

Optimal Learning for Drug Discovery in Ewing’s Sarcoma

Diana Negoescu ^{*} Peter Frazier [†]

Faculty Advisors: Professor Warren B. Powell [‡] Professor Jeffrey Toretzky [§]

Partner Organization: Lombardi Comprehensive Cancer Center - Georgetown University [¶]

Abstract

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 this molecule that best treats a given disease. The problem of choosing the 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. This algorithm uses correlations in our Bayesian prior belief between the performance of different alternatives (molecules) to dramatically reduce the number of molecular tests required, but has heavy computational requirements that limit the number of alternatives that can be considered to a few thousand. We develop computational improvements that allow the knowledge-gradient method to consider much larger sets of alternatives, and demonstrate the method on a problem with 87120 alternatives.

1 Introduction

Ewing’s sarcoma is a pediatric bone cancer primarily affecting adolescents. It has the least favorable prognosis of all primary musculoskeletal tumors (Iwamoto 2007), with an overall 5-year survival rate of 58%, and with only 18-30% of those children presenting metastatic disease at diagnosis surviving beyond this 5 year period. Current treatment usually consists of chemotherapy followed by surgery or radiation therapy, often causing damaging side effects. While the impact of the disease is mitigated by its relatively rarity, with approximately 200 cases diagnosed in the United States each year (Ries et al. 1999), this rarity limits funding for research that might lead to better treatments more quickly.

A medical research group led by Professor Jeffrey Toretzky at the Lombardi Cancer Center at Georgetown University has discovered a chemical compound that shows promise in treating

^{*}Undergraduate student, Princeton University, Department of Operations Research and Financial Engineering

[†]Graduate student, Princeton University, Department of Operations Research and Financial Engineering

[‡]Princeton University, Department of Operations Research and Financial Engineering

[§]Georgetown University, Lombardi Comprehensive Cancer Center

[¶]Contact: Jeffrey Toretzky, jat42@georgetown.edu

Ewing’s sarcoma. If successful as a drug, this compound would improve the patients’ prognosis and avoid many of the side effects caused by chemotherapy and radiation. This compound operates by preventing two proteins, RNA Helicase and EWS-FLI, from binding with each other, thus stopping the spread of the disease. The research group is now searching for variations of this compound that could block binding with even greater efficiency. Such searches are common in drug discovery, where medical researchers often begin with a single molecule or small group of molecules that show some promise for treating a disease, and then test many variations of those molecules to find one that produces the best results.

Variations on a compound are obtained by substituting atoms or groups of atoms at certain locations (sites) on the original chemical molecule with other atoms or groups of atoms (substituents) in order to improve the performance of the molecule in treating the disease. The number of possible variations increases exponentially in the number of sites and substituents, and 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 in order 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. While 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 in a drug discovery application and compare its performance to that of other policies. The dataset to which we apply the KGCB algorithm is a collection of narcotic analgetics for which data is publicly available. As a result of these computational improvements and the successful application of the KGCB method to this narcotic analgetics dataset, the medical research team at Georgetown University is now in the initial stages of using the KGCB algorithm to suggest which compounds to measure in their search for a treatment for Ewing’s sarcoma.

Although we designed the implementations of KGCB algorithm in this paper for use in drug discovery, they also may be applied to any problem in which beliefs can be approximated using a linear additive model. Such problems include:

- Choosing the shortest path through a network. The length along a path is exactly or approximately equal to the sum of the lengths along each link.
- Choosing members of a team in order to maximize performance. The performance of a team is approximately equal to the sum of the contributions of each member.
- Choosing features to include in a product. Both the value of a product to a consumer, and the cost to produce the product, are approximately equal to the sum of the values or costs of the features included.
- Choosing a collection of advertisements to include on a webpage. The number of click-throughs generated by a collection of advertisements is approximated by the number of click-throughs that would be received by each ad viewed individually.

Note that in each of these examples, we may optionally model combination effects with a linear model through terms specific to each combination.

Traditionally, 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 & 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 considers so-called Value of Information Procedures (VIP) (see, e.g., Chick & Inoue (2001)). VIP 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 & Miescke (1996), which uses a one-step analysis under an independent normal

prior. 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. (2007). Frazier et al. (2008) uses 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 that one has before the sample, and the knowledge that one has after it.

The knowledge-gradient approach of Gupta & Miescke (1996) is extended to correlated normal beliefs in Frazier et al. (2009) and its 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, often resulting in a dramatic improvement in measurement efficiency. While an abundant literature in ranking and selection treats beliefs that are independent across alternatives, correlated beliefs have received much less attention in that literature, and, to our knowledge, Frazier et al. (2009) is the first to consider them. Correlated priors have a rich history, however, within the literature on Bayesian global optimization, where Gaussian process priors model the similarity between values of a continuous function evaluated at nearby points. Then, similarly to 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. See Sasena (2002) for a review.

Preceding any Bayesian ranking and selection algorithm for drug discovery must be a prior belief on the values of the candidate compounds. Several such priors 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 would include those that test a compound’s ability to kill diseased cells, or to 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 Crum-Brown and Fraser published the first formulation of a quantitative relationship between ‘physiological activity’ and ‘chemical structure’ (Crum-Brown & Fraser (1868)). The modern origins of the field lie in the Free-Wilson model (Free & Wilson (1964)) and the Hansch Analysis (Hansch & Fujita (1964)). In both methods, multiple linear regression is used to correlate activity on a laboratory test with a collection of compound features. In the Free-Wilson model these features are indicator variables indicating the presence or absence of substituents at given sites, and in the Hansch Analysis these features are physiochemical properties such as molecular weight and lipophilicity. Many of the statistical methods have been applied to provide

QSAR methods, 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 Free-Wilson 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 classic algorithm. While the classic implementation manipulates a belief on the values of the *compounds*, our new implementations manipulate beliefs on the underlying contributions of the substituents at each site. Since the number of compounds is exponential in the number of substituents per site, maintaining a belief in this way results in substantial computational and memory savings, allowing the KGCB algorithm to be used on problems with even hundreds of thousands of candidate compounds.

The paper is organized as follows. In section 2, we describe the ranking and selection problem considered and the models describing relationships between the structure and value of an alternative. In Section 3 we review the classic implementation of the KGCB algorithm and in Sections 4 and 5 we describe two computational improvements to this implementation. In Section 6 we present the drug discovery problem and the computational and real-world performance of the KGCB algorithm on a narcotic analgetics data set. Section 7 presents our conclusions.

