Proceedings of the 2024 Winter Simulation Conference
H. Lam, E. Azar, D. Batur, S. Gao, W. Xie, S. R. Hunter, and M. D. Rossetti, eds.

ESTIMATING VALUE OF INFORMATION ARM ALLOCATION INDICES IN CONTEXTUAL RANKING AND SELECTION PROBLEMS

Andres Alban[1], Stephen E. Chick[2], and Spyros I. Zoumpoulis[3]

[1]Management Department, Frankfurt School of Finance & Management, Frankfurt, GERMANY
[2]Technology and Operations Management Area, INSEAD, Fontainebleau, FRANCE
[3]Decision Sciences Area, INSEAD, Fontainebleau, FRANCE

ABSTRACT

Contextual ranking & selection is attracting increasing attention in simulation and other fields. A successful approach to addressing related challenges uses arm allocation indices that compute the Bayesian expected value of information of one-step look-ahead policies. We recall recent work on such indices for linear contextual bandits that take advantage of structural information about the nature of the covariates that describe contexts. Such indices can be computed exactly with a finite number of contexts and no delay in observing outcomes, but may require Monte Carlo simulation otherwise. Our contribution is to describe and quantify the benefits of two variance reduction techniques (conditional Monte Carlo and common random numbers) to estimate such allocation indices for contextual ranking & selection problems when some covariates are continuous or outcomes are observed with delay. We find that both techniques significantly improve estimates and the speed of inference, but conditioning is particularly useful.

1 INTRODUCTION

The contextual Ranking & Selection (R&S) problem attempts to identify the best alternative as a function of one or several context variables (Shen et al. 2021; Xiong 2020; Li et al. 2020). The expected value of information (EVI) / knowledge gradient (KG) is a well-known heuristic to developing sampling policies for R&S problems (Chen et al. 2015). EVI methods rely on the computation of so-called allocation indices (hereafter, indices) that represent the expected improvements in value from observing one or more samples for a given context and alternative. Numerically stable techniques have been been developed to compute such indices accurately for R&S problems (Frazier et al. 2008) and certain contextual R&S problems (Pearce and Branke 2018; Ding et al. 2021; Alban et al. 2021).

This paper studies how we can estimate such indices, when some assumptions of previous work on contextual R&S do not hold, through Monte Carlo simulation. In particular, we study how the estimation is affected when context variables may be either discrete or continuous and when there is a delay in observing the outcome. Here, we assume that subjects arrive, their context is observed, and allocation decisions are made on the basis of choosing the arm with the largest index. After all outcomes are observed for a fixed sample size, an implementation decision is made. Here, an implementation decision is the choice of a function from covariates to finite set of alternatives (arms).

Alban et al. (2021) discuss two of the differences between contextual R&S applied to simulation optimization and to clinical trials: the potential for delays in observing outcomes while new allocations are being made, and the process in observing the context variables. Simulation optimization chooses the context to simulate that maximizes learning, while clinical trials may have random draws of the context from the population of patients as they enroll. Both of those effects can slow the speed of inference. While Alban et al. (2021) focused on the random contexts, here we account for delayed observations, in addition to random contexts, and allow for continuous covariate values. These extensions imply that Monte Carlo methods are useful for indices, and we also explore variance reduction techniques for those estimates here.

The contextual R&S literature has mainly studied the problem with finitely many values of the covariates (Gao et al. 2019; Cakmak et al. 2021). Shen et al. (2021) allows for continuous covariates and proposes a two-stage indifference zone procedure. Ding et al. (2021) and Pearce and Branke (2018) present procedures to estimate EVI indices with Monte Carlo simulation that are equivalent to the algorithm we study in this paper for a special case: when outcomes are observed prior to the next allocation, i.e., no delay.

When there is a delay in observing the outcomes, the sampling policy must choose an alternative to allocate before observing the outcome of the latest past samples, but knowing the context and chosen alternatives of those samples. We refer to those samples that are already in progress, but whose outcomes have not yet been observed, as samples in the *pipeline*. As an example of the effect of delays on a sampling policy, consider the alternative with the most uncertainty. If there are many samples in the pipeline that are sampling this alternative, the procedure needs to take into account the uncertainty that will be left after the samples in the pipeline are observed, and likely assign a higher value of information to another alternative that would be left with more uncertainty after the pipeline clears. We assume that all pipeline samples are eventually observed.

Wu and Frazier (2016) and Wang et al. (2020) study EVI procedures for the R&S problem when the sampling policy makes observations in parallel or in batches, i.e., the sampling policy selects a batch of alternatives to sample in parallel. Batch sampling is similar to delays in that the sampling policy chooses alternatives predicting the uncertainty that will be left after the samples in the batch are observed. However, they study the problem without covariates. Chick et al. (2017) study delays with only two alternatives and Chick et al. (2019) considers delays with multiple arms, both without covariates.

Alban et al. (2024) studied the same multi-arm contextual R&S problem that we pose in this paper and proposed a Monte Carlo (MC) algorithm to estimate those indices to create an arm allocation index called fEVI-MC (f for function estimation because in contextual R&S we learn a function of covariates to the best alternative; EVI for allocation indices based on the expected value of information; MC for the estimation of those indices). That algorithm estimates allocation indices while using two variance reduction techniques (VRTs): common random numbers (CRN) to sharpen contrasts between indices for each arm and conditional Monte Carlo to enable a numerical integration through a novel use of the $h(\cdot)$-function proposed by Frazier et al. (2009) for the correlated knowledge gradient (cKG).

Our aim in this paper is to assess the importance of each of those two VRTs with respect to the expected opportunity cost as a function of the sample size of a sequential learning experiment for contextual bandits, such as with clinical trials for precision medicine.

In Section 2, we introduce the preliminary model of a sequential learning experiment with covariates using a Bayesian linear regression to estimate effect of covariates and alternatives on the outcomes. Section 3 presents the fEVI policy and the definition of the fEVI-indices. In Section 4, we propose a generic Monte Carlo algorithm to estimate the indices that can use the two VRTs (conditioning and CRN) in combination or alone. Section 5 presents a simulation experiment, motivated by a clinical trial design in personalized medicine, to quantify the benefits of the VRTs. We show that both VRTs considered can reduce the expected opportunity cost of the policy that uses the indices from the Monte Carlo algorithm, particularly the VRT based on conditioning. We can recommend the fEVI-MC algorithm with both VRTs for contextual R&S applications, and identify parameter values for the algorithm that worked well.

