Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

CONTEXT-DEPENDENT RANKING AND SELECTION UNDER A BAYESIAN FRAMEWORK

Haidong Li

Department of Industrial Engineering and Management
Peking University
5 Yiheyuan Road
Beijing 100871, P. R. CHINA

Henry Lam

Department of Industrial Engineering and Operations Research
Columbia University
500 W. 120th Street
New York, NY 10027 USA

Zhe Liang

Department of Management Science and Engineering
Tongji University
1239 Siping Road
Shanghai 200092, P. R. CHINA

Yijie Peng

Department of Management Science and Information Systems
Peking University
5 Yiheyuan Road
Beijing 100871, P. R. CHINA

ABSTRACT

We consider a context-dependent ranking and selection problem. The best design is not universal but depends on the contexts. Under a Bayesian framework, we develop a dynamic sampling scheme for context-dependent optimization (DSCO) to efficiently learn and select the best designs in all contexts. The proposed sampling scheme is proved to be consistent. Numerical experiments show that the proposed sampling scheme significantly improves the efficiency in context-dependent ranking and selection.

1 INTRODUCTION

Simulation is a powerful tool for optimizing complex stochastic systems. We consider a simulation optimization problem of selecting the best design under different contexts. The mean performances of each design under each context are unknown and can only be estimated via simulation. The performance of each design depends on the contexts, and thus the best design is also context-dependent. For example, in patient-specific treatment regimen-making (Kim et al. 2011), therapies and patients can be regarded as designs and contexts, respectively. We aim to determine the most effective medical treatment for each patient. Other examples include movie recommendation (Liu et al. 2009) and automated asset management (Faloon and Scherer 2017).

For any fixed context, we aim to find the best design among a finite set of alternatives, which is referred to as ranking and selection (R&S) in the literature. R&S procedures intelligently allocate simulation replications to efficiently learn the best design. The probability of correct selection (PCS) is used as a measure to evaluate the efficiency of sampling procedure in R&S. In our problem, the best design is not universal but context-dependent. In this work, we develop a sampling scheme to allocate simulation replications for efficiently learning the best design in each context, and the worst-case probability of correct selection (PCS_W) under all contexts is used to measure the efficiency of our sampling scheme.

There are the frequentist and Bayesian branches in R&S. See Kim and Nelson (2006) and Chen et al. (2015) for overviews. Frequentist procedures (e.g., Rinott 1978, Kim and Nelson 2001, Luo et al. 2015) allocate simulation replications to guarantee a pre-specified PCS level, whereas Bayesian procedures (e.g. Chen et al. 2000, Chick and Frazier 2012, Gao et al. 2017) aim to either maximize the PCS or minimize the expected opportunity cost subject to a given simulation budget. Bayesian procedures usually achieve better performance than frequentist procedures under a given simulation budget, but they typically do not provide a guaranteed PCS. Peng et al. (2018) propose a stochastic control problem to formulate the sequential decision for the Bayesian framework of R&S. Using a value function approximation, they develop efficient sampling procedures to approximate the optimal policy for maximizing the posterior PCS.

The literature on context-dependent R&S is sparse relative to the actively studied R&S problem in simulation. Contexts are also known as the covariates, side information, or auxiliary quantities. To the best of our knowledge, the study of Shen et al. (2017) is the first research for this problem. They assume a linear relationship between the response of a design and the contexts, and develop sampling procedures to provide a guarantee on PCS for all contexts. Li et al. (2018) further extend the result in Shen et al. (2017) to high-dimensional contexts and general dependence between the mean performance of a design and the contexts. The aforementioned two studies adopt the Indifference Zone paradigm in the frequentist branch. Gao et al. (2019) adopt an optimal computing budget allocation (OCBA) approach in R&S, and solve the problem by identifying the rate-optimal budget allocation rule.

In our work, we consider the context-dependent R&S problem under a stochastic control framework, and formulate the sequential sampling decision as a stochastic dynamic programming problem. Under a Bayesian framework, we provide an efficient scheme to update the posterior information for each design-context pair and a dynamic sampling policy based on the sequentially updated posterior information to efficiently learn the performance of each design. The proposed sampling scheme is proved to be consistent and is demonstrated to perform well empirically.

The rest of the paper is organized as follows. In Section 2, we formulate the context-dependent R&S problem. Section 3 proposes an efficient dynamic sampling scheme to approximately solve the problem. Section 4 presents numerical results, and the last section concludes the paper and outlines future directions.

2 PROBLEM FORMULATION

Suppose there are n different designs. For $i = 1,\ldots,n$, the performance $y_i(\boldsymbol{x})$ of design i depends on a vector of context $\boldsymbol{x} = (x_1,\ldots,x_d)^\top$ for $\boldsymbol{x} \in \mathcal{X} \subseteq \mathbb{R}^d$. Each dimension of $\boldsymbol{x}$ could be a continuous variable, discrete variable, or categorical variable. In this study, we assume that $\mathcal{X}$ contains a finite number of m possible contexts $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_m$, and all the contexts are known upfront. Our objective is to correctly select the best design for a given value of $\boldsymbol{x}$ (see Figure 1 for an illustration), i.e., identify $\arg\max_i y_i(\boldsymbol{x})$. The offline optimization results can be used for online decision making. For example, in personalized medicine (Bertsimas et al. 2017), doctors use simulation to determine the best treatment regimen (design) for patients (context) with different biometric characteristics; in personalized movie recommendation (Zhang et al. 2016), we aim to recommend the most favorite movie (design) for the corresponding user (context).