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 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 $(\vartheta_1, \dots, \vartheta_M)'$. We further assume that our belief about ϑ is distributed according to a multivariate normal prior with mean vector μ^0 and positive semi-definite 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, \dots, x^{N-1} . The measurement decision x^n selects an alternative to test at time n from the set $\{1, \dots, M\}$. 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 $x^0, \hat{y}^1, x^1, \hat{y}^2, \dots, x^{n-1}, \hat{y}^n$. We write \mathbb{E}_n and Var_n to indicate $\mathbb{E}[\cdot | \mathcal{F}^n]$ and $\text{Var}[\cdot | \mathcal{F}^n]$, the conditional expectation and variance, respectively, taken with respect to \mathcal{F}^n . Then define $\mu^n := \mathbb{E}_n[\vartheta]$, and $\Sigma^n := \text{Cov}[\vartheta | \mathcal{F}^n]$. Conditionally on \mathcal{F}^n , our posterior belief 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, $\Pi := \{(x^0, \dots, x^{N-1}) : x^n \in \mathcal{F}^n\}$. We let π be a generic policy in Π and we write \mathbb{E}^π 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 0 or 1. 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 & Wilson (1964)) assumes that each attribute contributes additively to the value of the alternative. Denote by a_i the contribution of attribute i , and by s^x a vector of 0’s and 1’s, with a 1 for every attribute that is present in alternative x . Thus, $s_i^x = 1$ means that the i^{th} 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 alternative 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 meeting this

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

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

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

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

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

We suppose that we have an independent normal prior on ζ and a_1, \dots, a_k . Under this prior, 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

$$\begin{aligned} \Sigma_{ij}^0 = \text{Cov}(i, j) &= \text{Cov} \left(\sum_m a_m s_m^i + \zeta, \sum_{m'} a_{m'} s_{m'}^j + \zeta \right) \\ &= \text{Var}(\zeta) + \sum_{m, m'} s_m^i s_{m'}^j \text{Cov}(a_m, a_{m'}) = \text{Var}(\zeta) + \sum_{m \in \mathcal{L}_{ij}} \text{Var}(a_m), \end{aligned}$$

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

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, the terms b_1, \dots, b_M are normally distributed with mean 0 and a variance σ_b^2 . Additionally, 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

$$\begin{aligned} \Sigma_{ij}^0 = \text{Cov}(i, j) &= \text{Cov} \left(\sum_m a_m s_m^i + \zeta + b_i, \sum_{m'} a_{m'} s_{m'}^j + \zeta + b_j \right) \\ &= \text{Var}(\zeta) + \sum_{m, m'} s_m^i s_{m'}^j \text{Cov}(a_m, a_{m'}) + \sigma_b^2 \mathbf{1}_{\{i=j\}} \\ &= \text{Var}(\zeta) + \sum_{m \in \mathcal{L}_{ij}} \text{Var}(a_m) + \sigma_b^2 \mathbf{1}_{\{i=j\}}. \end{aligned}$$

Thus, under this general model, the covariance matrix of the values of the alternatives under our prior belief 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 Section 2.1. When σ_b^2 is non-zero but still significantly smaller than $\text{Var}(a_l)$, it models a situation in which the values of the alternatives are well-approximated by an additive structure, but also understands that there may be small deviations from it. When σ_b^2 is very large, and in particular much larger than the $\text{Var}(a_l)$ 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 non-zero but not too large, that this general model is likely to be most useful. It allows exploiting the additive structure of a problem to learn quickly, while simultaneously allowing 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, 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 in order to come close to finding a good alternative, we need to make at least as many measurements as the number of b 's. In such situations, we must either find other non-additive structure in the problem, or resign ourselves to making a number of measurements that is larger than the number of alternatives.

3 The Knowledge-Gradient 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

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

where $S^n := (\mu^n, \Sigma^n)$ parameterizes the posterior belief at measurement n . The knowledge-gradient (KG) factor, $\nu^{KG,n}$, 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 may be calculated using standard results

for normal sampling with a multivariate normal prior (see, e.g., Gelman et al. (2004)) as

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

where e_x is a column M -vector with a single 1 at index x and the rest 0s.

We now describe the time- n conditional distribution of μ^{n+1} , which 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

$$\begin{aligned}x^{KG}(s) &= \arg \max_x \mathbb{E} \left[\max_i (\mu_i^n + \tilde{\sigma}_i(\Sigma^n, x^n)Z) \mid S^n, x^n = x \right] - \max_i \mu_i^n \\ &= \arg \max_x h(\mu^n, \tilde{\sigma}(\Sigma^n, x)).\end{aligned}$$

Here, $h : \mathbb{R}^M \times \mathbb{R}^M \rightarrow \mathbb{R}$ 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 3.2. This algorithm in turn uses Algorithm 3.1 in an inner loop. In Algorithm 3.2, the components of p and q are sorted, and then some are dropped, resulting 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 $f(z) = \varphi(z) + z\phi(z)$, and where φ is the normal cumulative

distribution function and ϕ is the normal density.

Algorithm 3.1: CALCULATE c AND THE SET A (*Inputs* : p, q)

```

 $c_0 \leftarrow -\infty, c_1 \leftarrow +\infty, A \leftarrow \{1\}$ 
for  $i \leftarrow 1$  to  $M - 1$ 
  do
     $c_{i+1} \leftarrow +\infty$ 
    repeat
       $j \leftarrow A[\text{end}(A)]$ 
       $c_j \leftarrow (p_j - p_{i+1}) / (q_{i+1} - q_j)$ 
       $k \leftarrow A[\text{end}(A) - 1]$ 
      if  $c_j \leq c_k$ ,
        then
           $A \leftarrow A(1, \dots, \text{end}(A) - 1)$ 
           $\text{loopdone} \leftarrow \text{false}$ 
        else
           $\text{loopdone} \leftarrow \text{true}$ 
    until  $\text{loopdone}$ 
     $A \leftarrow (A, i + 1)$ 

```

Algorithm 3.2: COMPUTE h (*Inputs* : p, q)

```

sort  $(p_i, q_i)_{i=1}^M$  such that  $p_i$  are in non-decreasing order
and ties broken such that  $p_i \leq p_{i+1}$  if  $q_i = q_{i+1}$ 
for  $i \leftarrow 1$  to  $M - 1$ 
  do
    if  $q_i = q_{i+1}$ 
      then
        Remove entry  $i$  from  $(p_i, q_i)_{i=1}^M$ 
    Use algorithm 3.1 to compute  $c$  and  $A$  from  $p, q$ 
     $p \leftarrow p[A], q \leftarrow q[A], c \leftarrow (c[A], +\infty), M \leftarrow \text{length}(A)$ 
return  $(\log) \left( \sum_{i=1}^{M-1} (q_{i+1} - q_i) f(-|c_i|) \right)$ 

```

Using Algorithm 3.2, 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 computing 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 3.3.

Algorithm 3.3: KGCB₁ ALGORITHM (*Inputs* : μ^n, Σ^n)

```

for  $x \leftarrow 1$  to  $\text{length}(\mu^n)$ 
  do
     $p \leftarrow \mu^n$ 
     $q \leftarrow \tilde{\sigma}(\Sigma^n, x)$ 
     $\nu \leftarrow h(p, q)$  % use Algorithm 3.2
    if  $x = 1$  or  $\nu > \nu^*$ 
      then
         $\nu^* \leftarrow \nu, x^* \leftarrow x$ 

```

Within Algorithm 3.3, Algorithm 3.2 executes M times. Within one execution of Algorithm 3.2, the sort has complexity $O(M \log M)$ and Algorithm 3.1 has complexity $O(M)$. Thus,

the most computationally demanding step within Algorithm 3.2 is the sort, and the overall complexity of the KGCB algorithm as computed by Algorithm 3.3 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 3.3, 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 computing KGCB for large numbers of linearly related alternatives such as those encountered in drug discovery.

