Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

PROVABLY IMPROVING THE OPTIMAL COMPUTING BUDGET ALLOCATION ALGORITHM

Di Wu
Enlu Zhou

School of Industrial & Systems Engineering
Georgia Institute of Technology
755 Ferst Drive NW
Atlanta, GA 30332, USA

ABSTRACT

We boost the performance of the Optimal Computing Budget Allocation (OCBA) algorithm, a widely used and studied algorithm for Ranking and Selection (as known as Best Arm Identification) under a fixed budget. The proposed fully sequential algorithms, OCBA+ and OCBAR, are shown to have better performance both theoretically and numerically. Surprisingly, we reveal that in a two-design setting, a constant initial sample size in a family of OCBA-type algorithms (including the original OCBA) only amounts to a sub-exponential or even polynomial convergence rate of the probability of false selection (PFS). In contrast, our algorithms are guaranteed to converge exponentially fast, as is shown by a finite-sample bound on the PFS.

1 INTRODUCTION

In simulation optimization, the study of Ranking and Selection (R&S) mainly focuses on how to efficiently run simulations to identify the best design among a finite number of candidates (see, e.g., Chapter 17 in Henderson and Nelson 2006). R&S has two major formulations. The fixed confidence setting challenges us to achieve certain confidence level using the least possible simulation effort. The fixed budget setting, on the other hand, requires to maximize the probability of selecting the best design using a fixed budget of simulation runs. Our focus in this paper is on the fixed budget R&S. For fixed confidence, KN, BIZ and many other efficient procedures have been proposed in the literature and we refer the reader to Kim and Nelson 2001; Frazier 2014 for details and the references therein.

In fixed budget R&S, the Optimal Computing Budget Allocation (OCBA) algorithm in Chen et al. 2000 is considered as the one of the the most widely used algorithms. The static allocation ratios suggested by OCBA has been shown to be asymptotically optimal from a large deviations perspective (see Glynn and Juneja 2004). Meanwhile, its implementable version highlights a sequential style of budget allocation, which is critical to its finite-sample performance (e.g., Chen et al. 2006). The framework of OCBA has been extended to handle many applications (e.g., Lee et al. 2004; Chen and Lee 2010). Similar sequential allocation style has also been explored in subsequent works (e.g., Pasupathy et al. 2015). However, to the best of our knowledge, although the asymptotic properties of simple R&S procedures have been investigated (e.g., Jacobovic and Zuk 2017), the formal analysis and characterization of sequential OCBA's performance remains an open research problem.

In view of the current void in OCBA's theoretical performance guarantees, we are motivated to push the boundary by performing in-depth analysis on its behavior, and developing insights for improving its performance. Our contributions are outlined as follows.

1. We propose two fully sequential algorithms, OCBA+ and OCBAR, and provide a finite-sample bound on its probability of correct selection (PCS), guaranteeing an exponential convergence rate.

978-1-5386-6572-5/18/$31.00 ©2018 IEEE

2. It is revealed that for a two-design case, if OCBA, OCBA+ and OCBAR choose a constant initial sample size, then the PCS converges only at a *sub-exponential* (or even *polynomial*) rate.
3. Numerical experiments are conducted to show that both OCBA+ and OCBAR can achieve higher PCS than the original OCBA under the same budget.

2 PROBLEM SETUP

We briefly review the setup for the fixed budget R&S problem and lay down some notations. Given a set of K designs $\mathscr{I} = \{1, 2, \ldots, K\}$, the goal is to find (without loss of generality) the one with the highest expected performance. Each design's expected performance is unknown, and is typically evaluated through multiple simulation runs and approximated by the sample mean

$$\bar{X}_{i,N_i} = \frac{1}{N_i} \sum_{r=1}^{N_i} X_{ir},$$

where X_{ir} is a random variable representing design i's rth simulation output, and N_i is the number of times design i has been simulated. We will drop the subscript N_i when there is no ambiguity. The true and the observed best designs are denoted by

$$b = \arg\max_{i \in \mathscr{I}} \mu_i, \qquad \hat{b} := \arg\max_{i \in \mathscr{I}} \bar{X}_i,$$

respectively. We make the following standard assumptions to avoid technicalities, where $\mathscr{N}$ denotes a normal distribution.

Assumption 1

1. The best design is unique, i.e., b is a singleton.
2. $X_{ir} \sim \mathscr{N}(\mu_i, \sigma_i^2)$ and are independent across all $i \in \mathscr{I}$ and $r = 1, 2, \ldots$.

Then, under a fixed budget T of total simulation runs, the goal is to maximize the probability of correct selection (PCS), i.e.,

$$\text{PCS} := \mathbb{P}\left(\hat{b} = b\right) = \mathbb{P}\left\{\bigcap_{i \neq b} \{\bar{X}_b > \bar{X}_i\}\right\},$$

by carefully allocating the budget. Several algorithms have been proposed for solving this problem, and their performance is typically measured via the following approaches. The first approach is to take a large deviations (LD) perspective. Denoting by PFS the probability of false selection (defined as $1 - \text{PCS}$), it has been shown that many algorithms exhibit the behavior

$$-\lim_{T \to \infty} \frac{1}{T} \log \text{PFS}_{\mathscr{A}}(T, P) = R_{\mathscr{A}}(P),$$

where $\mathscr{A}$ is an algorithm, P is a problem instance, $\text{PFS}_{\mathscr{A}}(T, P)$ is the PFS of $\mathscr{A}$ under budget T and problem P, and $R_{\mathscr{A}}(\cdot) \geq 0$ is called the LD rate function. Asymptotically optimal algorithms can be derived by maximizing $R_{\mathscr{A}}$ (see, e.g., Glynn and Juneja 2004; Glynn and Juneja 2015), but it is an insufficient performance measure since it only focuses on the asymptotic performance. For example, although $Te^{-T}, T^2e^{-T}, T^3e^{-T}, \ldots$ all have the same LD rate, they behave quite differently for small values of T. The second approach is to approximate $\text{PFS}_{\mathscr{A}}(T, P)$ using tight bounds, but it could be very challenging for highly adaptive algorithms such as OCBA (see Section 3.1 for details). Another approach is to plot out the PCS curve and visually inspect how fast it converges to 1 as T increases. The main criticism, however, is that such results are problem-specific and may not represent the general performance very well. In this paper, we will improve OCBA from an LD perspective, where the improvement is substantiated by a finite-sample PFS bound combined with numerical results.

