Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

ENHANCED UPPER CONFIDENCE BOUND PROCEDURE FOR LARGE-SCALE RANKING AND SELECTION

Song Huang[1], Guangxin Jiang[1], Chenxi Li[1], and Ying Zhong[2]

[1]School of Management, Harbin Institute of Technology, Harbin, CHINA
[2]School of Management and Economics, University of Electronic Science and Technology of China, Chengdu, CHINA

ABSTRACT

With the rapid advancement of computing technology, there has been growing interest in effectively solving large-scale ranking and selection (R&S) problems. In this paper, we propose a new large-scale fixed-budget R&S procedure, namely the enhanced upper confidence bound (EUCB) procedure. The EUCB procedure incorporates variance information into the dynamic allocation of simulation budgets. It selects the alternative with the largest upper confidence bound. We prove that the EUCB procedure has sample optimality; that is, to achieve an asymptotically nonzero probability of correct selection (PCS), the total sample size required grows at the linear order with respect to the number of alternatives. We demonstrate the effectiveness of the EUCB procedure in numerical examples. In addition to achieving sample optimality under the PCS criterion, our numerical experiments also show that the EUCB procedure maintains sample optimality under the expected opportunity cost (EOC) criterion.

1 INTRODUCTION

Ranking and selection (R&S) is a classical and widely studied problem in the field of simulation optimization, where the goal is to identify the alternative with the largest mean performance from a finite set of alternatives by conducting simulation experiments and analyzing the resulting stochastic outputs. To address this problem, various procedures have been developed, which can be classified into two main categories. The first category is the fixed-precision procedure, which aims to guarantee a user-specified probability of correct selection (PCS) while minimizing the total simulation budget. Notable works in this category include the procedures proposed by Bechhofer (1954), Paulson (1964), Rinott (1978), Kim and Nelson (2001), and Hong (2006), among others. The second category is the fixed-budget procedure, which aims to optimally allocate a predetermined simulation budget across all alternatives—either statically or sequentially—to maximize a specific criterion, such as the PCS or expected opportunity cost (EOC). Prominent procedures in this category include those by Chen et al. (2000), Chick and Inoue (2001), and Frazier et al. (2008), among others. For more R&S procedures and applications, see (Kim and Nelson 2006; Chen et al. 2015; Hong et al. 2021; Fan et al. 2025) for surveys.

Traditionally, R&S procedures were developed to address small-scale problems, e.g., $k \leq 1,000$. This limitation arose because obtaining a reliable estimate of the best alternative required simulating each alternative many times. With a large number of alternatives, the computational constraints of a single processor made it impractical to solve such problems within a reasonable amount of time. However, in recent years, there is a growing need to solve large-scale R&S problems in practice. For example, in supply chain network design, each production or distribution stage may involve hundreds of participating firms, leading to an enormous number of possible network configurations (Farahani et al. 2014). Similarly, in reinforcement learning, the effectiveness of policies often relies on the values assigned to hyperparameters. When these hyperparameters are discretized, each combination of values can be considered an alternative.

979-8-3315-8726-0/25/$31.00 ©2025 IEEE

When the number of hyperparameters increases, the potential combinations grow rapidly (Luo et al. 2022). Therefore, developing efficient procedures for large-scale R&S problems has become increasingly important.

With the rapid advancement of computing technology, parallel computing environments are becoming increasingly prevalent and accessible to ordinary users. Consequently, designing efficient procedures to solve large-scale R&S problems in parallel computing environments has become a key focus in recent research. Several parallel fixed-precision procedures have been proposed, including the asymptotic parallel selection (APS) procedure by Luo et al. (2015), the good selection procedure (GSP) by Ni et al. (2017), the parallelized paulson's procedure (PPP) by Zhong et al. (2022), the knockout-tournament (KT) procedure by Zhong and Hong (2022), and the parallel adaptive survivor selection (PASS) procedure by Pei et al. (2024), etc. In the context of fixed-budget R&S problems, Hong et al. (2022) introduced the knockout-tournament framework and proposed the fixed-budget knockout-tournament (FBKT) procedure, which, for the first time, scaled the number of alternatives to the order of millions under the fixed-budget formulation. Moreover, Li et al. (2025) developed the explore-first greedy (EFG) procedure based on greedy principles, demonstrating superior performance in large-scale fixed-budget R&S problems.

Notice that the performance of fixed-budget procedures is typically evaluated by fixing the number of alternatives k and analyzing how quickly the PCS approaches 1 (or the probability of false selection decreases to 0) as the total sampling budget $N \to \infty$ (Wu and Zhou 2018a; Wu and Zhou 2018b). However, for large-scale fixed-budget R&S problems, an alternative evaluation framework has been proposed to better demonstrate the scalability and efficiency of procedures. Hong et al. (2022) introduced the concept of the minimal growth rate required for the total simulation budget N to ensure that the PCS does not diminish to zero as the number of alternatives k increases. This growth rate, termed the *rate for maintaining correct selection* (RMCS), serves as an important performance metric in large-scale problem settings, and a lower RMCS indicates a more efficient procedure. The authors further showed that the optimal lower bound for RMCS is $\mathcal{O}(k)$.

Building on this framework, Li et al. (2025) developed a remarkably simple greedy procedure for large-scale R&S problems, which allocates each new simulation budget to the alternative with the largest observed sample mean. Remarkably, the authors found that this simple procedure achieves the optimal RMCS $\mathcal{O}(k)$, and defined procedures that attain this optimal RMCS as exhibiting *sample optimality*. Despite its simplicity and sample optimality, the greedy procedure is inconsistent; that is, its PCS does not converge to one as $k \to \infty$, even when the total sampling budget grows faster than the order of k (e.g., on the order of $k \log k$). To address this issue, the authors incorporated an equal-allocation phase, inspired by (Hong et al. 2022), into the greedy framework, and proposed the explore-first greedy (EFG) procedure. In this procedure, a fixed proportion of the sampling budget is equally allocated to each alternative before implementing the greedy procedure, leading to substantial empirical improvements. However, the EFG procedure overlooks important statistical information from the simulation samples, such as the sample variance. In contrast, classic fixed-budget R&S procedures, such as Optimal Computing Budget Allocation (OCBA) procedure, have shown that effectively using variance information can significantly enhance performance. This raises a critical research question: how can variance information be reasonably and effectively incorporated into the EFG framework to enhance the PCS?