The performances $y_i(\boldsymbol{x}_j)$ are unknown and can only be learned via sampling. We assume that for each design and context, the simulation observations are i.i.d. normally distributed, i.e., $Y_{i,t}(\boldsymbol{x}_j) \sim N(y_i(\boldsymbol{x}_j), \sigma_i^2(\boldsymbol{x}_j))$, $i = 1,\ldots,n$, $j = 1,\ldots,m$, $t \in \mathbb{Z}^+$, and the replications are independent across different designs or different contexts. The variance $\sigma_i^2(\boldsymbol{x}_j)$ in the sampling distribution is assumed to be known in this study and the sample estimate is used as a plug-in for the true value in practice. Suppose the prior distribution of $y_i(\boldsymbol{x}_j)$ is $N(\mu_0, \sigma_0^2)$. By conjugacy (Gelman et al. 2014), the posterior distribution of $y_i(\boldsymbol{x}_j)$ is a normal distribution with the posterior mean:

$$\mu_{ij}^{(t)} = (\sigma_{ij}^2)^{(t)} \left[\frac{\sum_{h=1}^{t_{ij}} Y_{i,h}(\boldsymbol{x}_j)}{\sigma_i^2(\boldsymbol{x}_j)} + \frac{\mu_0}{\sigma_0^2} \right],$$

and the posterior variance:

$$(\sigma_{ij}^2)^{(t)} = 1 \Big/ \left[\frac{t_{ij}}{\sigma_i^2(\boldsymbol{x}_j)} + \frac{1}{\sigma_0^2} \right],$$

where t_{ij} is the number of samples allocated to estimate $y_i(\boldsymbol{x}_j)$ after allocating $t = \sum_{i=1}^n \sum_j^m t_{ij}$ samples. As for a normal distribution with unknown variance, the posterior distribution of $y_i(\boldsymbol{x}_j)$ has a normal-gamma conjugate prior (DeGroot 2005), and the corresponding analysis in context-dependent R&S problem is left for future research.

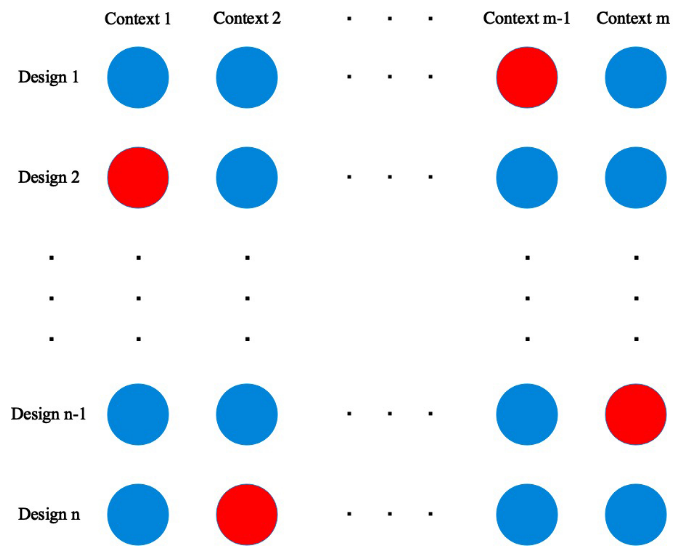

Figure 1: Selecting the best (red) design under each context.

Since sampling could be expensive, the total number of samples is usually limited. Moreover, when either n or m is relatively large, it would be practically infeasible to estimate all performances accurately for each design i and each context $\boldsymbol{x}_j$. Given the information of s allocated samples, the selection is to pick the designs with the largest posterior estimates in each context. Under a fixed context $\boldsymbol{x}_j$, the quality of the selection for the best design is measured by the probability of correct selection (PCS),

$$\text{PCS}(\boldsymbol{x}_j) = \mathbb{P}\left(y_{\langle 1 \rangle_{js}}(\boldsymbol{x}_j) > y_{\langle i \rangle_{js}}(\boldsymbol{x}_j), i \neq 1 \middle| \mathcal{E}_s \right),$$

where $\mathbb{P}\left(\cdot|\mathcal{E}_s\right)$ denotes the posterior probability, $\mathcal{E}_s$ is the information set of all s samples, and $\langle i \rangle_{js}$, $i = 1, \dots, n$, are the ranking indices for context $\boldsymbol{x}_j$ such that

$$\mu^{(s)}_{\langle 1 \rangle_{js} j} > \cdots > \mu^{(s)}_{\langle n \rangle_{js} j}.$$

In this study, we aim to provide the best design for all the $\boldsymbol{x}$ that might possibly appear, and therefore need a measure for evaluating the quality of the selection over the entire context space $\mathcal{X}$. Specifically, we adopt the worst-case probability of correct selection over $\mathcal{X}$:

$$\text{PCS}_{\text{W}} = \min_{\boldsymbol{x} \in \mathcal{X}} \text{PCS}(\boldsymbol{x}).$$

This measure has been used in contextual R&S (Gao et al. 2019), and is similar to the worst-case performance in robust optimization (Bertsimas et al. 2011) and R&S with input uncertainty (Gao et al. 2017, Fan et al. 2020).

