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

SELECTION OF THE BEST SYSTEM WITH AN OPTIMIZED CONTINUOUS VARIABLE

Yuhao Wang[1], Seong-Hee Kim[1], and Enlu Zhou[1]

[1]School of Industrial and Systems Eng., Georgia Institute of Technology, Atlanta, GA, USA

ABSTRACT

In this paper, we consider a generalized ranking and selection problem, where each system's performance depends on a continuous decision variable necessitating optimization. We focus on a fixed confidence formulation, aiming to find a near-optimal system alongside its corresponding decision variable, meeting specified error tolerances with prescribed confidence levels. To achieve this, we introduce a multi-stage optimization-pruning framework. This framework alternates between optimizing systems using stochastic gradient descent and evaluating their performance through feasibility checks. Our proposed approach offers computational savings by identifying sub-optimal systems before investing effort in optimizing them to the desired accuracy. We demonstrate its efficacy through a numerical study.

1 INTRODUCTION

Simulation optimization is a powerful tool in various industrial areas to select designs that can be modeled by complex simulation systems. Among the large body of study on simulation optimization, ranking and selection (R&S) is a typical class of problems with the goal to identify the best from a finite set of systems, whose performance is random and unknown and needs to be estimated through repeated simulations. In addition to the classic R&S (e.g., see Chen et al. (2000) and Kim and Nelson (2006)), R&S has been studied in various different settings, which include but are not limited to contextual R&S (Shen et al. 2021; Cakmak et al. 2024; Du et al. 2024), robust R&S (Gao et al. 2017; Fan et al. 2020), constrained R&S (Hunter and Pasupathy 2013; Pasupathy et al. 2014), R&S with input uncertainty (Corlu and Biller 2015; Wu and Zhou 2017; Xiao and Gao 2018; Song and Nelson 2019; Xiao et al. 2020; Xu et al. 2020) and data-driven R&S (Yuhao Wang and Enlu Zhou 2024; Wu et al. 2022; Kim and Song 2022; Wang and Zhou 2023).

Despite the variations in R&S methods discussed earlier, a commonality among them is that the simulation system does not include any additional decision variables. However, in many application problems, the system's performance is a random function of some decision variables. Take, for example, an inventory control problem where the manager seeks the optimal inventory policy. Candidate policies often have simple structures, such as order-up-to or (s,S) policies, for reasons of managerial simplicity and computational tractability. These policy types, such as order-up-to and (s,S) policies, can be viewed as separate systems. Within a specific system, say the order-up-to policy system, the performance depends on the chosen order-up-to level, which serves as the decision variable. The challenge then becomes identifying the system that performs best under its optimal decision variable. Notably, this inventory example could also be addressed using a classic R&S method by treating each policy with a specific value of the order-up-to level as an individual system. However, considering each value of the decision variable as a separate system requires discretizing the continuous decision variable, leading to an extremely large number of candidate systems. Moreover, it prevents the use of additional information, such as convexity and gradients, further increasing computational inefficiency. Motivated by these considerations, we explore a method for selecting the best system with an optimized continuous variable, where each system's performance is dependent on the continuous decision variable.

To address this problem, we adopt a fixed confidence approach aimed at identifying a system that is near-optimal, along with its corresponding near-optimal decision variable, within a predetermined error tolerance and confidence level. To achieve this, we develop a multi-stage optimization-pruning framework. This framework entails alternating between optimizing the decision variable for each system and comparing the performance of different systems across various stages. In relation to this topic, Si and Zheng (2022) also consider the same R&S problem involving a continuous decision variable within each system, but our approach differs significantly from theirs. Si and Zheng (2022) choose a fixed budget formulation with a pre-specified total computing budget. In contrast, our method uses the desired confidence level and error tolerance as input parameters, which determine the computing budget needed. Because of the difference in formulations, the methodology we develop also diverges considerably from that of Si and Zheng (2022).

The rest of this paper is organized as follows: Section 2 provides our problem. Section 3 gives the multi-stage framework to solve our problem. The optimization and pruning steps of the proposed framework are discussed in Sections 4 and 5, respectively. Numerical results are shown in Section 6, followed by concluding remarks in Section 7.