To bridge this gap, in this paper we propose a new large-scale fixed-budget procedure—the enhanced upper confidence bound (EUCB) procedure—by integrating the EFG procedure with the classic upper confidence bound (UCB) algorithm from the multi-armed bandit problem. The main challenge associated with integrating the UCB algorithm into the EFG framework is that UCB algorithm always allocate samples to underperforming alternative. Similar to the EFG procedure, EUCB begins by allocating a fixed portion of the total simulation budget equally across all alternatives. The remaining budget is then allocated sequentially according to a specific allocation rule. The key distinction lies in this rule: the EUCB procedure replaces the sample mean with the upper confidence bound of the sample mean. This adjustment allows for a better exploration-exploitation balance, as the procedure considers both the current average performance of each alternative and the potential opportunities from alternatives with greater uncertainty. Consequently, alternatives exhibiting either high average performance or significant uncertainty are prioritized for further sampling, thereby enhancing the exploration of high-potential alternatives.

The remainder of this paper is organized as follows. Section 2 presents the problem formulation and details of the EUCB procedure. The sample optimality of the proposed procedure is established in Section 3. Section 4 presents numerical experiments to validate the theoretical findings and evaluate the performance of our procedure for large-scale R&S problems, followed by conclusions in Section 5.

2 PROBLEM STATEMENT

Suppose that there are k alternatives, indexed by the set $\mathbb{K}=\{1,2,\ldots,k\}$. Let the random variable X_i denote the performance of each alternative $i\in\mathbb{K}$. Suppose that the best alternative is defined to have the largest mean performance, that is,

$$\text{the best alternative} = \arg\max_{i\in\mathbb{K}} \mathbb{E}\left[X_i\right].$$

Typically, we assume X_i follows a normal distribution with unknown mean μ_i and variance σ_i^2, and the observations $\{X_{i,j}: j=1,2,\ldots\}$ sampled from each alternative $i\in\mathbb{K}$ are independent and identically distributed (i.i.d.). Furthermore, without loss of generality, we assume that alternative 1 is the best alternative in the rest of this paper, which means $\mu_1>\mu_i,\ \forall i\in\mathbb{K}\backslash\{1\}$. We further define

$$\bar{X}_i(n_i)=\frac{1}{n_i}\sum_{j=1}^{n_i}X_{i,j} \quad\text{and}\quad \hat{S}_i^{\,2}(n_i)=\frac{1}{n_i-1}\sum_{j=1}^{n_i}\left(X_{i,j}-\bar{X}_i(n_i)\right)^2$$

as the sample mean and the sample variance of the first n_i observations of alternative i, and set $\hat{S}_i^{\,2}(1)=0,\ \forall i\in\mathbb{K}$. Following the convention in the fixed-budget R&S literature, our goal is to identify the best alternative, given a user-specified fixed total budget N. Then we further adopt $\hat{i}_\pi$ as the alternative ultimately selected by the procedure π. Therefore, the PCS of procedure π can be writen as

$$\text{PCS}_\pi=\Pr\left\{\hat{i}_\pi=1\right\}.$$

Following the definition of Hong et al. (2022) and Li et al. (2025), we present the following definition of the sample optimality.

Definition 1 A fixed-budget R&S procedure is **sample optimal** if there exists a positive contant $c>0$ such that

$$\liminf_{k\to\infty}\ \text{PCS}>0 \text{ for } N=ck. \tag{1}$$

This is a crucial metric for large-scale fixed-budget R&S procedures, as classical OCBA-type procedures require the total simulation budget to grow at least on the order of $k\log k$ to guarantee an asymptotically nonzero PCS, which becomes inefficient for large-scale problems. Definition 1 indicates that sample-optimal procedures only require the total budget N to grow linearly with k to maintain a non-zero PCS. The FBKT procedure (Hong et al. 2022) and EFG procedure (Li et al. 2025) are two representative sample-optimal fixed-budget procedures. In this paper, we introduce the EUCB procedure, a new sample-optimal algorithm that improves upon EFG by explicitly incorporating variance information to enhance selection efficiency and further improve PCS.

When addressing R&S problems, there is typically a trade-off to resolve before allocating samples. On the one hand, alternatives with larger sample means naturally deserve more samples to prioritize potential optimal solutions. On the other hand, alternatives with higher variances also require additional samples to improve estimation accuracy. However, none of these three sample-optimal procedures take variance into account in their sample allocation strategies. To address this, the EUCB procedure adopts a upper confidence bound that combines the sample mean with its sample standard deviation as an indicator for sample allocation, while the EFG procedure only relies on the sample mean for decision-making.

Building on this idea, the EUCB procedure, as detailed in Algorithm 1, requires three key inputs: the number of alternatives k, represented as $X_1, X_2, \ldots, X_k$, the total sampling budget N, and a first-stage sample size n_0 ($n_0 \geq 2$). The process starts by collecting n_0 independent observations from each alternative $i \in 1, \ldots, k$, denoted as $X_{i,1}, X_{i,2}, \ldots, X_{i,n_0}$. For each alternative i, the sample mean $\bar{X}_i(n_0)$ and the sample variance $\hat{S}_i^2(n_0)$ are calculated based on the first n_0 observations. Using these, the upper confidence bound of sample mean for each alternative is defined as $\hat{U}_i(n_0) = \bar{X}_i(n_0) + \sqrt{\hat{S}_i^2(n_0)/n_0}$. The number of samples allocated to each alternative, n_i, is then initialized accordingly.

