but unlike OCBA, it considers the change in posterior mean when estimating the improvement. One of the first contributions in VIP is Gupta and Miescke (1996), who use a one-step analysis under an independent normal prior distribution. Under this analysis, one chooses the measurement decision that would be optimal if only one additional sample were allowed. The independent normal case was then further analyzed by Frazier et al. (2008) and extended to the unknown variance case by Chick et al. (2010). Frazier

et al. (2008) use the term "knowledge-gradient" to refer to this one-step approach because the value of a single sample is the difference in implementation value between two different quantities of knowledge—the knowledge one has before the sample, and the knowledge one has after it. The knowledge-gradient approach of Gupta and Miescke (1996) is extended to correlated normal beliefs in Frazier et al. (2009) and their KGCB algorithm. With correlated beliefs, one may model how structural relationships between alternatives (such as those between chemically related molecules) cause the values of alternatives to relate to one another. These relationships allow learning about multiple alternatives from just a single measurement and often result in a dramatic improvement in measurement efficiency.

An abundant literature in ranking and selection treats beliefs that are a priori independent across alternatives, but correlated a priori beliefs have received much less attention. To our knowledge, Frazier et al. (2009) are the first to consider such beliefs for ranking and selection. There are, however, Bayesian approaches to ranking and selection using common random numbers (CRN). Correlated sampling with CRN induces correlated posteriors even if the prior is independent. CRN is useful for ranking and selection because sampling two systems with positively correlated noise allows one to estimate which system is better with more accuracy than is possible with independent noise (see, e.g., Ehrlichman and Henderson 2008, who also provide a more general technique than CRN for sampling with correlation). Chick and Inoue (2001a) propose a VIP approach for using CRN, whereas Fu et al. (2007) propose an OCBA-based approach. In the indifference-zone ranking-and-selection and multiple comparisons literature, several methods allow for CRN (Yang and Nelson 1991, Nelson and Matejcik 1995, Kim and Nelson 2001), but these methods are non-Bayesian and hence do not use prior distributions, correlated or otherwise. It is important to note that in the drug discovery problem we consider here, the samples are independent, and only the prior distribution is correlated, which differs from using an independent prior distribution with correlated sampling. If the prior is independent, then we must sample each alternative at least once to do well in all but certain special cases, regardless of whether samples are correlated. It is only with a correlated prior that one may sample fewer than the full set of alternatives and still reasonably hope to find an alternative whose value is close to best.

Although correlated priors have appeared infrequently in Bayesian ranking and selection, some frequentist methods use linear models to compare

alternatives' values without requiring that every alternative be sampled, and they achieve conceptually similar improvements to those with correlated priors. These methods even allow infinite numbers of alternatives, as would occur if we were tuning a set of continuous parameters to maximize the output of a simulation model. For the use of such techniques in multiple comparisons, see Hsu (1996, Chapter 7); for their use in multiarmed bandits; see Kleinberg (2005). The use of linear designs within response surface methods (Myers and Montgomery 2002) can also be understood in this way.

In contrast with their rarity within Bayesian ranking and selection, correlated prior distributions have a rich history within Bayesian global optimization (BGO), where Gaussian process prior distributions model the similarity between values of a continuous function evaluated at nearby points. Then, as in Bayesian ranking and selection, these beliefs are used to decide at which points to evaluate this function, with the ultimate goal of finding the function's global maximum. There are two predominant methods for choosing measurements in BGO: maximal probability of improvement (MPI, also called P^*) (Kushner 1964, Stuckman 1988, Perevozchikov 1991) and expected improvement (EI) (Močkus 1972, Močkus et al. 1978, Jones et al. 1998). These two methods are most frequently applied when measurements are free from noise, but MPI was extended to the noisy case in Calvin and Zilinskas (2005), and EI was extended to the noisy case in Huang et al. (2006). The expected improvement criterion is similar to the knowledge-gradient method except that EI considers the improvement in the function's value *only at the point measured*, whereas KG considers improvements *across the entire domain*. KG methods were compared with EI methods for a discretized BGO problem in Frazier et al. (2009). KG performed slightly better than EI in the noise-free case and significantly better in the noisy case. See Sasena (2002, §3.2) and Frazier (2011) for reviews of the BGO literature.

Within the medical community, the process of using statistical and computational tools to support drug discovery is called *virtual* or *in silico* screening. Although the literature on this subject has enjoyed massive growth in the past few years, only a small portion of it treats sequential designs. Most of it focuses instead on nonsequential prediction techniques using data collected in a single batch (see, e.g., Burbidge et al. 2001, Jorissen and Gilson 2005, Davies et al. 2006), despite the fact that sequential designs can be much more efficient (Warmuth et al. 2003). One exception is Bajorath (2002), which is an early review of virtual screening that differentiates between focused (which we call nonsequential) and sequential screening.

Previous mathematical work on sequential experimental design for drug discovery has been focused within the machine learning community, where it has been formulated as an active learning problem. Most frequently, active learning concerns itself with choosing which examples should be labeled to best train a classifier. Warmuth et al. (2003), with more mathematical details provided in Warmuth et al. (2002), take this approach, where compounds that perform well in a laboratory test are labeled as "active," and those that perform poorly are labeled "inactive." These labels are assumed to be available without noise. Taking an alternative approach, De Grave et al. (2008) consider a version of the drug discovery problem closer to the one we consider here, where one observes real-valued activity levels for compounds. Rather than searching for the single best compound, their goal is to find the k compounds with the largest activity levels where measurements are assumed to be free from noise. We also evaluate several empirical strategies, including versions of MPI and EI.

An important part of any Bayesian ranking-and-selection algorithm for drug discovery is the prior distribution on the values of the candidate compounds. Several such prior distributions have been developed in the medicinal chemistry literature, where the collection of statistical models that predict biological activity from chemical structure are known collectively as quantitative structure activity relationship (QSAR) models. In our case, the biological activity of interest is the compound's ability to treat a disease, as measured in a laboratory test. Examples of such laboratory tests include those that test a compound's ability to kill diseased cells or inhibit a protein interaction believed to be critical to the disease's progression.

The first attempt to quantify relationships between biological activity and chemical structure dates to 1868, when Brown and Fraser published the first formulation of a quantitative relationship between "physiological activity" and "chemical structure" (Brown and Fraser 1868). The modern origins of the field lie in the Free–Wilson model (Free and Wilson 1964) and the Hansch analysis (Hansch and Fujita 1964). In both methods, multiple linear regression is used to correlate activity on laboratory tests with collections of compound features. In the Free–Wilson model, these features are indicator variables that indicate the presence or absence of substituents at given sites. In the Hansch analysis, these features are physiochemical properties such as molecular weight and lipophilicity. More recently, support vector machines have been used together with much larger collections of compound features (see, e.g., Warmuth et al. 2003, Jorissen and Gilson 2005). Many other statistical methods have been applied to QSAR, including cluster analysis, pattern recognition, principle component

analysis, discriminant analysis, partial least squares, neural networks, and evolutionary algorithms. See Grover et al. (2000) for a review.

In this paper, we consider the Free–Wilson model. The Free–Wilson model assumes that the base molecule and each substituent contributes additively to the overall value of the compound. This contribution is assumed to be independent of the presence or absence of other substituents. We also consider a generalization of the Free–Wilson model that allows for deviations from perfect linearity. The structure of the Free–Wilson model allows us to derive two highly efficient implementations of the KGCB algorithm that can be used for sets of candidate compounds much larger than can be handled by the original algorithm. Although the original implementation manipulates a belief on the values of the *compounds*, our new implementations manipulate beliefs on the parameters of our linear model. Because the number of compounds is exponential in the number of substituents per site, maintaining a belief in this way reduces computation and memory requirements substantially and allows the KGCB algorithm to be used on problems with even hundreds of thousands of candidate compounds.

The rest of this paper is organized as follows. In §2, we describe the ranking-and-selection problem and the relationships between the structure and value of an alternative. In §3, we review the original implementation of the KGCB algorithm, and in §§4 and 5, we describe two computational improvements to this implementation. In §6, we present an empirical study of the KGCB algorithm on a previously published data set.

2. Model

We suppose that we have M alternatives and a budget of N measurements, and we wish to sequentially decide which alternatives to measure so that when we exhaust our budget of measurements, we have maximized our ability to find the best alternative. We assume that samples from testing alternative i are normally distributed with unknown mean ϑ_i and known variance λ_i and are conditionally independent of all other samples, given the unknown mean and the decision to sample that alternative. We write ϑ to indicate the column vector $(\vartheta_1, \ldots, \vartheta_M)'$. For a summary of all the symbols we use, please refer to Table 1. We further assume that our belief about ϑ is distributed according to a multivariate normal prior distribution with mean vector μ^0 and positive semidefinite covariance matrix Σ^0:

$$\vartheta \sim \mathcal{N}(\mu^0, \Sigma^0).$$

We assume that we have a budget of N sampling decisions, $x^0, x^1, \ldots, x^{N-1}$. The measurement

Table 1   Table of Notation

M	Number of alternatives (in drug discovery, total number of compounds)
N	Number of measurements in our budget
ϑ_i	Unknown mean of alternative i
λ_i	Known variance of alternative i
ϑ	Column vector $(\vartheta_1, \ldots, \vartheta_M)'$
μ^n, Σ^n	Mean and covariance of posterior distribution on ϑ after n measurements
x^i	Sampling decision at time i, $x^i \in \{1, \ldots, M\}$
ϵ^{n+1}	Measurement error of alternative x^n, $\epsilon^{n+1} \sim \mathcal{N}(0, \lambda_{x^n})$
$\hat{y}^{n+1}$	Sampling observation from measuring alternative x^n, $\hat{y}^{n+1} = \vartheta_{x^n} + \epsilon^{n+1}$
Π	Set of all possible policies $\{(x^0, \ldots, x^{N-1}): x^n \in \mathcal{F}^n\}$
S^n	Defined as the pair (μ^n, Σ^n)
e_x	Column vector of 0s with a 1 at position x
$\tilde{\sigma}(\Sigma, x)$	Defined as $\dfrac{\Sigma e_x}{\sqrt{\lambda_x + \Sigma_{xx}}}$
$\nu^{\mathrm{KG},n}$	Knowledge gradient factor, defined as $\max_x \mathbb{E}_n[\max_i \mu_i^{n+1} \mid S^n = s, x^n = x] - \max_i \mu_i^n$
x^{KG}	Measurement decision of the KGCB policy
a_i	Contribution of attribute (in drug discovery, substituent) i
s^x	Row vector of 0s and 1s. A 1 indicates that the ith attribute is present in alternative x
ζ	Value of the base alternative (with 0s in all locations)
$L(i)$	Location of substituent i
$A(l)$	Vector of substituents that may be placed at location l
k	Total number of attributes
b_x	Deviation term associated with alternative x
σ_b	Standard deviation of the b_x terms
α	Vector of attribute values $(\zeta, a_1, a_2, \ldots, a_k)$
X	Matrix in which each row is an alternative
$\tilde{x}^n$	Column vector of 0s and 1s encoding alternative x^n
γ^n	Defined as $\lambda_{x^n} + (\tilde{x}^n)^T C^n \tilde{x}^n$
$\hat{\epsilon}^{n+1}$	Defined as $y^{n+1} - (\theta^n)^T \tilde{x}^n$
θ, C	Mean and covariance of posterior distribution on α ($\alpha \sim \mathcal{N}(\theta, C)$)
θ^n, C^n	Mean and covariance of posterior distribution on α after n measurements

decision x^n selects an alternative to test at time n from the set $\{1, \ldots, M\}$. The measurement error ϵ^{n+1} is assumed to be normally distributed $\epsilon^{n+1} \sim \mathcal{N}(0, \lambda_{x^n})$ and independent conditional on x^n. Therefore, the resulting sample observation is $\hat{y}^{n+1} = \vartheta_{x^n} + \epsilon^{n+1}$. Through experiments, we try to learn the value of ϑ, which is assumed to be fixed throughout the duration of the trials.