2 PROBLEM STATEMENT

Consider the following problem.

$$\min_{k\in\mathscr{K},x_k\in\mathscr{X}_k} \quad f(k,x_k)=\mathbb{E}_\xi\left[F(k,x_k,\xi)\right], \tag{1}$$

where $\mathscr{K}=\{1,2,\ldots,K\}$ is a finite set of systems, x_k is a continuous decision variable that affects the (expected) output performance $f(k,x_k)$ for each $k\in\mathscr{K}$ and $\mathscr{X}_k$ is the domain of $f(k,\cdot)$. For a fixed x_k, one has access to a sequence of random simulation outputs $F(k,x_k,\xi_1),F(k,x_k,\xi_2),\ldots$. Due to the random simulation outputs, it is impossible to find the optimal system k^* along with its corresponding optimal decision variable x_k^* with probability 1 under finite samples. As a consequence, we want to achieve the following goal: with probability at least $1-\alpha$, find $\bar{k}$ and $\bar{x}_{\bar{k}}$ such that

$$f(\bar{k},\bar{x}_{\bar{k}})\le f(k^*,x_{k^*}^*)+\varepsilon,$$

where $\varepsilon>0$ represents the error tolerance of the decision maker. We say $(\bar{k},\bar{x}_{\bar{k}})$ is an ε-optimal system-decision solution and $\bar{k}$ is an ε-optimal system.

3 MULTI-STAGE FRAMEWORK

3.1 A Direct Two-Step Optimization-Simulation Approach

To address the problem outlined in (1), a straightforward strategy involves initially identifying a near-optimal decision variable $\bar{x}_k$ for each system $k\in\mathscr{K}$, using an optimization technique such as stochastic gradient descent (SGD). Subsequently, for each system $k\in\mathscr{K}$, an estimator $\widehat{f}_k$ of its expected performance $f(k,\bar{x}_k)$ under the decision variable $\bar{x}_k$ is obtained through simulation. Suppose that, with probability at least $1-\alpha_1$, $\bar{x}_k$ satisfies $f(k,\bar{x}_k)-f(k,x_k^*)\le\varepsilon_1,\forall k\in\mathscr{K}$. Furthermore, suppose that, with probability at least $1-\alpha_1'$, for each $k,i\in\mathscr{K}$, conditioned on $\bar{x}_k$ and $\bar{x}_i$, $|(\widehat{f}_k-\widehat{f}_i)-(f(k,\bar{x}_k)-f(i,\bar{x}_i))|\le\varepsilon_1'$. Then, if we pick $\bar{k}$ such that $\bar{k}=\arg\min_k\widehat{f}_k$, we have with probability at least $1-\alpha_1-\alpha_1'$,

$$f(\bar{k},\bar{x}_{\bar{k}})\le f(k^*,\bar{x}_{k^*})+\widehat{f}_{\bar{k}}-\widehat{f}_{k^*}+\varepsilon_1'\le f(k^*,x_{k^*}^*)+\widehat{f}_{\bar{k}}-\widehat{f}_{k^*}+\varepsilon_1+\varepsilon_1'\le f(k^*,x_{k^*}^*)+\varepsilon_1+\varepsilon_1'. \tag{2}$$

That is, $(\bar{k},\bar{x}_{\bar{k}})$ is an $(\varepsilon_1+\varepsilon_1')$-optimal system-decision solution with probability at least $1-\alpha_1-\alpha_1'$. By choosing $\alpha_1+\alpha_1'=\alpha$ and $\varepsilon_1+\varepsilon_1'=\varepsilon$, we obtain that $(\bar{k},\bar{x}_{\bar{k}})$ is an ε-optimal system-decision solution.

While the described methodology serves our objective, its practical application may suffer from computational inefficiency. Specifically, solving all systems to optimality or estimating their expected

performance within narrow error tolerance requires substantial optimization and simulation efforts. In fact, in order to obtain a near optimal system-decision solution, it is sufficient to first identify a near-optimal system and then only solve this near-optimal system to optimality with high accuracy level. However, since we do not know which system is the (near) optimal system, we need to solve for an optimal decision variable within one tolerance (ε_1) and estimate its expected performance conditioned on this decision variable within another tolerance (ε_1'). According to the second inequality in (2), we know that, with probability $1-\alpha_1-\alpha_1'$, for any two systems $k \neq i$,

$$\widehat{f}_k - \widehat{f}_i \leq f(k,x_k^*) - f(i,x_i^*) + \varepsilon_1 + \varepsilon_1'.$$