Following the first stage, the EUCB procedure enters an iterative sampling phase (Stage 2) that continues until the total number of samples allocated across all alternatives reaches the budget N. In each iteration, the alternative b with the highest upper confidence bound is selected, where $b = \arg\max_{i=1,\ldots,k} \hat{U}_i(n_i)$. A single additional observation is then sampled from alternative b. After that, the sample mean $\bar{X}_b(n_b+1)$, the sample variance $\hat{S}_b^2(n_b+1)$ and the upper confidence bound $\hat{U}_b(n_b+1)$ are updated. This process repeats, allocating one sample at a time to the alternative with the highest $\hat{U}_i(n_i)$, until the budget N is exhausted.

Once the sampling budget is fully utilized, the EUCB procedure concludes by selecting the alternative with the highest final sample mean as the best alternative.

Algorithm 1 Enhance Upper Confidence Bound (EUCB) procedure

Require: k alternatives $X_1, \ldots, X_k$, the total sampling budget N, and the first-stage sample size n_0, where $n_0 \geq 2$.

Stage 1

For all $i = 1, \ldots, k$, take n_0 independent observations $X_{i,1}, \ldots, X_{i,n_0}$ from alternative i, compute $\bar{X}_i(n_0) = \sum_{j=1}^{n_0} X_{i,j}/n_0$ and $\hat{S}_i^2(n_0) = \sum_{j=1}^{n_0}(X_{i,j} - \bar{X}_i(n_0))^2/(n_0-1)$, then let $\hat{U}_i(n_0) = \bar{X}_i(n_0) + \sqrt{\hat{S}_i^2(n_0)/n_0}$, set $n_i = n_0$.

Stage 2

while $\sum_{i=1}^k n_i < N$ **do**

  Let $b = \arg\max_{i=1,\ldots,k} \hat{U}_i(n_i)$ and take one observation x_{b,n_b+1} from alternative b;

  Update $\bar{X}_b(n_b+1) = \frac{1}{n_b+1}[n_b \bar{X}_b(n_b) + x_{b,n_b+1}]$, $\hat{S}_b^2(n_b+1) = \sum_{j=1}^{n_b+1}(X_{b,j} - \bar{X}_b(n_b+1))^2/n_b$ and $\hat{U}_b(n_b+1) = \bar{X}_b(n_b+1) + \sqrt{\hat{S}_i^2(n_b+1)/(n_b+1)}$, and let $n_b = n_b+1$;

end while

Select $\arg\max_{i\in\{1,\ldots,k\}} \bar{X}_i(n_i)$ as the best.

3 ANALYSIS OF THE ENHANCED UPPER CONFIDENCE BOUND PROCEDURE

In this section, we will show that the EUCB procedure is sample optimal, and all the proofs of the following lemmas and the theorem are included in Huang et al. (2025). First, we make the following assumption, which is usually used in the R&S procedure.

Assumption 1. *There exist constants $\delta > 0$ and $\bar{\sigma}^2 > 0$ such that $\mu_1 - \max_{i=2,\ldots,k} \mu_i \geq \delta$ and $\max_{i=1,\ldots,k} \sigma_i^2 \leq \bar{\sigma}^2$ even though $k \to \infty$.*

In Assumption 1, the constant δ is only required to exist, and its precise value may be challenging for users to specify. This contrasts with the traditional indifference-zone formulation, where δ is presumed to be known. Therefore, in the R&S literature, the assumption regarding the constant δ is a relatively weak condition. The presence of $\bar{\sigma}^2$ is to prevent the variance from becoming infinite as $k \to \infty$. And we prove the sample optimality of EUCB procedure in the following.

Let $Z_1, Z_2, \ldots$ be a sequence of normal random variables with mean μ and variance σ^2, $\bar{Z}(n) = \sum_{l=1}^n Z_l/n$ be the sample average and $S^2(n) = \sum_{l=1}^n (Z_l - \bar{Z}(n))^2/(n-1)$ be the sample variance of the first n observations.

Denote $\{\bar{Z}(n), n = 1,2,\dots\}$ as the running-average process and $\{U(n) = \bar{Z}(n) + \sqrt{S^2(n)/n}, n = 1,2,\dots\}$ as the running-average-upper-bound process. Besides, we let $S^2(1) = 0$. We will adapt two lemmas in Li et al. (2025) to prove the boundary-crossing time of the running-average process.

Lemma 1. *(Li et al. 2025) The running-average process $\{\bar{Z}(n), n = 1,2,\dots\}$ reaches its minimum in a finite number of observations almost surely, that is,* $\Pr\{\arg\min_{n\geq 1} \bar{Z}(n) < \infty\} = 1$.

We will establish the sample optimality using the boundary-crossing time. To simplify the notation, we let $\tilde{N}(x) = \inf\{n : \bar{Z}(n) < \mu + x\}$ and $\tilde{N}_U(x) = \inf\{n : U(n) < \mu + x\}$.

Lemma 2. *(Li et al. 2025) For any $x > 0$, $\Pr\{\tilde{N}(x) < \infty\} = 1$ and*

$$\mathbb{E}\left[\tilde{N}(x)\right] = \exp\left(\sum_{n=1}^{\infty} \frac{1}{n}\Phi\left(-\frac{\sqrt{n}x}{\sigma}\right)\right),$$

and furthermore, $\mathbb{E}[\tilde{N}(x)]$ is continuous and strictly decreasing in $x \in (0,\infty)$, where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

For any $x < 0$, $\mathbb{E}[\tilde{N}(x)] = \infty$ and

$$\Pr\left\{\tilde{N}(x) < \infty\right\} = 1 - \exp\left(-\sum_{n=1}^{\infty} \frac{1}{n}\Phi\left(\frac{\sqrt{n}x}{\sigma}\right)\right), \tag{2}$$

and furthermore, $\Pr\{\tilde{N}(x) < \infty\}$ is continuous and strictly increasing for $x \in (-\infty, 0)$.

When $x = 0$, $\Pr\{\tilde{N}(x) < \infty\} = 1$ and $\mathbb{E}[\tilde{N}(x)] = \infty$.

Lemma 1 provides the guarantee that the running-average process will reach its minimum in a finite number of observations almost surely and Lemma 2 provides the boundary-crossing time of the running-average-upper-bound process. By integrating these two lemmas, we derive the following boundary-crossing probability for the minimum of the running-average process.