4 A First Improvement

In this section, we present a first computational improvement to the standard implementation of KGCB that exponentially reduces the computational and storage requirements of the 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.

We first describe this improvement in the context of the Free-Wilson model from Section 2.1 in Sections 4.1 and 4.2, and then we extend this improvement to the general model from Section 2.2 in Section 4.3.

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 Section 2.1.

Let α be the vector of attribute values $\alpha = (\zeta, a_1, \dots, a_k)$ 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 Section 2.1 and we let X be a matrix comprised of rows representing the alternative. Each row of X is a vector of 0's and 1's of the same length as α , and each 1 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 1 in the first entry to indicate that the base molecule (whose value is ζ) is present, and then the subsequent entries contains a 1 for each substituent present. Thus, this row is a 1 followed by the vector s^x (defined in Section 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 ,

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

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

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

where e_i is, as before, a column vector of length the size of our alternative database, with a 1 on position i and zeros everywhere else. Thus, the belief induced on ϑ by (2) is

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

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

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 Section 3. Before providing this expression, we first introduce some additional notation. Let $\tilde{x}^n = (\tilde{x}_0^n; \tilde{x}_1^n; \dots; \tilde{x}_k^n)^T$ be a column vector of 0s and 1s describing the alternative $x^n \in \{1, \dots, M\}$ 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 1 for those attributes i present in alternative x^n and 0 otherwise. Additionally, 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)),

$$\begin{aligned} \theta^{n+1} &= \theta^n + \frac{\hat{\epsilon}^{n+1}}{\gamma^n} C^n \tilde{x}^n, \\ C^{n+1} &= C^n - \frac{1}{\gamma^n} (C^n \tilde{x}^n (\tilde{x}^n)^T C^n). \end{aligned} \quad (4)$$

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 Section 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 5 sites and 9 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 to

only 45 total substituents. In this case, Σ^n is $10^5 \times 10^5$ while C^n is only 46×46 (we add 1 to 45 to account for the base molecule).

4.2 Improved Implementation of KGCB

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 Section 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 , $\nu_x^{KG} = h(\mu^n, \tilde{\sigma}(\Sigma^n, x))$, only through $\tilde{\sigma}(\Sigma^n, x)$. This quantity is given by $\tilde{\sigma}(\Sigma^n, x) = \Sigma^n e_x / \sqrt{\lambda_x + \Sigma^n_{xx}}$, which depends only upon row x in Σ^n , and not 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

$$\Sigma^n_{x,\cdot} = (XC^nX^T)e_x = (XC^n)X^T_{x,\cdot}.$$

By calculating row x of Σ^n from C^n , then 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 4.1.

Algorithm 4.1: KGCB₂ ALGORITHM(*Inputs* : θ^n, C^n, X)

```

 $\mu^n \leftarrow X\theta^n$ 
 $B \leftarrow XC^n$ 
for  $x \leftarrow 1$  to  $M$ 
  do
     $\Sigma^n_{x,\cdot} \leftarrow BX^T_{x,\cdot}$ 
     $a \leftarrow \mu^n$ 
     $b \leftarrow \Sigma^n_{x,\cdot} / \sqrt{\lambda_x + \Sigma^n_{xx}}$ 
     $\nu \leftarrow h(a, b)$  % use Algorithm 3.2
    if  $x = 1$  or  $\nu > \nu^*$ 
      then
         $\nu^* \leftarrow \nu, x^* \leftarrow x$ 

```

Algorithm 4.1 is similar to Algorithm 3.3, with the main differences being the first operation, which retrieves the mean belief on alternatives from the mean of the belief on attributes, and the second and third operations (the latter being the first step inside the loop), 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.

This algorithm significantly improves upon Algorithm 3.3 because it computes only $\Sigma^n_{x,\cdot}$, 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 Extending this Improvement to the General Model

This 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 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 on b_x will have the same marginal distribution that it had under the prior, 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, \dots, a_k , and only those b_x terms for compounds we have measured, we can reconstruct our belief about those deviation terms we are not explicitly tracking as needed.

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

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 $x^n \in \{x^0, \dots, x^{n-1}\}$, 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 0 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 0s, 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 on α_n is $\mathcal{N}(\tilde{\theta}^{n-1}, \tilde{C}^{n-1})$. Additionally, let \tilde{x}^n be a column vector of 0s and 1s, with a 1 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) as

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

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

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

We now show how KG factors may be computed from a belief on α_n parameterized by θ^n and C^n . Since 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 to 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 0s and 1s whose each row corresponds to an alternative, and has a 1 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 1 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} . Thus, we can 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 x^{th} column of Σ^n is

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

and $\tilde{\sigma}(\Sigma^n, x^n)$ may be computed from this vector via $\tilde{\sigma}(\Sigma^n, x) = \Sigma^n e_x / \sqrt{\lambda_x + \Sigma_{xx}^n}$. The resulting method of computing the KGCB policy is summarized below in Algorithm 4.2.

Algorithm 4.2: GENERAL KGCB₂ ALGORITHM(*Inputs* : θ^n, C^n, X^n)

```

 $\mu^n \leftarrow X^n \theta^n$ 
for  $x \leftarrow 1$  to  $M$ 
  do
    { Compute  $\tilde{C}^n$  from  $C^n$  and  $x$ .
       $\Sigma_{x,\cdot}^n \leftarrow (\tilde{X}^n \tilde{C}^n)(\tilde{X}_{x,\cdot}^n)^T$ 
       $a \leftarrow \mu^n$ 
       $b \leftarrow \Sigma_{x,\cdot}^n / \sqrt{\lambda_x + \Sigma_{xx}^n}$ 
       $\nu \leftarrow h(a, b)$  % use Algorithm 3.2
      if  $x = 1$  or  $\nu > \nu^*$ 
        then
           $\nu^* \leftarrow \nu, x^* \leftarrow x$ 
    }
```

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 A Further Improvement for the Free-Wilson Model

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

This proposition states that the calculation of each KG factor, which would ordinarily require considering 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. Since the number of compounds grows exponentially with the number of substituents per site, and 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. While the KGCB policy still needs to calculate the KG factor corresponding to each possible measurement to find the largest one, faster calculation of each KG factor makes the overall computation much faster as well.

Before stating the proposition, recall that $\theta_j^n = \mathbb{E}_n[a_j]$ and $\theta_0^n = \mathbb{E}_n[\zeta]$ are the means of the time- n posterior belief 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,*

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

where $(\tilde{\sigma}_{xi}^n)^2 := \text{Var}_n [\theta_i^{n+1} | x^n = x]$ for $i > 0$ and $\theta_{-1}^n = \tilde{\sigma}_{x,-1}^n = 0$.