2 MODEL

We model the decision making process of an experimenter with a budget for T observations. The experimenter assigns sequentially subjects with covariates $\mathbf{X}_t \in \mathbb{R}^m$ for $t = 1, 2, \ldots, T$, where m is the dimension of the covariates vector. A subject may be a user of a website, a configuration of an engineering system, a patient in a clinical trial, or other contexts in a sequential learning problem. We assume that subjects arrive sequentially in equally spaced intervals. Each subject is assigned to an alternative $W_t \in \{1, 2, \ldots, n\}$, where the n is the number of available alternatives. After a fixed delay of $\Delta \geq 0$ subject arrivals, the experimenter observes the outcome Y_t. To make a choice of the alternative W_{t+1}, the experimenter uses

all information available to her, which includes prior information $\mathbf{K}_0$, the pairs of covariate values and alternatives for enlisted subjects $(\mathbf{X}_{t'}, W_{t'})$ for $t' = 1, 2, \ldots, t$, the outcomes that have already been observed $Y_{t'}$ for $t' = 1, 2, \ldots, t-\Delta$, and the covariates of the current subject $\mathbf{X}_{t+1}$.

After all subjects have been assigned and their outcomes have been observed, which happens at time $T+\Delta$, the experimenter selects a treatment strategy: a function that maps covariates to an alternative, to implement on a future population of subjects. The experimenter selects the treatment strategy that maximizes expected rewards for each covariate value. We now discuss how we learn from observations and summarize the information in a knowledge state that is a sufficient statistic.

2.1 Bayesian Linear Regression and Knowledge State

The outcomes are noisy observations that are normally distributed around the mean $r_{\boldsymbol{\mu}}(\mathbf{X}_t, W_t) = \mathbb{E}[Y_t \mid \boldsymbol{\mu}, \mathbf{X}_t, W_t]$, where $\boldsymbol{\mu}$ is the set of unknown parameters. We assume that $r_{\boldsymbol{\mu}}(\mathbf{x}, w)$ is a linear function of the interactions of the covariates and alternatives:

$$r_{\boldsymbol{\mu}}(\mathbf{x}, w) = \mu_{0,0}\xi_{0,0} + \sum_{i=1}^{n} \mathbb{1}_{w=i}\mu_{i,0}\xi_{i,0} + \sum_{l=1}^{m} x_l \mu_{0,l}\xi_{0,l} + \sum_{i=1}^{n}\sum_{l=1}^{m} \mathbb{1}_{w=i} x_l \mu_{i,l}\xi_{i,l}, \tag{1}$$

where $\mathbb{1}_a$ is the indicator function of event a and $\xi_{i,j} \in \{0,1\}$ induce the information about which terms are potentially active ($\xi_{i,j} = 1$) or are known to be inactive ($\xi_{i,j} = 0$). Inactive terms do not have any effect on the outcome of the subject, while potentially active terms are estimated from the trial data. We may choose some terms to be inactive so that (1) is not overparameterized or to include expert information about certain terms that are known not to affect the outcome.

To simplify notation and perform matrix and vector multiplication, the vector of unknown coefficients is assumed to have the following structure:

$$\boldsymbol{\mu} = \Big(\underbrace{\mu_{0,0}, \mu_{0,1}, \ldots, \mu_{0,m}}_{\text{associated with no alternative}} , \underbrace{\mu_{1,0}, \ldots, \mu_{1,m}}_{\text{associated with alternative 1}} , \ldots , \underbrace{\mu_{n,0}, \ldots, \mu_{n,m}}_{\text{associated with alternative } n} \Big)^\top.$$

Similarly, the vector of features that interacts alternative with covariates is given by the following use of the $\otimes$ operator:

$$w \otimes \mathbf{x} = \Big(\underbrace{\xi_{0,0}, \xi_{0,1}x_{0,1}, \ldots, \xi_{0,m}x_{0,m}}_{\text{associated with no alternative}}, \underbrace{\xi_{1,0}\mathbb{1}_{w=1}, \ldots, \xi_{1,m}\mathbb{1}_{w=1}x_m}_{\text{associated with alternative 1}}, \ldots, \underbrace{\xi_{n,0}\mathbb{1}_{w=n}, \ldots, \xi_{n,m}\mathbb{1}_{w=n}x_m}_{\text{associated with alternative } n} \Big).$$

With these notations, we can write

$$r_{\boldsymbol{\mu}}(\mathbf{x}, w) = (w \otimes \mathbf{x})\boldsymbol{\mu}.$$

We use Bayesian linear regression to learn the parameters $\boldsymbol{\mu}$. The prior distribution is given by

$$\boldsymbol{\mu} \sim \mathcal{N}(\boldsymbol{\theta}_0, \Sigma_0),$$

where $(\boldsymbol{\theta}_0, \Sigma_0)$ is the prior information. The noise is normally distributed with fixed sampling variance σ^2:

$$Y_t \mid \boldsymbol{\mu}, \mathbf{X}_t, W_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}((W_t \otimes \mathbf{X}_t)\boldsymbol{\mu}, \sigma^2).$$

The posterior parameters depend on data for subjects whose outcomes have already been observed. To keep track of covariates and assignments of alternatives for enlisted subjects whose outcomes have not been observed due to the delay, we define the *pipeline state* as follows:

$$\mathbf{J}_t := \begin{cases} \emptyset, & \text{for } \Delta = 0 \text{ or } t = 0 \text{ or } t = T+\Delta \\ (\mathbf{X}_1, W_1, \ldots, \mathbf{X}_t, W_t), & \text{for } 1 \le t \le \Delta \\ (\mathbf{X}_{t-\Delta+1}, W_{t-\Delta+1}, \ldots, \mathbf{X}_t, W_t), & \text{for } \Delta+1 \le t \le T \\ (\mathbf{X}_{t-\Delta+1}, W_{t-\Delta+1}, \ldots, \mathbf{X}_T, W_T), & \text{for } T+1 \le t \le T+\Delta-1. \end{cases}$$

The *knowledge state* $\mathbf{K}_t$ must also include the information about the pipeline: $\mathbf{K}_t := (\boldsymbol{\theta}_t, \Sigma_t, \mathbf{J}_t)$. The inference model has a conjugate prior distribution so that $\mu \mid \mathbf{K}_t \sim \mathcal{N}(\boldsymbol{\theta}_t, \Sigma_t)$ and we can update the parameters for $t > \Delta$ (when we observe outcomes) as follows (Powell and Ryzhov 2012, Sec. 8.2.2):

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \frac{Y_{t-\Delta} - (W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})\boldsymbol{\theta}_{t-1}}{\sigma^2 + (W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})\Sigma_{t-1}(W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})^\top} \Sigma_{t-1}(W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})^\top \tag{2a}$$

$$\Sigma_t = \Sigma_{t-1} - \frac{\Sigma_{t-1}(W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})^\top (W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})\Sigma_{t-1}}{\sigma^2 + (W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})\Sigma_{t-1}(W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})^\top}. \tag{2b}$$