We aim to provide a dynamic sampling scheme $\mathcal{A}_s$ to maximize the PCS_{W}:

$$\max_{\mathcal{A}_s} \min_{j=1,\ldots,m} \mathbb{P}\left(y_{\langle 1 \rangle_{js}}(\boldsymbol{x}_j) > y_{\langle i \rangle_{js}}(\boldsymbol{x}_j), i \neq 1 \middle| \mathcal{E}_s \right). \tag{1}$$

The dynamic sampling scheme $\mathcal{A}_s$ is a sequence of maps $\mathcal{A}_s(\cdot) = (A_1(\cdot), \ldots, A_s(\cdot))$. Based on sampling observations $\mathcal{E}_{t-1}$, $A_t(\mathcal{E}_{t-1}) \in \{(i,j) : 1 \leq i \leq n,\ 1 \leq j \leq m\}$ allocates the t-th sample to estimate the performance of design i in context $\boldsymbol{x}_j$. Similar to that in Peng et al. (2016) and Peng et al. (2018), the sequential sampling decision such as (1) can be formulated as a stochastic control (dynamic programming) problem. The expected payoff for a sampling scheme $\mathcal{A}_s$ can be defined recursively by

$$V_s(\mathcal{E}_s; \mathcal{A}_s) \triangleq \min_{j=1,\ldots,m} \mathbb{P}\left(y_{\langle 1 \rangle_{js}}(\boldsymbol{x}_j) > y_{\langle i \rangle_{js}}(\boldsymbol{x}_j), i \neq 1 | \mathcal{E}_s \right),$$

and for $0 \leq t < s$,

$$V_t(\mathcal{E}_t; \mathcal{A}_s) \triangleq \mathbb{E}\left[V_{t+1}(\mathcal{E}_t \cup \{Y_{i,t_{ij}+1}(\boldsymbol{x}_j)\}; \mathcal{A}_s) \middle| \mathcal{E}_t \right] \Big|_{(i,j)=A_{t+1}(\mathcal{E}_t)}.$$

Then, the optimal sampling scheme is well defined by

$$\mathcal{A}_s^* \triangleq \arg\max_{\mathcal{A}_s} V_0(\theta_0; \mathcal{A}_s),$$

where θ_0 is parameter in prior distribution. It is important to note that the definition of decision variable in our study is different from the one in R&S. For the R&S problem, the decision is to choose an alternative i in sampling, whereas our decision is to choose a design-context pair (i,j) in sampling.

3 DYNAMIC SAMPLING SCHEME

In principle, backward induction can be used to solve the stochastic dynamic programming problem, but it suffers from curse-of-dimensionality (Peng et al. 2018). To derive a dynamic sampling scheme with an analytical form, we adopt approximate dynamic programming (ADP) schemes which make dynamic decision based on a value function approximation (VFA) and keep learning the VFA with decisions moving forward.

By treating any t-th step as the last step, the value function in our problem is

$$V_t(\mathcal{E}_t) = \min_{j=1,\ldots,m} \mathbb{P}\left(y_{\langle 1 \rangle_{jt}}(\boldsymbol{x}_j) > y_{\langle i \rangle_{jt}}(\boldsymbol{x}_j), i \neq 1 \middle| \mathcal{E}_t \right). \tag{2}$$

Conditioned on $\mathcal{E}_t$, $y_i(\boldsymbol{x}_j)$ follows a normal distribution with mean $\mu_{ij}^{(t)}$ and variance $(\sigma_{ij}^2)^{(t)}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$. Therefore, the joint distribution of vector $(y_{\langle 1 \rangle_{jt}}(\boldsymbol{x}_j) - y_{\langle 2 \rangle_{jt}}(\boldsymbol{x}_j), \ldots, y_{\langle 1 \rangle_{jt}}(\boldsymbol{x}_j) - y_{\langle n \rangle_{jt}}(\boldsymbol{x}_j))$ follows a joint normal distribution with mean vector $(\mu_{\langle 1 \rangle_{jt}j}^{(t)} - \mu_{\langle 2 \rangle_{jt}j}^{(t)}, \ldots, \mu_{\langle 1 \rangle_{jt}j}^{(t)} - \mu_{\langle n \rangle_{jt}j}^{(t)})$ and covariance matrix $\Gamma' \Lambda \Gamma$, where $\Lambda \triangleq diag((\sigma_{\langle 1 \rangle_{jt}j}^2)^{(t)}, \ldots, (\sigma_{\langle n \rangle_{jt}j}^2)^{(t)})$, and

$$\Gamma \triangleq \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{bmatrix}_{n \times (n-1)},$$