Proof: We write

$$\begin{aligned} \max_{x'} \mu_{x'}^{n+1} &= \max_{x'} \mathbb{E}_{n+1} \left[\sum_j a_j s_j^{x'} + \zeta \right] = \max_{x'} \sum_i \theta_i^{n+1} s_i^{x'} + \theta_0^{n+1} \\ &= \max_{x'} \sum_{\ell} \left(\sum_i \theta_i^{n+1} \mathbf{1}_{\{L(i)=\ell\}} \right) + \theta_0^{n+1} = \sum_{\ell} \max_{i \in A(\ell) \cup \{-1\}} \theta_{\ell i}^{n+1} + \theta_0^{n+1}, \end{aligned}$$

where $\theta_{\ell-1}^{n+1}$ is defined to be 0, and where the crucial final step is due to the fact that 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

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

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

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

Since the joint distribution of $(\theta_i^{n+1})_{i \in A(\ell)}$ conditioned on \mathcal{F}_n and $x^n = x$ is the same as that of the \mathcal{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 random variable, and $\theta_{-1}^{n+1} = 0 = \theta_{-1}^n + \tilde{\sigma}_{x-1}^n Z$ almost surely, we may substitute for θ_i^{n+1} to obtain

$$\begin{aligned} \nu_x^{n,KG} &= \sum_{\ell} \mathbb{E}_n \left[\max_{i \in A(\ell) \cup \{-1\}} \theta_i^n + \tilde{\sigma}_{xi}^n Z \right] - \max_{i \in A(\ell) \cup \{-1\}} \theta_i^n \\ &= \sum_{\ell} h \left((\theta_i^n)_{i \in A(\ell) \cup \{-1\}}, (\tilde{\sigma}_{xi}^n)_{i \in A(\ell) \cup \{-1\}} \right). \quad \square \end{aligned}$$

To actually use the decomposition property inherent in this proposition, we must first compute $\tilde{\sigma}_{xi}^n$. We have $(\tilde{\sigma}_{xi}^n)^2 = \text{Var}_n [\mathbb{E}_{n+1}[a_i] | x^n = x]$ 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

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

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

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

We summarize the resulting implementation below in Algorithm 5.1.

Algorithm 5.1: KGCB₃ ALGORITHM(*Inputs* : θ^n, C^n, X)

```

for  $x \leftarrow 1$  to  $M$ 
   $\tilde{\sigma}_{x,\cdot}^n \leftarrow \frac{C^n \tilde{x}^n}{\sqrt{\lambda_{x^n} + (\tilde{x}^n)^T C^n \tilde{x}^n}}$ 
   $\nu_x \leftarrow 0$ 
  for  $\ell \leftarrow 1$  to  $L$ 
  do  $\left\{ \begin{array}{l} p \leftarrow (\theta_i^n)_{i \in A(\ell) \cup \{-1\}}$ 
       $q \leftarrow (\tilde{\sigma}_{xi}^n)_{i \in A(\ell) \cup \{-1\}}$ 
       $\nu_x \leftarrow \nu_x + h(p, q)$  % use Algorithm 3.2
    \end{array} \right.
  if  $x = 1$  or  $\nu_x > \nu^*$ 
  then  $\nu^* \leftarrow \nu_x, x^* \leftarrow x$ 

```

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 4.1 is $O(M^2 \ln M)$, which is equivalent to $O(l (M_l)^{2l} \ln (M_l))$. The computational complexity of Algorithm 5.1 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 to $O(l (M_l)^{2l} \ln (M_l))$ 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 that used the Free-Wilson model Katz et al. (1977). The paper contains a set of 6,7-Benzomorphans, which have 5 sites at which substituents can be attached. The molecule is shown in Figure 1. At site 1 there are 11 possible substituents (together with H), at site 2 there are 8 possible substituents, at site 3 there are 5 possible substituents, at site 4 there are 6 possible substituents and at site 5 there are 11 possible substituents. The collection of compounds resulting from choosing a substituent at each site contains $11 \times 8 \times 5 \times 6 \times 11 = 29040$ compounds from substituents at the 5 positions. Additionally, each compound can be charged positively, negatively or be neutral, which brings the number of compounds to 87120. The paper provides experimentally observed activity values for 99 of these compounds.

In the Section 6.1, we describe methods for choosing the prior on the values of this collection of compounds. Then, in Section 6.2, we present an empirical study of the performance of KGCB on this collection of compounds.

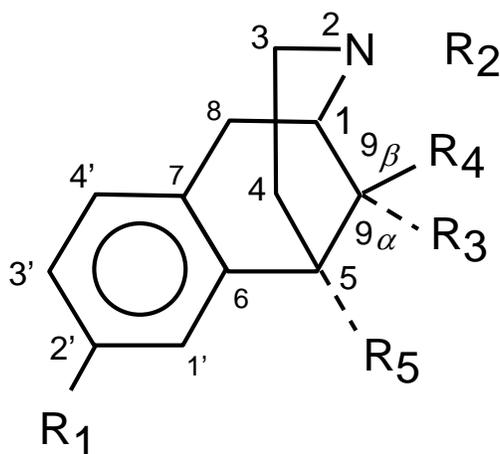


Figure 1: An example of substituent locations (marked with R_1, \dots, R_5). Source: Katz et al. (1977)

6.1 Setting up a prior on the substituent contributions

When choosing a prior distribution, one may be in contact with experienced practitioners who may be able to articulate their prior beliefs on the value 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 for the experts to articulate their 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 priors for the Free-Wilson model. Our method of choosing a prior 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 non-informative belief on the value ζ of the base molecule, which was subsequently altered by a laboratory measurement. This measurement provides an informative belief on the value of the base molecule that we will use within our prior whose mean is the value of the measurement, and whose variance is the variance of this measurement. A measurement of the base molecule is generally available in practice, since 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 is given by taking θ_0^0 and C_{00}^0 to be the value and measure-

ment variance, respectively, of our previous measurement of the base molecule, and 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 0. These values for θ^0 and C^0 specify a prior as discussed in Section 4.1, and the equivalent prior on compound values may be reconstructed as discussed in that section.

We now discuss how this method of setting a prior can be extended to the general model. To do so, 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, and, without loss of generality, we may assume that the mean of the population of deviation terms is 0 because we can take any non-zero mean and move it into the value of the base molecule. Then the prior, in the format of Section 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. There is no need to use different values of θ^0 or C^0 because, even in the general model, our θ^0 and C^0 do not encode beliefs on the deviation terms, as these beliefs are implicitly independent with mean 0 and variance σ_b^2 .

In order to use either of these methods for setting the prior, 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). This paper provides measured values for 99 of the possible 87120 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 of the various implementations of the Free-Wilson-based KGCB implementation on the set of 87120 compounds. Algorithm 5.1 is able to compute the KGCB decision and update the belief in less than 3 seconds, which is approximately 100 times faster than Algorithm 4.1. The standard implementation was so slow that we were unable to determine its runtime on the problem.

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 against two policies: a pure exploration policy that chooses compounds to test uniformly at random, and a one-factor-at-a-time (Montgomery (2005)) policy which begins by testing the base molecule and then replaces each substituent at a time, cycling through the same compounds.

Both policies update their beliefs using the same Bayesian updates as the KGCB algorithm. We emphasize that these belief updates use correlated beliefs, which give our pure exploration policy a substantial advantage over any policy that uses independent belief updates. With our version of pure exploration, even though we are choosing compounds to measure uniformly at random, we learn about the compounds through correlations the same way as we learn when we use the KGCB policy.