Lemma 3. *For any negative number x,*

$$\Pr\left\{\min_{n\geq 1} \bar{Z}(n) > \mu + x\right\} \geq \exp\left(-\sum_{n=1}^{\infty} \frac{1}{n}\Phi\left(\frac{\sqrt{n}x}{\sigma}\right)\right). \tag{3}$$

After getting the boundary-crossing probability, we relay the Chernoff bound of normal random variable and chi-squared random variable in Lemma 4 to get an the boundary-crossing time of running-average-upper-bound process in Lemma 5.

Lemma 4 (Chernoff Bound). *Let $X_1, X_2, \dots, X_n$ be a sequence of independent random variables, and let Y be a single random variable. The Chernoff bound provides the following tail probability bounds for specific distributions:*

(a) *If $X_1, X_2, \dots, X_n$ are normal random variables with mean μ and variance σ^2, and $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is their sample average, then for any $t > 0$,*

$$\Pr\{\bar{X} > \mu + t\} \leq \exp\left(-\frac{nt^2}{2\sigma^2}\right). \tag{4}$$

(b) *If Y follows a chi-squared distribution $Y \sim \chi^2_k$, where k is the degrees of freedom, then for any $\varepsilon > 0$,*

$$\Pr\{Y > (1+\varepsilon)k\} \leq \left((1+\varepsilon)e^{-\varepsilon}\right)^{k/2}. \tag{5}$$

Lemma 5. *Let $X_1, X_2, \dots, X_n$ be a sequence of normal random variables with mean μ and variance σ^2, let $\bar{X}$ be the sample average and $S^2 = \sum_{i=1}^{n}(X_i - \bar{X})^2/(n-1)$ be the sample variance of these n random variables. For any $t > 0$,when $n \geq \max\{5, \frac{16\sigma^2}{t^2}\}$, we have:*

$$\Pr\left\{\bar{X} + \sqrt{S^2/n} > \mu + t\right\} \leq 2\exp\left(-\frac{nt^2}{8\sigma^2}\right). \tag{6}$$

Building on Lemma 5 and some theorems in counting processes, we establish Lemma 6 as follows.

Lemma 6. *For any* $t > 0$, $\Pr\{\tilde{N}_U(t) < \infty\} = 1$ *and*

$$\mathbb{E}\left[\tilde{N}_U(t)\right] \leq \left\lceil \frac{16\sigma^2}{t^2} \right\rceil + \frac{2e^{t^2/8\sigma^2}}{e^{t^2/8\sigma^2} - 1} + 4. \tag{7}$$

Getting the above lemmas, we can prove the sample optimality of the EUCB procedure. We let $\tilde{N}_U(t;n_0) = \inf\{n \geq n_0 : U(n) < \mu + t\}$. We will introduce the relationship between $\tilde{N}_U(t)$ and $\tilde{N}_U(x;n_0)$, it is easy to know that $\tilde{N}_U(t;n_0) = \tilde{N}_U(t)$ when $n_0 = 1$, and for any $n_0 \geq 1$, we have

$$\mathbb{E}\left[\tilde{N}_U(t;n_0)\right] \leq \left\lceil \frac{16\sigma^2}{t^2} \right\rceil + \frac{2e^{t^2/8\sigma^2}}{e^{t^2/8\sigma^2} - 1} + n_0 + 4$$

Theorem 1 Suppose that Assumption 1 holds, if the total sampling budget N satisfies $N = ck$ and $c \geq \lceil \frac{16\bar{\sigma}^2}{t^2} \rceil + \frac{2e^{t^2/8\bar{\sigma}^2}}{e^{t^2/8\bar{\sigma}^2}-1} + n_0 + 4$. The PCS of EUCB procedure satisfies:

$$\liminf_{k\to\infty} \text{PCS} \geq \Pr\left\{\min_{n\geq 1} \bar{Z}(n) > \mu_1 - \delta_0\right\} = \exp\left(-\sum_{n=1}^{\infty} \frac{1}{n}\Phi\left(\frac{\sqrt{n}\delta_0}{\sigma}\right)\right) > 0,$$

where δ_0 is a positive constant satisfying $0 < \delta_0 < \delta$.

Theorem 1 demonstrates that EUCB procedure achieves sample optimality provided that the total sampling budget grows linearly with some constant factor.

4 EXPERIMENTS

In this section, we conduct some experiments to examine our theoretical results and test the performance of the proposed procedures. In Section 4.1, we show the sample optimality of the EUCB procedure. Besides, comparison with other procedures demonstrate the performance superiority of the EUCB procedure. In Section 4.2, we show the consistency of EUCB Procedure. Furthermore, we show the impact of the rate of samples in the first stage on the procedure's performance. Lastly, in Section 4.3, we test the sample optimality of EOC for existing large-scale fixed-budget procedures.

In this section, we adopt the same experimental setups as Li et al. (2025), including the four problem configurations.

- The slippage configuration of means with a common variance (SC-CV) under which

$$\mu_1 = 0.1,\ \mu_i = 0,\ i = 2,\ldots,k \text{ and } \sigma_i^2 = 1,\ i = 1,\ldots,k.$$

- The configuration with equally spaced means and a common variance (EM-CV) under which

$$\mu_1 = 0.1,\ \mu_i = -(i-1)/k,\ i = 2,\ldots,k \text{ and } \sigma_i^2 = 1,\ i = 1,\ldots,k.$$

- The configuration with equally spaced means and increasing variances (EM-IV) under which

$$\mu_1 = 0.1,\ \mu_i = -(i-1)/k,\ i = 2,\ldots,k \text{ and } \sigma_i^2 = 1 + (i-1)/k,\ i = 1,\ldots,k.$$

- The configuration with equally spaced means and decreasing variances (EM-DV) under which

$$\mu_1 = 0.1,\ \mu_i = -(i-1)/k,\ i = 2,\ldots,k \text{ and } \sigma_i^2 = 2 - (i-1)/k,\ i = 1,\ldots,k.$$