The posterior probability

$$\mathbb{P}\left(y_{\langle 1\rangle_{jt}}(\boldsymbol{x}_j) > y_{\langle i\rangle_{jt}}(\boldsymbol{x}_j), i \neq 1 \middle| \mathcal{E}_t\right) \tag{3}$$

is an integral of the multivariate standard normal density over a region encompassed by some hyperplanes. Under a given context, the above probability of correct selection is consistent with that in R&S literature, which allows us to use a similar approximation developed in Peng et al. (2018). As shown in Figure 2, the integral over a maximum tangent inner ball in the shadowed region can capture the main body of the integral over entire region due to the exponential decay of the normal density. It can be shown that the integral over the omitted domain decreases to zero exponentially as t goes to infinity. Therefore, we use the volume of the ball as an approximation for the PCS in (3). Specifically, we use a VFA for the value function (2) given by

$$\widetilde{V}_t(\mathcal{E}_t) = \min_{j=1,\ldots,m} \min_{i\neq 1} \frac{(\mu^{(t)}_{\langle 1\rangle_{jt}j} - \mu^{(t)}_{\langle i\rangle_{jt}j})^2}{(\sigma^2_{\langle 1\rangle_{jt}j})^{(t)} + (\sigma^2_{\langle i\rangle_{jt}j})^{(t)}}.$$

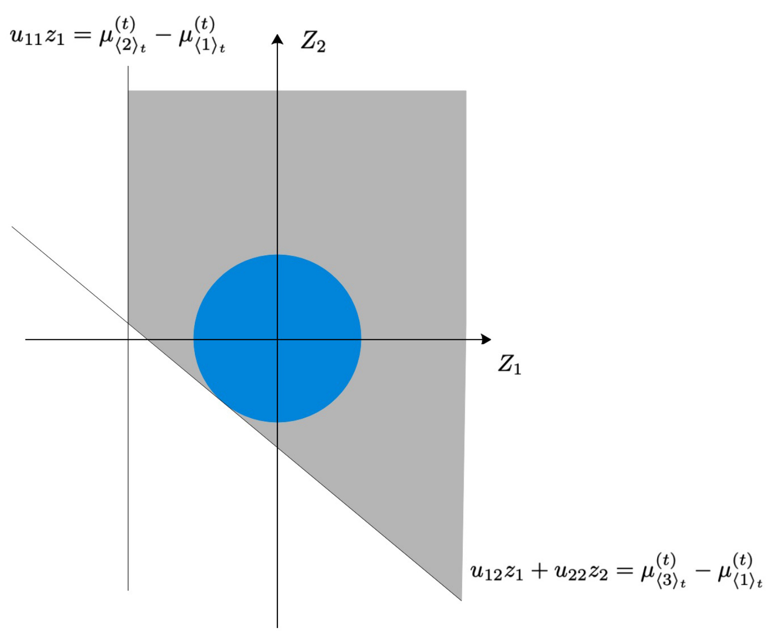

Figure 2: Approximation of PCS in (3).

At any step t, we treat the $(t+1)$-th step as the last step and try to maximize the expected VFA by allocating the $(t+1)$-th sample to a design-context pair (r,q):

$$\widetilde{V}_t(\mathcal{E}_t;(r,q)) \triangleq \mathbb{E}\left[\widetilde{V}_{t+1}(\mathcal{E}_t \cup \{Y_{r,t_{rq}+1}(\boldsymbol{x}_q)\}) \middle| \mathcal{E}_t\right].$$

Further, we apply a certainty equivalent approximation (Bertsekas 1995) to the VFA looking one-step ahead:

$$\mathbb{E}\left[\widetilde{V}_{t+1}(\mathcal{E}_t \cup \{Y_{r,t_{rq}+1}(\boldsymbol{x}_q)\}) \middle| \mathcal{E}_t\right] \approx \widetilde{V}_{t+1}\left(\mathcal{E}_t \cup \mathbb{E}\left[Y_{r,t_{rq}+1}(\boldsymbol{x}_q) \middle| \mathcal{E}_t\right]\right).$$

Then the VFA looking one-step ahead can be calculated, i.e.,

$$\widehat{V}_t(\mathcal{E}_t;(r,q)) \triangleq \min_{j=1,\dots,m} \min_{i\neq 1} \frac{(\mu^{(t)}_{\langle 1\rangle_{jt}j} - \mu^{(t)}_{\langle i\rangle_{jt}j})^2}{(\sigma^2_{\langle 1\rangle_{jt}j})^{(t;(r,q))} + (\sigma^2_{\langle i\rangle_{jt}j})^{(t;(r,q))}}, \tag{4}$$

where

$$(\sigma^2_{ij})^{(t;(r,q))} = \begin{cases} 1\Big/\left[\dfrac{t_{ij}+1}{\sigma_i^2(\boldsymbol{x}_j)} + \dfrac{1}{\sigma_0^2}\right] & , \quad (i,j)=(r,q), \\ (\sigma^2_{ij})^{(t)} & , \quad \text{otherwise.} \end{cases}$$

We propose the following dynamic sampling scheme for context-dependent optimization (DSCO):

$$A_{t+1}(\mathcal{E}_t) = \left\{(r^*,q^*)\middle|\widehat{V}_t(\mathcal{E}_t;(r^*,q^*)) = \max_{(r,q)} \widehat{V}_t(\mathcal{E}_t;(r,q))\right\}, \tag{5}$$

which maximizes VFA looking one-step ahead.

