Proceedings of the 2023 Winter Simulation Conference
C. G. Corlu, S. R. Hunter, H. Lam, B. S. Onggo, J. Shortle, and B. Biller, eds.

DATA-DRIVEN OPTIMAL ALLOCATION FOR RANKING AND SELECTION UNDER UNKNOWN SAMPLING DISTRIBUTIONS

Ye Chen

Statistical Sciences and Operations Research
Virginia Commonwealth University
1015 Floyd Avenue
Richmond, VA 23284, USA

ABSTRACT

Ranking and selection (R&S) is the problem of identifying the optimal alternative from multiple alternatives through sampling them. In the existing R&S literature, sampling distributions of the observations are usually assumed to be from some known parametric distribution families, even in works that consider input uncertainty. By contrast, this paper considers R&S under completely unknown sampling distributions. We for the first time propose a data-driven nonparametric tuning-free sequential budget allocation strategy that can asymptotically achieve the optimal allocation specified by large deviation analysis. Especially, we propose a new point estimation approach for estimating the optimal large deviation rates directly, which efficiently solves the challenge of estimating large deviation rate functions for lack of known sampling distributions.

1 INTRODUCTION

R&S is a classic mathematical framework of selecting the optimal alternative from multiple alternatives based on their performances, which are learned through sampling. The history of R&S dates back as early as Bechhofer (1954); see Hong and Nelson (2009), Chau et al. (2014) and Hong et al. (2021) for more details. R&S is widely applied in various backgrounds, including inventory management (Xu et al. 2010), agricultural plant breeding (Hunter and McClosky 2016), wind farm placement (Qu et al. 2015; Zhang and Song 2015) and material reliability testing (Chen et al. 2022). This paper considers R&S under unknown sampling distributions. Suppose there are $M > 2$ alternatives with unknown population means (performances) μ_x, $x = 1,...,M$. The optimal alternative $x^* = \arg\max_x \mu_x$ is unique. For convenience, suppose that $\mu_x \neq \mu_y$ for any $x \neq y$. For any x, we can collect independent samples $W_x \sim F_x$ with $\mathbb{E}(W_x) = \mu_x$, where F_x represents some *unknown* non-degenerate sampling distribution and is from the same distribution family (also *unknown*) for all x. Suppose the alternatives are independent, i.e, the samples of y carry no information about any $x \neq y$. At each stage n, let $x^{*,n}$ denote the estimate for x^*, and we say that "correct selection" occurs at time n if $x^{*,n} = x^*$. An allocation strategy refers to a sequence $(x^n)_{n\in\mathbb{N}}$, where x^n denotes the alternative selected for sampling at time n. Consequently, a sample $W_{x^n}^{n+1} \sim F_{x^n}$ is collected at time n. The goal is to improve the quality of $x^{*,n}$ by determining $(x^n)_{n\in\mathbb{N}}$.

The uncertainty about F_x has been relatively overlooked and related research is just gathering momentum recently. Most existing R&S literature designs allocation strategies by assuming a known parametric distribution family (mostly, normal) for F_x, and thus does not consider the uncertainty about F_x at all; some representative works include but not limited to, Jones et al. (1998), Kim and Nelson (2001), Bubeck et al. (2009), Qin et al. (2017), Salemi et al. (2019) and Eckman et al. (2020). Existing attempts for addressing the uncertainty about F_x can be summarized into two categories. One route is by designing allocation strategies under different distribution families of F_x. For example, Gao and Gao (2016) consider exponential sampling distributions. That said, by assuming a specific known F_x, these approaches only consider up to the uncertainty about the parameters instead of F_x itself. The other route is by considering R&S under input certainty, such as Corlu and Biller (2015), Song et al. (2015), Gao et al. (2017), Song and Nelson (2019), Fan et al. (2020), Xiao et al. (2020) and Wu et al. (2022). However, these approaches can only tackle the uncertainty about F_x to a limited extent, because they still rely on making certain assumptions of F_x, such as assuming it is from a finite set of known parametric distributions. Therefore, none of the above references can really handle R&S under completely unknown F_x.

To address the uncertainty about F_x, we propose a new methodology called Data-Driven Optimal Allocation (DDOA). Our approach has several advantages. First, DDOA is a nonparametric sequential allocation strategy that requires no knowledge of F_x at all. We also theoretically prove that DDOA can asymptotically recover the (static) optimal allocation specified by large deviation analysis (Glynn and Juneja 2004). The optimal allocation is derived to maximize the convergence rate of the probability of correct selection (PCS). However, it cannot be directly used to guide budget allocation because it depends on the distribution family of F_x as well as its parameters, both of which are generally unknown. Therefore, many recent works have been focusing on designing efficient sequential allocation strategies to recover the optimal allocation asymptotically; for example, see Chen and Lee (2010), Pasupathy et al. (2014), Chen et al. (2015), Hunter and Feldman (2015), Hunter and McClosky (2016), Zhang et al. (2016), Peng and Fu (2017), Shin et al. (2018), Chen and Ryzhov (2019) and Avci et al. (2023). That said, all these approaches cited above assume some known F_x (mostly, normal). Recently, Russo (2020) proposes a series of top-two algorithms in a Bayesian framework, which with extra tuning can achieve certain optimality criteria under some strict boundedness assumptions; Wang and Zhou (2022) proposes an allocation strategy that can achieve the optimal allocation derived by large deviation analysis for R&S under a specific parametric input certainty model; Chen and Ryzhov (2022) proposes a tuning-free allocation strategy that can recover the optimal allocation for general F_x, but still requires the particular distribution family of F_x to be known for deriving the large deviation rates in each specific case. By contrast, DDOA is completely data-driven and can recover the optimal allocation under unknown F_x.