In all four configurations, Alternative 1 is always the best alternative. In order to capture the sample optimality of R&S procedures, we set the number of alternatives as $k = 2^l$ with l ranging from 2 to 16, and for each k, we set the total sampling budget as $N = ck$, where c is a specific constant. In addition, we evaluate the performance of each procedure on a particular R&S problem based on 1,000 independent macro replications. We compare the EUCB procedure with the OCBA (Chen et al. 2000), ROCBA (Chen and Lee 2010), EFG (Li et al. 2025), FBKT, and FBKT$^+$ (Hong et al. 2022) procedures. Unless otherwise specified, we adopt the following parameter configurations in experiments. For the EUCB procedure, we allocate 80% of the total sample budget to the initial phase, setting $n_0 = 0.8c$, to estimate both the sample mean and sample variance. The remaining 20% of the budget is then distributed sequentially, guided by the maximum upper confidence bound criterion. In the ROCBA procedure, we assign 20 samples per round according to the OCBA rule, continuing this process until the entire sample budget is depleted. For the EFG procedure, we follow the original configuration from Li et al. (2025), dedicating 80% of the total sample budget to uniform exploration and the remaining 20% to the greedy phase. The FBKT procedure adheres to the settings outlined in its original paper, with the allocation rule parameter ϕ fixed at 3. Similarly, for the FBKT$^+$ procedure, 9% of the total sample budget is evenly distributed across all alternatives for initial estimation, while the remaining parameters are consistent with those of the standard FBKT procedure.

4.1 The Sample Optimality of EUCB procedure

We implement the six fixed-budget R&S procedures, setting the total budget $N = 100k$ across all four configurations and show the results in Figure 1. The experimental results validate the PCS sample optimality of the EUCB procedure while also confirming the efficacy of its variance-aware extension.

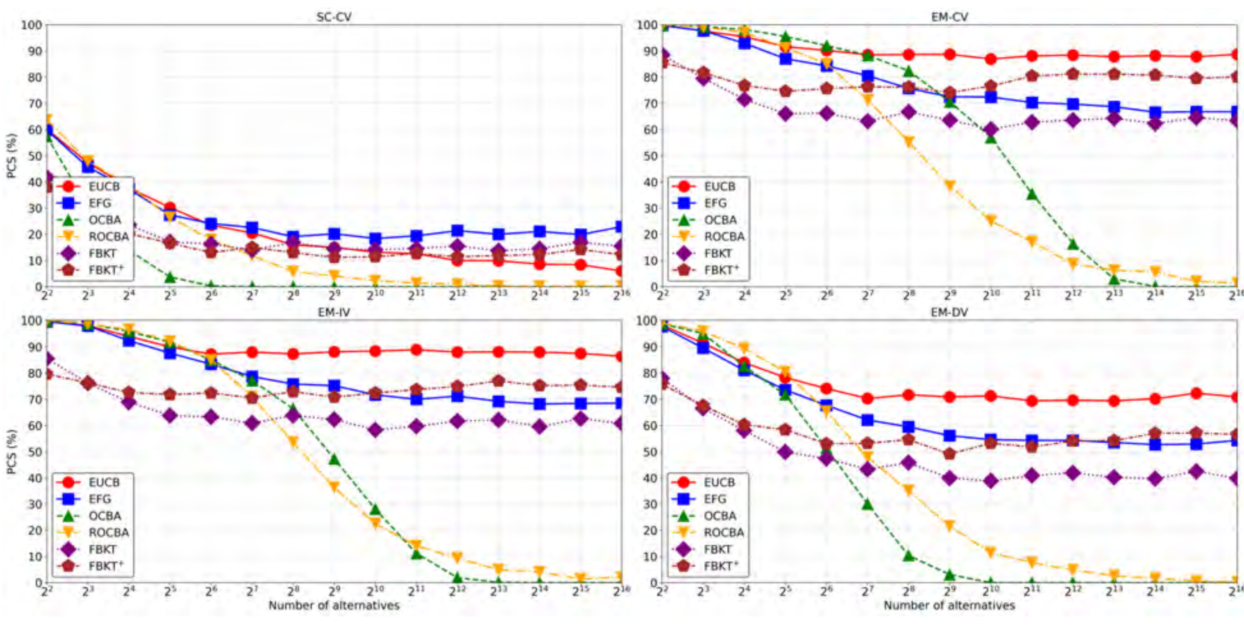

Figure 1: A comparison between EUCB, OCBA, ROCBA, EFG, FBKT, and FBKT$^+$ procedures.

From the results depicted in Figure 1, we observe that the estimated PCS of the EUCB procedure remains above non-zero levels despite exponential increase in the number of alternatives (k ranging from 2^2 to 2^{16}) across all configurations, demonstrating its sample optimality. Besides, under the EM-CV, EM-DV and EM-IV settings, as k increases, the PCS of the EUCB procedure consistently surpasses the PCS of the other five procedures by more than 10%.

The underlying insights derived from these results are straightforward. In the EUCB procedure, we address the uncertainty of sample mean by using its upper confidence bound as the criterion to allocate the simulation budget. This allocation strategy improves the overall accuracy of estimates in the set of alternatives. However, in the SC-CV configuration, when the total simulation budget is $N = 100k$, each alternative is allocated an average of 100 samples. Given a performance gap of 0.1 between the optimal and suboptimal alternatives and a uniform variance of 1 across all alternatives, the sample standard deviation of the mean estimates becomes the dominant factor influencing budget allocation. Consequently, this leads to the the inferior performance of the EUCB procedure compared to the EFG procedure. As the simulation budget increases, however, the impact of the sample standard deviation progressively diminishes, enabling the inherent algorithmic strengths of EUCB to become more pronounced.

Additionally, we observe a counterintuitive yet rational phenomenon: in the SC-CV configuration, the FBKT procedure consistently outperforms the FBKT$^+$ procedure. This discrepancy arises because the FBKT$^+$ procedure is designed to reserve a portion of the budget for estimating "seed candidates" and thereby prevent superior alternatives from being prematurely eliminated during early-stage encounters. However, in the SC-CV configuration, where all non-optimal alternatives are identical, such budget reservations become unnecessary. Consequently, the performance loss incurred by FBKT$^+$ becomes unjustifiable under this specific setting.