To enable the use of CRN to simulate allocation indices across arms, it will be useful to describe the distribution of the posterior means to be realized after the pipeline outcomes are observed relative to a standard normal random variable. It can be shown that (Alban et al. 2024)

$$\boldsymbol{\theta}_t \mid \mathbf{K}_{t-1}, \mathbf{X}_t, W_t \sim \mathcal{N}(\boldsymbol{\theta}_{t-1}, \tilde{\boldsymbol{\sigma}}_t \tilde{\boldsymbol{\sigma}}_t^\top),$$

where the *preposterior standard deviation* is given by

$$\tilde{\boldsymbol{\sigma}}_t = \frac{\Sigma_{t-1}(W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})^\top}{\sqrt{(\sigma^2 + (W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})\Sigma_{t-1}(W_{t-\Delta} \otimes \mathbf{X}_{t-\Delta})^\top)}}. \tag{3}$$

Therefore, by letting $Z_t \sim \mathcal{N}(0,1)$ represent the noise of the outcome of subject $t-\Delta$, we can show that the posterior mean is conditionally distributed as follows:

$$\boldsymbol{\theta}_t \mid \mathbf{K}_{t-1}, \mathbf{X}_t, W_t \sim \boldsymbol{\theta}_{t-1} + \tilde{\boldsymbol{\sigma}}_t Z_t. \tag{4}$$

Moreover, we can update the posterior variance matrix using the preposterior standard deviation:

$$\Sigma_t = \Sigma_{t-1} - \tilde{\boldsymbol{\sigma}}_t \tilde{\boldsymbol{\sigma}}_t^\top. \tag{5}$$

3 VALUE, POLICY, AND INDICES

A policy π is a mapping from the knowledge state and the covariate values of the current subject to a probability distribution over the available alternatives, i.e., $\pi(w \mid \mathbf{k}, \mathbf{x}) = \mathbb{P}(W_t = w \mid \mathbf{K}_{t-1} = \mathbf{k}, \mathbf{X}_t = \mathbf{x})$. The *implemented strategy* $\tilde{f}_{\boldsymbol{\theta}}(\mathbf{x})$ for a given posterior mean $\boldsymbol{\theta}$ maps the covariate values to the alternative that maximizes the expected outcome: $\tilde{f}_{\boldsymbol{\theta}}(\mathbf{x}) = \arg\max_w (w \otimes \mathbf{x})\boldsymbol{\theta}$.

We use expected value of information methods to design heuristic policies to maximize *value*:

$$V^\pi = \mathbb{E}^\pi \left[r_{\boldsymbol{\mu}}(\tilde{\mathbf{X}}, \tilde{f}_{\boldsymbol{\theta}_{T+\Delta}}(\tilde{\mathbf{X}})) \right], \tag{6}$$

where $\tilde{\mathbf{X}}$ is a random variable with the distribution of the covariates in the population, $F_{\tilde{x}}$.

The functional Expected Value of Information (fEVI) policy is a heuristic policy that samples the alternative with the largest one-step look-ahead value, which we will refer to as the fEVI-index.

$$\begin{aligned} \nu_t(\mathbf{x}, w) &= \mathbb{E}[(\tilde{f}_{\boldsymbol{\theta}_{t+\Delta+1}}(\tilde{\mathbf{X}}) \otimes \tilde{\mathbf{X}})\boldsymbol{\mu} \mid \mathbf{K}_t, \mathbf{X}_{t+1} = \mathbf{x}, W_{t+1} = w] - \mathbb{E}[(\tilde{f}_{\boldsymbol{\theta}_{t+\Delta}}(\tilde{\mathbf{X}}) \otimes \tilde{\mathbf{X}})\boldsymbol{\mu} \mid \mathbf{K}_t] \\ &= \underbrace{\mathbb{E}\left[\max_{\tilde{w}} (\tilde{w} \otimes \tilde{\mathbf{X}})\boldsymbol{\theta}_{t+\Delta+1} \mid \mathbf{K}_t, \mathbf{X}_{t+1} = \mathbf{x}, W_{t+1} = w \right]}_{\text{implement after one more subject and pipeline clears}} - \underbrace{\mathbb{E}\left[\max_{\tilde{w}} (\tilde{w} \otimes \tilde{\mathbf{X}})\boldsymbol{\theta}_{t+\Delta} \mid \mathbf{K}_t \right]}_{\text{implement after pipeline clears}}. \end{aligned} \tag{7}$$

This fEVI-index assesses the increment in the value of the optimal strategy selected after the data of the pipeline plus one subject is observed, relative to that if only pipeline data are observed.

The fEVI policy samples $W_{t+1} = \arg\max_{w \in \mathscr{W}} \nu_t(\mathbf{X}_{t+1}, w)$, $0 \le t \le T-1$. When covariates are discrete and $\Delta = 0$, Alban et al. (2021) show how to compute the indices using ideas from Frazier et al. (2009).

4 MONTE CARLO ALGORITHMS TO COMPUTE FEVI INDICES

The indices in (7) involve expectations that may be hard to numerically integrate with quadrature when some covariates are continuous or when the observation delay is positive. We now give unbiased estimators of those indices using Monte Carlo simulation in order to allow for continuous covariates and $\Delta > 0$.

Algorithm 1 (fEVI-MC) presents a procedure that determines four different Monte Carlo estimation techniques for the fEVI-indices. Each technique produces indices, and hence each represents a different selection procedure for the contextual R&S problem. The four different allocation indices are defined by two parameters (*hflag* and *CRNflag*, each valued true or false) that determine whether the algorithm uses each of two VRTs (conditioning using the $h(\cdot)$-function from Frazier et al. (2009) and CRN, respectively), which we describe in more detail below. The algorithm further depends on two parameters that determine the number of replications: η^{on} represents the number of sampled outcomes of experimental subjects and η^{off} represents the number of sampled subjects from the post-experiment population for each of the sampled experimental subjects. Overall, $\eta^{\text{on}} \times \eta^{\text{off}}$ samples of the indices are averaged to obtain the final estimate.

The first VRT is conditional Monte Carlo (or conditioning) using the $h(\cdot)$-function from Frazier et al. (2009). It conditions on the posterior parameters after clearing the pipeline and on the covariates in the population, followed by integrating over the outcome of the current subject using the $h(\cdot)$-function: $h(\mathbf{a},\mathbf{b}) = \mathbb{E}\left[\max_i a_i + b_i Z\right] - \max_i a_i$, where $Z \sim \mathcal{N}(0,1)$, and a_i and b_i are the ith entries of vectors $\mathbf{a}$ and $\mathbf{b}$, respectively. In addition to variance reduction, the conditional Monte Carlo of fEVI-MC has two advantages: 1) the estimates of the indices are guaranteed to be positive, and 2) it estimates the logarithm of the indices, which is numerically more stable. For notational convenience, Algorithm 1 is presented to compute the logarithm even when it does not use conditioning, but in that case it does not benefit from numerically stable methods such as those used in computing the logarithm of the $h(\cdot)$-function.