Algorithm 1 OCBA (Chen et al. (2000))

Input: N_0, Δ, T.
Initialization: Simulate each design N_0 times and compute $\bar{X}_i$ and S_i^2. $N_i \leftarrow N_0$. $T' \leftarrow N_0 K$.
while $\sum_{i \in \mathscr{I}} N_i \leq T$ **do**
  Compute $\hat{\alpha}_1, \ldots, \hat{\alpha}_K$ using (1) and plug-in estimates.
  for $i = 1, \ldots, K$ **do**
    Run $\max\{0, \lfloor \hat{\alpha}_i T' \rfloor - N_i\}$ replications for design i.
    $N_i \leftarrow \max\{N_i, \lfloor \hat{\alpha}_i T' \rfloor\}$. Update $\bar{X}_i$ and S_i.
  end for
  $T' \leftarrow T' + \Delta$.
end while
Output: $\hat{b} = \arg\max_{i \in \mathscr{I}} \bar{X}_i$.

3 IMPROVING THE OCBA ALGORITHM

The plan of this section is to first review the original OCBA algorithm, and then present the improved versions for comparison. Our purpose is to help the reader quickly grasp the motivations behind our algorithms, while results involving more technicalities are deferred to Section 4.

3.1 OCBA+ Algorithm

The original OCBA algorithm proposed in Chen et al. 2000 is summarized in Algorithm 1, and it is built on two key components. One is the inverse signal-to-noise ratio allocation rule,

$$\alpha_b = \sigma_b \sqrt{\sum_{i \neq b} \frac{\alpha_i^2}{\sigma_i^2}}, \qquad \frac{\alpha_i}{\alpha_j} = \frac{\sigma_i^2/\delta_{bi}^2}{\sigma_j^2/\delta_{bj}^2}, \quad i \neq j \neq b, \tag{1}$$

where $\alpha_i := N_i/T$ and $\delta_{bi} := \mu_b - \mu_i$. The ratios in (1) are derived through asymptotically maximizing an approximated PCS as $T \to \infty$. In practice, the μ_i's and σ_i's are unknown and thus are replaced by their estimates $\bar{X}_i$ and S_i. We let $\hat{\alpha}_i$ denote the estimated α_i by plugging in μ_i's and σ_i's in (1). This leads to the other more important component (see Chen et al. 2006 on its efficiency): sequential allocation, where at each iteration OCBA increases the available budget T' by Δ, updates the estimates timely, and matches the N_i's with the recomputed allocation ratios to the greatest possible extent. This adaptive scheme makes OCBA particularly robust against estimation noise. However, as we will show in Section 4.1, a major drawback of a constant N_0 is that it only leads to a polynomial convergence rate of PCS.

Algorithm 2 OCBA+

Input: $\alpha_0 \in (0,1), T$.
Initialization: Simulate each design $N_0 = \lfloor \alpha_0 T/K \rfloor$ times and compute $\bar{X}_i$ and S_i^2. $N_i \leftarrow N_0$.
while $\sum_{i \in \mathscr{I}} N_i \leq T$ **do**
  Compute $\hat{\alpha}_1, \ldots, \hat{\alpha}_K$ using (1) and plug-in estimates.
  Simulate design $l = \arg\max_i \hat{\alpha}_i / N_i$ once.
  $N_l \leftarrow N_l + 1$. Update $\bar{X}_l$ and S_l.
end while
Output: $\hat{b} = \arg\max_{i \in \mathscr{I}} \bar{X}_i$.

To fix this, we propose the OCBA+ algorithm, which is presented in Algorithm 2. We motivate OCBA+ as follows. The ratio α_0 is a constant between 0 and 1, which forces N_0 to grow proportionally with T and guarantees that the PFS converges exponentially fast (see Section 4.3). In addition, similar to the "most starving" version in Chen and Lee 2010, we get rid of Δ and modify Algorithm 1 into a fully sequential one: at each step we compute the ratio $\hat{\alpha}_i/N_i$ as a measure of how much design i's allocated budget deviates from the target ratios. Then, the next simulation replication goes to the design with the largest such ratio, since it is the least sampled (with respect to the ratios) and hence needs simulation the most. This way, we try to explicitly match that the ratios N_i/T's with those specified in (1). As we shall see in Sections 4.1 and 5, the fully sequential feature of OCBA+ not only facilitates theoretical analysis, but also improves OCBA's numerical performance.

3.2 OCBAR Algorithm

In addition to OCBA+, we propose another fully sequential algorithm, OCBAR ("R" meaning "Randomized"), which is summarized in Algorithm 3. Notice that OCBAR also requires N_0 to grow proportionally with T. This is because a constant N_0 will again result in a sub-exponential convergence rate of the PCS, as we shall see in Section 4.2. The only difference between OCBAR and OCBA+ is in step 5, where the next design to simulate is randomly sampled by using the ratios $(\hat{\alpha}_1,\dots,\hat{\alpha}_K)$ as a probability distribution. The rationale behind the randomized sampling strategy is twofold:

1. Similar to many randomized algorithms, the randomness in OCBAR is introduced for "exploring" the design space, especially when we are not confident about the designs' performance due to estimation error in early stages.
2. Using $(\hat{\alpha}_1,\dots,\hat{\alpha}_K)$ as a sampling distribution still achieves the optimal allocation ratios in (1) asymptotically as $T\to\infty$.