Second, DDOA is efficient and easy to implement. As noted in Chen and Ryzhov (2022), estimating the large deviation rates is quite challenging yet necessary for recovering the optimal allocation under unknown F_x, and there does not exist many related works. In fact, Glynn and Juneja (2004) is the only work that considers the exact same problem setting as in this paper. However, its approach needs to estimate the entire large deviation rate functions and find the optimal allocation by repeatedly solving concave programs based on the rate function estimates with brute force, thus is not really computationally tractable. There are also several recent works that consider the multi-armed bandits problem under nonparametric settings (e.g., Agrawal et al. 2020; Jourdan et al. 2022; Barrier et al. 2023), but none of them can recover the large-deviation-based optimal allocation in a tuning-free manner. Therefore, instead of estimating the entire large deviation rate functions, we propose a new point estimation approach for estimating only the optimal large deviation rates. These point estimates can be obtained quite efficiently by finding the zeroes of some monotonic functions, and DDOA determines $(x^n)_{n\in\mathbb{N}}$ simply based on these point estimates. Additionally, by adopting an auto-balancing approach (Chen and Ryzhov 2019), DDOA does not require any sort of tuning to achieve the optimal allocation. To the best of our knowledge, DDOA is the very first computationally tractable sequential allocation strategy with asymptotic optimality for R&S under unknown F_x.

The rest of the paper is organized as follows. Section 2 introduces some preliminary results about the optimal allocation derived by the large deviation analysis and provides a new expression for the optimal large deviation rates. Based on this new expression, Section 3 presents new point estimates for the optimal large deviation rates and the DDOA algorithm, and theoretically demonstrates the asymptotic optimality of DDOA. Finally, Section 4 conducts several numerical experiments to illustrate the empirical performance of DDOA, with Section 5 concluding the paper.

2 PRELIMINARIES

2.1 Optimal Allocation Based on Large Deviation Analysis

Depending on how $(x^n)_{n\in\mathbb{N}}$ is determined, an allocation strategy can be categorized into two types. Let $\mathscr{F}^n = \mathscr{B}\left(x^0, W^1_{x^0}, x^1, W^2_{x^1}, ..., x^{n-1}, W^n_{x^{n-1}}\right)$ denote the Borel σ-algebra of all information collected by time n. If all sampling decisions are made at stage 0, i.e., $x^n \in \mathscr{F}^0$ for all n, then an allocation strategy is called static. By contrast, if the sampling decisions are made adaptively based on the collected information, i.e., $x^n \in \mathscr{F}^n$ for all n, then an allocation strategy is called sequential. Let $N^n_x = \sum_{m=0}^{n-1} 1^x_m$ be the number of times that alternative x is chosen for sampling up to time n, where 1^x_m is a binary indicator that is equal to 1 if $x^m = x$ and 0 otherwise. At each stage n, for any x, one can (sequentially) compute the mean of all its available samples

$$\theta^n_x = \frac{1}{N^n_x}\sum_{m=0}^{n-1} 1^x_m W^{m+1}_x.$$

Then, one can estimate x^* with $x^{*,n} = \arg\max_x \theta^n_x$, the alternative that has the largest sample mean at stage n. We summarize the main result of Glynn and Juneja (2004) below to show how the optimal allocation is derived by using the large deviation analysis to maximize the convergence rate of PCS. The interested readers may refer to that paper for further details.

Consider a static allocation strategy that allocates α_x proportion of the total samples to alternative x, where $\alpha_x > 0$ for all x and $\sum_x \alpha_x = 1$. Note that the large deviation analysis only covers distributions that have moment generating functions, and so does our proposed framework. For any x, let $\Psi_x(\gamma) = \log \mathbb{E}\left(e^{\gamma W_x}\right)$ be its log-moment generating function. The large deviation rate function $I_x(u)$ is given by the Fenchel-Legendre transform of $\Psi_x(\gamma)$,

$$I_x(u) = \sup_\gamma \gamma u - \Psi_x(\gamma).$$

The large deviation analysis of Glynn and Juneja (2004) is established based on the following assumption.

Assumption 1 For all x, $\Psi_x(\gamma)$ exists. Furthermore, $I_x(u) < \infty$ for $\min_z \mu_z \le u \le \max_z \mu_z$.

Between x^* and any $x \neq x^*$, the probability of having $\theta^n_{x^*} \le \theta^n_x$ is known to have an exponential decay rate as the number of total samples increases,

$$-\lim_{n\to\infty} \frac{1}{n} \log P\left(\theta^n_{x^*} \le \theta^n_x\right) \quad = \quad \Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x),$$

where $\Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x) = \inf_u \alpha_{x^*} I_{x^*}(u) + \alpha_x I_x(u)$. Then, the convergence rate of PCS can be characterized by

$$-\lim_{n\to\infty} \frac{1}{n} \log(1-\mathrm{PCS}) \quad = \quad \min_{x\neq x^*} \Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x).$$

Therefore, optimizing the convergence rate of PCS is equivalent to maximizing $\min_{x\neq x^*}\Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x)$.

Let $u_{x^*,x}$ be the solution to

$$\alpha_{x^*}\frac{\partial I_{x^*}}{\partial u}(u) + \alpha_x \frac{\partial I_x}{\partial u}(u) = 0. \tag{1}$$

Then, $\Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x)$ can be expressed as

$$\Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x) \quad = \quad \alpha_{x^*} I_{x^*}(u_{x^*,x}) + \alpha_x I_x(u_{x^*,x}). \tag{2}$$

It follows that the optimality conditions that α_x satisfies to maximize $\min_{x\neq x^*}\Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x)$ can be expressed in two parts:

- Total balance condition: $\sum_{x\neq x^*} \frac{I_{x^*}(u_{x^*,x})}{I_x(u_{x^*,x})} = 1$;
- Individual balance condition: $\Gamma_{x^*,x}(\alpha_{x^*},\alpha_x) = \Gamma_{x^*,y}(\alpha_{x^*},\alpha_y), \quad \forall\ x,y \neq x^*$.

These optimality conditions can be expressed in closed forms if F_x is known. For example, if $F_x \sim \mathcal{N}\left(\mu_x, \sigma_x^2\right)$ is from the normal distribution family, then the optimality conditions become

$$\begin{aligned}
\frac{\alpha_{x^*}}{\sigma_{x^*}^2} &= \sum_{x\neq x^*} \frac{\alpha_x}{\sigma_x^2}, \\
\frac{(\mu_x - \mu_{x^*})^2}{\frac{\sigma_x^2}{\alpha_x} + \frac{\sigma_{x^*}^2}{\alpha_{x^*}}} &= \frac{(\mu_y - \mu_{x^*})^2}{\frac{\sigma_y^2}{\alpha_y} + \frac{\sigma_{x^*}^2}{\alpha_{x^*}}}, \quad \forall\ x,y \neq x^*.
\end{aligned}$$

One can see that the optimality conditions still depend on the unknown parameters of F_x even if F_x is known. However, under unknown F_x, neither the large deviation rate functions nor the optimality conditions will be available. Therefore, we give a new perspective below for expressing the optimal large deviation rates, which inspires an efficient point estimation procedure for estimating them.