If system k is better than system i (i.e., $f(k,x_k^*) - f(i,x_i^*) < 0$), in order to identify that system k is better than system i, it is sufficient to set $\varepsilon_1 + \varepsilon_1' < f(i,x_i^*) - f(k,x_k^*)$. This leads to the following inequality: $\widehat{f}_k - \widehat{f}_i \leq \varepsilon_1 + \varepsilon_1' - (f(i,x_i^*) - f(k,x_k^*)) < 0$. Given that the system selection process entails choosing the system associated with the lowest estimator, indicated by $\bar{k} = \arg\min_k \hat{f}_k$, system i will consequently be excluded as the optimal choice. When the difference $f(i,x_i^*) - f(k,x_k^*)$ significantly exceeds ε, a substantial reduction in computational resources allocated for optimizing $f(i,\cdot)$ and estimating $f(i,\bar{x}_i)$ can be achieved by opting for comparatively larger values of $\varepsilon_1 + \varepsilon_1'$ relative to ε instead of setting the sum equal to ε. This strategic adjustment facilitates a more efficient allocation of computational efforts while maintaining the accuracy of selection of the best system.

As an illustrative example, consider a set of two systems, denoted as $\mathscr{K} = \{1,2\}$ with expected performances described by $f(1,x_1) = x_1^2 + 10$ and $f(2,x_2) = x_2^2$. The objective is to identify an $\varepsilon = 1$-optimal system-decision solution, which is $(2,\bar{x}_2)$, with a confidence level of at least $1-\alpha$. Employing a direct approach necessitates determining ε_1 and ε_1' such that $\varepsilon_1 + \varepsilon_1' = \varepsilon = 1$ and solving both $f(1,x_1)$ and $f(2,x_2)$ to at least 1-optimal since $\varepsilon_1 \leq \varepsilon = 1$. However, by selecting $\varepsilon_1 = 5$ and $\varepsilon_1' = 4$, it can be demonstrated with a probability of at least $1-\alpha$ that:

$$\widehat{f}_2 - \widehat{f}_1 \leq f(2,\bar{x}_2) - f(1,\bar{x}_1) + \varepsilon_1' \leq f(2,x_2^*) - f(1,x_1^*) + \varepsilon_1 + \varepsilon_1' = 0 - 10 + 5 + 4 = -1 < 0.$$

This identifies system 2 as the optimal system despite the larger tolerances ε_1 and ε_1'. Once we identify system 2 as the optimal system with $\varepsilon_1 = 5$ and $\varepsilon_1' = 4$, we can spend the remaining computational resources to optimally solve only system 2 by re-setting $\varepsilon_1 = 1$. This example illustrates that when there is a large performance differences between systems, exhaustive optimization across all systems or precise estimation of their expected performances becomes unnecessary. This approach enables a more efficient allocation of computational resources by focusing efforts on systems with a higher likelihood of being optimal.

3.2 Multi-Stage Optimization-Pruning Framework

Addressing the issue of sampling inefficiency inherent in the aforementioned direct approach, We propose a multi-stage Pruning framework that utilizes a sequence of progressively narrowing error tolerances. Specifically, the framework operates on two sequences of tolerances: $\varepsilon_1 > \varepsilon_2 > \ldots > \varepsilon_N$ and $\varepsilon_1' > \varepsilon_2' > \ldots > \varepsilon_N'$, where ε_t is the error tolerance for solution $\vec{x}_k^t$ at stage t and ε_t' is the tolerance for estimation $\widehat{f}_k^t - \widehat{f}_i^t$ at stage t, which satisfies $\varepsilon_N' + \varepsilon_N = \varepsilon$. The number of tolerance levels, N, such as $N = \lceil \log_2 K \rceil$, where K represents the total number of systems, plays the role of balancing computational efficiency and accuracy of the optimization and simulation process. The methodology begins with a comparatively larger tolerance value, which is then gradually narrowed down to the target tolerance. At any given stage $t \leq N$, the process involves optimizing each of the "remaining systems" (i.e., those not yet classified as sub-optimal) to find a solution that is ε_t-optimal. By discarding systems demonstrated to be inferior relative to others early on, computation resources can be saved for later stages. This multi-stage optimization-pruning process efficiently identifies the optimal system by progressively concentrating efforts on the most promising candidates, thereby enhancing overall efficiency.