We define a filtration $(\mathcal{F}^n)_{n=0}^N$ as the sequence of sigma-algebras generated by the samples observed by time n and the identities of their originating alternatives. More explicitly, $\mathcal{F}^n$ is the sigma-algebra generated by $x^0, \hat{y}^1, x^1, \hat{y}^2, \ldots, x^{n-1}, \hat{y}^n$. We write $\mathbb{E}_n$ and Var_n to indicate $\mathbb{E}[\cdot \mid \mathcal{F}^n]$ and $\mathrm{Var}[\cdot \mid \mathcal{F}^n]$, the conditional expectation and variance, respectively, taken with respect to $\mathcal{F}^n$. Then, define $\mu^n := \mathbb{E}_n[\vartheta]$, and $\Sigma^n := \mathrm{Cov}[\vartheta \mid \mathcal{F}^n]$. Conditionally on $\mathcal{F}^n$, our posterior belief distribution on ϑ is multivariate normal with mean vector μ^n and covariance matrix Σ^n.

We define Π to be the set of all possible policies satisfying our sequential requirement; that is, $\Pi :=$

$\{(x^0, \ldots, x^{N-1}): x^n \in \mathcal{F}^n\}$. We let π be a generic policy in Π, and we write $\mathbb{E}^\pi$ to indicate the expectation taken when the policy is π.

After exhausting the budget of N measurements, we select the alternative with the highest posterior mean. Our goal is to choose a measurement policy maximizing expected reward, which can be written as

$$\sup_{\pi \in \Pi} \mathbb{E}^\pi \left[\max_i \mu_i^N \right].$$

We now describe how the prior mean vector μ^0 and covariance matrix Σ^0 are chosen according to the Free–Wilson model. Although we apply this work to drug discovery, where we think of "compounds," "sites," and "substituents," we describe the model in the generic language of "alternatives," "dimensions," and "attributes." These generic terms facilitate the observation that the Free–Wilson model is simply a linear model whose explanatory variables are zero or one. To define them, we suppose the existence of several dimensions and state that each attribute may be associated with only one dimension. Each alternative is obtained through a specific choice of which single attribute, if any, is present in each dimension. In the context of drug discovery, an alternative is a compound, a dimension is a site, and an attribute is a substituent, where we consider the same atom or chemical group substituted at two different sites to be two different attributes.

2.1. The Free–Wilson Model

The Free–Wilson model (Free and Wilson 1964) assumes that each attribute contributes additively to the value of the alternative. Denote by a_i the contribution of attribute i, and denote by s^x a vector of zeros and ones with a one for every attribute that is present in alternative x. Thus, $s_i^x = 1$ means that the ith attribute is present in alternative x. We denote by ζ the value of the base alternative, which is the alternative obtained from taking $s_i^x = 0$ over every attribute i. Let $L(i)$ denote the dimension associated with attribute i, and let k denote the total number of attributes. We restrict the s^x to specify at most one attribute associated with each dimension. That is, we require of each x that $\sum_i s_i^x 1_{\{L(i)=l\}} \leq 1$ for each dimension l. Furthermore, we allow any s^x that meets this specification.

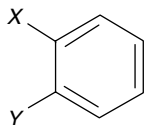To illustrate these notions for drug discovery, we constructed a small toy example in Figure 1. This

compound (alternative) has two sites (dimensions) X and Y, at which atoms or groups of atoms (attributes) can be attached to create new compounds. Assume there are two substituents, A and B, which can be placed at site X and two substituents, C and D, which can be placed at site Y. Then, the original compound without any substituents attached (having hydrogen atoms at sites X and Y) can be represented by the vector $s^{00} = (0, 0, 0, 0)$, where the first two zeros correspond to site X and the last two zeros correspond to site Y. The compound that results from attaching nothing at site X (keeping the hydrogen atom) and substituent C at site Y is represented by the vector $s^{0C} = (0, 0, 1, 0)$. The compound that results from attaching substituent B at site X and substituent C at site Y is represented by the vector $s^{BC} = (0, 1, 1, 0)$. For this example, k is 4 (A, B, C, D), $L(1) = L(A) = 0$ (attribute A is absent), $L(2) = L(B) = X$, $L(3) = L(C) = Y$, and $L(4) = L(D) = 0$.

The Free–Wilson model assumes that each attribute contributes additively to the value of the alternative, which may be expressed as

$$\vartheta_x = \sum_i a_i s_i^x + \zeta.$$

To exemplify, going back to the example in Figure 1, the value of the alternative having attribute B at dimension X and attribute C at dimension Y is then

$$\vartheta_{BC} = a_A \cdot 0 + a_B \cdot 1 + a_C \cdot 1 + a_D \cdot 0 + \zeta = a_B + a_C + \zeta.$$

Under this model, if we sample alternative x, having attributes given by $s_1^x, \ldots, s_k^x$, the sample value would be of the form

$$\hat{y}_x = a_1 s_1^x + \cdots + a_k s_k^x + \zeta + \epsilon,$$

where $\epsilon \sim \mathcal{N}(0, \lambda_x)$ is independent measurement noise.

We suppose that we have an independent normal prior distribution on ζ and $a_1, \ldots, a_k$. Under this prior distribution, the mean μ_i of our belief about the value of alternative i is

$$\mu_i^0 = \mathbb{E}[\zeta] + \sum_m s_m^i \mathbb{E}[a_m],$$

and the covariance Σ_{ij}^0 between the values of alternatives i and j is

$$\Sigma_{ij}^0 = \mathrm{Cov}(i, j) = \mathrm{Cov}\left(\sum_m a_m s_m^i + \zeta, \sum_{m'} a_{m'} s_{m'}^j + \zeta \right)$$

$$= \mathrm{Var}(\zeta) + \sum_{m, m'} s_m^i s_{m'}^j \mathrm{Cov}(a_m, a_{m'})$$

$$= \mathrm{Var}(\zeta) + \sum_{m \in \mathcal{L}_{ij}} \mathrm{Var}(a_m),$$

Figure 1 A Toy Example Illustrating the Notions of "Attributes," "Dimensions," and "Alternatives"

where $\mathscr{L}_{ij} = \{l \in \{1, \ldots, k\} \mid s_l^i = s_l^j = 1\}$ is the set of attributes common to alternatives i and j.

The Free–Wilson model assumes that the contribution of each substituent is linear and additive. In practice, it is easy to envision substituents that interact with each other. For example, there might be a set of synergistic substituents for which the compound containing all of them has a much larger value than compounds containing only a subset of them. Such behaviors may be handled with models that are linear in the parameters but which include deviation terms. One such model is presented in the next section.

2.2. The General Model
We now generalize the Free–Wilson model, which assumes a perfectly additive structure, to allow some deviation from perfect additivity. These deviations are specified by terms b_x, with one such term for each alternative x. The resulting model is still a linear model but with more terms. Under this model the value of alternative x is

$$\vartheta_x = \sum_i a_i s_i^x + b_x + \zeta.$$

Under our prior distribution, the terms $b_1, \ldots, b_M$ are normally distributed with mean 0 and a variance σ_b^2. In addition, they are independent of each other as well as from the a_i and ζ terms. Under this structure, the covariance Σ_{ij}^0 between the values of alternatives i and j is

$$\Sigma_{ij}^0 = \mathrm{Cov}(i,j) = \mathrm{Cov}\left(\sum_m a_m s_m^i + \zeta + b_i, \sum_{m'} a_{m'} s_{m'}^j + \zeta + b_j\right)$$

$$= \mathrm{Var}(\zeta) + \sum_{m,m'} s_m^i s_{m'}^j \mathrm{Cov}(a_m, a_{m'}) + \sigma_b^2 \mathbf{1}_{\{i=j\}}$$

$$= \mathrm{Var}(\zeta) + \sum_{m \in \mathscr{L}_{ij}} \mathrm{Var}(a_m) + \sigma_b^2 \mathbf{1}_{\{i=j\}}.$$

Thus, under this general model, the covariance matrix of the values of the alternatives under our prior distribution is obtained by taking the covariance matrix from the Free–Wilson model and adding σ_b^2 to the diagonal entries. Since $\mathbb{E}[b_x] = 0$ for all x, the mean μ^0 of our belief about the values of the alternatives is the same as it is under the Free–Wilson model.

We see a spectrum of behaviors from this model over different values of σ_b^2. When $\sigma_b^2 = 0$, the model is identical to the Free–Wilson model in §2.1. When σ_b^2 is nonzero but still significantly smaller than $\mathrm{Var}(a_l)$, it models a situation in which the values of the alternatives are well approximated by an additive structure, but we understand that there may be small deviations from it. When σ_b^2 is very large (in particular, much larger than the $\mathrm{Var}(a_l)$ terms), the deviation from additivity is quite large, correlations between the alternatives are very weak, and our belief is similar to one that is independent across alternatives.

It is in this second regime, where σ_b^2 is nonzero but not too large, that this general model is likely to be most useful. It allows us to exploit the additive structure of a problem to learn quickly while it simultaneously allows flexibility in its understanding that the alternatives' true values may not obey this structure perfectly.

In the third regime, where σ_b^2 is large, we essentially have an independent prior distribution, and learning the value of one alternative teaches us almost nothing about the value of the other alternatives. This makes it very difficult to learn in situations with large numbers of alternatives because to come close to finding a good alternative, we need to make at least as many measurements as the number of Bs. In such situations, we must either find another nonadditive structure in the problem or resign ourselves to making more measurements than the number of alternatives.

Although the problem is inherently quite difficult when σ_b^2 is large and often requires many measurements to find good alternatives, the KGCB policy that we present may still work well *relative to the difficulty of the problem*. Indeed, the case of independent beliefs was considered in Frazier et al. (2008), and the KG policy was shown to work well compared with other state-of-the-art policies.