2.2 A New Expression for the Optimal Large Deviation Rates

For any u, let $\gamma = g_x(u)$ denote the solution to

$$\frac{\partial}{\partial \gamma}(\gamma u - \Psi_x(\gamma)) = u - H_x(\gamma) = 0, \tag{3}$$

where $\frac{\partial}{\partial \gamma}\Psi_x(\gamma)$ is denoted by $H_x(\gamma)$ for convenience. In a word, g_x is the inverse function of H_x, and its existence is guaranteed by the monotonicity of H_x. One can also see that a closed-form expression for g_x may only exist if F_x is known. Nonetheless, we can rewrite the rate function $I_x(u)$ as

$$I_x(u) = g_x(u)u - \Psi_x(g_x(u)). \tag{4}$$

Consequently, we also have

$$\begin{aligned}
\frac{\partial I_x}{\partial u}(u) &= g_x(u) + u\frac{\partial}{\partial u}g_x(u) - \frac{\partial}{\partial g_x}\Psi_x(g_x(u))\frac{\partial}{\partial u}g_x(u) \\
&= g_x(u) + u\frac{\partial}{\partial u}g_x(u) - u\frac{\partial}{\partial u}g_x(u) \qquad (5)\\
&= g_x(u), \qquad (6)
\end{aligned}$$

where (5) follows from (3). Then, since $u_{x^*,x}$ is the solution to (1), we have

$$\begin{aligned}
0 &= \alpha_{x^*}\frac{\partial I_{x^*}}{\partial u}(u_{x^*,x}) + \alpha_x \frac{\partial I_x}{\partial u}(u_{x^*,x}) \\
&= \alpha_{x^*} g_{x^*}(u_{x^*,x}) + \alpha_x g_x(u_{x^*,x}), \qquad (7)
\end{aligned}$$

where (7) follows from (6). For convenience, denote $\gamma^*_{x,x^*} = g_x(u_{x^*,x})$ and $\gamma^*_{x^*,x} = g_{x^*}(u_{x^*,x})$. Then, from (7), we have

$$\alpha_{x^*}\gamma^*_{x^*,x} + \alpha_x \gamma^*_{x,x^*} = 0. \tag{8}$$

Note that from (3), we also have

$$H_{x^*}(\gamma^*_{x^*,x}) = u_{x^*,x} = H_x(\gamma^*_{x,x^*}). \tag{9}$$

Combining (7) and (9), we have that

$$H_{x^*}(\gamma^*_{x^*,x}) - H_x\left(-\frac{\alpha_{x^*}}{\alpha_x}\gamma^*_{x^*,x}\right) \quad = \quad 0. \tag{10}$$

We call $I_{x^*}(u_{x^*,x})$, $I_x(u_{x^*,x})$ and $n\Gamma_{x^*,x}(\alpha_{x^*},\alpha_x)$ the optimal large deviation rates, because together they characterize the optimal allocation. Then, we can express them in terms of $\gamma^*_{x^*,x}$ and γ^*_{x,x^*} by

$$\begin{aligned}
I_{x^*}(u_{x^*,x}) &= \gamma^*_{x^*,x}H_{x^*}(\gamma^*_{x^*,x}) - \Psi_{x^*}(\gamma^*_{x^*,x}), && (11)\\
I_x(u_{x^*,x}) &= \gamma^*_{x,x^*}H_{x^*}(\gamma^*_{x^*,x}) - \Psi_x(\gamma^*_{x,x^*}), && (12)\\
n\Gamma_{x^*,x}(\alpha_{x^*},\alpha_x) &= N^n_{x^*}I_{x^*}(u_{x^*,x}) + N^n_x I_x(u_{x^*,x}) &&\\
&= N^n_{x^*}\left[\gamma^*_{x^*,x}H_{x^*}(\gamma^*_{x^*,x}) - \Psi_{x^*}(\gamma^*_{x^*,x})\right] + N^n_x\left[\gamma^*_{x,x^*}H_{x^*}(\gamma^*_{x^*,x}) - \Psi_x(\gamma^*_{x,x^*})\right] &&\\
&= -N^n_{x^*}\Psi_{x^*}(\gamma^*_{x^*,x}) - N^n_x\Psi_x(\gamma^*_{x,x^*}), && (13)
\end{aligned}$$

where (11)-(12) hold due to (4) and (9), and (13) holds due to (8) and (9). Note that we proportionally scale $\Gamma_{x^*,x}(\alpha_{x^*},\alpha_x)$ by n in (13), which allows us to work with N^n_x directly. From (11)-(13), one can see that the optimal large deviation rates can be efficiently estimated if we can construct simple point estimates for $\gamma^*_{x^*,x}$ and γ^*_{x,x^*}, and the complexity of such point estimation is tremendously reduced in contrast to estimating the entire rate functions. We further note that similar analysis has been conducted in the existing literature, such as Glynn and Juneja (2004) and Li et al. (2018). Nonetheless, none of them proposes to express the optimal large deviation rates in the same way as this paper does.