Algorithm 3 OCBAR

1: **Input**: $\alpha_0\in(0,1),T$.
2: **Initialization:** Simulate each design $N_0=\lfloor\alpha_0 T/K\rfloor$ times and compute $\bar{X}_i$ and S_i^2. $N_i\leftarrow N_0$.
3: **while** $\sum_{i\in\mathscr{I}}N_i\le T$ **do**
4:     Compute $\hat{\alpha}_1,\dots,\hat{\alpha}_K$ using (1) and plug-in estimates.
5:     Independently draw a design l from the distribution specified by $\hat{\alpha}_1,\dots,\hat{\alpha}_K$.
6:     Simulate design l once. $N_l\leftarrow N_l+1$. Update $\bar{X}_l$ and S_l.
7: **end while**
8: **Output:** $\hat{b}=\arg\max_{i\in\mathscr{I}}\bar{X}_i$.

4 CONVERGENCE ANALYSIS

In this section, we focus on a two-design case and reveal that if the initial sample size N_0 is chosen as a constant (independent of T), then OCBA, OCBA+ and OCBAR can only achieve a sub-exponential convergence rate of PCS, essentially due to the error in variance estimation. In particular, the convergence rate of OCBA and OCBA+ with a constant initial sample size is at most polynomial. Furthermore, we show that OCBA+ and OCBAR can achieve an exponential convergence rate. Although it is known that any static allocation achieves exponential convergence rate (see Glynn and Juneja 2004), our results do not imply adaptive algorithms such as OCBA, which only achieve sub-exponential convergence rate, are inferior than static allocations. Instead, abundant empirical results in the literature have shown that adaptive algorithms usually work better. The reason is, exactly as we have mentioned at the end of Section 2, that the convergence rate is only asymptotic and thus has limited implication on finite-sample performance. However, within the same class of OCBA-type algorithms, we show that OCBA+ and OCBAR are improved versions of the original OCBA in terms of both asymptotic convergence rate theoretically (in Section 4) and finite-sample performance empirically (in Section 5).

4.1 Constant N_0 in OCBA+ and OCBA

To keep the main idea uncluttered, we study the special case of $K=2$ and analyze the convergence rate for OCBA+. Note that in the two-design case, the allocation rule in (1) reduces to $\alpha_1/\alpha_2=\sigma_1/\sigma_2$. Later we will discuss the possibility of extending our result to the original OCBA with $K\geq 3$.

Theorem 1 Let Assumption 1 hold, and suppose that $K=2$ and $\mu_1>\mu_2$. If OCBA+ chooses a constant $N_0\geq 2$, then there exists some constant $C>0$ such that

$$\mathrm{PFS}(T)\geq \frac{C}{(T-N_0)^{N_0-1}},\quad \forall T\geq 2N_0. \tag{2}$$

Theorem 1 is somewhat surprising since it is sharply different from the common practice of OCBA using a constant initial sample size. As an extreme case, for $N_0=2$ the PFS converges as slowly as $1/T$. Before proving Theorem 1, we first derive a few intermediate results that will come in handy. The key is to characterize the distributions of $\bar{X}_{1,N_1}$ and $\bar{X}_{2,N_2}$ conditional on N_1. Notice that for OCBA+, N_1 and N_2 are influenced by the sequentially updated S_1,S_2,N_1,N_2 and a complicated feedback mechanism. Fortunately, the normal distribution enjoys the following nice property.

Lemma 1 Let $\bar{X}_n$ and S_n^2 be the sample mean and sample variance of n i.i.d. normal random variables, respectively. Then, for all $n\geq 2$, $\bar{X}_n$ is independent of $(S_2^2,S_3^2,\ldots,S_n^2)$.

Proof. For any $2\leq k\leq n$, S_k^2 is a function of the deviations $(\bar{X}_k-X_1,\ldots,\bar{X}_k-X_k)$. Thus, it suffices to show that

$$\bar{X}_n\perp((\bar{X}_2-X_1,\bar{X}_2-X_2),\ldots,(\bar{X}_n-X_1,\ldots,\bar{X}_n-X_n)),$$

where $\perp$ denotes independence and the right hand side (RHS) is denoted by Y_n. Note that $(\bar{X}_n,Y_n)$ is a linear transformation of $(X_1,\ldots,X_n)$ and hence are jointly normal, the result follows from

$$\mathrm{Cov}(\bar{X}_n,\bar{X}_k-X_j)=0,\quad \forall 2\leq k\leq n, j\leq k, \tag{3}$$

which can be verified easily through a direct computation. □

Lemma 1 generalizes a classical result that $\bar{X}_n$ and S_n^2 are independent for normal distribution. Its implication for our context is given by the following corollary.

Corollary 2 For OCBA+ with $K=2$, $(\bar{X}_{1,N_1},\bar{X}_{2,N_2})\mid N_1\sim(Z_1,Z_2)$ almost surely, where Z_1 and Z_2 are independent $\mathcal{N}(\mu_1,\sigma_1^2/N_1)$ and $\mathcal{N}(\mu_2,\sigma_2^2/(T-N_1))$ random variables, respectively.

Proof. For $N_0\leq k\leq T-N_0$, let

$$Y_{1,k}:=(S_{1,N_0},\ldots,S_{1,k}),\quad Y_{2,k}:=(S_{2,N_0},\ldots,S_{2,k}).$$

Note that $N_1=k$ if and only if $(Y_{1,k},Y_{2,T-k})$ falls into some set $A_k(T)$. Furthermore, following the proof of Lemma 1, it can be shown that $(\bar{X}_{1,k},\bar{X}_{2,T-k})\perp(Y_{1,k},Y_{2,T-k})$ since $(X_{1,1},X_{1,2},\ldots)\perp(X_{2,1},X_{2,2},\ldots)$. Thus, for any $N_0\leq k\leq T-N_0$,

$$\begin{aligned}&\mathbb{P}(\bar{X}_{1,N_1}\leq x,\bar{X}_{2,N_2}\leq y\mid N_1=k)\\=&\mathbb{P}(\bar{X}_{1,k}\leq x,\bar{X}_{2,T-k}\leq y\mid(Y_{1,k},Y_{2,T-k}\in A_k(T)))\\=&\mathbb{P}(\bar{X}_{1,k}\leq x,\bar{X}_{2,T-k}\leq y).\end{aligned}$$

Since $\bar{X}_{1,k}\perp\bar{X}_{2,T-k}$, the conclusion follows. □

Corollary 2 guarantees that for OCBA+, conditional on N_1, the joint distribution of the two designs's means are still independent normal random variables. In addition to the above results, we also need the following lemma regarding a tail bound of sample standard deviation.