To be specific, let $\mathscr{R}_t$ be the remaining set at the beginning of stage t, with $\mathscr{R}_1 = \mathscr{K}$ representing the initial set of all systems. At the start of each stage t, with initial solutions $\vec{x}_k^{t-1}$ for each system k within $\mathscr{R}_t$, a predetermined number L_k (which depends on ε_t and α_t) of SGD steps are executed for system k. This process yields updated solutions $\vec{x}_k^t$ for all $k \in \mathscr{R}_t$, ensuring that, with probability α_t,

$$f(k,\vec{x}_k^t) \leq f(k,x_k^*) + \varepsilon_t, \quad \forall k \in \mathscr{R}_t.$$

This procedure is referred to as the **Optimization** step at stage t.

Following the **Optimization** step at stage t, we next construct an estimator $\widehat{f}_k^t - \widehat{f}_i^t$ that approximates the difference $f(k,\vec{x}_k^t) - f(i,\vec{x}_i^t)$ for all pairs of systems $i \neq k$ within the remaining set $\mathscr{R}_t$. This estimator is constructed by executing a number of simulation runs, which guarantee the estimation achieves an error tolerance of ε_t' with a specified probability $1-\alpha_t'$.

With a combined probability of $1-\alpha_t-\alpha_t'$, it follows that for every pair of distinct systems k and i within the remaining set $\mathscr{R}_t$,

$$f(k,x_k^*) - f(i,x_i^*) \geq f(k,\vec{x}_k^t) - f(i,\vec{x}_i^t) - \varepsilon_t \geq \widehat{f}_k^t - \widehat{f}_i^t - \varepsilon_t - \varepsilon_t',$$

for all $k \neq i \in \mathscr{R}_t$. Therefore, if for any system $i \in \mathscr{R}_t$, the condition $\widehat{f}_k^t - \widehat{f}_i^t \geq \varepsilon_t + \varepsilon_t'$ is met, system k can be eliminated from $\mathscr{R}_t$ as it is inferred to be no better than system i. In cases where no such system i meets this criterion with respect to system k, then system k is retained within $\mathscr{R}_t$. This process is denoted as the **Pruning** step in stage t.

The multi-stage **Optimization-Pruning** Framework is outlined in Algorithm 1. Details regarding the procedures involved in the **Optimization** and **Pruning** steps, including their input parameters, are elaborated in Sections 4 and 5, respectively.

Algorithm 1 Multi-stage Optimization-Pruning Framework

1: **Input:** System set $\mathscr{K}$, sequences of tolerance $(\varepsilon_1,\ldots,\varepsilon_N)$ and $(\varepsilon_1',\ldots,\varepsilon_N')$, initial point for decision variable $\vec{x}_k^0$ for $k \in \mathscr{K}$, input parameter r_0 for **Pruning**.
2: **Initialize:** $\mathscr{R}_1 \leftarrow \mathscr{K}$. $t \leftarrow 1$.
3: **while** $t < N$ and $|\mathscr{R}_t| > 1$ **do**
4: For each $k \in \mathscr{R}_t$, update the decision variable to $\vec{x}_k^t$ by running **Optimization**$\big(\mathscr{R}_t, \{\vec{x}_k^{t-1}\}_{k\in\mathscr{R}_t}, \varepsilon_t\big)$.
5: $\mathscr{R}_{t+1} \leftarrow$ **Pruning**$\big(\mathscr{R}_t, \{\vec{x}_k^t\}_{k\in\mathscr{R}_t}, \varepsilon_t, \varepsilon_t', r_0\big)$.
6: $t \leftarrow t+1$.
7: **end while**
8: **if** $t < N$ **then**
9: Let k be the unique system in $\mathscr{R}_t$. Update the decision variable to $\vec{x}_k^N$ by running **Optimization**$\big(\mathscr{R}_t, \{\vec{x}_k^{t-1}\}_{k\in\mathscr{R}_t}, \varepsilon\big)$.
10: **end if**
11: **Output:** Any $(k,\vec{x}_k^N)$ for $k \in \mathscr{R}_t$ as the ε-optimal system-decision solution.