3 MAIN RESULTS

3.1 Point Estimation of the Optimal Large Deviation Rates

We construct our point estimates for the optimal large deviation rates below. First, we estimate the log-moment generating function $\Psi_x(\gamma)$ and its derivative $H_x(\gamma)$ by their sample average estimators $\Psi^n_x(\gamma)$ and $H^n_x(\gamma)$, respectively, which are given by

$$\begin{aligned}
\Psi^n_x(\gamma) &= \log\left(\frac{1}{N^n_x}\sum_{m=0}^{n-1}1^x_m e^{\gamma W^{m+1}_x}\right), && (14)\\
H^n_x(\gamma) &= \frac{\partial}{\partial\gamma}\Psi^n_x(\gamma) = \frac{\sum_{m=0}^{n-1}1^x_m W^{m+1}_x e^{\gamma W^{m+1}_x}}{\sum_{m=0}^{n-1}1^x_m e^{\gamma W^{m+1}_x}}. && (15)
\end{aligned}$$

Note that our estimate for x^* is $x^{*,n}$ at each stage n. Then, for any $x \neq x^{*,n}$, we estimate $\gamma^*_{x^*,x}$ by $\gamma^n_{x^{*,n},x}$, which is the solution to

$$H^n_{x^{*,n}}(\gamma) - H^n_x\left(-\frac{N^n_{x^{*,n}}}{N^n_x}\gamma\right) \quad = \quad 0, \tag{16}$$

and estimate γ^*_{x,x^*} by

$$\gamma^n_{x,x^{*,n}} = -\frac{N^n_{x^*}}{N^n_x}\gamma^n_{x^{*,n},x}. \tag{17}$$

Consequently, from (9), we estimate $u_{x^*,x}$ by

$$u^n_{x^{*,n},x} = H^n_{x^{*,n}}\left(\gamma^n_{x^{*,n},x}\right) = H^n_x\left(\gamma^n_{x,x^{*,n}}\right). \tag{18}$$

Let $I_x^n(\gamma) \triangleq \gamma H_x^n(\gamma) - \Psi_x^n(\gamma)$. Then, from (11)-(13), we estimate $I_{x^*}(u_{x^*,x})$, $I_x(u_{x^*,x})$ and $n\Gamma_{x^*,x}(\alpha_{x^*}, \alpha_x)$, respectively, by

$$
\begin{aligned}
I_{x^{*,n}}^n\left(\gamma_{x^{*,n},x}^n\right) &= \gamma_{x^{*,n},x}^n H_{x^{*,n}}^n\left(\gamma_{x^{*,n},x}^n\right) - \Psi_{x^{*,n}}^n\left(\gamma_{x^{*,n},x}^n\right), && (19)\\
I_x^n\left(\gamma_{x,x^{*,n}}^n\right) &= \gamma_{x,x^{*,n}}^n H_x^n\left(\gamma_{x,x^{*,n}}^n\right) - \Psi_x^n\left(\gamma_{x,x^{*,n}}^n\right), && (20)\\
\Gamma_{x^{*,n},x}^n &= N_{x^{*,n}}^n I_{x^{*,n}}^n\left(\gamma_{x^{*,n},x}^n\right) + N_x^n I_x^n\left(\gamma_{x,x^{*,n}}^n\right) \\
&= -N_{x^{*,n}}^n \Psi_{x^{*,n}}^n\left(\gamma_{x^{*,n},x}^n\right) - N_x^n \Psi_x^n\left(\gamma_{x,x^{*,n}}^n\right). && (21)
\end{aligned}
$$

Two immediate observations follow. First, the point estimates given in (19)-(21) requires no knowledge of the distribution family of F_x. Second, the key of this point estimation procedure is to solve (16) for $\gamma_{x^{*,n},x}^n$, which is inspired by (1) and (8). In other words, instead of estimating the entire rate functions, (19)-(21) only focus on estimating the optimal large deviation rates, i.e., certain values of the rate functions that characterize the optimal allocation. This point estimation approach significantly distinguishes from any related work in the existing literature that considers R&S under unknown F_x. For example, Glynn and Juneja (2004) estimates the entire rate functions first and then makes them satisfy certain optimizing conditions by spending considerable computational effort on solving some complex programs, and Xiao et al. (2020) briefly discusses a similar point estimation approach but for a purpose other than estimating the optimal large deviation rates and does not conduct any related theoretical analysis or numerical experiments.

3.2 The DDOA Algorithm

In this section, we propose a fully sequential nonparametric tuning-free approach that can achieve the optimality conditions asymptotically under unknown F_x. First of all, we update the definition of $x^{*,n}$ to clear any potential ambiguity. As before, let $x^{*,n} = \arg\max_x \theta_x^n$ if $\arg\max_x \theta_x^n$ is unique; if $\arg\max_x \theta_x^n$ is not unique, then let $x^{*,n}$ be the alternative that has the smallest sample size N_y^n among $y \in \arg\max_x \theta_x^n$. This new definition of $x^{*,n}$ accounts for the situation where there may be more than one alternative having the largest sample mean when F_x is not continuous.

We present the DDOA algorithm in Algorithm 1. Several key features of DDOA can be observed. First, DDOA is initialized by inputting two constants, ρ and ν, as one's prior belief of μ_x for all x, where $\rho < \nu$. For ease of notation, let $x^n = n \bmod M + 1$ for $n = 0, 1, ..., 2M-1$, then assign $W_{x^n}^{n+1} = \rho$ for $n = 0, ..., M-1$ and $W_{x^n}^{n+1} = \nu$ for $n = M, ..., 2M-1$. The two unequal constants ρ and ν initialize H_x^n in (15) for all x and guarantee that (16) in Step 2 is solvable and has a unique negative root.

Second, DDOA is computationally efficient. It only uses the point estimates of the optimal large deviation rates that are proposed in Section 3.1 and does not require estimating the entire rate functions, thus it is computationally tractable and practically applicable. In addition, DDOA also adopts an auto-balancing approach to determine x^n in Step 3 and 4 by reverse-engineering the total balance condition and the individual balance condition, respectively. Such an auto-balancing approach has been adopted in several recently-proposed sequential selection algorithms for R&S under known F_x (e.g., Chen and Ryzhov 2019), with the aim of recovering the optimal allocation while requiring no extra tuning such as solving nonlinear equation systems or concave programs.

Finally, DDOA does not require any knowledge or make any assumptions of F_x, which is essentially different from the vast majority of allocation strategies in the existing literature. Even compared with the latest sequential selection approaches for R&S under known F_x (e.g., Chen and Ryzhov 2022), DDOA only requires extra computational effort in Step 2 for solving (16), which actually can be considered a very minimal cost of not knowing F_x. Interestingly, the LHS of (16) actually can be proved to be monotonic, thus (16) can be solved quite efficiently, for example, by the bisection method (Sikorski 1982). In summary, DDOA is a completely ready-to-use tool that requires no essential effort other than running the algorithm itself.