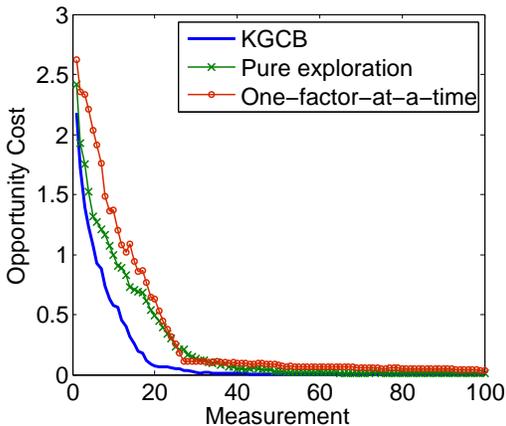
We considered the one-factor-at-a-time policy even though it is often regarded as a poor policy because it neglects interaction between factors (Montgomery (2005)) because, in the perfectly linear Free-Wilson model these interaction terms do not exist. By using a policy like one-factor-at-a-time that neglects these terms in tests that do not include them, we provide an advantage to the policy against which KGCB competes. In cases like the general model in which interaction terms exist, a factorial design would perhaps be more appropriate (Montgomery (2005)), but such designs are infeasible in problems with many dimensions like the drug discovery problem considered here.

In the next section (Section 6.2.1) we present numerical results from the Free-Wilson model described in Section 2.1, and in Section 6.2.2 we present numerical results from the general model described in Section 2.2.

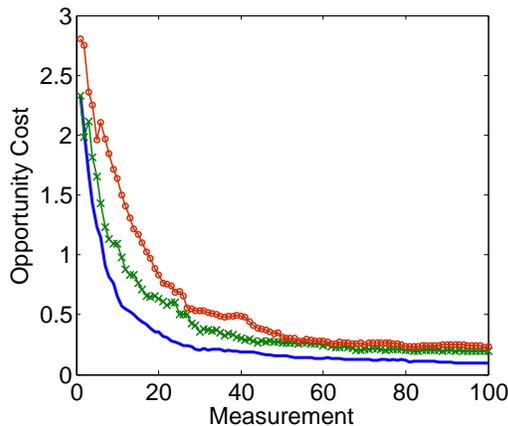
6.2.1 Results using the Free-Wilson model

We begin by describing our results using the simple, linear-additive model described in Section 2.1. We assume that we have a budget of 100 measurements, and at each time, we plot the opportunity cost of each policy, defined as the difference between the actual highest compound value and the true value of the compound which is best according to the policy’s posterior belief.

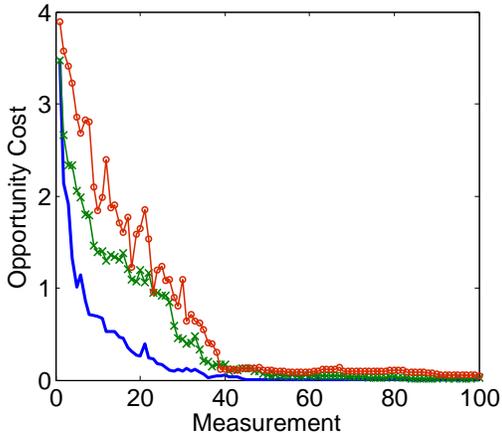
Figure 2: Opportunity cost for varying data set sizes and measurement noise.



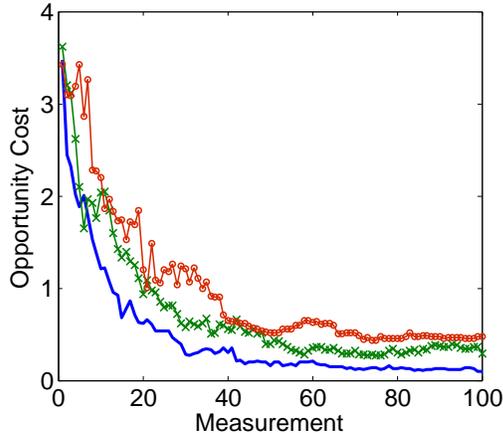
2a: Average opportunity cost over 100 runs using a data set of 2640 compounds and a noise standard deviation of 0.1.



2b: Average opportunity cost over 100 runs using a data set of 2640 compounds and a noise standard deviation of 0.5.



2c: Average opportunity cost over 10 runs using a data set of 87120 compounds and a noise standard deviation of 0.1.



2d: Average opportunity cost over 10 runs using a data set of 87120 compounds and a noise standard deviation of 0.5.

Thus, once the opportunity cost is 0, the policy has found the best compound.

Figure 2 presents the results using a data set of 2640 compounds and one of 87120 compounds, for two levels of noise: a noise standard deviation of 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 2640 compounds data set, we average over 100 runs, and for the 87120 compounds data set we average over 10 runs. As Figure 2 shows, the average opportunity cost for KGCB is always lower than the average opportunity cost of the other two policies. Among the other two policies, pure exploration seems to outperform one-factor-at-a-time (OFAAT), due to the fact that OFAAT cycles through the same set of compounds when measuring, whereas pure exploration chooses at random among all compounds. In general, a

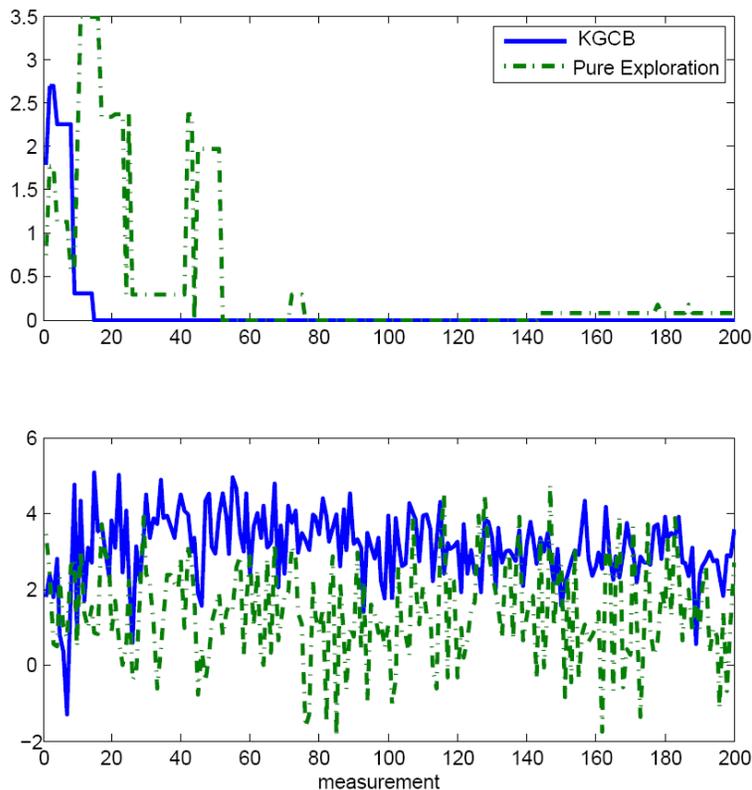


Figure 3: A sample path of 200 measurements on a data set of 1000 compounds showing opportunity cost (top) and the true values of the compounds tested at each time step (bottom).

higher level of noise leads to worse performance by all policies, but relative to the other policies, KGCB performs best. Since these initial experiments showed that pure exploration is more competitive than OFAAT, our following experiments compare KGCB only to pure exploration.