The second VRT is to use CRN for common draws of outcomes from the pipeline subjects and of covariates of the future population of subjects to be treated, in computing the indices for each arm. Although we use CRN across arms for a given t, independent draws are made for each time t.

Algorithm 1 can use both, either, or neither of the two VRTs, for a total of four possible combinations of the VRTs, by activating the VRTs using the parameters *hflag* for conditioning and *CRNflag* for CRN.

The parameters η^{on} and η^{off} represent the number of samples from random variables for the Monte Carlo estimation. When conditioning, η^{on} represents the number of samples of the outcomes from the pipeline, i.e, the samples $Y_{t'}$ for $t' = t-\Delta+1,\ldots,t$ (the algorithm samples the noise $\hat{\mathbf{Z}}_i$ instead of $Y_{t'}$). These are from the subjects that have been assigned to an alternative but whose outcome has not been observed yet due to the delay. Otherwise, η^{on} represents the number of samples of the outcomes from the pipeline, but also includes the outcome of the subject whose allocation decision is being made, Y_{t+1}. Conditioning avoids the sampling of Y_{t+1} and numerically integrates conditional on the other sampled variables using the $h(\cdot)$-function from the cKG. Regardless of the VRTs, η^{off} represents the number of samples of the covariates from the population, $\tilde{\mathbf{X}}$, whose samples are denoted in the algorithm by $\hat{\mathbf{X}}_i$. Finally, the algorithm averages $\eta^{\text{on}} \times \eta^{\text{off}}$ unbiased samples of the fEVI-indices.

When conditioning and when there is no delay, the allocation of the replications between η^{on} and η^{off} is inconsequential, as long as the product of the two is the same. However, without conditioning, fEVI-MC benefits from η^{on}, rather than η^{off}, being larger, because the sample space of the outcome Y_{t+1} is more widely explored. When there is a delay, η^{on} and η^{off} balance more (random) exploration of the outcomes of experimental subjects (larger η^{on}) and more exploration of the covariates in the population (larger η^{off}).

Algorithm 1 here is equivalent to Algorithm 1 of Alban et al. (2024) when it uses both conditioning and CRN. The algorithms of Pearce and Branke (2018) and Ding et al. (2021) do not account for delayed observations, but are equivalent to Algorithm 1 with both VRTs for the special case of no delay, i.e., $\Delta = 0$. When there is no delay and there are a finite number of covariate values, Alban et al. (2021) shows how to compute the fEVI-indices exactly (up to numerical error), and we refer to this procedure as the exact computation of the fEVI-indices (or the fEVI policy) when applicable.

Algorithm 1 *f*EVI-MC: Monte Carlo estimates of *f*EVI-indices. The *hflag* parameter controls whether we use conditioning with the $h(\cdot)$-function. The *CRNflag* parameter controls whether we use CRN.

function *f*EVI-MC($t, \mathbf{K}_t, \mathbf{X}_{t+1}; \eta^{\text{on}}, \eta^{\text{off}}, hflag, CRNflag$)
  $\Delta'_t \leftarrow \min\{t, \Delta\}$ ▷ Compute length of pipeline.
  $\tau \leftarrow \max\{t, \Delta\}$ ▷ Compute the time step when we will next observe an outcome
  Compute $\tilde{\boldsymbol{\sigma}}_{\tau+i}$ for $i = 1, 2, \ldots, \Delta'_t$ ▷ Prepost std dev in (3) for pipeline
  for all $j = 1, 2, \ldots, \eta^{\text{on}}$ **do** ▷ Compute offline rewards for η^{on} replications
    $\hat{\mathbf{Z}}_i \sim \mathcal{N}(0,1)$ for $i = 1, 2, \ldots, \Delta'_t + 1$ ▷ Noise vector of the outcomes of length $\Delta'_t + 1$
    $\hat{\boldsymbol{\theta}} \leftarrow \boldsymbol{\theta}_t + \sum_{i=1}^{\Delta'_t} \tilde{\boldsymbol{\sigma}}_{\tau+i} \hat{\mathbf{Z}}_i$ ▷ Repeated application of (4) to simulate a posterior mean $\boldsymbol{\theta}_{t+\Delta}$
    $\hat{\mathbf{X}}_i \sim F_{\tilde{x}}$ for $i = 1, 2, \ldots, \eta^{\text{off}}$ ▷ Sample η^{off} post-trial covariates
    for all $w \in \mathscr{W}$ **do** ▷ For each alternative that could be assigned
      if not *CRNflag* **then** ▷ Without CRN, resample random variables for each alternative
        $\hat{\mathbf{Z}}_i \sim \mathcal{N}(0,1)$ for $i = 1, 2, \ldots, \Delta'_t + 1$
        $\hat{\boldsymbol{\theta}} \leftarrow \boldsymbol{\theta}_t + \sum_{i=1}^{\Delta'_t} \tilde{\boldsymbol{\sigma}}_{\tau+i} \hat{\mathbf{Z}}_i$
        $\hat{\mathbf{X}}_i \sim F_{\tilde{x}}$ for $i = 1, 2, \ldots, \eta^{\text{off}}$
      end if
      $W_{t+1} \leftarrow w$ ▷ Compute index assuming that the subject is assigned to alternative w…
      Compute $\tilde{\boldsymbol{\sigma}}_{\tau+\Delta'_t+1}$
      for all $i = 1, 2, \ldots, \eta^{\text{off}}$ **do** ▷ For each subject sampled from the population…
        if *hflag* **then**
          $\mathbf{a} \leftarrow ((1, 2, \ldots, n) \otimes \hat{\mathbf{X}}_i)\hat{\boldsymbol{\theta}}$ ▷ Vector of means for n alternatives with covariates $\hat{\mathbf{X}}_i$
          $\mathbf{b} \leftarrow ((1, 2, \ldots, n) \otimes \hat{\mathbf{X}}_i)\tilde{\boldsymbol{\sigma}}_{\tau+\Delta'_t+1}$ ▷ Prepost std dev for n alternatives w/ covariates $\hat{\mathbf{X}}_i$
          $\log(\hat{v}_{w,j,i}) \leftarrow \log(h(\mathbf{a}, \mathbf{b}))$ ▷ Conditional EVI, where $\log(h(\mathbf{a}, \mathbf{b}))$ is the v …
        else ▷ … for cKG computed in Algorithm 2 in Frazier et al. (2009).
          $\hat{\boldsymbol{\theta}} \leftarrow \hat{\boldsymbol{\theta}} + \tilde{\boldsymbol{\sigma}}_{\tau+\Delta'_t+1} \hat{\mathbf{Z}}_{\Delta'_t+1}$ ▷ Posterior mean
          $\log(\hat{v}_{w,j,i}) \leftarrow \log\left((\tilde{f}_{\hat{\boldsymbol{\theta}}}(\hat{\mathbf{X}}_i) \otimes \hat{\mathbf{X}}_i)\hat{\boldsymbol{\theta}}\right)$ ▷ Realized reward for a given posterior
        end if
      end for i
    end for w
  end for j
  for all $w \in \mathscr{W}$ **do**
    $\log(\hat{v}_w) \leftarrow \log\left((1/(\eta^{\text{on}}\eta^{\text{off}}))\sum_{j=1}^{\eta^{\text{on}}}\sum_{i=1}^{\eta^{\text{off}}} v_{w,j,i}\right)$ ▷ Index for w, ave. over param uncertainty
  end for
  return $W_{t+1} = \arg\max_{w \in \mathscr{W}} \log(\hat{v}_w)$ ▷ Pick alternative with largest *f*EVI-MC-index
end function ▷ (break ties uniformly at random)