DSCO uses the information on the posterior means and variances of context-dependent performances of design. We note that $(\mu^{(t)}_{\langle 1\rangle_{jt}j} - \mu^{(t)}_{\langle i\rangle_{jt}j})^2$ and $\left((\sigma^2_{\langle 1\rangle_{jt}j})^{(t;(r,q))} + (\sigma^2_{\langle i\rangle_{jt}j})^{(t;(r,q))}\right)$ are the squared mean and variance of the posterior distribution of the difference in performances of design, respectively. Therefore, equation (4) can be rewritten as

$$\min_{j=1,\dots,m} \min_{i\neq 1} 1/c_v^2(i,j),$$

where $c_v(i,j)$ is the coefficient of variation (CV, or sometimes called noise-signal ratio) of the posterior $(y_{\langle 1\rangle_{jt}}(\boldsymbol{x}_j) - y_{\langle i\rangle_{jt}}(\boldsymbol{x}_j))$. DSCO minimizes the maximum of $c_v(i,j)$, which is intuitively reasonable since large $c_v(i,j)$ implies high difficulty in comparing $y_{\langle 1\rangle_{jt}}(\boldsymbol{x}_j)$ and $y_{\langle i\rangle_{jt}}(\boldsymbol{x}_j)$ from the posterior information. DSCO sequentially allocates each sample to estimate performance of design to reduce the CV, focusing on the worst-case context and the pair most difficult to rank. Under the Bayesian framework, consistency means that for any possible realization of the true means for different alternatives from the prior distribution, the selected best alternative based on the posterior means of different alternatives would eventually be the true best alternative for almost every sample path in the sampling probability space as the simulation budget goes to infinity (Frazier and Powell 2011). DSCO is proved to be consistent in the following theorem.

Theorem 1 The proposed DSCO is consistent, i.e., $\forall j = 1,\dots,m$,

$$\lim_{t\to+\infty} \langle 1\rangle_{jt} = \langle 1\rangle_j \quad a.s.$$

where $\langle i\rangle_j,\ i=1,\dots,n$, are the ranking indices for context $\boldsymbol{x}_j$ such that $y_{\langle 1\rangle_j}(\boldsymbol{x}_j) > \cdots > y_{\langle n\rangle_j}(\boldsymbol{x}_j)$ and thus $\langle 1\rangle_j$ denotes the true best design in context $\boldsymbol{x}_j$.

Proof. We only need to prove that each $y_i(\boldsymbol{x}_j)$ will be sampled infinitely often a.s. following DSCO scheme, and the consistency will follow by the law of large numbers. Suppose $y_i(\boldsymbol{x}_j)$ is only sampled finitely often and $y_r(\boldsymbol{x}_q)$ is sampled infinitely often. Therefore, there exists a finite number N_0 such that $y_i(\boldsymbol{x}_j)$ will stop receiving replications after the sampling number t exceeds N_0. Thus we have

$$\lim_{t\to+\infty} (\sigma^2_{ij})^{(t)} > 0,\ \lim_{t\to+\infty} (\sigma^2_{rq})^{(t)} = 0.$$

By noticing that

$$\lim_{t\to+\infty} \left[(\sigma^2_{rq})^{(t)} - (\sigma^2_{rq})^{(t;(r,q))}\right] = 0,$$

$$\lim_{t\to+\infty}\left[(\sigma_{ij}^2)^{(t)}-(\sigma_{ij}^2)^{(t;(i,j))}\right]>0,$$

we have

$$\lim_{t\to+\infty}\left[\widehat{V}_t(\mathcal{E}_t;(r,q))-\widetilde{V}_t(\mathcal{E}_t)\right]=0\ \ a.s.,$$

$$\lim_{t\to+\infty}\left[\widehat{V}_t(\mathcal{E}_t;(i,j))-\widetilde{V}_t(\mathcal{E}_t)\right]>0\ \ a.s.,$$

which contradicts with the sampling rule (5) that the design-context pair with the largest $\widehat{V}_t(\mathcal{E}_t;(i,j))$ is sampled. Therefore, the proposed DSCO scheme must be consistent. □