Algorithm 1 DDOA Algorithm

Step 0: Initialize with $n = 2M$ and $N_x^n = 2$ by inputting two constant estimates, ρ and ν, of μ_x for all x, where $\rho < \nu$.
Step 1: If $\arg\max_x \theta_x^n$ is not unique, assign $x^n = x^{*,n}$ and proceed directly to Step 5.
Step 2: If $\arg\max_x \theta_x^n$ is unique, find $\gamma_{x^{*,n},x}^n$ by solving (16) for all $x \neq x^{*,n}$, then compute $I_{x^{*,n}}^n\left(\gamma_{x^{*,n},x}^n\right)$, $I_x^n\left(\gamma_{x,x^{*,n}}^n\right)$ and $\Gamma_{x^{*,n},x}^n$ for all $x \neq x^{*,n}$ using (19)-(21).
Step 3: Check whether

$$\sum_{x\neq x^{*,n}} \frac{I_{x^{*,n}}^n\left(\gamma_{x^{*,n},x}^n\right)}{I_x^n\left(\gamma_{x,x^{*,n}}^n\right)} > 1. \tag{22}$$

Step 4: If (22) holds, assign $x^n = x^{*,n}$. Otherwise, assign

$$x^n = \arg\min_{x\neq x^{*,n}} \Gamma_{x^{*,n},x}^n. \tag{23}$$

Step 5: Collect new sample $W_{x^n}^{n+1}$, update sample means. Increment n by 1 and return to Step 1.

3.3 Optimality of DDOA

In this section, we theoretically prove that DDOA achieves the optimality conditions asymptotically. Since DDOA is proposed for R&S under unknown F_x, our theoretical analysis is quite general and does not rely on any specific features of F_x, such as whether it has parameters other than μ_x. The entire analysis is highly technical and consists of multiple intermediate results. Therefore, we summarize the main results below and only describe some key steps of the derivation. All our statements are supposed to hold almost surely. Whence, a suitable set of measure 0 should be assumed discarded in the following context so that there is no need to keep repeating the qualification "almost surely".

The main technical difficulty to learn the asymptotics of DDOA comes from characterizing the behavior of $\gamma_{x^{*,n},x}^n$. Therefore, any convergence analysis relying on estimating the entire large deviation rate functions cannot be applied here, such as Glynn and Juneja (2004). It is well known that Ψ_x^n and H_x^n converge uniformly to Ψ_x and H_x respectively on well-defined finite closed intervals; for example, see Feuerverger (1989). However, because the proposed point estimates of the optimal large deviation rates depend on $\gamma_{x^{*,n},x}^n$, one is not able to directly use the uniform convergence of Ψ_x^n and H_x^n to analyze the asymptotics of DDOA. On the one hand, $\gamma_{x^{*,n},x}^n$ is the root of equation (16), and thus is affected by the ratio $\frac{N_{x^{*,n}}^n}{N_x^n}$, which is further determined by the allocation strategy. Consequently, it is unclear if all $\gamma_{x^{*,n},x}^n$ belong to a finite closed interval. On the other hand, $\gamma_{x^{*,n},x}^n$ also affects the allocation strategy x^n, as can be seen from Step 3-4 of Algorithm 1. To resolve this dilemma, it is necessary to learn how $\gamma_{x^{*,n},x}^n$ behaves. First, Theorem 1 shows that DDOA is consistent, i.e., $N_x^n \to \infty$ for all x, which implies that the PCS will converge to 1 as the total sample size $n \to \infty$.

Theorem 1 Under Assumption 1 and Algorithm 1, $N_x^n \to \infty$ for all x.

By the law of large numbers, Theorem 1 implies that $x^{*,n} = x^*$ for all large enough n. As discussed above, because $\frac{N_{x^*}^n}{N_x^n}$ affects the behavior of $\gamma_{x^*,x}^n$, we need to learn how $\frac{N_{x^*}^n}{N_x^n}$ behaves first. Whence, Theorem 2 proves that DDOA actually samples all alternatives at an equivalent rate, i.e., $N_x^n = \Theta(n)$ for all x.

Theorem 2 Under Assumption 1 and Algorithm 1, for any $x \neq y$, $\limsup_{n\to\infty} \frac{N_x^n}{N_y^n} < \infty$.

Theorem 2 implies $\frac{N_x^n}{N_y^n}$ is uniformly bounded for all $x \neq y$, thus $\gamma_{x^*,x}^n$ is also uniformly bounded. Consequently, we only need to consider Ψ_x^n and H_x^n on a finite closed interval U. Whence, we can use a positive decreasing sequence (β^n) to uniformly bound the convergence rate of Ψ_x^n and H_x^n., i.e., for all x and all $\gamma \in U$, $|\Psi_x^n(\gamma) - \Psi_x(\gamma)| = O(\beta^n)$ and $|H_x^n(\gamma) - H_x(\gamma)| = O(\beta^n)$. The rate (β^n) varies for different F_x and can be expressed in closed forms when F_x is known. In general, it is known that $\lim_{n\to\infty} \beta^n = 0$ and $\lim_{n\to\infty}(\sqrt{n}\beta^n)^{-1} = 0$. The interested readers may refer to Feller (1968) and Feuerverger (1989) for further details. Based on these results, we can learn how fast DDOA samples each alternative. Essentially, it is proved the number of samples that can be allocated to x^* between two samples of any suboptimal alternatives (not necessarily the same one) is $O(n\beta^n)$ and vice versa. Consequently, Theorem 3 shows that DDOA asymptotically achieves the total balance condition. The proof is established by noting that the LHS of (22) can only cross 1 when DDOA switches from sampling x^* to some $x \neq x^*$ and vice versa, and the margin by which it can rise above or fall below 1 is $O(\beta^n)$.

Theorem 3 Under Assumption 1 and Algorithm 1,

$$\lim_{n\to\infty} \sum_{x\neq x^*} \frac{I_{x^*}^n\left(\gamma_{x^*,x}^n\right)}{I_x^n\left(\gamma_{x,x^*}^n\right)} = 1.$$

Now we can prove an even stronger result about DDOA's sampling rates: between two samples of the same suboptimal alternative, the number of samples that can be allocated to x^* or any other suboptimal alternative is $O(n\beta^n)$. This implies that the margin by which the ratio $\frac{\Gamma_{x^*,y}^n}{\Gamma_{x^*,z}^n}$ can rise above or fall below 1 is also $O(\beta^n)$. Consequently, Theorem 4 proves that DDOA asymptotically achieves the individual balance condition.