5 SIMULATION RESULTS

We consider eight alternatives ($n = 8$), one categorical covariate with four categories, and one real-valued covariate. The labeling $\boldsymbol{\xi}$ is such that 10 coefficients of the linear regression are active and need to be estimated. The experiments were performed using the Julia language (Bezanson et al. 2017) on a 12th Gen Intel Core i7 processor. This setup was also used for numerical experiments of Alban et al. (2024).

We aim to compare Algorithm 1 under the four combinations of the VRTs (both conditioning and CRN, only conditioning, only CRN, and none) in terms of their computation time, accuracy, and expected opportunity cost (EOC). The EOC is sometimes called the expected regret. Moreover, we provide some guidance about the number of replications required.

5.1 Scenario Where Exact Computation is Feasible

We first explore a scenario where we can compute the indices exactly (up to numerical error) using the algorithm of Alban et al. (2021) by assuming that the real-valued covariate is discrete with three possible values ($0, 1, 2$ with probabilities $1/4, 1/2, 1/4$) and no delay ($\Delta = 0$). This numerical experiment allows us to compare the fEVI-MC algorithm to the exact computation, which will provide evidence of the usefulness of the Monte Carlo method and the benefits of the VRTs.

We run 5,000 replications of a trial with a horizon $T = 600$ where we use the fEVI policy to make allocation decisions. At each time step, we compute the indices with the four versions of the fEVI-MC algorithm, each with $\eta^{\text{on}} = 5, 10, 20, 50$, while keeping $\eta^{\text{off}} = 1$.

Table 1 shows the computation time, fraction of replications in which the fEVI-MC would have selected the same arm as the fEVI policy, and EOC. We observe that the computation time increases close to linearly with η^{on}. CRN saves between 6-7% of computation time when also conditioning and 13-14% otherwise. Conditioning requires repeated computation of the $h(\cdot)$-function, which increases the computation time compared to the versions without conditioning. This increase does not exceed a factor of two in this experiment. The exact computation requires 12 evaluations of the $h(\cdot)$-function, one for each context (a context here is a combination of the two covariates, which have three and four possible values). For $\eta^{\text{on}} = 5$, the fEVI-MC algorithm with conditioning only performs five evaluations of the $h(\cdot)$-function and obtains a lower computation time. Due to other operations of the algorithm, mainly random number generation, the algorithm requires more computation time when $\eta^{\text{on}} = 10$ despite using fewer evaluations of the $h(\cdot)$-function.

The fraction of arm choices that are the same as the arm choices of fEVI is substantially higher when conditioning. CRN provides an increase in addition to conditioning. Without conditioning, CRN also provides a small increase in terms of selecting the arm with the highest fEVI-index. For the range of η^{on} considered here, the versions with conditioning obtain an increase of 5-6% in the fraction of decision that select the arm with the highest fEVI-index as η^{on} increases, while the versions without conditioning only obtain an increase of 2%. To raise the performance in this metric of the version without conditioning to match the versions with conditioning, an increase of orders of magnitude in η^{on} would be required.

A more direct measure of performance of the algorithms is their ability to obtain larger value, or equivalently, a lower EOC, defined as the difference between the value obtained by the policy and an oracle that has perfect information about the coefficients:

$$\text{EOC}^{\pi}(T) := \mathbb{E}^{\pi}\left[\max_{w \in \mathscr{W}}(w \otimes \tilde{\mathbf{X}}_1)\boldsymbol{\mu} - (\tilde{f}_{\boldsymbol{\theta}_{T+\Delta}}(\tilde{\mathbf{X}}_1) \otimes \tilde{\mathbf{X}}_1)\boldsymbol{\mu} \mid \boldsymbol{\theta}, \Sigma\right]. \tag{8}$$

While the algorithms may not always select an arm with the largest fEVI-index, they often select an arm with a "good" EVI, ultimately leading to high value (low EOC) of the policy. We again observe that both versions with conditioning obtain a lower (better) EOC than the versions without. However, we do not observe a significant improvement due to CRN in addition to conditioning. Both versions with conditioning obtain an EOC that is statistically equivalent to that of the exact computation.

Figure 1 (left panel) shows the fraction of arm choices equal those of fEVI for the four versions of the fEVI-MC algorithm ($\eta^{\text{on}} = 50$). The versions that do not use CRN have a low fraction for $T = 0$ but for $T = 1$ have a jump to higher fractions. Both versions with conditioning have a relatively constant (with some noise) fraction for the whole horizon, while the versions without conditioning have a decreasing fraction until around $T = 50$ – the indices decrease exponentially, and so does the difference between indices, making accurate identification of the highest index more difficult as the sample size increases. The conditioning VRT is able to handle small values by estimating the logarithm of the indices as discussed in Section 4. Without conditioning, estimating exponentially small indices becomes impractical as the sample size increases, making arm choice effectively at random (Branke et al. 2007; Frazier et al. 2009).

Figure 1 (right panel) shows the EOC for the four versions of the fEVI-MC algorithm ($\eta^{\text{on}} = 50$ for all), in addition to the fEVI policy, as a function of the sample size T. For all sample sizes, both versions with

Table 1: Computation time (as a multiple of CPU time required for the exact fEVI computation, 0.102ms), fraction of estimates that lead to the same decisions as the exact fEVI, and EOC at sample size $T = 600$, for different versions of Monte Carlo algorithms for estimating fEVI indices. The CRN VRT denotes use of the same pipeline outcomes and covariates from the population across each alternative index. The $h(\cdot)$ VRT denotes use of the $h(\cdot)$-function of cKG (Frazier et al. 2009) to integrate over outcomes of subject $t+1$, conditional on sampled pipeline outcomes and population covariates.