3. The KG Algorithm with Correlated Beliefs
The knowledge-gradient with correlated beliefs (KGCB) policy, as introduced in Frazier et al. (2009), measures the alternative that attains the maximum in

$$\nu^{\mathrm{KG},n} = \max_x \mathbb{E}_n\left[\max_i \mu_i^{n+1} \Big| S^n = s, x^n = x\right] - \max_i \mu_i^n, \quad (1)$$

where $S^n := (\mu^n, \Sigma^n)$ parameterizes the posterior belief distribution at measurement n. The knowledge-gradient (KG) factor $\nu^{\mathrm{KG},n}$ represents the incremental value obtained from measuring a particular alternative x.

After each alternative is measured, we obtain a posterior distribution on ϑ that depends on which alternative was measured x^n, its sampled value $\hat{y}^{n+1}$, and our belief on ϑ prior to sampling, which is parameterized by μ^n and Σ^n. This posterior distribution may be calculated using standard results for normal sampling with a multivariate normal prior distribution (see, e.g., Gelman et al. 2004) as

$$\mu^{n+1} = \mu^n + \frac{\hat{y}^{n+1} - \mu_x^n}{\lambda_x + \Sigma_{xx}^n} \Sigma^n e_x, \quad \Sigma^{n+1} = \Sigma^n - \frac{\Sigma^n e_x e_x' \Sigma^n}{\lambda_x + \Sigma_{xx}^n},$$

where e_x is a column M-vector with a single one at index x and the rest zeros.

We now describe the time n conditional distribution of μ^{n+1} that allows us to compute (1). This distribution is multivariate normal, with mean given by the tower property as $\mathbb{E}_n[\mu^{n+1}] = \mu^n$ and covariance given by the conditional variance formula as $\tilde{\sigma}(\Sigma^n, x^n)\tilde{\sigma}(\Sigma^n, x^n)'$, where

$$\tilde{\sigma}(\Sigma, x) := \frac{\Sigma e_x}{\sqrt{\lambda_x + \Sigma_{xx}}}.$$

The details of this computation are given in Frazier et al. (2009).

Thus, the time n conditional distribution of μ^{n+1} is the same as that of $\mu^{n+1} = \mu^n + \tilde{\sigma}(\Sigma^n, x^n)Z$, where Z is any independent one-dimensional standard normal random variable. This allows us to rewrite (1) as

$$x^{\mathrm{KG}}(s) = \operatorname*{argmax}_x \mathbb{E}\Big[\max_i(\mu_i^n + \tilde{\sigma}_i(\Sigma^n, x^n)Z)\Big| S^n, x^n = x\Big]$$
$$-\max_i \mu_i^n$$
$$= \operatorname*{argmax}_x h(\mu^n, \tilde{\sigma}(\Sigma^n, x)).$$

Here, $h: \mathbb{R}^M \times \mathbb{R}^M \to \mathbb{R}$ is defined by $h(p, q) = \mathbb{E}[\max_i p_i + q_i Z] - \max_i p_i$, where p and q are deterministic M vectors, and again, Z is any one-dimensional standard normal random variable.

Frazier et al. (2009) provides a method for computing $h(p, q)$, which is summarized in Algorithm 2. This algorithm in turn uses Algorithm 1 in an inner loop. In Algorithm 2, the components of p and q are sorted and then some are dropped, which results in new vectors p' and q' of length M' and a sequence c'. These quantities are then used to calculate $h(p, q)$ via

$$h(p, q) = \sum_{i=1}^{M'-1} (q'_{i+1} - q'_i) f(-|c'_i|),$$

where the function f is given by $f(z) = \Phi(z) + z\phi(z)$ and where Φ is the normal cumulative distribution function and ϕ is the normal density.

Algorithm 1 (Compute sequence c and undominated set A)

Require: Inputs p and q.
1: $c_0 \leftarrow -\infty$, $c_1 \leftarrow +\infty$, $A \leftarrow \{1\}$
2: **for** $i = 1$ to $M - 1$ **do**
3:     $c_{i+1} \leftarrow +\infty$
4:     **repeat**
5:         $j \leftarrow A[\mathrm{end}(A)]$
6:         $c_j \leftarrow (p_j - p_{i+1})/(q_{i+1} - q_j)$
7:         **if** $\mathrm{length}(A) \neq 1$ **then**
8:             $k = A[\mathrm{end}(A) - 1]$
9:             **if** $c_j \leq c_k$ **then**
10:                $A \leftarrow A(1, \ldots, \mathrm{end}(A) - 1)$
11:                loopdone $\leftarrow$ false
12:             **else**
13:                loopdone $\leftarrow$ true

14:             **end if**
15:         **end if**
16:     **until** loopdone
17:     $A \leftarrow (A, i+1)$
18: **end for**

Algorithm 2 (Compute helper function h)

Require: Inputs p and q.
1: Sort $(p_i, q_i)_{i=1}^M$ such that p_i are in nondecreasing order and ties broken such that $p_i \leq p_{i+1}$ if $q_i = q_{i+1}$
2: **for** $i = 1$ to $M - 1$ **do**
3:     **if** $q_i = q_{i+1}$ **then**
4:         Remove entry i from $(p_i, q_i)_{i=1}^M$
5:     **end if**
6: **end for**
7: Use Algorithm 1 to compute c and A from p, q
8: $p \leftarrow p[A]$, $q \leftarrow q[A]$, $c \leftarrow (c[A], +\infty)$,
    $M \leftarrow \mathrm{length}(A)$
9: $h(p, q) \leftarrow \log(\sum_{i=1}^{M-1}(q_{i+1} - q_i) f(-|c_i|))$.

Using Algorithm 2, we can compute $h(\mu^n, \tilde{\sigma}(\Sigma^n, x))$ for any vectors μ^n and $\tilde{\sigma}(\Sigma^n, x)$. This then allows us to compute the KG factor via (1) for each alternative, the largest of which gives the measurement decision of the KGCB policy. This is summarized in Algorithm 3.

Algorithm 3 (KGCB$_1$ algorithm)

Require: Inputs μ^n and Σ^n.
1: **for** $x = 1$ to $\mathrm{length}(\mu^n)$ **do**
2:     $p \leftarrow \mu^n$
3:     $q \leftarrow \tilde{\sigma}(\Sigma^n, x)$
4:     $\nu \leftarrow h(p, q)$ % *use Algorithm 2*
5:     **if** $x = 1$ or $\nu > \nu^*$ **then**
6:         $\nu^* \leftarrow \nu$, $x^* \leftarrow x$
7:     **end if**
8: **end for**

Within Algorithm 3, Algorithm 2 executes M times. Within one execution of Algorithm 2, the sort has complexity $O(M \log M)$ and Algorithm 1 has complexity $O(M)$. Thus, the most computationally demanding step within Algorithm 2 is the sort, and the overall complexity of the KGCB algorithm as computed by Algorithm 3 is $O(M^2 \log M)$.

In drug discovery, families of molecules often contain tens or hundreds of thousands of compounds, which makes this algorithm for computing KGCB computationally infeasible. Thus, even though we might very much wish to use the KGCB method to reduce the number of *physical* measurements that need to be taken, the *computational* requirements of actually computing the KGCB measurement decisions under the standard algorithm (Algorithm 3) preclude doing so in most cases. The next section describes a first improvement to the standard algorithm that dramatically reduces the computational requirements and allows us to compute KGCB for large numbers of

linearly related alternatives such as those encountered in drug discovery.

4. The KGCB$_2$ Algorithm

In this section, we present the first of two computational improvements to the standard implementation of KGCB that exponentially reduce the computational and storage requirements of the standard algorithm. The essential idea behind this improvement is to maintain a belief on the attributes themselves instead of on the (much larger) set of alternatives. This greatly improves the efficiency of the KGCB algorithm.

In the subsections that follow, we first describe this improvement in the context of the Free–Wilson model from §2.1, and then we extend the improvement to the general model from §2.2.

4.1. Beliefs on Attributes

In this section we describe how one may maintain a belief on attributes rather than on alternatives in the Free–Wilson model of §2.1.

Let α be the vector of attribute values $\alpha = (\zeta, a_1, \ldots, a_k)$ containing the value of the base molecule and of each substituent, where chemically identical substituents at different locations are given different indices. We assume the linear additive model for modeling structure–value relationships from §2.1, and we let X be a matrix comprised of rows representing the alternative. Each row of X is a vector of zeros and ones of the same length as α, and each one indicates an attribute that is present in the alternative. The value of this attribute is the corresponding component in α. In the context of drug discovery, this row contains a single one in the first entry to indicate that the base molecule (whose value is ζ) is present, and then the subsequent entries contain a one for each substituent present. Thus, this row is a one followed by the vector s^x (defined in §2.1) corresponding to the molecule x being represented. With these definitions, the true value of the alternatives is $\vartheta = X\alpha$.

Any multivariate normal belief on α induces a multivariate normal belief on ϑ. If we have a multivariate normal belief on α with the $(k+1)$-dimensional mean vector θ and the $(k+1) \times (k+1)$ covariance matrix C,

$$\alpha \sim \mathcal{N}(\theta, C), \tag{2}$$

we then have the mean of the values of the alternatives given by $\mathbb{E}[\vartheta] = X\theta$. The covariance between the values of the alternatives is given by

$$\begin{aligned}
\text{Cov}(\vartheta_i, \vartheta_j) &= \text{Cov}\left(\sum_k X_k^i \alpha_k, \sum_k X_k^j \alpha_k\right) \\
&= \sum_{k,k'} X_k^i X_{k'}^j \text{Cov}(\alpha_k, \alpha_{k'}) = \sum_{k,k'} X_k^i X_{k'}^j C_{k,k'} \\
&= e_i^T XCX^T e_j,
\end{aligned}$$

where e_i is, as before, a column vector of length the size of our alternative database, with a one on position i and zeros everywhere else. Thus, the belief induced on ϑ by (2) is

$$\vartheta \sim \mathcal{N}(X\theta, XCX^T). \tag{3}$$

Having described how a generic multivariate normal prior distribution on α induces a multivariate normal prior distribution on ϑ, we begin with a prior distribution on α with mean vector θ^0 and covariance matrix C^0. Thus, the parameters of the induced prior distribution on ϑ are $\mu^0 = X\theta^0$ and $\Sigma^0 = XC^0X'$. We similarly define θ^n and C^n be the mean vector and covariance matrix, respectively, of the posterior belief distribution on α after n measurements. This posterior belief distribution is also multivariate normal, and we have $\mu^n = X\theta^n$ and $\Sigma^n = XC^nX'$.