4 OPTIMIZATION STEP

We use the stochastic mirror descent, a typical SGD algorithm, from Lan (2020). Let $g(k,x_k) \in \partial f(k,x_k)$ represent a sub-gradient of $f(k,x_k)$ with respect to x_k and $G(k,x_k,\xi) \in \partial F(k,x_k,\xi)$ be a sub-gradient of $F(k,x_k,\xi)$ with respect to x_k. Let $\|\cdot\|_2$ denote the L^2-norm. Then we make the following assumption:

Assumption 1

1. $f(k,\cdot)$ and $F(k,\cdot,\xi)$ are convex for every system k and almost every ξ.
2. $\mathbb{E}_\xi[G(k,x_k,\xi)] = g(k,x_k)$.

3. There exists $\sigma_{G,k} > 0$, such that $G(k, x_k, \cdot)$ is $\sigma_{G,k}$-sub-Gaussian random variable. In particular, a random variable X is σ-sub-Gaussian if

$$\mathbb{E}[e^{\lambda^2 X^2}] \leq e^{\lambda^2 \sigma^2}, \quad \forall |\lambda| \leq \frac{1}{\sigma}.$$

4. There exists $M_k > 0$, $\|G(k, x_k)\|_2 \leq M_k$.
5. There exists $D_k > 0$, $D_k^2 = \sup_{x,y \in \mathscr{X}_k} \frac{\|x-y\|_2^2}{2} < \infty$.

Consider a fixed system k. Since now we focus on conducting SGD within a stage, we drop the superscript t for notational simplicity, and let $x_{k,\ell}$ denote the current solution for system k at iteration ℓ. When employing SGD, at each iteration one has access to a stochastic sub-gradient $G(k, x_{k,\ell}, \xi_{k,\ell})$, which is utilized to execute a single step of the SGD update. To be specific, let $\gamma_{k,\ell}$ be the step size. Then, the update rule is

$$x_{k,\ell+1} = \arg\min_{x \in \mathscr{X}_k} \gamma_{k,\ell} < G(k, x_{k,\ell}, \xi_{k,\ell}), x > + \frac{\|x - x_{k,\ell}\|_2^2}{2}.$$

Let L_k denote the number of SGD iterations required to achieve the desired tolerance and confidence level. Then, to determine L_k, we utilize the concentration results presented in Lan (2020):

Lemma 1 (Proposition 4.1 in Lan (2020)) Suppose Assumption 1 holds. For a fixed system k, let $(x_{k,\ell})_{\ell \geq 1}$ be the sequence given by the stochastic mirror descent algorithm. If we choose a constant step size $\gamma_{k,\ell} = \sqrt{\frac{D_k^2}{L_k}\left(M_k^2 + \sigma_{G,k}^2\right)}$ $\forall \ell = 1, \ldots, L_k$ and let $\bar{x}_{k,L_k} = \frac{1}{L_k}\sum_{\ell=1}^{L_k} x_{k,\ell}$. Then, $\forall \lambda \geq 0$,

$$\mathbb{P}\left(f(k, \bar{x}_{k,L_k}) - f(k, x_k^*) \geq \frac{3D_k}{\sqrt{L_k}}\left(\sqrt{M_k^2 + \sigma_{G,k}^2} + \lambda \sigma_{G,k}\right)\right) \leq \mathrm{e}^{-\lambda} + \mathrm{e}^{-\frac{\lambda^2}{3}}.$$

Hence, after completing a sufficient number of SGD iterations at the beginning of each stage, we can then simulate under the averaged solution, which is the mean of all solutions generated by the SGD iterations. The number of SGD iterations for system k to guarantee ε_t accuracy with confidence level $1 - \alpha_t$ can be obtained by first calculating

$$\lambda = \min\left\{\lambda' : \mathrm{e}^{-\lambda'} + \mathrm{e}^{-\frac{(\lambda')^2}{3}} \leq \alpha_t\right\},$$

and then setting

$$L_k = \left\lceil \frac{9D_k^2}{\varepsilon_t^2}\left(\sqrt{M_k^2 + \sigma_{G,k}^2} + \lambda_k \sigma_{G,k}\right)^2 \right\rceil.$$

The nominal error α_t is set to $\frac{\alpha}{2N|\mathscr{R}_t|}$, considering $|\mathscr{R}_t|$ as the count of remaining systems at stage t. The details of the **Optimization** step is presented in Algorithm 2.