VRTs	η^{on}	η^{off}	Computation time	Fraction of arm choices equal to those of fEVI	EOC at $T = 600$
Exact fEVI	-	-	1	1	1.86e-03 ± 1.45e-04
$h(\cdot)$ and CRN	5	1	0.9156	0.9417 ± 0.0001	2.12e-03 ± 1.57e-04
	10	1	1.7713	0.9822 ± 0.0001	1.76e-03 ± 1.33e-04
	20	1	3.5182	0.9950 ± 0.0000	2.03e-03 ± 1.55e-04
	50	1	8.7615	0.9966 ± 0.0000	1.97e-03 ± 1.49e-04
$h(\cdot)$	5	1	0.9735	0.8215 ± 0.0002	1.83e-03 ± 1.37e-04
	10	1	1.8940	0.8417 ± 0.0002	1.80e-03 ± 1.39e-04
	20	1	3.7439	0.8587 ± 0.0002	1.63e-03 ± 1.32e-04
	50	1	9.3170	0.8807 ± 0.0002	1.87e-03 ± 1.43e-04
CRN	5	1	0.4642	0.4323 ± 0.0003	3.63e-03 ± 2.36e-04
	10	1	0.8999	0.4397 ± 0.0003	2.93e-03 ± 1.93e-04
	20	1	1.7743	0.4479 ± 0.0003	3.15e-03 ± 2.11e-04
	50	1	4.3925	0.4596 ± 0.0003	2.86e-03 ± 1.91e-04
None	5	1	0.5309	0.4297 ± 0.0003	3.13e-03 ± 2.14e-04
	10	1	1.0353	0.4345 ± 0.0003	2.83e-03 ± 1.93e-04
	20	1	2.0442	0.4390 ± 0.0003	3.12e-03 ± 2.01e-04
	50	1	5.0785	0.4454 ± 0.0003	3.02e-03 ± 2.15e-04

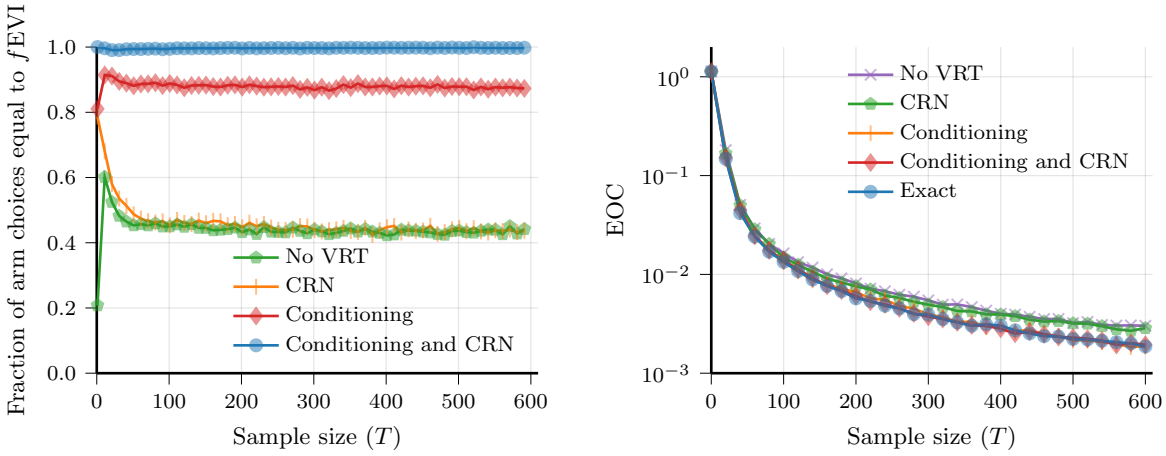

Figure 1: Fraction of arm choices equal to those of fEVI (left panel) and EOC (right panel) in a scenario where indices can be computed exactly. Only results with $\eta^{\text{on}} = 50$ (used in the four versions of fEVI-MC) are shown, because no substantial differences are observed for other values of η^{on}.

conditioning have approximately equal performance as the fEVI policy. The versions without conditioning move farther apart from the fEVI policy as T increases. Both versions with conditioning require ≈400 observations to obtain the same EOC that the other two versions obtain with 600 observations. We do not show the results for other values of η^{on}, as we do not observe significant differences.

We highlight the following *takeaways* from this example. The conditional Monte Carlo VRT substantially increases the accuracy of the Monte Carlo estimates with a modest increase in computation time. The arm with the highest fEVI-index is selected with probability over 80% when conditioning but less than 50% otherwise. CRN provides an additional increase in the fraction of choices equal to those of fEVI (10-12%) but is small compared to the increase provided by conditional Monte Carlo. The increase in fraction of arm choices with the highest fEVI-index due to CRN does not translate into a lower EOC. While the accuracy of the estimates may increase with VRTs, those improvements do not necessarily translate into obtaining lower EOC. The fEVI-MC algorithm with VRTs may obtain statistically the same EOC as fEVI, with lower computation time (in our experiment with $\eta^{\text{on}} = 5$). Finally, the additional computation time for conditioning, due to the $h(\cdot)$-function computations, does not exceed a factor of two, which can be justified for (monetary or computationally) expensive experiments by the decrease in the EOC. (This example suggests that two thirds of the samples are necessary to obtain the same EOC.)

5.2 Scenario with a Continuous Covariate and a Delay

We now study a scenario where the real-valued covariate is continuous and where a positive delay ($\Delta > 0$) is present. A positive delay decreases the value because it forces the policy to make allocation decisions with less information at hand.

Alban et al. (2024) show that the fEVI-MC algorithm with conditioning and CRN is able to obtain the statistically equivalent EOC with a delay ($\Delta = 50$) and with no delay ($\Delta = 0$) for large sample sizes (above 150 in our experiments). Similarly, fEVI-MC without VRTs can obtain a statistically equivalent EOC with a delay and with no delay for large sample sizes, but the EOC remains significantly higher than that with VRTs (data not shown).

Table 2 shows the computation time and EOC at $T = 600$ with a delay $\Delta = 50$, which is a moderate delay (8.3% of the horizon considered here). The table shows results for the different VRTs, $\eta^{\text{on}} = 5, 10$ and $\eta^{\text{off}} = 5, 10$. Due to the delay and continuous covariates, obtaining the exact indices by quadrature would be computationally challenging.

Both versions of the fEVI-MC algorithm that use conditioning obtain a statistically lower EOC than the two versions that do not use conditioning. We observe a statistically significant difference in EOC between the version with both VRTs and the version with only conditioning for $\eta^{\text{on}} = \eta^{\text{off}} = 10$. However, as we discuss below, this observation may be due to noise. Unlike the results in Table 1, we observe a significant and large decrease in EOC when using CRN compared to not using any VRTs, suggesting that CRN is particularly useful in scenarios with delay.