There exists a recursive expression for θ^n and C^n that is similar to the recursive expression for μ^n and Σ^n given in §3. Before providing this expression, we first introduce some additional notation. Let $\tilde{x}^n = (\tilde{x}_0^n; \tilde{x}_1^n; \ldots; \tilde{x}_k^n)^T$ be a column vector of zeros and ones describing the alternative $x^n \in \{1, \ldots, M\}$ that was measured at iteration n, where $\tilde{x}_0^n = 1$ represents the presence of the base alternative and $\tilde{x}_i^n$ is one for those attributes i present in alternative x^n and zero otherwise. For instance, if the alternative we choose to measure at iteration 1 is the alternative in the toy example in Figure 1 having attribute B at site X and attribute C at site Y, then $\tilde{x}^1 = (1; 0; 1; 1; 0)^T = (1; s^{BC})^T$. In addition, define $\hat{\epsilon}^{n+1} = y^{n+1} - (\theta^n)^T \tilde{x}^n$ and $\gamma^n = \lambda_{x^n} + (\tilde{x}^n)^T C^n \tilde{x}^n$. Then, the following updating equations result from standard expressions for normal sampling of linear combinations of attributes (see, e.g., Powell 2007):

$$\theta^{n+1} = \theta^n + \frac{\hat{\epsilon}^{n+1}}{\gamma^n} C^n \tilde{x}^n,$$

$$C^{n+1} = C^n - \frac{1}{\gamma^n}\left(C^n \tilde{x}^n (\tilde{x}^n)^T C^n\right). \tag{4}$$

When the number of substituents is large, maintaining θ^n and C^n through this recursive expression is much more efficient than maintaining μ^n and Σ^n through the recursive expression in §3. This is because the dimension of Σ^n is equal to the number of alternatives, which grows exponentially with the number of substituents per site. For a set of compounds with five sites and nine possible substituents at each site (to make 10 possible choices at each site, including the possibility of attaching no substituent), the number of possible compounds is 10^5 compared with only 45 total substituents. In this case, Σ^n is $10^5 \times 10^5$, whereas C^n is only 46×46 (we add 1 to 45 to account for the base molecule).

4.2. Implementation of the KGCB$_2$ Algorithm

Because θ^n and C^n are much smaller and easier to maintain than μ^n and Σ^n, there is significant computational and storage savings to be achieved by calculating the KGCB policy from these inputs directly rather than from μ^n and Σ^n.

To accomplish this, we recall from §4.1 that μ^n and Σ^n may be written in terms of θ^n and C^n as $\mu^n = X\theta^n$ and $\Sigma^n = XC^nX^T$. We also recall that Σ^n enters into the computation of the KG factor for alternative x, $\nu_x^{\text{KG}} = h(\mu^n, \tilde{\sigma}(\Sigma^n, x))$, only through $\tilde{\sigma}(\Sigma^n, x)$. This quantity is given by $\tilde{\sigma}(\Sigma^n, x) = \Sigma^n e_x / \sqrt{\lambda_x + \Sigma_{xx}^n}$, which depends only on row x in Σ^n and not on the entire matrix. To facilitate describing these computations, we define $A_{x,\cdot}$ to be row x from generic matrix A. Then, the required row of Σ^n may be computed from C^n as

$$\Sigma_{x,\cdot}^n = (XC^nX^T)e_x = (XC^n)X_{x,\cdot}^T.$$

By calculating row x of Σ^n from C^n, calculating $\tilde{\sigma}(\Sigma^n, x)$ from this row vector, and then computing the KG factor from $\tilde{\sigma}(\Sigma^n, x)$ and μ^n, we obtain the KG factor while completely avoiding any computations with matrices of size $M \times M$. This is summarized in Algorithm 4.

Algorithm 4 (KGCB$_2$ algorithm)

Require: Inputs θ^n, C^n, and X.
1: $\mu^n \leftarrow X\theta^n$
2: $B \leftarrow XC^n$
3: **for** $x = 1$ to M **do**
4:     $\Sigma_{x,\cdot}^n \leftarrow BX_{x,\cdot}^T$
5:     $a \leftarrow \mu^n$
6:     $b \leftarrow \Sigma_{x,\cdot}^n / \sqrt{\lambda_x + \Sigma_{xx}^n}$
7:     $\nu \leftarrow h(a, b)$ % *use Algorithm 2*
8:     **if** $x = 1$ or $\nu > \nu^*$ **then**
9:         $\nu^* \leftarrow \nu$, $x^* \leftarrow x$
10:     **end if**
11: **end for**

Algorithm 4 is similar to Algorithm 3; the main differences are Step 1, which retrieves the mean belief on alternatives from the mean of the belief on attributes, and Steps 2 and 4, which together retrieve the covariance of the belief on alternatives from the covariance of the belief on attributes. The B matrix used in the second and third operations caches XC^n, which does not depend on the measurement x being considered in the loop.

Algorithm 4 significantly improves on Algorithm 3 because it computes only $\Sigma_{x,\cdot}^n$, a column vector, instead of the full matrix Σ^n. This is particularly significant when there are many alternatives. If there were 10^6 alternatives, we would create a vector of size $10^6 \times 1$ instead of a matrix of size $10^6 \times 10^6$.

4.3. Extension to the General Model

This first computational improvement, which has been described thus far in the context of the Free–Wilson model, can also be implemented for the general model.

To do so, we must maintain a belief about the b_x terms together with our belief about ζ and the a_i terms. Although the number of b_x terms is equal to the number of compounds, which is generally very large, we need only to maintain a mean vector and covariance matrix for only those b_x terms corresponding to alternatives that we have *measured*. If we have not measured a particular alternative x by time n, then our posterior belief distribution on b_x will have the same marginal distribution that it had under the prior distribution and will remain independent of ζ, the a_i terms, and all other deviation terms. Thus, by explicitly maintaining a belief about ζ, $a_1, \ldots, a_k$, and only those b_x terms for compounds we have measured, we can reconstruct our belief as needed about those deviation terms that we are not explicitly tracking.

Toward this end, let us define a vector α_n that contains $\alpha = (\zeta, a_1, \ldots, a_k)$ and the b_x for x ranging over the unique alternatives in $x^0, \ldots, x^{n-1}$. This vector plays the role that α plays in §4.1. Let θ^n and C^n be the mean and variance, respectively, of our time n posterior belief distribution on α_n. Note that $\alpha_0 = \alpha$, $\theta^0 = (\mathbb{E}[\zeta], \mathbb{E}[a_1], \ldots, \mathbb{E}[a_k])$, and C^0 is a diagonal matrix whose diagonal is $(\mathbb{E}[\zeta], \mathbb{E}[a_1], \ldots, \mathbb{E}[a_k])$.

Before providing recursive expressions for θ^n and C^n, C^{n-1}, we first define two quantities, $\tilde{\theta}^{n-1}$ and $\tilde{C}^{n-1}$. If we have previously measured alternative x^n, so $x^n \in \{x^0, \ldots, x^{n-1}\}$, then let $\tilde{\theta}^{n-1} = \theta^{n-1}$ and $\tilde{C}^{n-1} = C^{n-1}$. If we have not previously measured x^n, then let $\tilde{\theta}^{n-1}$ be the column vector obtained by appending a scalar zero to θ^{n-1}, and let $\tilde{C}^{n-1}$ be the $(1+k+n) \times (1+k+n)$ matrix obtained from the $(k+n) \times (k+n)$ matrix C^{n-1} by adding one extra row and column after the last row and column of C^{n-1}. This extra row and column is all zeros, except for the diagonal entry, which is σ_b^2. These quantities $\tilde{\theta}^{n-1}$ and $\tilde{C}^{n-1}$ are constructed so that our time $n-1$ posterior belief distribution on α_n is $\mathcal{N}(\tilde{\theta}^{n-1}, \tilde{C}^{n-1})$. In addition, let $\tilde{x}^n$ be a column vector of zeros and ones, with a one at exactly those indices of α^{n+1} for which the alternative x^n contains the corresponding base alternative, substituent, or deviation term. We also define $\hat{\epsilon}^{n+1} = y^{n+1} - (\tilde{\theta}^n)^T \tilde{x}^n$ and $\gamma^n = \lambda_{x^n} + (\tilde{x}^n)^T \tilde{C}^n \tilde{x}^n$. With these definitions, we may update θ^{n+1} and C^{n+1} recursively from θ^n and C^n (via $\tilde{\theta}^n$ and $\tilde{C}^n$, respectively) as

$$\theta^{n+1} = \tilde{\theta}^n + \frac{\hat{\epsilon}^{n+1}}{\gamma^n} \tilde{C}^n \tilde{x}^n, \tag{5}$$

$$C^{n+1} = \tilde{C}^n - \frac{1}{\gamma^n} (\tilde{C}^n \tilde{x}^n (\tilde{x}^n)^T \tilde{C}^n). \tag{6}$$

These updating equations allow us to maintain our belief about α_n in a computationally efficient way, which is analogous to the way in which we were able to recursively maintain our belief about α in §4.1.

We now show how KG factors may be computed from a belief on α_n parameterized by θ^n and C^n. Because the KG factor is given by $\nu_x = h(\mu^n, \tilde{\sigma}(\Sigma^n, x))$, it is enough to compute μ^n and $\tilde{\sigma}(\Sigma^n, x)$ efficiently (without computing the much larger matrix Σ^n) and then use the standard implementation of h. The first term μ^n does not depend on x and is given by

$$\mu^n = X^n \theta^n,$$

where X^n is a $M \times |\alpha_n|$ matrix of zeros and ones whose each row corresponds to an alternative and has a one for the base alternative and each substituent and deviation term from α_n contained within the alternative. To compute the second term $\tilde{\sigma}^n(\Sigma^n, x)$, fix $x^n = x$ and the corresponding α_{n+1} and $\tilde{C}^n$ resulting from this choice of x^n. Let $\tilde{X}^n$ be a $M \times |\alpha_{n+1}|$ matrix that is similar to X^n, except that it maps alternatives to components of α_{n+1} rather than α_n. That is, each row of $\tilde{X}^n$ corresponds to an alternative and has a one for the base alternative and each substituent and deviation term from α_{n+1} contained within the alternative. Then, observe that the beliefs about those b_x not included in α_{n+1} will not change as a result of measuring x^n, and so $\tilde{\sigma}(\Sigma^n, x^n)$, which is the standard deviation of the change in beliefs about the values of the alternatives, is not affected by these deviation terms not included in α_{n+1}. We can thus compute $\tilde{\sigma}(\Sigma^n, x^n)$ by dropping these left-out deviation terms. In such a model in which these deviation terms outside α_{n+1} have been left out of the model, the xth column of Σ^n is

$$\Sigma^n_{x, \cdot} = (\tilde{X}^n \tilde{C}^n)(\tilde{X}^n_{x, \cdot})^T,$$

and $\tilde{\sigma}(\Sigma^n, x^n)$ may be computed from this vector via $\tilde{\sigma}(\Sigma^n, x) = \Sigma^n e_x / \sqrt{\lambda_x + \Sigma^n_{xx}}$. The resulting method of computing the KGCB policy is summarized in Algorithm 5.