5 FULLY SEQUENTIAL PRUNING STEP

In this section, we elaborate on the **Pruning** step introduced in Section 3. Recall in Section 3 we want to construct estimator $\widehat{f}_k^t, \forall k \in \mathscr{R}_t$ such that the absolute difference between the estimator $\widehat{f}_k^t - \widehat{f}_i^t$ and the actual performance difference $(f(k, \vec{x}_k^t) - f(i, \vec{x}_i^t))$ does not exceed ε_t' for any pair of distinct systems $k \neq i$ within $\mathscr{R}_t$. Such estimator can be obtained through Monte Carlo simulation with a number of samples determined by leveraging some concentration inequality of $F(k, \vec{x}_k^t, \cdot), k \in \mathscr{R}_t$, similar as how to determine the number of SGD in Section 4.

Nonetheless, such a static approach that pre-determines the number of simulations at the beginning can be inefficient in terms of samples, especially when compared to a fully sequential approach that determines

Algorithm 2 Optimization

Input: Remaining set $\mathscr{R}_t$; initial solution x_k^{t-1} for $k \in \mathscr{R}_t$; and tolerance ε_t.
Initialize: $x_k \leftarrow x_k^{t-1}\ \forall k \in \mathscr{R}_t$; $\alpha_t \leftarrow \frac{\alpha}{2N|\mathscr{R}_t|}$; and $\lambda = \min\left\{\lambda' : \mathrm{e}^{-\lambda'} + \mathrm{e}^{-\frac{(\lambda')^2}{3}} \leq \alpha_t\right\}$.
For each $k \in \mathscr{R}$, let

$$L_k \leftarrow \left\lceil \frac{9D_k^2}{\varepsilon_t^2}\left(\sqrt{M_k^2+\sigma_{G,k}^2}+\lambda\sigma_{G,k}\right)^2\right\rceil$$

Set $z_k \leftarrow \frac{x_{k,0}}{L_k}$, $x_k \leftarrow x_{k,0}$.
for each $k \in \mathscr{K}$ **do**
  $\gamma_k \leftarrow \sqrt{\frac{D_k^2}{L_k\left(M_k^2+\sigma_{G,k}^2\right)}}$.
  for ℓ from 1 to L_k **do**
    Obtain a stochastic gradient $G(k,x_k,\xi)$.
    $x_k \leftarrow \arg\min_{x\in\mathscr{X}_k} \gamma_k < G(k,x_k,\xi), x > + \frac{\|x-x_k\|_2^2}{2}$
    $z_k \leftarrow z_k + \frac{x_k}{L_k}$
  end for
end for
Output: $\{z_k\}_{k\in\mathscr{R}_t}$

whether more simulation is needed sequentially. Specifically, the reason of constructing the estimator $\widehat{f}_k^t - \widehat{f}_i^t$ is to infer the value of $f(k,\vec{x}_k^t) - f(i,\vec{x}_i^t)$. We are interested in whether the following two inequalities are satisfied: for any two systems $i \neq k$ and $i,k \in \mathscr{R}_t$,

$$f(i,\vec{x}_i^t) - f(k,\vec{x}_k^t) \leq \varepsilon_t' \tag{3}$$

$$f(i,\vec{x}_i^t) - f(k,\vec{x}_k^t) \geq -\varepsilon_t' \tag{4}$$

To see why (3) and (4) are of interest, if there exists $i \neq k$ such that (3) does not hold, then we know with probability $1-\alpha_t$,

$$f(i,x_i^*) \geq f(i,\vec{x}_i^t) - \varepsilon_t > f(k,\vec{x}_k^t) + \varepsilon_t' - \varepsilon_t \geq f(k,x_k^*) + \varepsilon_t' - \varepsilon_t. \tag{5}$$

Suppose $\varepsilon_t' > \varepsilon_t$, which can be easily satisfied as the tolerance sequences are chosen by the decision maker. We obtain $f(i,x_i^*) > f(k,x_k^*)$, which implies system k is superior to system i. Hence, there is no need to run more simulations or SGD iterations for system i, and we remove system i from the remaining set, i.e., let $\mathscr{R}_t = \mathscr{R}_t \backslash \{i\}$. Otherwise, if (3) holds for a fixed i and all $k \in \mathscr{R}_t$, we obtain

$$f(i,\vec{x}_i^t) - f(k,x_k^*) \leq f(i,\vec{x}_i^t) - f(k,\vec{x}_k^t) + \varepsilon_t \leq \varepsilon_t' + \varepsilon_t.$$