Figure 2 shows the EOC as a function of the sample size T for $\Delta = 20$ (left panel) and $\Delta = 50$ (right panel) for $\eta^{\text{on}} = \eta^{\text{off}} = 10$. For $\Delta = 50$, we observe that the version with both VRTs has a down-tick at the end of the horizon, while the version with only conditioning has an up-tick, which can explain the statistical difference observed in Table 2. Both figures illustrate that, in a scenario with moderate delays, both versions with conditioning are superior and statistically equivalent. The version with only CRN is significantly better than the version with no VRTs.

We highlight the following main *takeaways*. The fEVI-MC algorithm with conditioning is able to effectively deal with delays, achieving approximately the same EOC for large sample sizes as when there is no delay and outperforming the versions of fEVI-MC that do not use conditioning. CRN plays a more important role in reducing EOC when the delay is positive than when the delay is zero.

5.3 Binning a Continuous Covariate in a Setting without Delay

We now consider a scenario with a continuous covariate but no delay. Here, we can obtain an estimate of the fEVI-indices by first binning the continuous covariate into finitely many bins and then using the exact algorithm. We aim to compare the EOC of the fEVI-MC algorithm to the exact algorithm with binning and provide insights into the benefits when Monte Carlo estimates are most beneficial.

Table 2: Computation time (relative to the computation time corresponding to the first row of the table, which equals 1.075ms) and EOC for the four versions of the fEVI-MC algorithm with $\Delta = 50$ and different values of η^{on}, η^{off}.

VRTs	η^{on}	η^{off}	Computation time	EOC at $T = 600$
$h(\cdot)$ and CRN	5	5	1.0000	2.19e-03 ± 1.61e-04
	5	10	1.4541	2.09e-03 ± 1.54e-04
	10	5	1.4578	1.98e-03 ± 1.56e-04
	10	10	2.6669	1.76e-03 ± 1.36e-04
$h(\cdot)$	5	5	1.0437	2.06e-03 ± 1.53e-04
	5	10	1.5399	2.27e-03 ± 1.66e-04
	10	5	1.5658	2.09e-03 ± 1.54e-04
	10	10	2.8562	2.29e-03 ± 1.69e-04
CRN	5	5	0.7748	2.97e-03 ± 1.97e-04
	5	10	0.9892	2.90e-03 ± 2.01e-04
	10	5	0.9954	2.47e-03 ± 1.83e-04
	10	10	1.5544	2.69e-03 ± 1.99e-04
None	5	5	0.8136	4.74e-03 ± 2.85e-04
	5	10	1.0526	5.03e-03 ± 3.05e-04
	10	5	1.0731	4.22e-03 ± 2.70e-04
	10	10	1.7086	4.28e-03 ± 2.77e-04

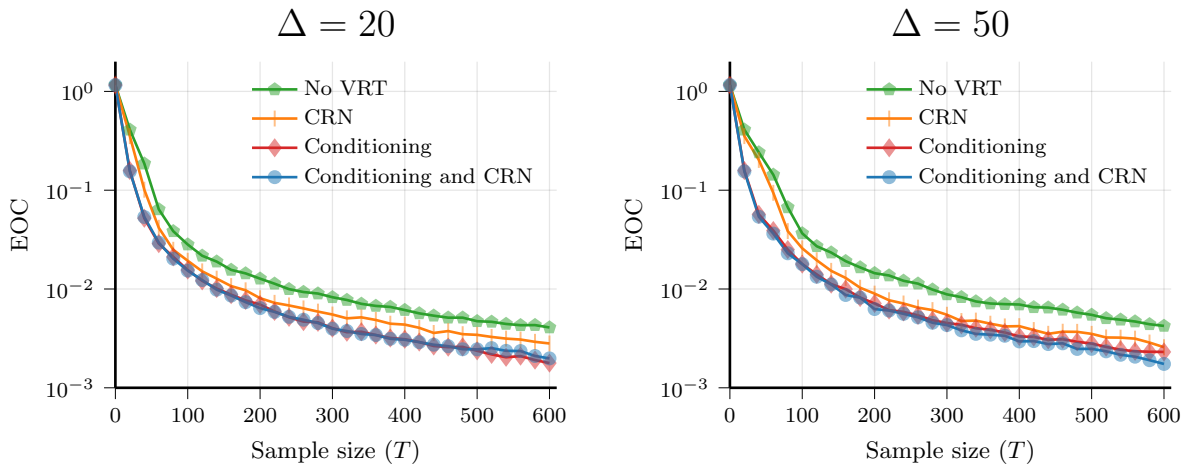

Figure 2: EOC with a continuous covariate and a delay (left $\Delta = 20$, right $\Delta = 50$). Here, $\eta^{\text{on}} = \eta^{\text{off}} = 10$.

Figure 3 shows the EOC for the four versions of the fEVI-MC algorithm and the exact fEVI algorithm that bins the covariates into three and five bins. The algorithm with three bins obtains a higher EOC than the fEVI-MC algorithm with no VRTs, and with five bins statistically equivalent to the fEVI-MC with CRN and no VRTs. The versions of fEVI-MC with conditioning obtain a lower EOC. The computation time of the exact algorithm increases linearly with the number of bins because each bin requires the computation of the $h(\cdot)$-function. A linear increase is not a major concern when only one covariate is binned, as we do here. However, binning several covariates can cause large increases in computation time because the number of bins will increase exponentially with the number of covariates. The computation time for the Monte Carlo algorithms is similar to those reported in Table 1, and the computation time of binning into three bins is similar to the exact computation in Table 1. Thus, fEVI-MC may obtain a lower EOC (compared to binning covariates and then using the exact fEVI) with shorter computation time in this experiment with $\eta^{\text{on}} = 5$.

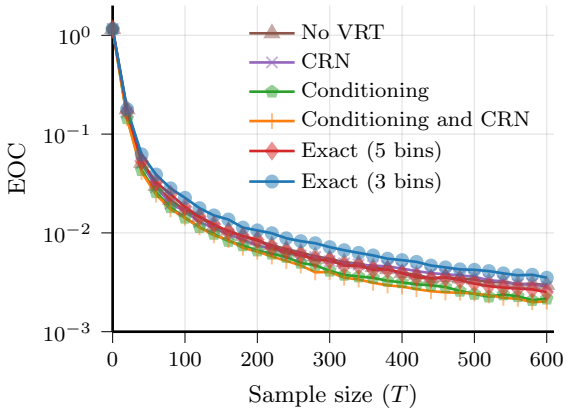

Figure 3: EOC with a continuous covariate.

6 DISCUSSION

Computing the fEVI-indices of Bayesian selection procedures for contextual R&S problems with continuous covariates or a delay in observing outcomes requires a challenging numerical integration over a high-dimensional space. Monte Carlo simulation is a very effective method to estimate such multi-dimensional integrals. We find that Monte Carlo methods with appropriate VRTs can obtain good estimates of the indices, and we provide a policy that can obtain an EOC at least comparable to an exact computation of the indices. Moreover, we find that it can properly handle scenarios with a moderate delay, which are the scenarios where sequential sampling policies are most useful. In particular, the VRT that conditions on the pipeline outcomes and the population covariates and uses the $h(\cdot)$-function from the cKG algorithm provides a significant improvement in the estimates of the fEVI-indices and improves the EOC.