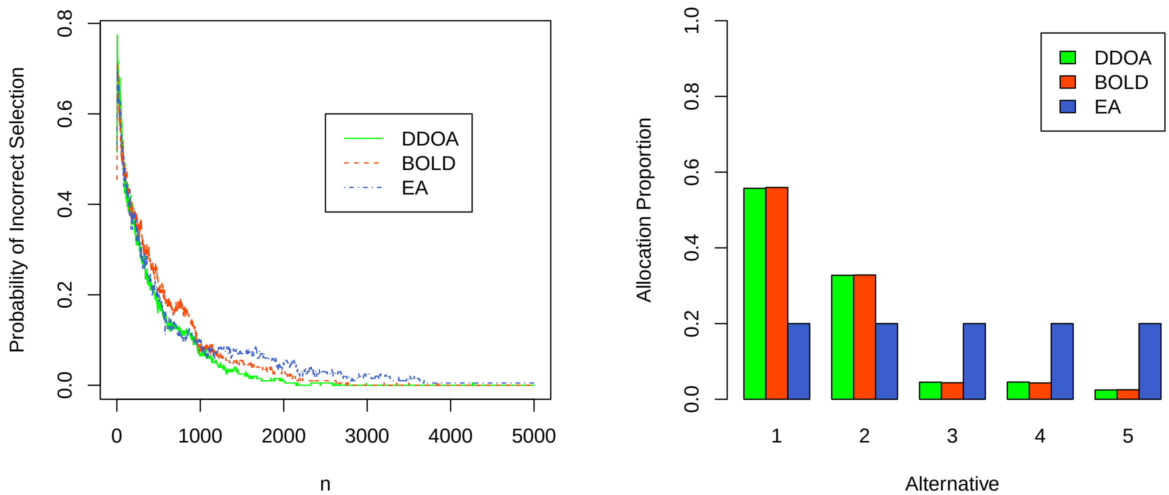
Theorem 4 Under Assumption 1 and Algorithm 1, for any $y, z \neq x^*$,

$$\lim_{n\to\infty} \frac{\Gamma_{x^*,y}^n}{\Gamma_{x^*,z}^n} = 1.$$

4 NUMERICAL ILLUSTRATION

In this section, we conduct several numerical experiments to illustrate DDOA's empirical performance. These experiments cover both continuous and discrete F_x. To the best of our knowledge, the existing R&S literature does not have any other computationally tractable approaches for completely unknown F_x. Therefore, we use EA (Equal Allocation) and BOLD (Chen and Ryzhov 2022) as our benchmark. EA is a naive approach to take when F_x is completely unknown, as it simply samples each alternative with an equal proportion. At the same time, BOLD is a sequential allocation strategy that can achieve the optimal allocation asymptotically under known general F_x. We don't consider OCBA-type methods (Chen and Lee 2010) because, while requiring extra computational effort for solving nonlinear equation systems, they can only approximate the large-deviation-based optimal allocation and are not self-adaptive for different F_x (Gao and Gao 2016). The empirical performance of BOLD has been thoroughly studied in Chen and Ryzhov (2022) and compared with other well-known sequential allocation strategies under known F_x. Therefore, we choose BOLD as a representative for sequential allocation strategies under known F_x. In the following experiments, there are $M = 5$ alternatives and we run 200 macro-replications with 5000 total samples as our budget. To compare the performance of each method, we report the probability of incorrect collection (1-PCS, i.e., the proportion of the 200 macro-replications in which the method does not make the correct selection), as well as the allocation proportion (averaged over 200 macro-replications) achieved by each method when all budget is exhausted.

In Figure 1, $F_x \sim \mathcal{N}\left(\mu_x, \sigma_x^2\right)$ is from the normal distribution family, where μ_x and σ_x are set the same as in Chen and Ryzhov (2017), i.e., $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5) = (0.5, 0.4, 0.3, 0.2, 0.1)$ and $(\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5) = (1, 0.6, 0.6, 1, 1)$. In Figure 2, $F_x \sim \text{Bernoulli}(\mu_x)$ is from the Bernoulli distribution family, where $(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5) = (0.5, 0.4, 0.3, 0.2, 0.1)$. The bisection method is used to solve (16) in Step 2 of DDOA in both experiments. For BOLD, F_x is given as known inputs (σ_x is replaced by the corresponding sample standard deviation estimate in Figure 1), thus the uncertainty is only about the unknown parameters. Figure 1(a) and Figure 2(a) show the PCS of all three methods. We can see that the PCS of DDOA converges almost as fast as the PCS of BOLD and is even slightly better in Figure 1(a). Considering the fact that DDOA knows nothing of F_x and needs to spend extra effort on estimating the optimal large deviation rates, its empirical performance is quite competitive. We can also see that both DDOA and BOLD significantly outperform EA in terms of PCS. However, BOLD's performance is based on the additional advantage of a known F_x while DDOA's is not. Figure 1(b) and 2(b) show the allocation proportions of each method to all five alternatives. We can see that DDOA achieves approximately the same allocation as BOLD, which demonstrates DDOA's ability to recover the optimal allocation for both continuous and discrete F_x.



(a) Probability of Incorrect Selection.