Lemma 2 Let S_n^2 be the variance estimate of n i.i.d. normal samples with variance σ^2. Then, for any $0 < x < \sigma$,

$$\mathbb{P}(S_n \leq \sigma - x) \leq \exp\left\{-\frac{(n-1)}{4}\left[1-\left(\frac{\sigma-x}{\sigma}\right)^2\right]^2\right\}. \tag{4}$$

Proof. According to Lemma 1 in Laurent and Massart 2000, if $X \sim \chi^2(n)$, then

$$\mathbb{P}(X - n \leq -2\sqrt{nx}) \leq e^{-x}, \quad \forall x > 0,$$

Since $(n-1)S_n/\sigma^2 \sim \chi^2(n-1)$, (4) can be derived by using a change of variable. □

We now present the proof of theorem 1, where the main idea is to look at an "extreme" event E where design 1 gets "frozen" after initialization (i.e., it will not be simulated after the initial N_0 replications), and $\bar{X}_{1,N_0}$ falls below $\mu_2 - \eta$ for some $\eta > 0$. Note that we only simulate design 1 if $S_1/N_1 > S_2/N_2$. Thus, event E suggests that $S_{1,N_0}/N_1 < S_{2,k}/N_2$ for all $k = N_0, \dots, T - N_0$. In that case, $\bar{X}_2$ will converge to μ_2 while $\bar{X}_1$ never gets updated, resulting in a higher chance of false selection. The rest is to bound the probability of E from below and show that E is not too rare.

Proof of Theorem 1. Choose any $\eta > 0$ and $\varepsilon \in (0, \sigma_2)$. We construct the following events.

$$E_1 := \{\bar{X}_{1,N_0} \leq \mu_2 - \eta\}, \qquad E_{2,\varepsilon} := \{S_{2,k} > \sigma_2 - \varepsilon, \forall k \geq N_0\},$$

$$E_{3,\varepsilon}(T) := \left\{\frac{S_{1,N_0}}{N_0} < \frac{\sigma_2 - \varepsilon}{T - N_0}\right\}, \qquad E_4(T) := \left\{\bar{X}_{2,T-N_0} > \mu_2 - \frac{\eta}{2}\right\}.$$

By Corollary 2, these events are mutually independent for all $T \geq 2N_0$. Moreover, if they occur simultaneously, then we have a false selection. We lower bound the probability of each event as follows.

1. E_1: Since N_0 is constant, we have $\mathbb{P}(E_1) > C_1$ for some $C_1 > 0$.
2. $E_{2,\varepsilon}$: By a union bound,

$$\mathbb{P}(E_{2,\varepsilon}) \geq 1 - \sum_{k=N_0}^{\infty} \mathbb{P}(S_{2,k} \leq \sigma_2 - \varepsilon).$$

Apply (4) in Lemma 2 to each term in the sum and we have

$$\mathbb{P}(S_{2,k} \leq \sigma_2 - \varepsilon) \leq \exp\left(-\frac{1}{4}\left[1-\left(\frac{\sigma_2-\varepsilon}{\sigma_2}\right)^2\right]^2 (k-1)\right) \leq C_\varepsilon \exp(-\gamma_\varepsilon k),$$

where $\gamma_\varepsilon := \frac{1}{4}\left[1-\left(\frac{\sigma_2-\varepsilon}{\sigma_2}\right)^2\right]^2$ and $C_\varepsilon := e^{\gamma_\varepsilon}$. Thus, for any $l \geq 2$,

$$\sum_{k=l}^{\infty} \mathbb{P}(S_{2,k} \leq \sigma_2 - \varepsilon) \leq \frac{C_\varepsilon e^{-\gamma_\varepsilon l}}{1 - e^{-\gamma_\varepsilon}}, \tag{5}$$

so there exists $L \geq 2$ such that

$$\sum_{k=L}^{\infty} \mathbb{P}(S_{2,k} \leq \sigma_2 - \varepsilon) < \frac{1}{4}. \tag{6}$$

If $L \leq N_0$, then $\mathbb{P}(E_{2,\varepsilon}) \geq \frac{3}{4}$. Otherwise if $L > N_0$, then there exists $\tilde{\varepsilon} \in (0, \sigma_2)$ such that $\sum_{k=N_0}^{L-1} \mathbb{P}(S_{2,k} < \sigma_2 - \tilde{\varepsilon}) < \frac{1}{4}$. Take $\varepsilon' := \max\{\varepsilon, \tilde{\varepsilon}\}$, then (6) still holds because its LHS becomes smaller, and we have $\mathbb{P}(E_{2,\varepsilon'}) \geq 1 - \frac{1}{4} - \frac{1}{4} = \frac{1}{2}$.