Algorithm 5 (General KGCB$_2$ algorithm)

  Require: Inputs θ^n, C^n, and X^n
  1: $\mu^n \leftarrow X^n \theta^n$
  2: **for** $x = 1$ to M **do**
  3:     Compute $\tilde{C}^n$ from C^n and x
  4:     $\Sigma^n_{x, \cdot} \leftarrow (\tilde{X}^n \tilde{C}^n)(\tilde{X}^n_{x, \cdot})^T$
  5:     $a \leftarrow \mu^n$
  6:     $b \leftarrow \Sigma^n_{x, \cdot} / \sqrt{\lambda_x + \Sigma^n_{xx}}$
  7:     $\nu \leftarrow h(a, b)$ % use Algorithm 2
  8:     **if** $x = 1$ or $\nu > \nu^*$ **then**
  9:         $\nu^* \leftarrow \nu$, $x^* \leftarrow x$
 10:    **end if**
 11: **end for**

With these expressions, we may compute the KG factor for alternative x in the general model without explicitly computing a covariance matrix Σ^n. The dimensionality of the objects θ^n and C^n that we must retain in memory is the sum of $1 + k$ together with the number of unique alternatives measured. When the number of measurements that can be made is much smaller than the number of alternatives—as it is in drug discovery—this is a significant savings.

5. The KGCB$_3$ Algorithm for the Free–Wilson Model

We now return from the general model to consider only the Free–Wilson model from §2.1. When using this model, we may compute the KGCB policy with even greater efficiency than has been described thus far. This further improvement has at its heart Proposition 1.

This proposition states that the calculation of each KG factor, which would ordinarily require us to consider the entire family of molecules, may be decomposed into the sum of a set of much smaller and easier-to-compute KG factors. Each of these smaller KG factors is the KG factor that would result from a base molecule with substituents only at a single site. Because the number of compounds grows exponentially with the number of substituents per site, and because of the complexity of computing a KG factor scales as $O(s \log s)$, where s is the number of substituents considered, this decomposition dramatically reduces the computational effort required to compute a KG factor. Although the KGCB policy still needs to calculate the KG factor that corresponds to each possible measurement to find the largest one, faster calculation of each KG factor makes the overall computation much faster as well.

Before we state Proposition 1, recall that $\theta^n_j = \mathbb{E}_n[a_j]$ and $\theta^n_0 = \mathbb{E}_n[\zeta]$ are the means of the time n posterior belief distribution on the value of substituent j and the base molecule, respectively, and that $L(j)$ is the site at which substituent j may be placed. Additionally, let $A_\ell \in \{i: L(i) = \ell\}$ so that A_ℓ is a vector containing those substituents that may be placed at site ℓ.

PROPOSITION 1. *Under the Free–Wilson model,*

$$\nu^{n, KG}_x = \sum_\ell h\big((\theta^n_i)_{i \in A(\ell) \cup \{-1\}}, (\tilde{\sigma}^n_{xi})_{i \in A(\ell) \cup \{-1\}}\big),$$

where $(\tilde{\sigma}^n_{xi} t)^2_n := \mathrm{Var}_n[\theta^{n+1}_i | x^n = x]$ *for* $i > 0$ *and* $\theta^n_{-1} = \tilde{\sigma}^n_{x, -1} = 0$.

PROOF. We write

$$\max_{x'} \mu^{n+1}_{x'} = \max_{x'} \mathbb{E}_{n+1}\left[\sum_j a_j s^{x'}_j + \zeta\right] = \max_{x'} \sum_i \theta^{n+1}_i s^x_i + \theta^{n+1}_0$$

$$= \max_{x'} \sum_\ell \left(\sum_i \theta^{n+1}_i \mathbf{1}_{\{L(i) = \ell\}}\right) + \theta^{n+1}_0$$

$$= \sum_\ell \max_{i \in A(\ell) \cup \{-1\}} \theta^{n+1}_{\ell i} + \theta^{n+1}_0,$$

where $\theta_{\ell-1}^{n+1}$ is defined to be 0, and where the crucial final step occurs because the maximum over x' is achieved by the x' that places the substituent with the largest estimated value in each site (or no substituent in a site, corresponding to $i = -1$, if all the estimates at that site are negative). Substituting this expression for $\max_{x'} \mu_{x'}^{n+1}$ and noting that the tower property implies $\mathbb{E}_n[\theta_0^{n+1}] = \mathbb{E}_n[\mathbb{E}_{n+1}[\zeta]] = \mathbb{E}_n[\zeta] = \theta_0^n$, we obtain

$$\mathbb{E}_n\left[\max_{x'}\mu_{x'}^{n+1}\right] = \mathbb{E}_n\left[\sum_\ell \max_{i \in A(\ell) \cup \{-1\}} \theta_i^{n+1} + \theta_0^{n+1}\right]$$

$$= \mathbb{E}_n\left[\sum_\ell \max_{i \in A(\ell) \cup \{-1\}} \theta_i^{n+1}\right] + \theta_0^n.$$

Then, noting that $\max_{x'} \mu_{x'}^n = \sum_\ell \max_{i \in A(\ell) \cup \{-1\}} \theta_{\ell i}^n + \theta_0^n$ by an argument similar to the one for $\max_{x'} \mu_{x'}^{n+1}$, we have

$$\nu_x^{n,\text{KG}} = \left(\sum_\ell \mathbb{E}_n\left[\max_{i \in A(\ell) \cup \{-1\}} \theta_i^{n+1} \Big| x^n = x\right] + \theta_0^n\right)$$

$$- \left(\sum_\ell \max_{i \in A(\ell) \cup \{-1\}} \theta_i^n + \theta_0^n\right)$$

$$= \sum_\ell \mathbb{E}_n\left[\max_{i \in A(\ell) \cup \{-1\}} \theta_i^{n+1} \Big| x^n = x\right] - \max_{i \in A(\ell) \cup \{-1\}} \theta_i^n.$$

Because the joint distribution of $(\theta_i^{n+1})_{i \in A(\ell)}$ conditioned on $\mathscr{F}_n$ and $x^n = x$ is the same as that of the $\mathscr{F}_n$ conditional distribution of $(\theta_i^n + \tilde{\sigma}_{xi}^n Z)_{i \in A(\ell)}$, where Z is a one-dimensional standard normal variable, and because $\theta_{-1}^{n+1} = 0 = \theta_1^n + \tilde{\sigma}_{xi}^n Z$ almost surely, we substitute for θ_i^{n+1} and

$$\nu_x^{n,\text{KG}} = \sum_\ell \mathbb{E}_n\left[\max_{i \in A(\ell) \cup \{-1\}} \theta_i^n + \tilde{\sigma}_{xi}^n Z\right] - \max_{i \in A(\ell) \cup \{-1\}} \theta_i^n$$

$$= \sum_\ell h\left((\theta_i^n)_{i \in A(\ell) \cup \{-1\}}, (\tilde{\sigma}_{xi}^n)_{i \in A(\ell) \cup \{-1\}}\right). \quad \square$$

To actually use the decomposition property inherent in this Proposition 1, we must first compute $\tilde{\sigma}_{xi}^n$. We have $(\tilde{\sigma}_{xi}^n)^2 = \text{Var}_n[\mathbb{E}_{n+1}[a_i] \mid x^n = x]$, which, using the conditional variance formula and the fact that $\mathbb{E}_n[\text{Var}_{n+1}[a_i]] = \text{Var}_{n+1}[a_i]$, becomes

$$(\tilde{\sigma}_{xi}^n)^2 = \text{Var}_n[a_i] - \text{Var}_{n+1}[a_i \mid x^n = x].$$

We have $\text{Var}_n[a_i] = C_{ii}^n$, and by (4),

$$\text{Var}_{n+1}[a_i] = e_i^T C^{n+1} e_i = e_i^T\left[C^n - \frac{1}{\gamma^n} C^n \tilde{x}^n (\tilde{x}^n)^T C^n\right] e_i$$

$$= C_{ii}^n - \frac{e_i^T C^n \tilde{x}^n (\tilde{x}^n)^T C^n e_i}{\lambda_{x^n} + (\tilde{x}^n)^T C^n \tilde{x}^n}.$$

Therefore,

$$(\tilde{\sigma}_{xi}^n)^2 = \frac{e_i^T C^n \tilde{x}^n (\tilde{x}^n)^T C^n e_i}{\lambda_{x^n} + (\tilde{x}^n)^T C^n \tilde{x}^n},$$

which gives the vector $\tilde{\sigma}_{x,\cdot}^n$, with i ranging over *all* the substituents (not just those corresponding to a particular site) as

$$\tilde{\sigma}_{x,\cdot}^n = \frac{C^n \tilde{x}^n}{\sqrt{\lambda_{x^n} + (\tilde{x}^n)^T C^n \tilde{x}^n}}. \qquad (7)$$

We summarize the resulting implementation below in Algorithm 6.

Algorithm 6 (KGCB$_3$ algorithm)
 **Require**: Inputs θ^n, C^n, and X
  1: **for** $x = 1$ to M **do**
  2:     $\tilde{\sigma}_{x,\cdot}^n \leftarrow \dfrac{C^n \tilde{x}^n}{\sqrt{\lambda_{x^n} + (\tilde{x}^n)^T C^n \tilde{x}^n}}$
  3:     $\nu_x \leftarrow 0$
  4:     **for** $\ell = 1$ to L **do**
  5:       $p \leftarrow (\theta_i^n)_{i \in A(\ell) \cup \{-1\}}$
  6:       $q \leftarrow (\tilde{\sigma}_{xi}^n)_{i \in A(\ell) \cup \{-1\}}$
  7:       $\nu_x \leftarrow \nu_x + h(p, q)$ % *use Algorithm 2*
  8:     **end for**
  9:     **if** $x = 1$ or $\nu_x > \nu^*$ **do**
 10:       $\nu^* \leftarrow \nu_x$, $x^* \leftarrow x$
 11:     **end if**
 12: **end for**

If there are l dimensions with M_l substituents that can be placed at each dimension, the number of total alternatives is $M = (M_l)^l$. The computational complexity of Algorithm 4 is $O(M^2 \ln M)$, which is equivalent to $O(l(M_l)^{2l} \ln(M_l))$. The computational complexity of Algorithm 6, on the other hand, is $O((M_l)^l l(M_l) \ln(M_l))$ because the outer loop executes $(M_l)^l$ times while the inner loop executes l times and takes $O((M_l) \ln(M_l))$. Thus the efficient implementation has a computational complexity of $O(l(M_l)^{l+1} \ln(M_l))$ compared with $O(l(M_l)^{2l} \ln(M_l))$ for the previous implementation.