One can also prove the proposed DSCO achieves the following asymptotically optimal sampling ratio obtained in Gao et al. (2019) as the number of samples goes to infinity:

$$\frac{t^2_{\langle 1\rangle_j j}}{\sigma^2_{\langle 1\rangle_j}(\boldsymbol{x}_j)}=\sum_{i=2}^{n}\frac{t^2_{\langle i\rangle_j j}}{\sigma^2_{\langle i\rangle_j}(\boldsymbol{x}_j)},\ j=1,\ldots,m,$$

$$\frac{(y_{\langle 1\rangle_j}(\boldsymbol{x}_j)-y_{\langle i\rangle_j}(\boldsymbol{x}_j))^2}{\sigma^2_{\langle 1\rangle_j}(\boldsymbol{x}_j)/t_{\langle 1\rangle_j j}+\sigma^2_{\langle i\rangle_j}(\boldsymbol{x}_j)/t_{\langle i\rangle_j j}}=\frac{(y_{\langle 1\rangle_{j'}}(\boldsymbol{x}_{j'})-y_{\langle i'\rangle_{j'}}(\boldsymbol{x}_{j'}))^2}{\sigma^2_{\langle 1\rangle_{j'}}(\boldsymbol{x}_{j'})/t_{\langle 1\rangle_{j'} j'}+\sigma^2_{\langle i'\rangle_{j'}}(\boldsymbol{x}_{j'})/t_{\langle i'\rangle_{j'} j'}},\ i,i'=2,\ldots,n,\ j,j'=1,\ldots,m.$$

The proof is rather technical and will be included in later work.

4 NUMERICAL RESULTS

In this section, we conduct numerical experiments to test the performance of different sampling procedures for context-dependent R&S problems. The proposed DSCO is compared with the equal allocation (EA), two-stage indifference-zone (IZ) procedure in Shen et al. (2017), and the contextual optimal computing budget allocation (C-OCBA) in Gao et al. (2019):

- EA: This procedure allocates the same number of samples to any design-context pair.
- IZ: At first stage, this procedure takes n_0 independent samples of each design-context pair and calculates sample variances S_{ij}^2; At second stage, take $\max\{\lceil h^2 S_{ij}^2/\delta^2\rceil - n_0, 0\}$ additional independent samples of design i in context $\boldsymbol{x}_j$, where h and δ are IZ parameters.
- C-OCBA: This procedure sequentially allocates each sample to achieve the asymptotically optimal sampling ratio.

In all numerical examples, the statistical efficiency of the sampling procedures is measured by the $\mathrm{PCS_W}$ estimated by 10,000 independent experiments. The $\mathrm{PCS_W}$ is reported as a function of the sampling budget in each experiment.

Example 1: 10×10 **Design-Context Pairs**

We test our proposed DSCO in a synthetic case with 10 designs and 10 contexts. The performances of each design for each context are generated as follows:

$$y_i(\boldsymbol{x}_j)\sim N(50,3^2),\ i=1,\ldots,10,\ j=1,\ldots,10,$$

which means there is no clear performance clustering structure and all design-context pairs belong to a common cluster. For design i and context $\boldsymbol{x}_j$, samples are drawn independently from a normal distribution $N(y_i(\boldsymbol{x}_j),\sigma_i^2(\boldsymbol{x}_j))$, where $\sigma_i(\boldsymbol{x}_j)\sim U(8,12)$. We set the number of initial replications as $n_0=5$ for each design-context pair.

In Figure 3, we can see that DSCO and C-OCBA perform better than IZ and EA, which could be attributed to the reason that EA utilizes no sample information and IZ only utilizes sample variances while the other two sampling procedures utilize the information in the posterior means and variances. In order

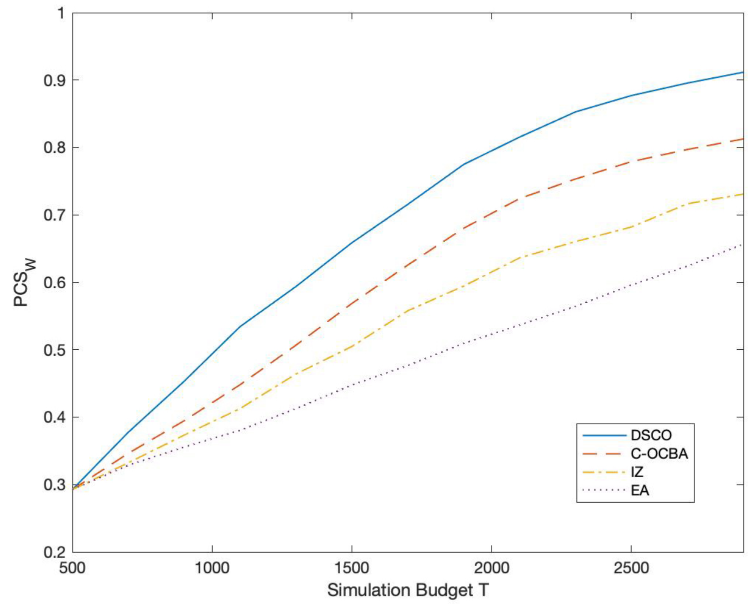

Figure 3: $\mathrm{PCS_W}$ of the four sampling procedures in Example 1.

to attain $\mathrm{PCS_W} = 80\%$, the number of samples consumed by DSCO is 2000, while EA, IZ and C-OCBA require more than 2800 samples. That is to say DSCO reduces the sampling budget by more than 28%. Compared with C-OCBA, the performance enhancement of DSCO could be attributed to that it has a theoretical support for the finite-sample performance in its derivation via VFA.