3. $E_{3,\varepsilon}(T)$: Take ε to be the ε' we just defined. Since $(N_0-1)S_{1,N_0}^2/\sigma_1^2 \sim \chi^2(N_0-1)$,

$$\mathbb{P}(E_{3,\varepsilon}(T)) = \mathbb{P}\left(\chi^2(N_0-1) \le \frac{B}{(T-N_0)^2}\right), \tag{7}$$

where $B := (N_0-1)N_0^2(\sigma_2-\varepsilon)^2/\sigma_1^2$. Notice that for i.i.d. $\mathscr{N}(0,1)$ random variables $Z_1,\ldots,Z_k$ and $x>0$,

$$\mathbb{P}(\chi^2(k) \le x) \ge \mathbb{P}\left(\cap_{i=1}^k \left\{Z_i^2 \le x/k\right\}\right) = \left[\mathbb{P}\left(|Z_1| \le \sqrt{x/k}\right)\right]^k, \tag{8}$$

where by inspecting the shape of normal p.d.f., we have

$$\mathbb{P}(|Z_1| \le t) \ge \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \cdot 2t \ge Kt \tag{9}$$

for some $K>0$ and all t small. Apply (8) and (9) to (7) to get

$$\mathbb{P}(E_{3,\varepsilon}(T)) \ge \left(K\sqrt{\frac{\frac{B}{(T-N_0)^2}}{N_0-1}}\right)^{N_0-1} \ge \frac{C_3}{(T-N_0)^{N_0-1}}$$

for some $C_3>0$ and all $T \ge 2N_0$.

4. $E_4(T)$: Since $\mathbb{P}(E_4(T)) \uparrow 1$ as $T \uparrow \infty$, there exists some $C_2>0$ such that for all $T \ge 2N_0$, we have $\mathbb{P}(E_4(T)) \ge C_2$.

Letting $E(T) := E_1 \cap E_{2,\varepsilon'}(T) \cap E_{3,\varepsilon'}(T) \cap E_4(T)$, the above lower bounds imply that

$$\begin{aligned}
\mathrm{PFS}(T) &= \mathbb{P}(\bar{X}_{1,N_1} < \bar{X}_{2,N_2}) \\
&\ge \mathbb{P}(\bar{X}_{1,N_1} < \bar{X}_{2,N_2} \mid E(T))\,\mathbb{P}(E(T)) \\
&\ge 1 \cdot C_1 \cdot \frac{1}{2} \cdot \frac{C_3}{(T-N_0)^{N_0-1}} \cdot C_2 \ge \frac{C}{(T-N_0)^{N_0-1}}
\end{aligned}$$

for some $C>0$ and all $T \ge 2N_0$. This completes the proof. □

From Theorem 1's proof, we see that OCBA+ with a constant N_0 has a polynomial convergence rate because $\mathbb{P}(E_{3,\varepsilon}(T))$ decreases only polynomially fast as $T \to \infty$, which is essentially due to (9), i.e., as $t \to 0$, $\mathbb{P}(|Z_1| \le t)$ converges to 0 only at a linear rate. It may be possible to extend the same idea to OCBA when $K \ge 3$, where we need to construct a similar but more complicated event on which design b gets "frozen" after initialization. Specifically, we may consider when $\bar{X}_{b,N_0} < \min_{i \in \mathscr{I}} \mu_i - \eta$ for some $\eta > 0$, and S_{b,N_0} is arbitrarily small as T increases. However, the main difficulty here is that $E_{3,\varepsilon}(T)$ may not be so easily characterized and the independence suggested by Lemma 1 might not be as useful.

4.2 Constant N_0 in OCBAR

We have seen in Section 4.1 that a constant N_0 will slow down OCBA and OCBA+'s convergence rate to polynomial. In this section, we show for OCBAR that randomization does not help get around this issue. Similarly, we focus on a two-design case and later discuss how to extend to $K \ge 3$.

Theorem 3 Let Assumption 1 hold, and suppose that $K=2$ and $\mu_1 > \mu_2$. If OCBAR chooses a constant $N_0 \ge 2$, then

$$-\lim_{T\to\infty} \frac{1}{T} \log \mathrm{PFS}(T) = 0.$$

Proof. We follow the same idea as in Theorem 1's proof, which is to construct an event on which design 1 gets "frozen" after initialization. Choose $\eta > 0, \varepsilon > 0$, and consider the following event.

$$A := \{\bar{X}_{1,N_0} < \mu_2 - 2\eta, S_{1,N_0} \leq \varepsilon\},$$

which occurs with a positive probability. Let l_t denote the tth design sampled/simulated after initialization, where $t = 1, 2, \ldots, T - 2N_0$. Note that

$$\text{PFS}(T) \geq \mathbb{P}(\{\bar{X}_{2,T} \geq \mu_2 - \eta\} \cap \{l_t = 2, \forall t = 1, \ldots, T - 2N_0\} \mid A)\mathbb{P}(A).$$

We then have

$$\begin{aligned}
&\mathbb{P}(\{\bar{X}_{2,T} \geq \mu_2 - \eta\} \cap \{l_t = 2, \forall t = 2, \ldots, T - 2N_0\} \mid A) \\
\geq &\underbrace{\mathbb{P}(l_t = 2, \forall t = 1, \ldots, T \mid A)}_{P_1(T)} - \underbrace{\mathbb{P}(\bar{X}_{2,T} < \mu_2 - \eta \mid A)}_{P_2(T)}.
\end{aligned}$$

Write $P_1(T)$ as $\mathbb{P}(l_t = 2, \forall t \mid A)$ for short. Notice that

$$\begin{aligned}
\mathbb{P}(l_t = 2, \forall t \mid A) =& \mathbb{P}(l_t = 2, \forall t \mid S_{1,N_0} \leq \varepsilon) \\
=& \frac{\mathbb{P}(\{l_t = 2, \forall t\} \cap \{S_{1,N_0} \leq \varepsilon\})}{\mathbb{P}(S_{1,N_0} \leq \varepsilon)} \\
=& \frac{\int_0^\varepsilon \mathbb{P}(l_t = 2, \forall t \mid S_{1,N_0} = x) f_{S_{1,N_0}}(x) dx}{\mathbb{P}(S_{1,N_0} \leq \varepsilon)},
\end{aligned} \tag{10}$$

where f_{1,N_0} denotes the p.d.f. of S_{1,N_0}. In addition, by the definition of Algorithm 3,

$$\mathbb{P}(l_t = 2, \forall t \mid S_{1,N_0} = x) = \mathbb{E}\left[\prod_{t=N_0}^{T-1} \frac{S_{2,t}}{S_{2,t} + x}\right],$$

which implies that for all $x \in [0, \varepsilon]$,

$$\mathbb{P}(l_t = 2, \forall t \mid S_{1,N_0} = x) \geq \mathbb{E}\left[\prod_{t=N_0}^{T-1} \frac{S_{2,t}}{S_{2,t} + \varepsilon}\right]. \tag{11}$$