6. Empirical Study

We simulate the performance of the KGCB algorithm using data from a previously published QSAR study of narcotic analgetics by Katz et al. (1977) that uses the Free–Wilson model. Their paper contains a set of 6,7-benzomorphans, which have five sites at which substituents can be attached. The molecule is shown in Figure 2. At the site labeled R_1, there are 11 possible substituents (together with H); at site R_2, there are 8 possible substituents; at site R_3, there are 5 possible substituents; at site R_4, there are 6 possible substituents; and at site R_5, there are 11 possible substituents. The collection of compounds resulting from choosing a substituent at each site contains $11 \times 8 \times 5 \times 6 \times 11 = 29{,}040$ compounds. In addition, each compound can be charged positively or negatively or be neutral, which brings the number of compounds to 87,120. Katz et al. (1977) provide experimentally observed activity values for 99 of these compounds.
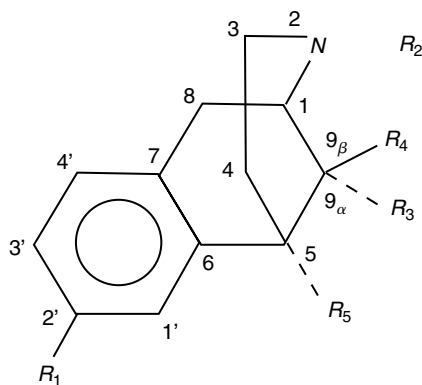
Figure 2 **An Example of Substituent Locations (Marked with** $R_1, \ldots, R_5$**)**

In §6.1, we describe methods for choosing the prior distribution on the values of this collection of compounds. In §6.2, we present an empirical study of the performance of KGCB on this collection of compounds and compare it against two other policies. In §6.3, we discuss other policies that we did not compare but could be included in a more extensive empirical study.

6.1. Prior Distribution on the Substituent Contributions

When choosing a prior distribution, one may be working with experienced practitioners who may be able to articulate their prior distributions on the values of compounds or substituents. In many cases, however, even if one is quite experienced in drug discovery or is working with those who are quite experienced, it may be difficult to articulate prior beliefs. In such cases, it is useful to have a method for setting the prior distribution from some other source, such as previous measurements of families of molecules that may be completely different from the one to which drug discovery effort is being applied. We now present one such method that may be used with either the Free–Wilson model or the general model.

We first discuss prior distributions for the Free–Wilson model. Our method of choosing a prior distribution supposes that there exists a large population of substituent values in nature and that nature has drawn independently at random from this population the particular substituent values to create the family of compounds being investigated. We suppose that we may approximate the distribution of values within this population with a normal distribution whose mean and variance we can estimate. A method of estimation is discussed below.

Our method also supposes that we began with a noninformative belief on the value ζ of the base molecule. This belief was subsequently updated by a laboratory measurement that provided an informative belief on the value of the base molecule that we will use within our prior distribution. The mean of this belief is the value of this measurement, and its variance is the variance of the measurement noise. A measurement of the base molecule is generally available in practice, because a molecule is usually chosen to have its space of chemical derivatives searched because it performed well on an initial screening of many chemically dissimilar molecules.

Given these suppositions, our prior distribution is given by taking θ_0^0 and C_{00}^0 to be the value and measurement variance, respectively, of our previous measurement of the base molecule, and by taking θ_i^0 and C_{ii}^0 for $i > 0$ to be the estimated mean and variance, respectively, of the population of substituents. Our independence supposition causes the off-diagonal terms of C^0 to be zero. These values for θ^0 and C^0 specify a prior distribution as discussed in §4.1, and the equivalent prior distribution on compound values may be reconstructed also as discussed in §4.1. This provides a prior distribution appropriate for the Free–Wilson model.

We now discuss how this method of setting a prior distribution can be extended to the general model. We suppose, in a manner quite similar to our supposition about substituent values, that there exists a large population of deviation terms from which nature has drawn independently at random to create the family of compounds being considered. Let σ_b^2 be the estimated variance of this population. Without loss of generality, we may assume that the mean of the population of deviation terms is zero because we can take any nonzero mean and move it into the value of the base molecule. Then the prior distribution, in the format of §4.3, is given by taking the same θ^0 and C^0 as just described for the Free–Wilson model and using this value of σ_b^2 in subsequent updates.

To use either of these methods for setting the prior distribution, one needs access to a measurement of the base molecule (with that measurement's variance), estimates of the mean and variance of the population of substituents, and, when using the general model, an estimate of the variance of the population of deviation terms. These estimates may be obtained from measurements of a reasonably large family of compounds that may be quite dissimilar from the one being investigated. Using these measurements, one may use linear regression to estimate the substituent values and deviation terms present in this family of compounds. The substituent values and deviation terms that result will certainly be a very small subset of the entire population of substituent values and deviation terms present in nature, but if one is willing to suppose that they are representative of what one may encounter in nature, then the population means

and variances of the observed values may be taken as estimators of the means and variances of the overall populations. Because one might make this assumption that the observed values are "representative," it is better if one has observations from multiple families of molecules, and it is also better if the observed values are from families that are more similar to the one to which drug discovery effort is being applied.

We followed exactly this method of estimation using the data in Katz et al. (1977), who provided measured values for 99 of the possible 87,120 compounds. The measurement technique used is quite accurate, and it is reasonable to take the measured values as the "true" values. By fitting a multiple linear regression to this data set, we obtain substituent values and deviation terms. From these values, we estimate the mean and variance of the population of substituents as 0.31 and 0.47, respectively, and the variance of the population of deviation terms as $\sigma_b^2 = 0.15$.

6.2. Simulation Results

Before discussing the ability of KGCB to discover good molecules with relatively few measurements in simulation, we first describe the computational costs and memory requirements of the various implementations of the Free–Wilson-based KGCB implementation on the set of 87,120 compounds. Algorithm 6 is able to compute the KGCB decision and update the belief in less than three seconds, which is approximately 100 times faster than Algorithm 4. We were unable to run the standard implementation on this data set of 87,120 compounds because it stores and manipulates a covariance matrix on the compounds with dimension $87,120 \times 87,120$, which requires 14 GB of memory (if only the upper triangular part of the symmetric matrix is stored and a four-byte single-precision floating point format is used). This memory requirement could not be met by the hardware available to us. In contrast, Algorithms 6 and 4 manipulate a covariance matrix on the attributes with dimension 43×43, requiring only 3.6 KB of storage (15 KB if the full matrix is stored in double precision).

In our simulations, we observed the number of measurements required by the KGCB algorithm to find good compounds among collections of compounds of various sizes. In these simulations, we compared KGCB with two policies: a pure exploration policy that chooses compounds to test uniformly at random, and a policy from the frequentist design of experiments (DOE) literature based on a linear model, called one-factor-at-a-time (OFAAT) (Montgomery 2005). This policy begins by testing the base molecule and then sequentially tests each compound created by adding one substituent to the base molecule.

The OFAAT policy neglects interaction between factors (Montgomery 2005). In most applications, this

is detrimental to performance; however, under the perfectly linear Free–Wilson model, there are no interactions between terms, and thus neglecting interactions provides an advantage to OFAAT. In the general model, where factors interact via the deviation terms b_x, a factorial design would perhaps be more appropriate (Montgomery 2005), but such designs are infeasible in problems such as our drug discovery problem that have many dimensions.

In our simulations, we maintain beliefs for each competing policy (exploration and OFAAT) using the same updates as in the KGCB algorithm, and we use these to choose implementation decisions. This allows the other policies to take advantage of the linear structure of the problem in generating implementation decisions, even if they do not in generating measurement decisions. We emphasize that these belief updates use correlated beliefs that give our pure exploration policy a substantial advantage over any policy that uses independent belief updates. With our version of pure exploration, even though it measures uniformly at random, it learns about the compounds through correlations in the same way that KGCB does.

Although we compare only against pure exploration and the DOE-based OFAAT policy, other policies from the literature that more completely exploit the structure of the problem might have performed better. Although the measurement decisions of OFAAT were chosen with the knowledge of the linear structure of the problem, they do not exploit the objective function (which rewards information about the best alternative rather than information about all alternatives), and so other DOE-based policies might perform better. Pure exploration takes advantage of the linear structure of the problem only when making its implementation decisions and not when making its measurement decisions. Thus it should be viewed only as a benchmark representing nonadaptive learning. In §6.3, we discuss other policies from the literature that might be applied to this problem and could be included in a more extensive empirical study.

Section 6.2.1 presents numerical results from the Free–Wilson model described in §2.1, and §6.2.2 presents numerical results from the general model described in §2.2.

6.2.1. Results Using the Free–Wilson Model. We begin by describing results for the simple, linear-additive model described in §2.1. We consider the opportunity cost, defined as the difference between the true value of the compound that is actually best and the true value of the compound that is best according to the policy's posterior belief distribution. Thus, once the opportunity cost is zero, the policy has found the best compound. Although we consider the

opportunity cost, which is a linear function of biological activity, the value of a compound to a practitioner is often a nonlinear function of activity, e.g., a case where a compound may only be used as a drug when its activity exceeds a threshold.

Figure 3 plots the average opportunity cost as a function of the number of measurements using two data sets (one with 2,640 compounds and the other with 87,120 compounds) and two values for the noise standard deviation (0.1 and 0.5). Using the measured values published in Katz et al. (1977), we generate a truth for these compounds by fitting a linear regression to the data, taking the true means θ to be the predicted values from the regression and the standard deviation of the noise in our measurements to be the standard deviation of the residuals from the regression. For the 2,640 compounds data set, we average over 100 runs, and for the 87,120 com-

pounds data set, we average over 10 runs. Averaging over more runs removes more of the variance introduced by Monte Carlo sampling of the truths, causing the curves for 2,640 compounds to be smoother than those for 87,120 compounds. As Figure 3 shows, KGCB outperforms the other policies, with an average opportunity cost that is always lower.