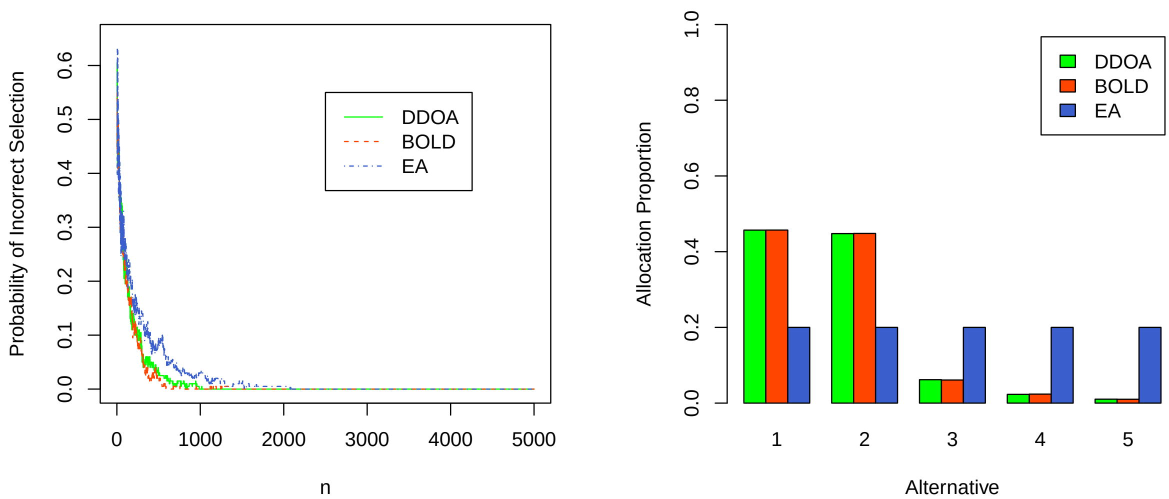
(b) Empirical Allocations.

Figure 1: Illustration with Normal Sampling Distributions.

The goal of these experiments is not to show that DDOA has the best finite-time performance in each particular case, since learning the finite-time empirical performance of sequential allocations is not the focus of this paper. That said, they still numerically demonstrate that DDOA can achieve the optimal allocation asymptotically despite not knowing F_x at all, and that DDOA also has competitive finite-time performance, even compared to approaches with known F_x as inputs. More importantly, as a data-driven nonparametric tuning-free approach, DDOA can be directly applied in all situations with guaranteed empirical performance. By contrast, existing approaches proposed for known F_x usually require nontrivial effort to apply, such as deriving closed-form selection criteria or large deviation rate functions, let alone the risk of misspecifying F_x in practice.

We further note that recovering the optimal allocation asymptotically does not mean that a sequential allocation strategy can achieve the optimal convergence rate specified by large deviation analysis for a static allocation, as discussed in Glynn and Juneja (2018) and Wu and Zhou (2018). Nonetheless, as noted in Garivier and Kaufmann (2016) and Chen and Ryzhov (2022), achieving the optimal allocation is substantial to the performance of sequential allocation strategies for R&S under known F_x. Therefore, DDOA demonstrates that the optimal allocation can be achieved by a simple computationally-tractable sequential allocation strategy under completely unknown F_x, and its logic of addressing the uncertainty of

F_x can also be applied to design sequential allocation strategies for achieving new optimality conditions that may be derived by future frameworks.



(a) Probability of Incorrect Selection.

(b) Empirical Allocations.

Figure 2: Illustration with Bernoulli Sampling Distributions.

5 CONCLUSION

In this paper, we have proposed the DDOA algorithm, a simple nonparametric tuning-free sequential approach for R&S under unknown sampling distributions. Furthermore, we have theoretically proved that DDOA achieves the optimal allocation derived by large deviation analysis and numerically studied its empirical performance. To the best of our knowledge, DDOA is the very first computationally tractable sequential allocation strategy with rigorous theoretical analysis that completely addresses the uncertainty about sampling distributions. Especially, we have proposed a new efficient point estimation procedure for estimating the optimal large deviation rates, which significantly simplifies the difficulty of estimating large deviation rate functions.

There are plenty of directions of future work. One potential direction is to further improve the computational efficiency of DDOA. As noted earlier, DDOA relies on point estimates of the optimal large deviation rates, which can be obtained quite efficiently by finding the zeros of some monotonic functions. The computation of such point estimates can also be performed in parallel among the alternatives. However, such computation still slows down gradually as the sample size builds up. Whence, one may look forward to accelerating the computation by sequentially approximating the zeros of those monotonic functions, or even by constructing more efficient estimation procedures that are fundamentally different. Another future direction is to study the convergence rate of PCS under sequential allocation strategies, such as DDOA. Particularly, one may wish to characterize the optimal convergence rate that can be achieved by sequential allocation strategies and design efficient sequential allocation strategies that are able to achieve this optimal convergence rate.

REFERENCES

Agrawal, S., S. Juneja, and P. Glynn. 2020. "Optimal δ-Correct Best-Arm Selection for Heavy-Tailed Distributions". In *Algorithmic Learning Theory*, 61–110. PMLR.

Avci, H., B. L. Nelson, and A. Wächter. 2023. "Getting to "Rate-Optimal" in Ranking & Selection". *arXiv preprint arXiv:2302.02254*.

Barrier, A., A. Garivier, and G. Stoltz. 2023. "On Best-Arm Identification with a Fixed Budget in Non-Parametric Multi-Armed Bandits". In *International Conference on Algorithmic Learning Theory*, 136–181. PMLR.

Bechhofer, R. 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.

Bubeck, S., R. Munos, and G. Stoltz. 2009. "Pure Exploration in Multi-Armed Bandits Problems". In *Algorithmic Learning Theory: 20th International Conference, ALT 2009, Porto, Portugal, October 3-5, 2009. Proceedings 20*, 23–37. Springer.

Chau, M., M. C. Fu, H. Qu, and I. O. Ryzhov. 2014. "Simulation Optimization: A Tutorial Overview and Recent Developments in Gradient-Based Methods". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 21–35. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, C.-H., S. E. Chick, L. H. Lee, and N. A. Pujowidianto. 2015. "Ranking and Selection: Efficient Simulation Budget Allocation". In *Handbook of Simulation Optimization*, edited by M. C. Fu, 45–80. New York: Springer-Verlag.

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

Chen, Y., and I. O. Ryzhov. 2017. "Rate-Optimality of the Complete Expected Improvement Criterion". In *Proceedings of the 2017 Winter Simulation Conference*, edited by W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, 2173–2182. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Chen, Y., and I. O. Ryzhov. 2019. "Complete Expected Improvement Converges to an Optimal Budget Allocation". *Advances in Applied Probability* 51(1):209–235.

Chen, Y., and I. O. Ryzhov. 2022. "Balancing Optimal Large Deviations in Sequential Selection". *Management Science* 69(6):3457–3473.

Chen, Y., Q. Zhang, M. Li, and W. Cai. 2022. "Sequential Selection for Accelerated Life Testing via Approximate Bayesian Inference". *Naval Research Logistics (NRL)* 69(2):336–351.