4.2 Additional Properties comparison between the EFG Procedure and EUCB Procedure

4.2.1 The consistency of EUCB Procedure

In this subsection, we verify the consistency of the EUCB procedure and the EFG procedure in Figure 2. We consider a fixed number of alternatives $k = 2^{12} = 4096$ and we vary the value of c from 100 to 900 for the total sampling budget $N = ck$ to conduct this experiment.

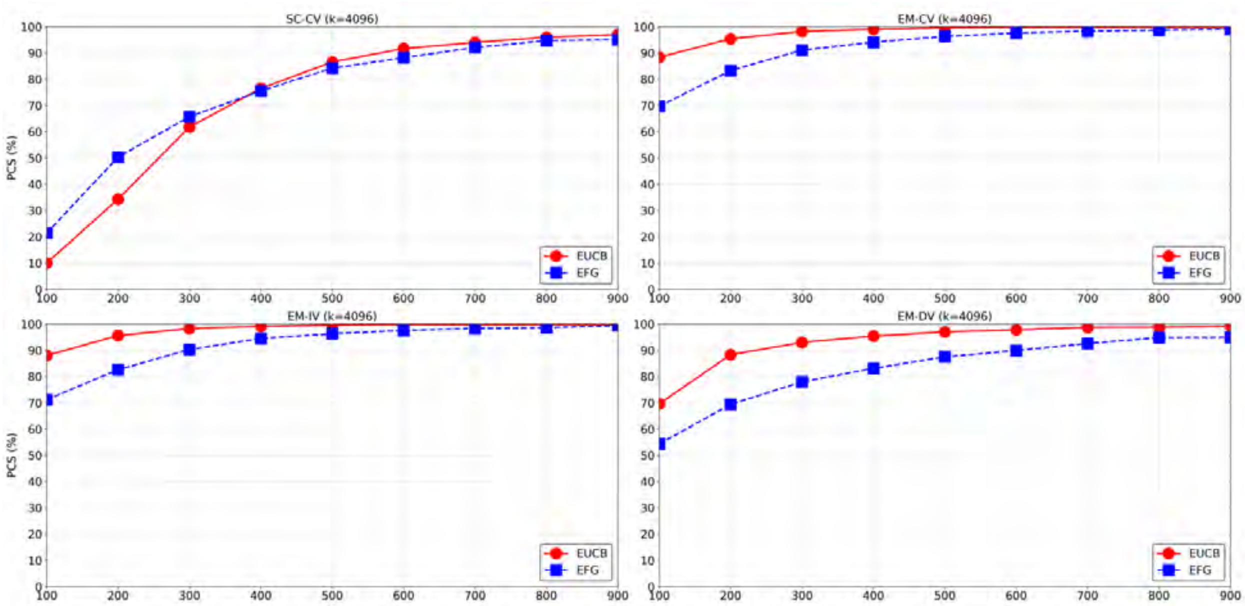

Figure 2: Estimated PCS of the EUCB procedure and the EFG procedure for different c.

The experiment yields the following findings. First, both the EFG procedure and the EUCB procedure are consistent. As shown in Figure 2, under each configuration, as the total sampling budget increases from $100k$ to $900k$, the PCS of both the EFG procedure and the EUCB procedure rise to nearly 100%. Second, we find that under the SC-CV configuration, although the PCS of the EUCB procedure is lower than that of the EFG procedure when N is small, it surpasses the PCS of the EFG procedure as N increases. This experimental result further validates the explanation presented in Figure 1 regarding the inferior performance of the EUCB procedure compared to EFG under SC-CV configuration. In this setting, the PCS of the EUCB procedure converges to 100% more rapidly than that of the EFG procedure.

4.2.2 Budget allocation between the first stage and second stage

To understand the impact of budget allocation between the first and second stages on the PCS of both the EUCB and EFG procedures, we use the proportion of samples allocated in the first stage relative to the total sample size, defined as $p = n_0/c$, to measure the budget allocation. In the following experiment, we let $k = 2^{12} = 4096$, $N = 200k$, and vary the value of p from 0.05 to 1. Then, we plot the estimated PCS against different p for four different settings in Figure 3.