Plug (11) into (10) and we have

$$\mathbb{P}(l_t = 2, \forall t \mid A) \geq \mathbb{E}\left[\prod_{t=N_0}^{T-1} \frac{S_{2,t}}{S_{2,t} + \varepsilon}\right],$$

and it follows by Jensen's inequality that

$$\begin{aligned}
\lim_{T\to\infty} \frac{1}{T} \log \mathbb{P}(l_t = 2, \forall t \mid A) \geq& \lim_{T\to\infty} \frac{1}{T} \log \mathbb{E}\left[\prod_{t=N_0}^{T-1} \frac{S_{2,t}}{S_{2,t} + \varepsilon}\right] \\
\geq& \lim_{T\to\infty} \frac{1}{T} \sum_{t=N_0}^{T-1} \mathbb{E}\left[\log\left(\frac{S_{2,t}}{S_{2,t} + \varepsilon}\right)\right].
\end{aligned} \tag{12}$$

Furthermore, we have the following claim.

Claim 1

$$\lim_{t\to\infty}\mathbb{E}\left[\log\left(\frac{S_{2,t}}{S_{2,t}+\varepsilon}\right)\right]=\log\left(\frac{\sigma_2}{\sigma_2+\varepsilon}\right),\quad \forall\varepsilon>0.$$

Proof of Claim 1. This can be shown using the Dominated Convergence Theorem. □

Apply Claim 1 to (12) to get

$$R_1(\varepsilon):=-\lim_{T\to\infty}\frac{1}{T}\log P_1(T)\le\log\left(\frac{\sigma_2+\varepsilon}{\sigma_2}\right). \tag{13}$$

On the other hand, $P_2(T)$ involves a tail event and by the independence of $\{\bar{X}_{2,T}<\mu_2-\eta\}$ and A,

$$R_2:=-\lim_{T\to\infty}\frac{1}{T}\log P_2(T)=-\lim_{T\to\infty}\frac{1}{T}\log\mathbb{P}\left(\bar{X}_{2,T}<\mu_2-\eta\right)=\frac{\eta^2}{2\sigma_2^2}. \tag{14}$$

We can find $\tilde{\varepsilon}>0$ such that $R_1(\tilde{\varepsilon})<R_2$, so for $\varepsilon=\tilde{\varepsilon}$ there exists $T_{\tilde{\varepsilon}}\in\mathbb{Z}^+$ such that $P_1(T)-P_2(T)>0$ for all $T\ge T_{\tilde{\varepsilon}}$. Taking $\varepsilon=\tilde{\varepsilon}$ and combining all these together,

$$\limsup_{T\to\infty}-\frac{1}{T}\log\mathrm{PFS}(T)\le-\lim_{\substack{T\ge T_\varepsilon\\ T\to\infty}}\frac{1}{T}\log\left\{[P_1(T)-P_2(T)]\mathbb{P}(A)\right\}=\log\left(\frac{\sigma_2+\varepsilon}{\sigma_2}\right).$$

Take the limit on the RHS as $\varepsilon\downarrow 0$ and the conclusion follows. □

Theorem 3 implies a sub-exponential convergence rate. It is promising to extend the proof technique to $K\ge 3$: we only need to study an event where (i) all suboptimal designs have arbitrarily small initial variance estimates and are “frozen” after initialization; (ii) one suboptimal design has an initial mean estimate that is higher than $\mu_1+2\eta$ for some $\eta>0$; (iii) the best design's final estimate is lower than $\mu_1+\eta$. The final step is to check if such an event is still only sub-exponentially rare.

Theorems 1 and 3 both suggest that it may be sensible to avoid using a constant initial sample size N_0 when designing algorithms for fixed budget R&S, since it will likely result in a sub-exponential PFS convergence rate. Setting all the maths aside, an intuitive explanation is as follows. After initialization, there is always a nonzero probability that some designs' mean and variance get severely underestimated/overestimated. If an algorithm highly depends on those estimates, then it may be “tricked” into undersampling or even “freezing” some designs, forbidding a timely correction of their estimation error. The problem with a constant N_0 is that it will result in a constant probability of being “tricked”. However, if we force the initial sample size N_0 to grow proportionally with T, then the probability of such an event will decrease exponentially fast as $T\to\infty$.

4.3 Exponential Convergence Rate of OCBA+ and OCBAR

With a linearly increasing N_0, we show that the PFS converges exponentially fast for OCBA+ and OCBAR. Our PFS bound, though crude, fills the long-standing void of finite-sample guarantee for OCBA. The key idea is that $N_0\le N_i\le T$ for each design i, so if for all $r=N_0,\dots,T$, the $\bar{X}_{i,r}$'s are close enough to their true values, then a correct selection can be guaranteed regardless of the exact value of N_i. Considering the complement of this event yields an exponential upper bound on the PFS.

Theorem 4 Let Assumption 1 hold and suppose $\mu_1>\mu_2\ge\cdots\ge\mu_K$. Then, for OCBA+ and OCBAR (with $N_0=\lfloor\alpha_0T/K\rfloor$), there exists some positive constants $C_1,\dots,C_K$ (independent of T) such that

$$\mathrm{PFS}(T)\le C_1\exp\left(-\frac{\delta^2\alpha_0T}{8\sigma_1^2K}\right)+\sum_{i=2}^{K}C_i\exp\left(-\frac{\bar{\delta}_i^2\alpha_0T}{2\sigma_i^2K}\right),\quad\forall T\ge KN_0, \tag{15}$$

where $\delta:=\mu_1-\mu_2$ and $\bar{\delta}_i=\mu_2-\mu_i+\frac{\delta}{2}$ for $i=2,\dots,K$.