In further experiments, we randomly selected sets of compounds from the full set of 87120. Figure 3 shows a sample path of 200 measurements on a set of randomly selected 1000 compounds, giving the opportunity cost at every iteration on the top panel and the true value of the compound that is measured at each step on the bottom panel. While pure exploration fails to find the best compound in the first 200 measurements, KGCB manages to do so in about 15 measurements, which is a significant improvement if we think about the amount of time and money each measurement requires. As the bottom panel shows, part of KGCB’s success in this sample path was the fact that it measured two very good compounds early on, at steps 10 and 15, while pure exploration was measuring mediocre compounds for a long period of time, which prevented it from discovering the best compound quickly enough.

To better assess the difference in performance between KGCB and pure exploration, we also ran 15 sample paths and plotted, on the same graph, the mean and standard deviation of the mean for KGCB and pure exploration. Figure 4 shows the resulting plot, where each error bar is of length twice the standard deviation of the mean in each direction (computed as

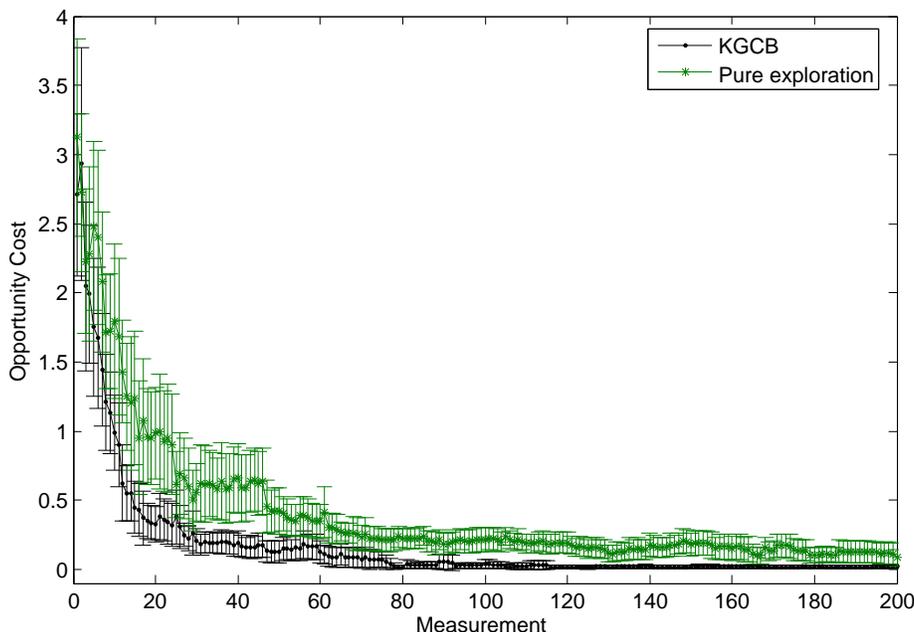


Figure 4: Mean and standard deviation of the mean using 15 sample paths of 10000 compounds each, using a noise standard deviation of 0.38.

standard deviation/ $\sqrt{n-1}$, where n , the number of samples, is 15 in our case.) As Figure 4 shows, the estimated mean opportunity cost for the KGCB policy is lower than that of the pure exploration policy.

To get a sense of the distribution of the relative performance of KGCB vs pure exploration, we ran our code on 75 sets of randomly selected 10000 compounds out of the entire data set of 87120 compounds, with a budget of 200 measurements. For each measurement and sample path, we compute the difference between the pure exploration opportunity cost and the KGCB opportunity cost, and then for each measurement, we plot the mean difference and standard deviation of the difference in opportunity cost across the 75 sample paths. The results are shown in Figure 5. As illustrated in Figure 5, the mean difference is always positive after the first measurement, and it has a maximum at about the tenth measurement, suggesting that the learning rate is faster on average for KGCB than for pure exploration, and the value of using the KGCB policy is maximized at early measurements.

Numerical evidence indicates that KGCB has better average performance than pure exploration. For more numerical work using the Free-Wilson model, please see the appendix.

6.2.2 Numerical results for the general structural model

We now present the results using the general model described in Section 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), and for each compound x , we generate b_x independently from a normal distribution with

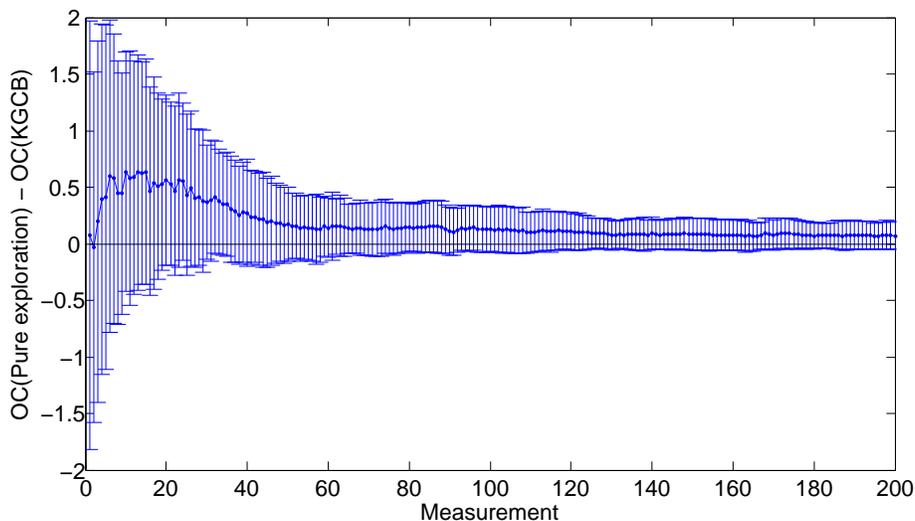


Figure 5: Mean and standard deviation of difference in opportunity cost between pure exploration and KGCB using 75 sample paths of 10000 compounds each and a noise standard deviation of 0.38.

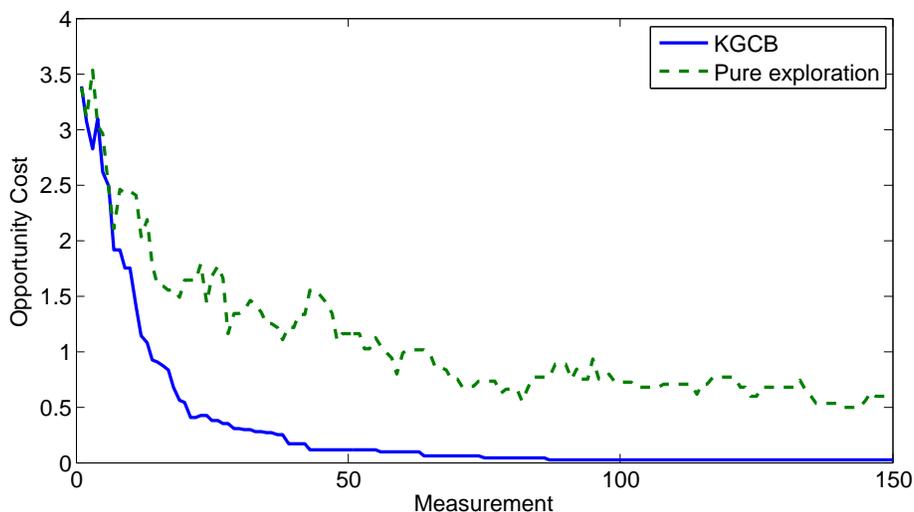


Figure 6: Average opportunity cost over 10 runs using different data sets of 1000 compounds.

mean 0 and variance σ_b^2 (where σ_b^2 was obtained from the fit as well, as described in Section 6.1). We then combine these values according to the model to obtain true values for the compounds.