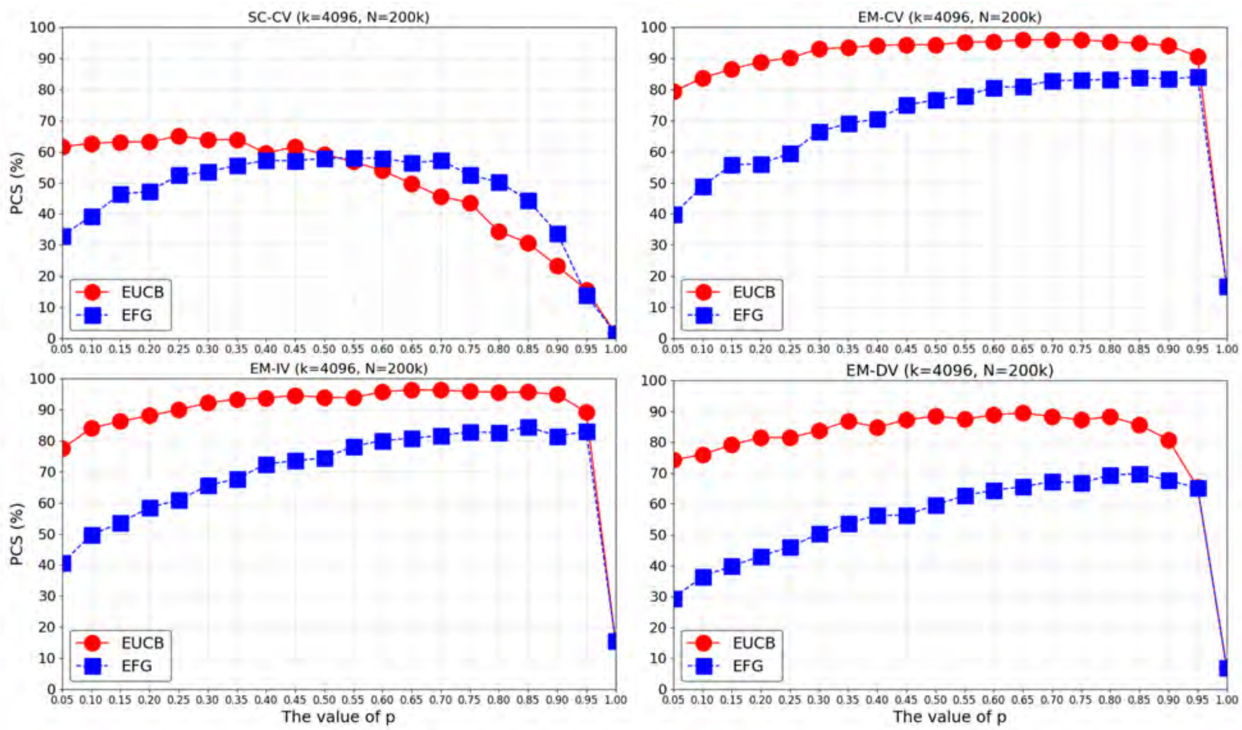

Figure 3: Estimated PCS of EUCB procedure and EFG procedure for different $p = n_0/c$.

From Figure 3, we get some findings regarding the PCS. Beyond the SC-CV configuration, the EUCB procedure demonstrates superior performance to the EFG procedure at any identical p-value across other configurations, further substantiating its advantages in R&S problems. Moreover, under other configurations, the EUCB procedure exhibits enhanced robustness as the p-value varies, demonstrating superior performance stability compared to EFG procedure. Besides, a small value of p seems a good selection for EUCB procedure since the EUCB procedure with a sufficiently small p achieves comparable performance to the optimal performance of the EFG procedure.

4.3 The sample optimality of EOC

Unlike the PCS, which employs the probability of correctly selecting the best alternative as the evaluation criterion, EOC adopts the expected difference between the true mean of the ultimately selected alternative and that of the actual best alternative as its assessment standard. This approach bears conceptual similarity to the notion of regret in multi-armed bandit problems. In this paper, we define the EOC of procedure π as follows:

$$\mathrm{EOC}_\pi = \mathbb{E}[\mu_1 - \mu_{\hat{i}_\pi}].$$

While the experiments in Section 4.1 have demonstrated the PCS sample optimality of the EUCB procedure, we further investigate its EOC sample optimality. In this subsection, we conduct empirical evaluations of the EOC sample optimality for four procedures previously identified as PCS sample-optimal. The experimental results are illustrated in Figure 4.

As demonstrated in Figure 4, the R&S procedures with PCS sample optimality also exhibit sample optimality under the EOC metric, closely aligning with the PCS results presented in the previous subsection. Moreover, the EUCB procedure maintains a significant performance advantage over other sample-optimal procedures.

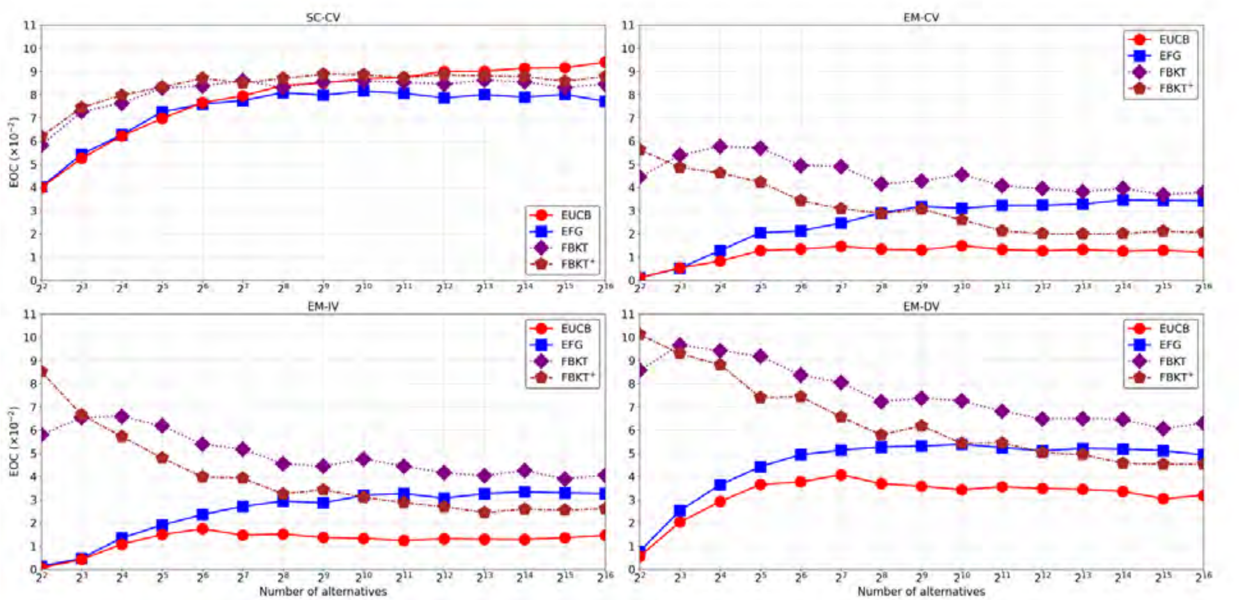

Figure 4: A EOC comparison between EUCB, EFG, FBKT, and FBKT$^+$ procedures.

5 CONCLUSION

In this paper, we introduce the EUCB procedure to address the challenges of large-scale R&S problems by incorporating variance information into the dynamic allocation of simulation budget. We establish the sample optimality of EUCB through a boundary-crossing perspective. Compared to the EFG procedure, the EUCB procedure explicitly accounts for the variance of alternatives, which only increases the computation of the sample sum of squares, leading to a significant improvement in the PCS as demonstrated in our numerical experiments. Furthermore, simulation results show that EUCB achieves sample optimality not only in terms of PCS but also with respect to EOC. The numerical experiment also suggests that EUCB is consistent and exhibits robustness to the proportion of budget allocated during the initial exploration phase.