Proof. Note that if the event

$$\mathscr{E} := \bigcap_{r=N_0}^{T} \left\{ \left\{ \bar{X}_{1,r} \geq \mu_1 - \frac{\delta}{2} \right\} \bigcap \left\{ \bigcap_{i \neq 1} \{ \bar{X}_{i,r} \leq \mu_i + \bar{\delta}_i \} \right\} \right\}$$

occurs, then we have a correct selection. Apply a Gaussian tail bound and we have

$$\begin{aligned} \mathrm{PFS}(T) \leq \mathbb{P}(\mathscr{E}^c) &\leq \sum_{r=N_0}^{T} \left[\mathbb{P}\left(\bar{X}_{1,r} < \mu_1 - \frac{\delta}{2} \right) + \sum_{i=2}^{K} \mathbb{P}\left(\bar{X}_{i,r} > \mu_i + \bar{\delta}_i \right) \right] \\ &\leq \sum_{r=N_0}^{\infty} \mathbb{P}\left(\bar{X}_{1,r} < \mu_1 - \frac{\delta}{2} \right) + \sum_{i=2}^{K} \sum_{r=N_0}^{\infty} \mathbb{P}\left(\bar{X}_{i,r} > \mu_i + \bar{\delta}_i \right) \\ &\leq \sum_{r=N_0}^{\infty} \exp\left(-\frac{\delta^2 r}{8\delta_1^2} \right) + \sum_{i=2}^{K} \sum_{r=N_0}^{\infty} \exp\left(-\frac{\bar{\delta}_i^2 r}{2\sigma_i^2} \right). \end{aligned}$$

The conclusion follows from evaluating the geometric sums. □

Theorem 4 applies to a wide range of algorithms that starts by simulating each design $\lfloor \alpha_0 T/K \rfloor$ times. In particular, the original OCBA with $N_0 = \lfloor \alpha_0 T/K \rfloor$ will also achieve an exponential convergence rate. In Section 5, we use numerical experiments to demonstrate the benefits of OCBA+ and OCBAR's fully sequential feature by comparing them with OCBA with a linearly growing N_0.

5 NUMERICAL RESULTS

We compare the performance of OCBA+, OCBAR, OCBA, and a revised OCBA with $N_0 = \lfloor \alpha_0 T \rfloor$, which we call "OCBA2". The problem parameters are listed as follows, where "Slippage Configuration" refers to the least favorable configuration where all the suboptimal designs have the same mean.

1. *Ten designs A*: $\mu = [1, 1.1, 1.2, \ldots, 1.8, 5], \sigma = [5, 5, \ldots, 5, 20]$.
2. *Ten designs B*: $\mu = [1, 1.1, 1.2, \ldots, 1.8, 5], \sigma = [20, 20, \ldots, 20, 5]$.
3. *Equal variances*: $\mu = [1, 2, \ldots, 10], \sigma_i = 10, \forall i = 1, 2 \ldots, 10$.
4. *Increasing variances*: $\mu = [1, 2, \ldots, 10], \sigma = [6, 7, 8, \ldots, 15]$.
5. *Slippage Configuration A*: $\mu = [1, 1, 1, 1, 2], \sigma = [2, 2, 2, 2, 10]$.
6. *Slippage Configuration B*: $\mu = [1, 1, 1, 1, 2], \sigma = [10, 10, 10, 10, 2]$.

The algorithm parameters are $\alpha_0 = 0.2$ for OCBA+, OCBAR and OCBA2, $N_0 = 10$ for OCBA, and $\Delta = 20$ for OCBA and OCBA2. We would like to see how fast the PCS converges to 1 when T ranges from 200 to 4,000 (with an increment of 200). All four algorithms are run for 10,000 independent replications to estimate the PCS using common random numbers, i.e., they share the same X_{ir}'s for all $i = 1, \ldots, K$ and $r = 1, 2, \ldots$. The PCS curves are gathered in Figure 1. We have the following observations.

1. In all six cases, it can be seen that OCBA+, OCBAR and OCBA2 all attain higher PCS than OCBA under every fixed T. In particular, in the first instance, OCBA takes as much as four times the budget of OCBAR to achieve a 0.95 PCS.
2. In Figure 1 (a), (b), (e) and (f), OCBA clearly suffers from a slower convergence rate compared with our algorithms.
3. The advantage of our algorithms is less obvious in Figure 1 (c) and (d), where no design's variance is significantly higher/lower than the others'.
4. Based on observations 2 and 3, we may conclude that OCBA+ and OCBAR are more likely to achieve a large improvement if the designs have a wide range of variances. Also, which of these two performs better appears to be problem-specific.

5. Finally, OCBA+ and OCBAR almost always outperform OCBA2, but the improvement is much smaller compared with that over OCBA. This suggests that (i) a constant N_0 is the main bottleneck of OCBA; (ii) a fully sequential allocation strategy can facilitate theoretical analysis as well as further boost the performance.

REFERENCES

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.

Chen, C.-H., D. He, and M. Fu. 2006. "Efficient Dynamic Simulation Allocation in Ordinal Optimization". *IEEE Transactions on Automatic Control* 51(12):2005–2009.

Chen, C.-H., and L. H. Lee. 2010. *Stochastic Simulation Optimization: An Optimal Computing Budget Allocation*. 1st ed. River Edge, NJ, USA: World Scientific Publishing Co., Inc.

Frazier, P. I. 2014. "A Fully Sequential Elimination Procedure for Indifference-Zone Ranking and Selection with Tight Bounds on Probability of Correct Selection". *Operations Research* 62(4):926–942.

Glynn, P., and S. Juneja. 2004. "A Large Deviations Perspective on Ordinal Optimization". In *Proceedings of the 2004 Winter Simulation Conference*, edited by J. Smith et al., 577–585. Piscataway, New Jersey: IEEE.