To test the performance of KGCB under this model, we randomly selected 10 different sets of 1000 compounds each and ran the KGCB algorithm using the prior described in Section 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 6.

As with the Free-Wilson model, the KGCB policy performs significantly better than pure exploration. 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 1000 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 1000 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.

7 Conclusions

Drug discovery is a long and expensive process that requires synthesizing and testing many molecules in order to find one that is efficient in treating disease. Our simulation results show that the KGCB policy reduces the number of molecules that need to be tested in this process, saving time and money. Furthermore, since budgets are limited and a search for a new drug is declared a failure if it does not find a good molecule within this budget, more efficient search procedures may cause the success of drug discovery efforts that otherwise would have failed. Previous implementations of the KGCB policy required too much computational effort to be practically applicable to large problems like drug discovery, and the new implementations presented here overcome this barrier. Although further mathematical effort is needed, particularly in creating sequential search methods that use models from medicinal chemistry beyond Free-Wilson, we believe that this effort will significantly improve our ability to discover new drugs.

The Georgetown University team has just started to test a long sequence of compounds, and are planning to use the KGCB policy to help choose compounds to test. As a starting point, they have decided to use the Free-Wilson model for modeling structure-activity relationships and an informative prior obtained from their initial measurements. Future improvements might involve using a different QSAR model such as the Hansch model or the partial least squares method. Although the KGCB method can benefit all forms of drug discovery, for the time being we hope it will help the research team at Georgetown find a cure for Ewing’s sarcoma.

Acknowledgments

We would like to thank Jeffrey Toretzky, Sivanesan Dakshanamurthy, and the many others at the Lombardi Comprehensive Cancer Center at Georgetown University for providing their substantial expertise and for teaching us about the broader context that motivated this paper. We would also like to thank the Go4TheGoal foundation for introducing us to drug discovery, and for encouraging us to pursue this work.

References

Chen, C., Dai, L. & Chen, H. (1996), ‘A gradient approach for smartly allocating computing budget for discrete event simulation’, *Simulation Conference Proceedings, 1996. Winter* pp. 398–405.

- Chick, S., Branke, J. & Schmidt, C. (2007), New myopic sequential sampling procedures. Submitted to *INFORMS J. on Computing*.
- Chick, S. & Inoue, K. (2001), 'New two-stage and sequential procedures for selecting the best simulated system', *Operations Research* **49**(5), 732–743.
- Crum-Brown, A. & Fraser, T. (1868), *On the connection between chemical constitution and physiological action. Part 1. On the physiological action of salts of the ammonium bases, derived from Strychnia, Brucia, Thebia, Codeia, Morphia, and Nicotia.*
- Frazier, P., Powell, W. B. & Dayanik, S. (2008), 'A knowledge-gradient policy for sequential information collection', *SIAM Journal on Control and Optimization* **47**(5).
- Frazier, P., Powell, W. B. & Dayanik, S. (2009), 'The knowledge-gradient policy for correlated normal rewards', *To appear in INFORMS Journal on Computing*.
- Free, S. & Wilson, W. (1964), 'Contribution to structure-activities studies', *J Med Chem* **7**, 395–399.
- Gelman, A., Carlin, J. & Stern, H. (2004), *Bayesian Data Analysis*, CRC Press.
- Grover, M., Singh, B., Bakshi, M. & Singh, S. (2000), 'Quantitative structure-property relationships in pharmaceutical research - Part 1', *Pharmaceutical Science and Technology Today* **3**(1), 28–35.
- Gupta, S. & Miescke, K. (1996), 'Bayesian look ahead one-stage sampling allocations for selection of the best population', *Journal of statistical planning and inference* **54**(2), 229–244.
- Hansch, C. & Fujita, T. (1964), 'p- σ - π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure', *Journal of the American Chemical Society* **86**(8), 1616–1626.
- Iwamoto, Y. (2007), 'Diagnosis and Treatment of Ewing's Sarcoma', *Japanese Journal of Clinical Oncology*.
- Katz, R., Osborne, S. & Ionescu, F. (1977), 'Application of the Free-Wilson technique to structurally related series of homologues. Quantitative structure-activity relationship studies of narcotic analgetics.', *J Med Chem* **20**(11), 1413–9.
- Kim, S. & Nelson, B. (2006), 'Selecting the best system', *Handbooks in Operations Research and Management Science: Simulation*.
- Montgomery, D. (2005), 'Design and Analysis of Experiments (6th edn.) John Wiley and Sons', New York, NY.
- Powell, W. B. (2007), *Approximate Dynamic Programming: Solving the curses of dimensionality*, John Wiley and Sons, New York.
- Ries, L. A. G., Smith, M. A., Gurney, J., Linet, M., Tamra, T., Young, J. L. & Bunin, G. (1999), 'Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975-1995', *National Cancer Institute, SEER Program, NIH Pub* pp. 99–4649.
- Sasena, M. (2002), Flexibility and efficiency enhancements for constrained global design optimization with kriging approximations, PhD thesis, University of Michigan.
- Swisher, J., Jacobson, S. & Yucesan, E. (2003), 'Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey', *ACM Transactions on Modeling and Computer Simulation* **13**(2), 134–154.

8 Appendix: More Numerical Results

We first present the results from running the code on data sets of randomly selected 10000 compounds. Figure 7 shows four such sample paths of the new implementation. As the figure shows, the performance of the KGCB algorithm is encouraging, as it is usually able to find the best compound by the 75th measurement, and gets very close to finding it by the 25th measurement. In most sample paths, KGCB does at least as well as pure exploration - the policy in which we choose compounds to measure completely at random, but update the belief in the same way as we do for the KGCB policy.

We are interested also in the typical performance of the KGCB policy. That is, if we are able to perform only one sample path (which is generally the case in practice), we are interested in knowing the probability that the KGCB policy performs better than an exploration policy.

To better get a sense of the distribution of the relative performance of KGCB versus pure exploration, we have made a box and whiskers plot of the differences between opportunity costs at every 10 measurements. As shown in Figure 8, except for the first measurement, all the other measurements have the lower quartile cutoff at or above 0, which suggests that the probability that KGCB is better than pure exploration, on any measurement of any sample path, is higher than 1/2.