Example 2: 30×30 **Design-Context Pairs**

In this example, our proposed DSCO is tested in a larger synthetic case with 30 designs and 30 contexts. The performances of each design for each context are generated as follows:

$$y_i(\boldsymbol{x}_j) \sim N(50, 15^2),\ i = 1, \ldots, 30,\ j = 1, \ldots, 30.$$

For design i and context $\boldsymbol{x}_j$, samples are drawn independently from a normal distribution $N(y_i(\boldsymbol{x}_j), \sigma_i^2(\boldsymbol{x}_j))$, where $\sigma_i(\boldsymbol{x}_j) \sim U(4, 6)$. We set the number of initial replications as $n_0 = 5$ for each design-context pair.

Figure 4 illustrates the performance of the four sampling procedures. In the presence of more alternative designs and contexts, larger sampling budget is needed to reach the same $\mathrm{PCS_W}$ level. Similar to Example 1, DSCO remains as the most efficient sampling procedure among the four, and C-OCBA is better than IZ and EA. Comparing the result in this example with that in Example 1, we can see that the advantage of DSCO is more significant when the numbers of designs and contexts become larger. In order to attain $\mathrm{PCS_W} = 90\%$, DSCO consumes less than 30,000 samples, while EA, IZ and C-OCBA require more than 45,000 samples. That is to say DSCO reduces the sampling budget by more than 33%. Moreover, the gap between the $\mathrm{PCS_W}$ of DSCO and those of EA, IZ, and C-OCBA widens as the sampling budget increases in the experiment.

5 CONCLUSION

This paper studies a sample allocation problem for context-dependent R&S. We propose an efficient sampling procedure named DSCO, which maximizes the worst-case probability of correct selection over the entire context space. Numerical experiments demonstrate that DSCO is significantly more efficient than the other tested sampling procedures. Future research includes the asymptotic analysis for the sampling ratio of the

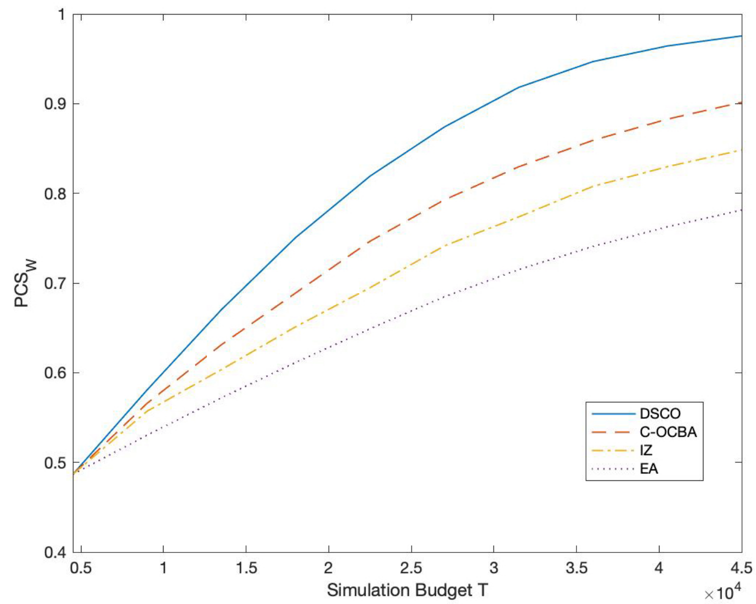

Figure 4: PCS$_W$ of the four sampling procedures in Example 2.

proposed sequential sampling procedure. Considering the performance clustering between different designs and contexts could also be an interesting future work.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation of China (NSFC) under Grants 71901003, 91846301, 71790615, 71720107003, 71690232, and 61603321, by the National Science Foundation under Awards ECCS-1462409, CMMI-1462787, CAREER CMMI-1834710 and IIS-1849280, and by the scholarship from China Scholarship Council (CSC) under the Grant CSC.

REFERENCES

Bertsekas, D. P. 1995. *Dynamic Programming and Optimal Control*, Volume 1. Belmont, Massachusetts: Athena Scientific.

Bertsimas, D., D. B. Brown, and C. Caramanis. 2011. “Theory and Applications of Robust Optimization”. *Society for Industrial and Applied Mathematics Review* 53(3):464–501.

Bertsimas, D., N. Kallus, A. M. Weinstein, and Y. D. Zhuo. 2017. “Personalized Diabetes Management Using Electronic Medical Records”. *Diabetes Care* 40(2):210–217.

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. New York: Springer Science & Business Media.

Chen, C.-H., J. Lin, E. Yücesan, and S. E. Chick. 2000. “Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization”. *Discrete Event Dynamic Systems* 10(3):251–270.

Chick, S. E., and P. Frazier. 2012. “Sequential Sampling with Economics of Selection Procedures”. *Management Science* 58(3):550–569.

DeGroot, M. H. 2005. *Optimal Statistical Decisions*, Volume 82. Hoboken, New Jersey: John Wiley & Sons, Inc.

Faloon, M., and B. Scherer. 2017. “Individualization of Robo-Advice”. *The Journal of Wealth Management* 20(1):30–36.

Fan, W., L. J. Hong, and X. Zhang. 2020. "Distributionally Robust Selection of the Best". *Management Science* 66(1):190–208.