OFAAT performs worse than both KGCB and pure exploration. Although OFAAT is a first-order orthogonal design and is hence optimal with respect to a minimal variance criterion (Myers and Montgomery 2002, §7.3.1), we consider a different criterion: the ability to find the *best* compound. Also, OFAAT measures compounds containing single substituents, whereas KGCB and pure exploration measure compounds containing several. Although this provides OFAAT more information per measurement about each substituent measured, the way that KGCB and
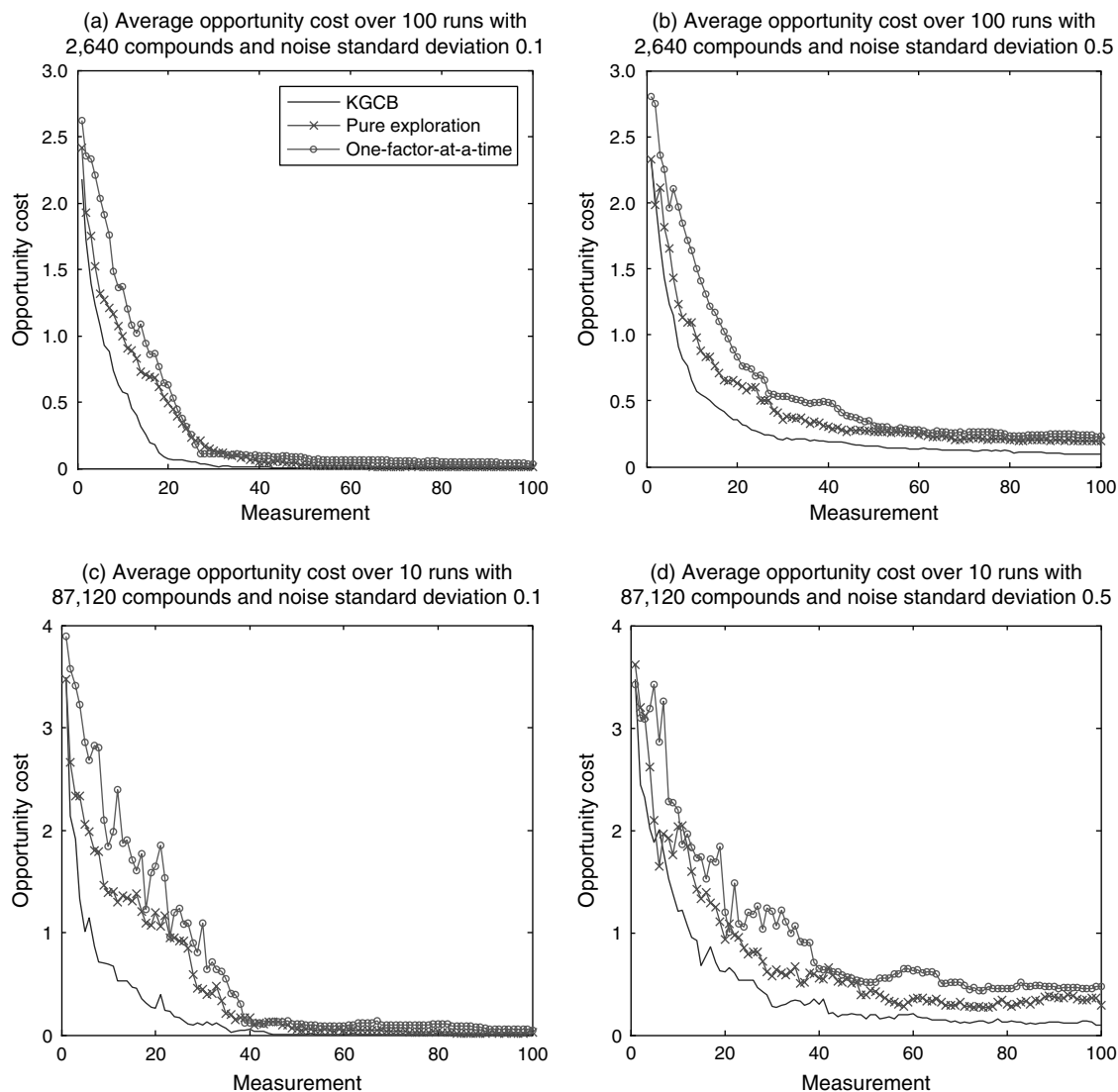

Figure 3   Average Opportunity Cost Under the Free–Wilson Model

pure exploration learn a little bit about more substituents seems to be better in this application.

KGCB also performs better than pure exploration. Although pure exploration is able to use the full Bayesian machinery to guide its implementation decisions and thus take advantage of the linear structure of the model, its measurement decisions are chosen purely at random. In contrast, KGCB is able to take information gleaned in the early measurements to guide its later measurements to more informative compounds.

To further compare the performance of KGCB and pure exploration, we randomly selected sets of compounds from the full set of 87,120 and performed more extensive analysis on these smaller problems. Figure 4(a) shows the expected opportunity cost of KGCB and pure exploration on a single set of compounds. To generate this figure, we randomly selected a single set of 10,000 compounds, ran 15 independent sample paths each for KGCB and pure exploration on these compounds, and estimated the expected opportunity cost as a function of the number of measurements. An estimate and its standard error is plotted for each policy, where each error bar is twice the standard error in each direction. Figure 4(a) shows that the estimated expected opportunity cost for the KGCB policy is lower than that of the pure exploration policy.

Figure 4(b) shows the distribution of the difference in opportunity cost over multiple sets of compounds. To generate this figure, we randomly selected 75 sets containing 10,000 compounds each. For each set of compounds, we ran KGCB and pure exploration independently and observed the difference in opportunity cost as a function of the number of measurements n. When this difference is positive, pure exploration has a higher opportunity cost, and KGCB performs better. We then plot, for each n, the mean and standard deviation of this population of 75 differences. This characterizes the overall distribution of the difference in performance between KGCB and pure exploration across multiple sets of compounds. Figure 4(b) shows that the mean difference is always positive after the first measurement and has a maximum near the tenth measurement. This suggests that the learning rate is faster for KGCB than for pure exploration, and the value of using the KGCB policy is maximized when the measurement budget is small.

For more numerical work using the Free–Wilson model, please see the Online Supplement (available at http://joc.pubs.informs.org/ecompanion.html).

6.2.2. Results Using the General Structural Model.
We now present the results using the general model described in §2.2. To simulate a true set of compound values in this set of experiments, we take the fitted values of a_i and ζ from Katz et al. (1977),


(a) Opportunity cost of KGCB and exploration

(b) Difference in opportunity cost
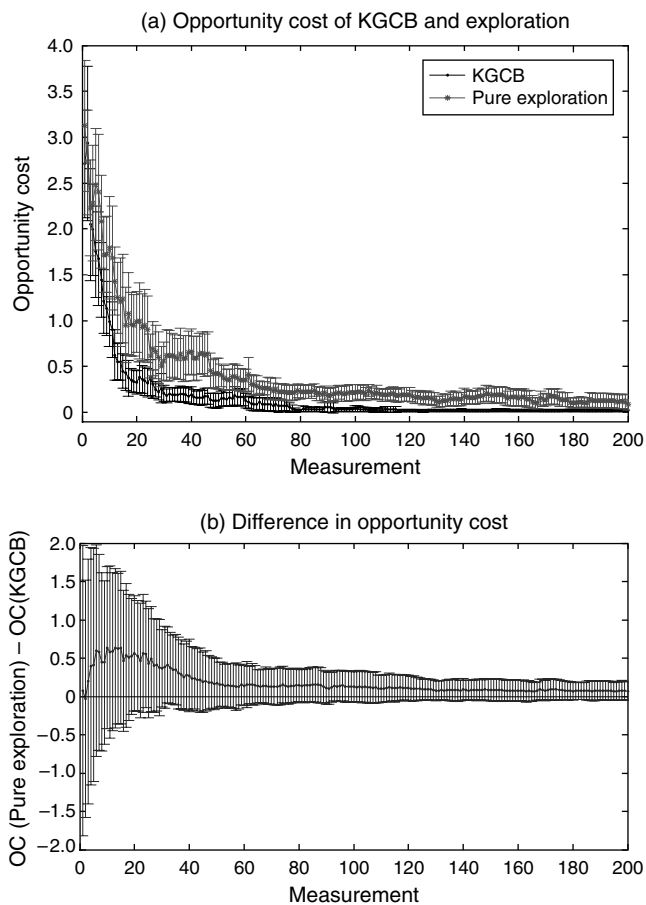
Figure 4 **(a) Average Opportunity Cost of KGCB and Pure Exploration, Estimated Using 15 Sample Paths; (b) Difference in Average Opportunity Cost Between KGCB and Pure Exploration, Estimated Using 75 Sample Paths**

Note. Both figures use problems with 10,000 compounds, noise standard deviation of 0.38, and the Free–Wilson model.

and for each compound x, we generate b_x independently from a normal distribution with mean 0 and variance σ_b^2 (where σ_b^2 was obtained from the fit as well, as described in §6.1). We then combine these values according to the model to obtain true values for the compounds.

We compared KGCB only to the pure exploration policy, because OFAAT is not designed to handle the deviation terms present in the general model, and it did not perform as well as pure exploration in the deviation-free Free–Wilson case. To test the performance of KGCB under this model, we randomly selected 10 different sets of 1,000 compounds each and ran the KGCB algorithm using the prior distribution described in §6.1. For each sample path, we perform 150 measurements comparing KGCB to pure exploration, and we plot the average opportunity cost over the 10 runs just as in our previous experiments. The results are shown in Figure 5.

As with the Free–Wilson model, the KGCB policy performs significantly better than pure exploration.
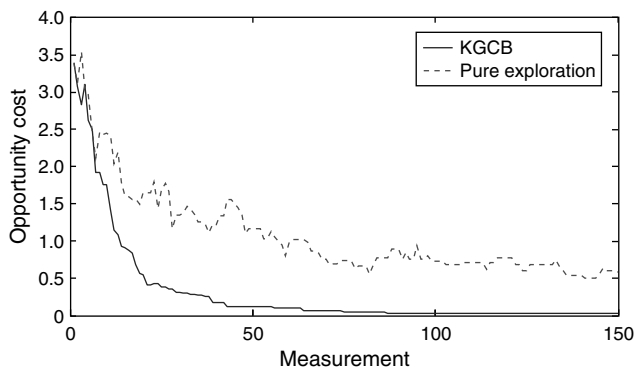
Figure 5 Average Opportunity Cost Over 10 Runs Using Different Data Sets of 1,000 Compounds

Opportunity cost decreases more slowly in this test than it did under tests that assumed the Free–Wilson model. This is because this test adds deviation terms to each molecule, which makes it harder for any policy to find the best compound. Although this test was performed with 1,000 compounds and the general model contains a different deviation term for each one, the KGCB policy finds a very good compound with many fewer than 1,000 measurements. This is because σ_b^2 is smaller than $\text{Var}_0(a_l)$, which makes learning the a_l terms more important than learning the b_x deviation terms and allows a policy to find a compound with small opportunity cost without learning all the deviation terms.

6.3. Other Related Measurement Policies

We have compared the KGCB policy against two other policies: pure exploration and OFAAT. In this section, we discuss other policies that a more extensive empirical study might include and discuss their relationship to KGCB. We note that KGCB has been compared against many of these other policies in other problem settings and has outperformed them there.

First, the BGO literature contains several relevant policies. First, there are the expected improvement (EI; Močkus 1972) and maximum probability of improvement (MPI; Kushner 1964) policies, both of which were discussed in §1. In addition, there is the *interval estimation* (IE) policy (Kaelbling 1993), which combines exploration with exploitation by measuring the compound for which a particular linear combination of the posterior mean and posterior standard deviation is largest. IE was applied to BGO in Cox and John (1997), where it was called SDO, or sequential design for optimization.