Having tested the KGCB policy on data sets of 10000 compounds, we increased the size of the data sets further to 25000 compounds. Figure 9 shows four sample paths from running the KGCB and pure exploration policies on four randomly selected data sets of 25000 compounds. The rate of convergence for these plots is slower than for 10000 compounds, but the KGCB policy still manages to get reasonably close to the best compounds after about 50 measurements.

Figure 10 shows the mean opportunity cost for nine sample paths using nine different randomly chosen data sets of 25000 compounds. Although not as impressive as the mean plot for 10000 compounds, this plot also attests that there is value in using the KGCB policy as opposed to a pure exploration policy.

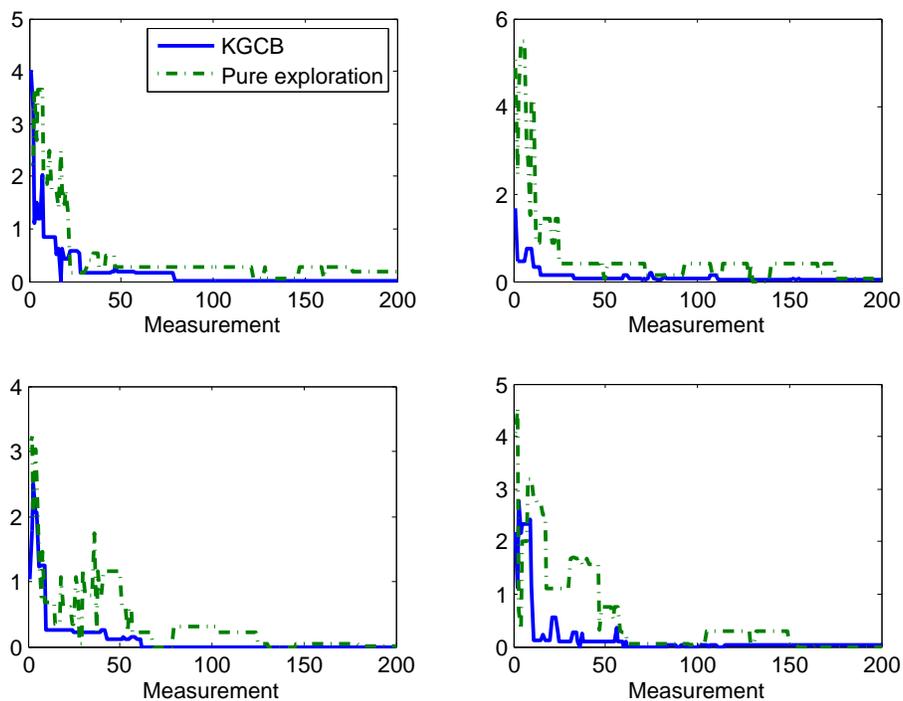


Figure 7: Four sample paths using data sets of 10000 compounds and a noise standard deviation of 0.38.

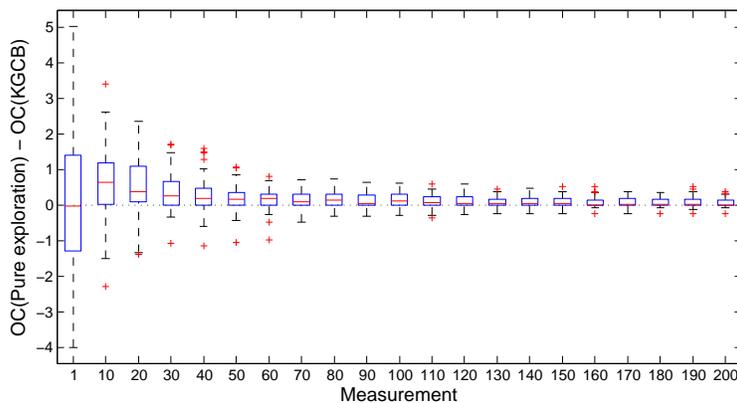


Figure 8: Distribution of difference between opportunity costs between pure exploration and KGCB using 75 sample paths of 10000 compounds each and a noise standard deviation of 0.38.

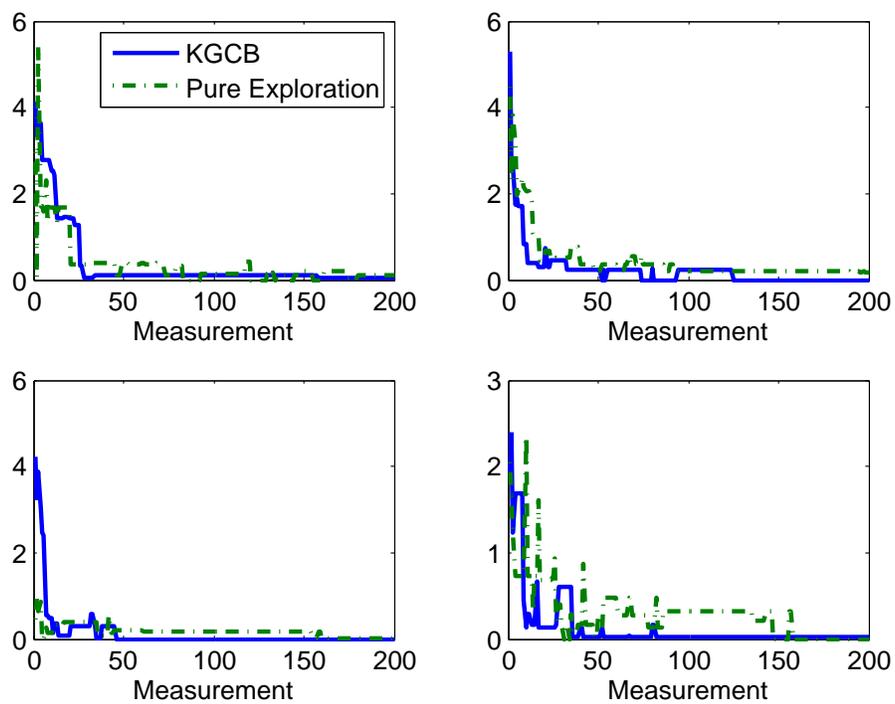


Figure 9: Four sample paths using data sets of 25000 compounds and a noise standard deviation of 0.38.

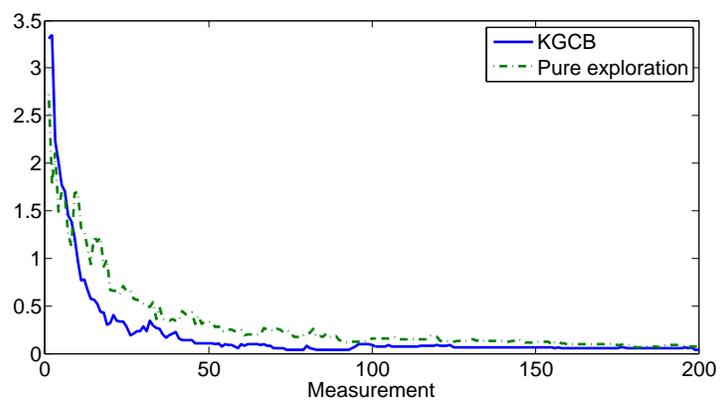


Figure 10: Average over nine runs of sample paths using data sets of 25000 compounds and a noise standard deviation of 0.38.