Frazier, P. I., and W. B. Powell. 2011. "Consistency of Sequential Bayesian Sampling Policies". *Society for Industrial and Applied Mathematics Journal on Control and Optimization* 49(2):712–731.

Gao, S., W. Chen, and L. Shi. 2017. "A New Budget Allocation Framework for the Expected Opportunity Cost". *Operations Research* 65(3):787–803.

Gao, S., J. Du, and C.-H. Chen. 2019. "Selecting the Optimal System Design under Covariates". In *Proceedings of the 15th IEEE International Conference on Automation Science and Engineering*, 547–552. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Gao, S., H. Xiao, E. Zhou, and W. Chen. 2017. "Robust Ranking and Selection with Optimal Computing Budget Allocation". *Automatica* 81:30–36.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2014. *Bayesian Data Analysis*. Boca Raton, Florida: CRC Press.

Kim, E. S., R. S. Herbst, I. I. Wistuba, J. J. Lee, G. R. Blumenschein, A. Tsao, D. J. Stewart, M. E. Hicks, J. Erasmus, S. Gupta et al. 2011. "The Battle Trial: Personalizing Therapy for Lung Cancer". *Cancer Discovery* 1(1):44–53.

Kim, S.-H., and B. L. Nelson. 2001. "A Fully Sequential Procedure for Indifference-Zone Selection in Simulation". *Association for Computing Machinery Transactions on Modeling and Computer Simulation* 11(3):251–273.

Kim, S.-H., and B. L. Nelson. 2006. "Selecting the Best System". In *Handbooks in Operations Research and Management Science: Simulation*, Volume 13, 501–534. Amsterdam, Netherlands: Elsevier.

Li, X., X. Zhang, and Z. Zheng. 2018. "Data-Driven Ranking and Selection: High-Dimensional Covariates and General Dependence". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, 1933–1944. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Liu, A., Y. Zhang, and J. Li. 2009. "Personalized Movie Recommendation". In *Proceedings of the 17th Association for Computing Machinery International Conference on Multimedia*, 845–848. New York, New York: Association for Computing Machinery.

Luo, J., L. J. Hong, B. L. Nelson, and Y. Wu. 2015. "Fully Sequential Procedures for Large-Scale Ranking-and-Selection Problems in Parallel Computing Environments". *Operations Research* 63(5):1177–1194.

Peng, Y., C.-H. Chen, M. C. Fu, and J.-Q. Hu. 2016. "Dynamic Sampling Allocation and Design Selection". *INFORMS Journal on Computing* 28(2):195–208.

Peng, Y., E. K. Chong, C.-H. Chen, and M. C. Fu. 2018. "Ranking and Selection as Stochastic Control". *IEEE Transactions on Automatic Control* 63(8):2359–2373.

Rinott, Y. 1978. "On Two-Stage Selection Procedures and Related Probability-Inequalities". *Communications in Statistics - Theory and Methods* 7(8):799–811.

Shen, H., L. J. Hong, and X. Zhang. 2017. "Ranking and Selection with Covariates for Personalized Decision Making". https://arxiv.org/abs/1710.02642v2, accessed 15^{th} July, 2020.

Zhang, S., Z. Jin, and J. Zhang. 2016. "The Dynamical Modeling and Simulation Analysis of the Recommendation on the User-Movie Network". *Physica A: Statistical Mechanics and its Applications* 463:310–319.

AUTHOR BIOGRAPHIES

HAIDONG LI is a Ph.D. candidate in the Department of Industrial Engineering and Management, Peking University, Beijing, China. He received his B.S. Degree from the Department of Engineering Mechanics at Peking University. His research interests include simulation optimization and network analysis. He is a visiting scholar under the supervision of Prof. Henry Lam during 2019-2020. His email address is haidong.li@pku.edu.cn.

HENRY LAM is an Associate Professor in the Department of Industrial Engineering and Operations Research at Columbia University. He received his Ph.D. degree in statistics from Harvard University in 2011, and was on the faculty of Boston University and the University of Michigan before joining Columbia in 2017. His research focuses on Monte Carlo simulation, uncertainty quantification, risk analysis, and stochastic and robust optimization. His email address is henry.lam@columbia.edu.

ZHE LIANG is a Professor in the Department of Management Science and Engineering, Tongji University, Shanghai, China. He received the B.E. degree in Computer Engineering from National University of Singapore, Singapore, Singapore, in 2001, and the Ph.D. degree in Industrial and Systems Engineering from Rutgers, The State University of New Jersey, New Brunswick, United States, in 2011, respectively. His research interests include combinatorial optimization, mathematical programming, and scheduling and management in aviation, logistics, ports, transportation, and telecommunications. His email address is liangzhe@tongji.edu.cn.

YIJIE PENG is an Assistant Professor in the Department of Management Science and Information Systems in Guanghua School of Management at Peking University, Beijing, China. He received the B.S. degree in mathematics from Wuhan University, Wuhan, China, in 2007, and the Ph.D. degree in management science from Fudan University, Shanghai, China, in 2014, respectively. His research interests include stochastic modeling and analysis, simulation optimization, machine learning, data analytics, and healthcare. His email address is pengyijie@pku.edu.cn.