In our experiments, we find that a small number of replications $(\eta^{\text{on}}, \eta^{\text{off}})$ for the fEVI-MC algorithm with VRTs are sufficient to obtain a statistically equivalent EOC. In particular, we find that the required computation time of the fEVI-MC algorithm is comparable, and potentially even shorter, than an exact computation of the fEVI policy, while maintaining a statistically equivalent EOC. Conditioning using the $h(\cdot)$-function increases computation time but reduces the EOC, justifying the use of conditioning for expensive experiments. CRN does not necessarily improve the EOC of the fEVI-MC policy, but did not deteriorate it, and it reduces the computation time of the indices. Thus, we would recommend using both VRTs (conditioning and CRN) for use in practice. A further exploration of how to pick $\eta^{\text{on}}, \eta^{\text{off}}$ in general is an area of future work.

Some modified versions of our model, such as when the outcomes are not normally distributed, may not be amenable to the conditioning VRT as we present it here. In such scenarios, the fEVI-MC algorithm without conditioning may still be appropriate. Without conditioning, we find that CRN can provide a substantial improvement in EOC for the fEVI-MC policy, particularly when the delay is positive. Future work can pursue how CRN, and potentially other VRTs, can improve the estimation of fEVI-indices under non-Gaussian models. Related work discusses techniques to handle heteroscedastic or unknown variances (Alban et al. 2024).

ACKNOWLEDGMENTS

We thank the European Union's support through the MSCA-ESA-ITN project (676129), and discussions with Drs. A.P.J. Vlaar, W.J. Wiersinga, F. Uhel, B. Scicluna, and N. van Mourik (Amsterdam University Medical Center). Chick acknowledges the support of Dr. Simba Gill and Sabi Dau to the INSEAD Healthcare Management Initiative and of the Novartis Chair of Healthcare Management at INSEAD.

REFERENCES

Alban, A., S. Chick, and S. Zoumpoulis. 2024. "Learning Personalized Treatment Strategies with Predictive and Prognostic Covariates in Adaptive Clinical Trials". SSRN, https://ssrn.com/abstract=4160045.

Alban, A., S. E. Chick, and S. I. Zoumpoulis. 2021. "Expected value of information methods for contextual ranking and selection: clinical trials and simulation optimization". In *Proc. 2021 Winter Simulation Conference*, edited by S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, 1–12. IEEE, Inc.

Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah. 2017. "Julia: A fresh approach to numerical computing". *SIAM review* 59(1):65–98.

Branke, J., S. Chick, and C. Schmidt. 2007. "Selecting a Selection Procedure". *Management Science* 53(12):1916–1932.

Cakmak, S., E. Zhou, and S. Gao. 2021. "Contextual ranking and selection with Gaussian processes". In *Proc. 2021 Winter Simulation Conference*, edited by S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, 1–12. IEEE, Inc.

Chen, C.-H., S. E. Chick, L. H. Lee, and N. A. Pujowidianto. 2015. "Ranking and selection: Efficient simulation budget allocation". In *Handbook of Simulation Optimization*, 45–80. Springer.

Chick, S. E., M. Forster, and P. Pertile. 2017. "A Bayesian decision theoretic model of sequential experimentation with delayed response". *Journal of the Royal Statistical Society. Series B* 79(5):1439–1462.

Chick, S. E., N. Gans, and O. Yapar. 2019. "Sequential, Value-Based Designs for Certain Clinical Trials with Multiple Arms Having Correlated Rewards". In *Proc. 2019 Winter Simulation Conference*, 1032–1043. IEEE.

Ding, L., L. J. Hong, H. Shen, and X. Zhang. 2021. "Technical note – Knowledge Gradient for Selection with Covariates: Consistency and Computation". *Naval Research Logistics* 69(3):496–507.

Frazier, P. I., W. B. Powell, and S. Dayanik. 2008. "A knowledge-gradient policy for sequential information collection". *SIAM Journal on Control and Optimization* 47(5):2410–2439.

Frazier, P. I., W. B. Powell, and S. Dayanik. 2009. "The knowledge-gradient policy for correlated normal beliefs". *INFORMS Journal on Computing* 21(4):599–613.

Gao, S., J. Du, and C.-H. Chen. 2019. "Selecting the optimal system design under covariates". In *IEEE 15th International Conference on Automation Science and Engineering (CASE)*, 547–552. IEEE.

Li, H., H. Lam, Z. Liang, and Y. Peng. 2020. "Context-Dependent Ranking and Selection Under a Bayesian Framework". In *Proc. 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 2060–2070. IEEE, Inc.

Pearce, M. and J. Branke. 2018. "Continuous multi-task Bayesian optimisation with correlation". *European Journal of Operational Research* 270(3):1074–1085.

Powell, W. B. and I. O. Ryzhov. 2012. *Optimal learning*. John Wiley & Sons.

Shen, H., L. J. Hong, and X. Zhang. 2021. "Ranking and selection with covariates for personalized decision making". *INFORMS Journal on Computing* 33(3):1500–1519.

Wang, J., S. C. Clark, E. Liu, and P. I. Frazier. 2020. "Parallel Bayesian Global Optimization of Expensive Functions". *Operations Research* 68(6):1850—1865.

Wu, J. and P. Frazier. 2016. "The parallel knowledge gradient method for batch Bayesian optimization". In *Advances in Neural Information Processing Systems*, 3126–3134.

Xiong, S. 2020. "Personalized optimization and its implementation in computer experiments". *IISE Transactions* 52(5):528–536.

AUTHOR BIOGRAPHIES

ANDRES ALBAN is an Assistant Professor at the Frankfurt School of Finance & Management. His research interests revolve around stochastic simulation and optimization to support decision-making in healthcare settings. His email address is a.alban@fs.de and his website is https://www.frankfurt-school.de/en/Person/0000008791291~~/andres-alban.

STEPHEN E. CHICK is a Professor of Technology and Operations Management and the Novartis Chair of Healthcare Management at INSEAD. He works in the areas of simulation analysis, sequential optimization, health care management, and Bayesian inference. His email address is stephen.chick@insead.edu and his website is https://www.insead.edu/faculty/stephen-e-chick.

SPYROS I. ZOUMPOULIS is an Associate Professor of Decision Sciences at INSEAD. He works on developing prescriptive analytics and algorithms for personalization and experimentation. His email address is spyros.zoumpoulis@insead.edu and his website is https://www.insead.edu/faculty-personal-site/spyros-zoumpoulis.