Glynn, P., and S. Juneja. 2015. "Ordinal Optimization-Empirical Large Deviations Rate Estimators, and Stochastic Multi-Armed Bandits". *arXiv preprint arXiv:1507.04564*.

Henderson, S. G., and B. L. Nelson. 2006. *Handbooks in Operations Research and Management Science: Simulation*. 1st ed, Volume 13. New York, NY, USA: Elsevier.

Jacobovic, R., and O. Zuk. 2017. "On the Asymptotic Efficiency of Selection Procedures for Independent Gaussian Populations". *Electronic Journal of Statistics* 11(2):5375–5405.

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 (TOMACS)* 11(3):251–273.

Laurent, B., and P. Massart. 2000. "Adaptive Estimation of a Quadratic Functional by Model Selection". *Annals of Statistics* 28(5):1302–1338.

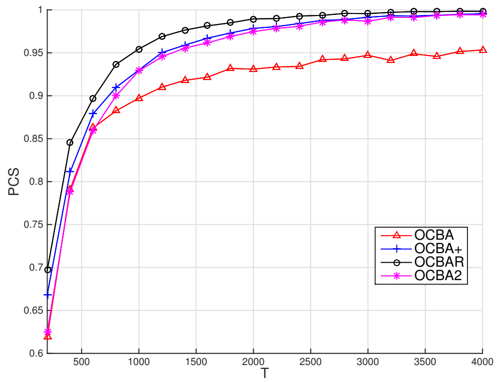
Lee, L. H., E. P. Chew, S. Teng, and D. Goldsman. 2004. "Optimal Computing Budget Allocation for Multi-Objective Simulation Models". In *Proceedings of the 2004 Winter Simulation Conference*, edited by J. Smith et al., 586–594. Piscataway, New Jersey: IEEE.

Pasupathy, R., S. R. Hunter, N. A. Pujowidianto, L. H. Lee, and C.-H. Chen. 2015. "Stochastically Constrained Ranking and Selection via SCORE". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 25(1):1.

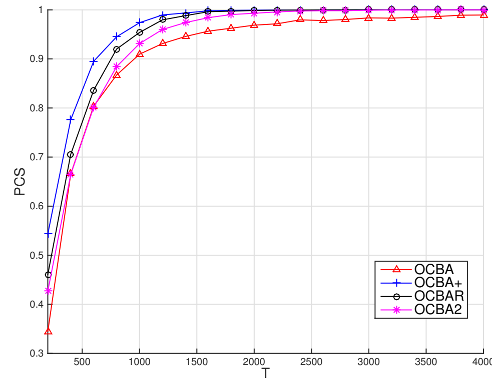
AUTHOR BIOGRAPHIES

DI WU is a Ph.D. candidate in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology. He received his M.S. in control science and engineering from Tsinghua University, China, in 2014. His research interests include simulation optimization, risk quantification, and online learning. His e-mail address is dwu80@gatech.edu.

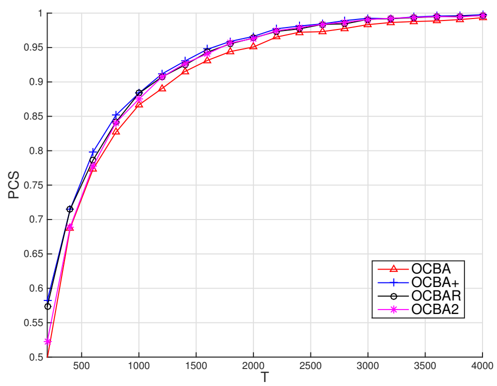
ENLU ZHOU is an Associate Professor in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology. She received the B.S. degree with highest honors in electrical engineering from Zhejiang University, China, in 2004, and received the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 2009. Her research interests include stochastic control and simulation optimization. Her email address is enlu.zhou@isye.gatech.edu and her web page is http://enluzhou.gatech.edu/.



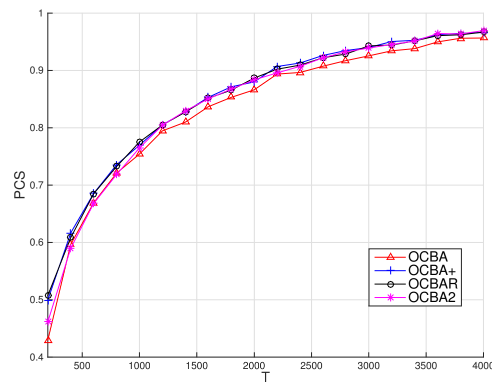
(a) Ten designs A.



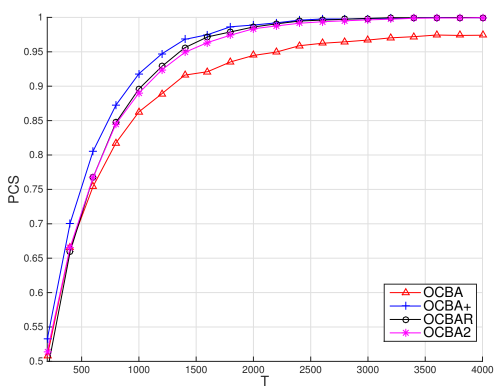
(b) Ten designs B.



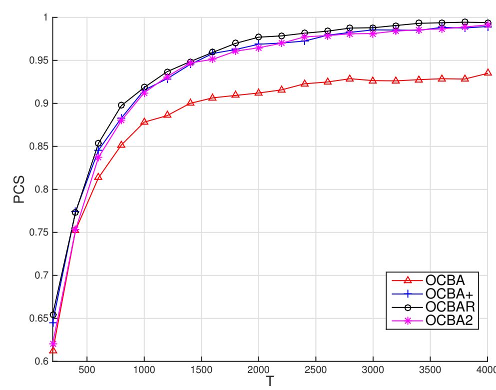
(c) Equal variances.



(d) Increasing variances.



(e) Slippage Configuration A.



(f) Slippage Configuration B.

Figure 1: Comparison of PCS for different algorithms.