There are some interesting directions for future research. One promising direction is to extend the analysis to the Probability of Good Selection (PGS), which could provide deeper insights into the procedure's performance in scenarios where selecting near-optimal alternatives is acceptable. Additionally, while this paper establishes sample optimality with respect to PCS, a formal proof of sample optimality for the EOC metric remains an open question that merits further investigation. Besides, we will test the robustness and efficiency of the EUCB Procedure in real-world scenarios.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China [Grants 72293562, 72121001, 72101047, and 72171060].

REFERENCES

Bechhofer, R. E. 1954. "A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances". *The Annals of Mathematical Statistics* 25(1):16–39.

Chen, C.-H., S. E. Chick, L. H. Lee, and N. A. Pujowidianto. 2015. "Ranking and selection: Efficient simulation budget allocation". In *Chapter 3 in Handbook of simulation optimization*, edited by M. C. Fu, 45–80. New York: Springer.

Chen, C.-H., and L. H. Lee. 2010. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*, Volume 1. Singapore: World Scientific.

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 K. Inoue. 2001. "New two-stage and sequential procedures for selecting the best simulated system". *Operations Research* 49(5):732–743.

Fan, W., L. J. Hong, G. Jiang, and J. Luo. 2025. "Review of Large-Scale Simulation Optimization". *Journal of the Operations Research Society of China* Forthcoming.

Farahani, R. Z., S. Rezapour, T. Drezner, and S. Fallah. 2014. "Competitive supply chain network design: An overview of classifications, models, solution techniques and applications". *Omega* 45:92–118.

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.

Hong, L. J. 2006. "Fully sequential indifference-zone selection procedures with variance-dependent sampling". *Naval Research Logistics* 53(5):464–476.

Hong, L. J., W. Fan, and J. Luo. 2021. "Review on ranking and selection: A new perspective". *Frontiers of Engineering Management* 8(3):321–343.

Hong, L. J., G. Jiang, and Y. Zhong. 2022. "Solving Large-Scale Fixed-Budget Ranking and Selection Problems". *INFORMS Journal on Computing* 34(6):2930–2949.

Huang, S., C. Li, G. Jiang, and Y. Zhong. 2025. "Select the Best System with Variance Information for Large-Scale Ranking and Selection". Working papaer.

Kim, S.-H., and B. L. Nelson. 2001. "A fully sequential procedure for indifference-zone selection in simulation". *ACM 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*, edited by S. G. Henderson and B. L. Nelson, Volume 13, Chapter 17, 501–534. Amsterdam: North-Holland.

Li, Z., W. Fan, and L. J. Hong. 2025. "The (Surprising) Sample Optimality of Greedy Procedures for Large-Scale Ranking and Selection". *Management Science* 71(2):1238–1259.

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.

Luo, X., L. Li, L. Zhao, and J. Lin. 2022. "Dynamic intra-cell repositioning in free-floating bike-sharing systems using approximate dynamic programming". *Transportation Science* 56(4):799–826.

Ni, E. C., D. F. Ciocan, S. G. Henderson, and S. R. Hunter. 2017. "Efficient ranking and selection in parallel computing environments". *Operations Research* 65(3):821–836.

Paulson, E. 1964. "A sequential procedure for selecting the population with the largest mean from k normal populations". *The Annals of Mathematical Statistics* 35(1):174–180.

Pei, L., B. L. Nelson, and S. R. Hunter. 2024. "Parallel adaptive survivor selection". *Operations Research* 72(1):336–354.

Rinott, Y. 1978. "On two-stage selection procedures and related probability-inequalities". *Communications in Statistics-Theory and methods* 7(8):799–811.

Wu, D., and E. Zhou. 2018a. "Analyzing and Provably Improving Fixed Budget Ranking and Selection Algorithms". *arXiv preprint* arXiv:1811.12183:all.

Wu, D., and E. Zhou. 2018b. "Provably Improving the Optimal Computing Budget Allocation Algorithm". In *2018 Winter Simulation Conference (WSC)*, 1921–1932 https://doi.org/10.1109/WSC.2018.8632199.

Zhong, Y., and L. J. Hong. 2022. "Knockout-tournament procedures for large-scale ranking and selection in parallel computing environments". *Operations Research* 70(1):432–453.

Zhong, Y., S. Liu, J. Luo, and L. J. Hong. 2022. "Speeding up Paulson's procedure for large-scale problems using parallel computing". *INFORMS Journal on Computing* 34(1):586–606.

AUTHOR BIOGRAPHIES

SONG HUANG is a Ph.D student in the School of Management at Harbin Institute of Technology, His research interests include simulation optimization and machine learning. His email address is 23b910018@stu.hit.edu.cn.

GUANGXIN JIANG is a professor in the School of Management at Harbin Institute of Technology. He earned his Ph.D. in Applied Mathematics from Tongji University. His research interests include stochastic models and simulation, machine learning, financial engineering and risk management, FinTech, etc. His email address is gxjiang@hit.edu.cn. His website is http://homepage.hit.edu.cn/jiangguangxin/.

CHENXI LI is an undergraduate student in the School of Management at Harbin Institute of Technology. He is expected to earn his bachelor's degree in June 2025. His research interests include simulation optimization, reinforcement learning, and chronic disease management. His email address is 2021113331@stu.hit.edu.cn.

YING ZHONG is an associate professor in the School of Management and Economics at University of Electronic Science and Technology of China. He earned his Ph.D. from the Department of Management Sciences at City University of Hong Kong. His research interests include large-scale stochastic simulation optimization, reinforcement learning, and parallel computing. His email address is yzhong4@uestc.edu.cn. His website is https://www.mgmt.uestc.edu.cn/szxq.jsp?urltype=news.NewsContentUrl&wbtreeid=1097&wbnewsid=3243.