Corlu, C. G., and B. Biller. 2015. "Subset Selection for Simulations Accounting for Input Uncertainty". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 437–446. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Eckman, D. J., M. Plumlee, and B. L. Nelson. 2020. "Revisiting Subset Selection". In *Proceedings of the 2020 Winter Simulation Conference*, edited by K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, 2972–2983. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

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

Feller, W. 1968. "An Extention of the Law of the Iterated Logarithm to Variables without Variance". *Journal of Mathematics and Mechanics* 18(4):343–355.

Feuerverger, A. 1989. "On the Empirical Saddlepoint Approximation". *Biometrika* 76(3):457–464.

Gao, F., and S. Gao. 2016. "Optimal Computing Budget Allocation with Exponential Underlying Distribution". In *Proceedings of the 2016 Winter Simulation Conference*, edited by T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, 682–689. 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.

Garivier, A., and E. Kaufmann. 2016. "Optimal Best Arm Identification with Fixed Confidence". In *Conference on Learning Theory*, 998–1027. PMLR.

Glynn, P., and S. Juneja. 2018. "Selecting the Best System and Multi-Armed Bandits". *arXiv preprint arXiv:1507.04564*.

Glynn, P. W., and S. Juneja. 2004. "A Large Deviations Perspective on Ordinal Optimization". In *Proceedings of the 2004 Winter Simulation Conference*, edited by R. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 577–585. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

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., and B. L. Nelson. 2009. "A Brief Introduction to Optimization via Simulation". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. Rosetti, R. Hill, B. Johansson, A. Dunkin, and R. Ingalls, 75–85. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Hunter, S. R., and G. Feldman. 2015. "Optimal Sampling Laws for Bi-Objective Simulation Optimization on Finite Sets". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 3749–3757. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Hunter, S. R., and B. McClosky. 2016. "Maximizing Quantitative Traits in the Mating Design Problem via Simulation-Based Pareto Estimation". *IIE Transactions* 48(6):565–578.

Jones, D., M. Schonlau, and W. Welch. 1998. "Efficient Global Optimization of Expensive Black-Box Functions". *Journal of Global Optimization* 13(4):455–492.

Jourdan, M., R. Degenne, D. Baudry, R. de Heide, and E. Kaufmann. 2022. "Top Two Algorithms Revisited". *arXiv preprint arXiv:2206.05979*.

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.

Li, J., W. Liu, G. Pedrielli, L. H. Lee, and E. P. Chew. 2018. "Optimal Computing Budget Allocation to Select the Nondominated Systems—A Large Deviations Perspective". *IEEE Transactions on Automatic Control* 63(9):2913–2927.

Pasupathy, R., S. R. Hunter, N. A. Pujowidianto, L. H. Lee, and C.-H. Chen. 2014. "Stochastically Constrained Ranking and Selection via SCORE". *ACM Transactions on Modeling and Computer Simulation* 25(1):1:1–1:26.

Peng, Y., and M. C. Fu. 2017. "Myopic Allocation Policy with Asymptotically Optimal Sampling Rate". *IEEE Transactions on Automatic Control* 62(4):2041–2047.

Qin, C., D. Klabjan, and D. Russo. 2017. "Improving the Expected Improvement Algorithm". In *Advances in Neural Information Processing Systems*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Volume 30, 5381–5391. Red Hook, New York: Curran Associates, Inc.

Qu, H., I. O. Ryzhov, M. C. Fu, and Z. Ding. 2015. "Sequential Selection with Unknown Correlation Structures". *Operations Research* 63(4):931–948.

Russo, D. 2020. "Simple Bayesian Algorithms for Best-Arm Identification". *Operations Research* 68(6):1625–1647.

Salemi, P. L., E. Song, B. L. Nelson, and J. Staum. 2019. "Gaussian Markov Random Fields for Discrete Optimization via Simulation: Framework and Algorithms". *Operations Research* 67(1):250–266.

Shin, D., M. Broadie, and A. Zeevi. 2018. "Tractable Sampling Strategies for Ordinal Optimization". *Operations Research* 66(6):1693–1712.

Sikorski, K. 1982. "Bisection is Optimal". *Numerische Mathematik* 40(1):111–117.

Song, E., and B. L. Nelson. 2019. "Input–Output Uncertainty Comparisons for Discrete Optimization via Simulation". *Operations Research* 67(2):562–576.

Song, E., B. L. Nelson, and L. J. Hong. 2015. "Input Uncertainty and Indifference-Zone Ranking & Selection". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 414–424. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Wang, Y., and E. Zhou. 2022. "Data-driven Optimal Computing Budget Allocation under Input Uncertainty". *arXiv preprint arXiv:2209.11809*.

Wu, D., Y. Wang, and E. Zhou. 2022. "Data-Driven Ranking and Selection under Input Uncertainty". *Operations Research (to appear)*.

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

Xiao, H., F. Gao, and L. H. Lee. 2020. "Optimal Computing Budget Allocation for Complete Ranking with Input Uncertainty". *IISE Transactions* 52(5):489–499.

Xu, J., B. L. Nelson, and L. J. Hong. 2010. "Industrial Strength COMPASS: A Comprehensive Algorithm and Software for Optimization via Simulation". *ACM Transactions on Modeling and Computer Simulation* 20(1):1–29.

Zhang, Q., and Y. Song. 2015. "Simulation Selection for Empirical Model Comparison". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, 3777–3788. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Zhang, S., L. H. Lee, E. P. Chew, J. Xu, and C.-H. Chen. 2016. "A Simulation Budget Allocation Procedure for Enhancing the Efficiency of Optimal Subset Selection". *IEEE Transactions on Automatic Control* 61(1):62–75.

AUTHOR BIOGRAPHIES

YE CHEN is an Assistant Professor of Statistical Sciences and Operations Research at Virginia Commonwealth University. He received a Ph.D. in Statistics from the University of Maryland in 2018. His research interests include stochastic optimization, statistical learning, applied probability and transportation science. He was a finalist in the Best Theoretical Paper Award competition at the 2016 Winter Simulation Conference and a winner of the 2022 INFORMS TSL Best Paper Award. His email address is ychen24@vcu.edu.