As described in §1, De Grave et al. (2008) also consider sequential designs for drug discovery and actually apply all three of these policies to a drug discovery problem (IE is referred to as the optimistic strategy). However, that empirical study assumes that

measurements are free from noise, and the versions of these BGO policies they employ require modification for the noisy case before they can be applied to the problem we consider here. In addition to these three BGO policies, De Grave et al. (2008) also consider three other policies: pure exploration; *largest value*, which measures the untested compound with the largest estimated value; and *largest variance*, which measures the compound whose value has the largest posterior variance.

Frazier et al. (2008) compare KG to IE, largest variance, largest value, and pure exploration on the ranking-and-selection problem with an independent normal prior distribution. In this problem, because the prior distribution is independent and measurements of one alternative do not affect the other alternatives, KG can also be understood as a generalization of EI to the noisy case. (KGCB does not reduce to EI when beliefs are correlated, even if measurements are noise-free.) Frazier et al. (2009) compare KG to EI in a discretized version of the BGO problem, as well as a version of EI generalized to handle noise called sequential kriging optimization (Huang et al. 2006). In both of these studies, KG performed as well or better than these other methods; it found better alternatives with fewer measurements but required more computation for each decision.

There are also policies from the DOE literature that might be applied to this problem. In the numerical study of the Free–Wilson model, we compared it against OFAAT. One could compare against other first-order orthogonal designs or to a response surface method (Myers and Montgomery 2002) using such designs. One could also compare against screening methods such as the sequential bifurcation method described in Kleijnen (2008). With such screening methods, it may be possible to determine which of the substituents have the largest values with fewer measurements than there are substituents. The sequential bifurcation method cannot be directly applied because it requires a knowledge of the sign of the effect of each factor, but it may be possible to adapt this method or other similar methods to the problem of drug discovery.

7. Conclusions

Drug discovery is a long and expensive process for which sophisticated sequential procedures like KGCB are well suited. Such procedures have the potential to reduce the costs involved and to increase the number of drug discovery efforts that find success within budget. Our new algorithms for calculating the KGCB policy allow efficient implementation on large data sets, and our empirical study demonstrates that KGCB performs much better than the pure exploration and one-factor-at-a-time policies. Extensions

and improvements to this work that we feel would be of value include more extensive empirical comparisons between KGCB and other policies, as well as implementations of the KGCB algorithm based on other QSAR models.

Acknowledgments

References

Bajorath, J. 2002. Integration of virtual and high-throughput screening. *Nature Rev. Drug Discovery* **1**(11) 882–894.

Brown, A. C., T. R. Fraser. 1868. On the connection between chemical consitution and physiological action. Part 1. On the physiological action of salts of the amonium bases, derived from Strychnia, Brucia, Thebia, Codeia, Morphia, and Nicotia. *J. Anatomy Physiol.* **2**(2) 224–242.

Burbidge, R., M. Trotter, B. Buxton, S. Holden. 2001. Drug design by machine learning: Support vector machines for pharmaceutical data analysis. *Comput. Chemistry* **26**(1) 5–14.

Calvin, J. M., A. Zilinskas. 2005. One-dimensional global optimization for observations with noise. *Comput. Math. Appl.* **50**(1–2) 157–169.

Chen, C.-H., H.-C. Chen, L. Dai. 1996. A gradient approach for smartly allocating computing budget for discrete event simulation. *Proc. 28th Conf. Winter Simul.* IEEE Computer Society, Washington, DC, 398–405. http://portal.acm.org/citation.cfm?id=256696.

Chick, S. E., J. Branke, C. Schmidt. 2010. Sequential sampling to myopically maximize the expected value of information. *INFORMS J. Comput.* **22**(1) 71–80.

Chick, S. E., K. Inoue. 2001a. New procedures to select the best simulated system using common random numbers. *Management Sci.* **47**(8) 1133–1149.

Chick, S. E., K. Inoue. 2001b. New two-stage and sequential procedures for selecting the best simulated system. *Oper. Res.* **49**(5) 732–743.

Cox, D. D., S. John. 1997. SDO: A statistical method for global optimization. N. M. Alexandrov, M. Y. Hussaini, eds. *Multidisciplinary Design Optimization: State of the Art.* SIAM, Philadelphia, 315–329.

Davies, J. W., M. Glick, J. Jenkins. 2006. Streamlining lead discovery by aligning in silico and high-throughput screening. *Current Opinion Chemical Biol.* **10**(4) 343–351.

De Grave, K., J. Ramon, L. De Raedt. 2008. Active learning for high throughput screening. *Discovery Science.* Lecture Notes in Computer Science, Vol. 5255. Springer, Berlin, 185–196.

Ehrlichman, S. M. T., S. G. Henderson. 2008. Comparing two systems: Beyond common random numbers. *Proc. 40th Winter Simul. Conf.*, IEEE Computer Society, Washington, DC, 245–251.

Frazier, P. I. 2011. Learning with dynamic programming. J. J. Cochran, L. A. Cox Jr., P. Keskinocak, J. P. Kharoufeh, J. C. Smith, eds. *Wiley Encyclopedia of Operations Research and Management Science.* John Wiley & Sons, New York.

Frazier, P. I., W. B. Powell, S. Dayanik. 2008. A knowledge-gradient policy for sequential information collection. *SIAM J. Control Optim.* **47**(5) 2410–2439.

Frazier, P. I., W. B. Powell, S. Dayanik. 2009. The knowledge-gradient policy for correlated normal rewards. *INFORMS J. Comput.* **21**(4) 599–613.

Free, S. M., W. J. Wilson. 1964. A mathematical contribution to structure-activity studies. *J. Medicinal Chemistry* **7**(4) 395–399.

Fu, M. C., J.-Q. Hu, C.-H. Chen, X. Xiong. 2007. Simulation allocation for determining the best design in the presence of correlated sampling. *INFORMS J. Comput.* **19**(1) 101–111.

Gelman, A. B., J. B. Carlin, H. S. Stern, D. B. Rubin. 2004. *Bayesian Data Analysis*, 2nd ed. CRC Press, Boca Ration, FL.

Grover, M., B. Singh, M. Bakshi, S. Singh. 2000. Quantitative structure-property relationships in pharmaceutical research—Part 1. *Pharmaceutical Sci. Tech. Today* **3**(1) 28–35.

Gupta, S. S., K. J. Miescke. 1996. Bayesian look ahead one-stage sampling allocations for selection of the best population. *J. Statist. Planning Inference* **54**(2) 229–244.

Hansch, C., T. Fujita. 1964. p-σ-π Analysis: A method for the correlation of biological activity and chemical structure. *J. Amer. Chemical Soc.* **86** 1616–1626.

Hsu, J. C. 1996. *Multiple Comparisons: Theory and Methods.* CRC Press, Boca Raton, FL.

Huang, D., T. T. Allen, W. I. Notz, N. Zeng. 2006. Global optimization of stochastic black-box systems via sequential Kriging meta-models. *J. Global Optim.* **34**(3) 441–466.

Jones, D. R., M. Schonlau, W. J. Welch. 1998. Efficient global optimization of expensive black-box functions. *J. Global Optim.* **13**(4) 455–492.

Jorissen, R. N., M. K. Gilson. 2005. Virtual screening of molecular databases using a support vector machine. *J. Chemical Inform. Modeling* **45**(3) 549–561.

Kaelbling, L. P. 1993. *Learning in Embedded Systems.* MIT Press, Cambridge, MA.

Katz, R., S. F. Osborne, F. Ionescu. 1977. Application of the Free–Wilson technique to structurally related series of homologues: Quantitative structure-activity relationship studies of narcotic analgetics. *J. Medicinal Chemistry* **20**(11) 1413–1419.

Kim, S.-H., B. L. Nelson. 2006. Selecting the best system. S. G. Henderson, B. L. Nelson, eds. *Handbooks in Operations Research and Management Science: Simulation*, Vol. 13. Elsevier, Amsterdam, 501–534.

Kim, S.-H., B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM Trans. Model. Comput. Simul.* **11**(3) 251–273.

Kleijnen, J. 2008. *Design and Analysis of Simulation Experiments.* Springer, New York.

Kleinberg, R. D. 2005. Online decision problems with large strategy sets. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.

Kushner, H. J. 1964. A new method of locating the maximum of an arbitrary multipeak curve in the presence of noise. *J. Basic Engrg.* **86**(1) 97–106.

Močkus, J. 1972. On Bayesian methods for seeking the extremum. *Automatics Comp.* **4**(1) 53–62 (in Russian).

Močkus, J., V. Tiesis, A. Zilinskas. 1978. The application of Bayesian methods for seeking the extremum. L. C. W. Dixon, G. P. Szegö, eds. *Towards Global Optimisation*, Vol. 2. North-Holland, Amsterdam, 117–129.

Montgomery, D. 2005. *Design and Analysis of Experiments*, 6th ed. John Wiley & Sons, New York.

Myers, R., D. Montgomery. 2002. *Response Surface Methodology: Process and Product Optimization Using Designed Experiments.* John Wiley & Sons, Hoboken, NJ.

Nelson, B. L., F. J. Matejcik. 1995. Using common random numbers for indifference-zone selection and multiple comparisons in simulation. *Management Sci.* **41**(12) 1935–1945.

Perevozchikov, A. G. 1991. The approximation of generalized stochastic gradients of random regular functions. *Comput. Math. Math. Phys.* **31**(5) 28–33.

Powell, W. B. 2007. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, Hoboken, NJ.

Sasena, M. J. 2002. Flexibility and efficiency enhancements for constrained global design optimization with Kriging approximations. Ph.D. thesis, University of Michigan, Ann Arbor.

Stuckman, B. E. 1988. A global search method for optimizing nonlinear systems. *IEEE Trans. Systems, Man Cybernetics* **18**(6) 965–977.

Swisher, J. R., S. H. Jacobson, E. Yücesan. 2003. Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey. *ACM Trans. Model. Comput. Simul.* **13**(2) 134–154.

Warmuth, M. K., G. Ratsch, M. Mathieson, J. Liao, C. Lemmen. 2002. Active learning in the drug discovery process. T. G. Dietterich, S. Becker, Z. Ghahramani, eds. *Advances in Neural Information Processing Systems*. MIT Press, Cambridge, MA, 1449–1456.

Warmuth, M. K., J. Liao, G. Ratsch, M. Mathieson, S. Putta, C. Lemmen. 2003. Active learning with support vector machines in the drug discovery process. *J. Chemical Inform. Comput. Sci.* **43**(2) 667–673.

Yang, W. C.-N., B. L. Nelson. 1991. Using common random numbers and control variates in multiple-comparison procedures. *Oper. Res.* **39**(4) 583–591.