Setting $k = k^*$, then we conclude that $(i,\vec{x}_i^t)$ is an $(\varepsilon_t + \varepsilon_t')$-optimal system-decision solution. Similarly, if there exists $i \neq k$ such that (4) does not hold, then k is identified as a sub-optimal system and we let $\mathscr{R}_t = \mathscr{R}_t \backslash \{k\}$. Otherwise, if for a fixed k and all $i \in \mathscr{R}_t$ such that (4) hold, we have $(k,\vec{x}_k^t)$ is an $(\varepsilon_t + \varepsilon_t')$-optimal system-decision solution. That is, if a pair of systems, i and k, are feasible with respect to (3) and (4), then both systems remain $\mathscr{R}_t$. However, if either (3) or (4) is infeasible, one system is eliminated.

To efficiently determine the feasibility with respect to (3) and (4) in a sequential manner, we draw upon the methodology outlined in Zhou et al. (2022), which provides a feasibility determination procedure for a constraint with threshold, q_t, and an indifference-zone (IZ) parameter, τ_t. Note that Zhou et al. (2022) name the IZ parameter τ_t tolerance level, which is different from our tolerance ε_t or ε_t'. To avoid

confusion, we call τ_t the acceptance level. To apply the methodology in Zhou et al. (2022), we set threshold $q_t = \frac{\varepsilon_t' + \varepsilon_t}{2}$ and acceptance level $\tau_t = \frac{\varepsilon_t' - \varepsilon_t}{2}$. We construct a confidence interval of the target value $f(i,\vec{x}_i^t) - f(k,\vec{x}_k^t)$. The confidence interval shrinks as more samples are available and possesses a crucial property: with high probability, the true target value remains within this interval throughout the entire sampling process, adhering to an acceptance level of τ_t. When q_t reaches the upper bound of the confidence interval, we declare (3) as feasible. Based on the statistical guarantee of Zhou et al. (2022), it holds with high probability that $f(i,\vec{x}_i^t) - f(k,\vec{x}_k^t) \leq q_t + \tau_t = \varepsilon_t'$. Conversely, if q_t reaches the lower bound of the confidence interval, we declare (3) as infeasible. In this case, it holds with high probability that $f(i,\vec{x}_i^t) - f(k,\vec{x}_k^t) \geq q_t - \tau_t = \varepsilon_t$. Following similar arguments in (5), we then obtain $f(i,x_i^*) \geq f(k,x_k^*)$ and, consequently, we can remove i from $\mathscr{R}_t$. The feasibility of (4) can be determined in a similar way. Based on the outcome of this feasibility check, we can then proceed with the **Pruning** step as discussed previously, effectively eliminating sub-optimal systems from consideration.

Let $y_k^t = f(k,\vec{x}_k^t), \forall k \in \mathscr{K}$ be the expected outcome for system k at decision variable $\vec{x}_k^t$ and $Y_k^t = F(k,\vec{x}_k^t,\xi)$ be corresponding stochastic simulation outputs. Furthermore, let $Y_{k,1}^t, Y_{k,2}^t, \ldots$ denote a sequence of samples of Y_k^t at current solution $\vec{x}_k^t$ for all $k \in \mathscr{K}$. We make the following assumption for the validity of the **Pruning** step:

Assumption 2

1. The random simulation outputs follow a multivariate normal distribution:
$$\begin{bmatrix} Y_1^t \\ Y_2^t \\ \vdots \\ Y_K^t \end{bmatrix} \sim \mathscr{N}\left(\begin{bmatrix} y_1^t \\ y_2^t \\ \vdots \\ y_K^t \end{bmatrix}, \Sigma(\bar{\mathbf{x}}^t) \right), \tag{6}$$
where $\bar{\mathbf{x}}^t = (\vec{x}_1^t, \ldots, \vec{x}_K^t)$ and $\Sigma(\bar{\mathbf{x}}^t)$ is a $K \times K$ covariance matrix of $\bar{\mathbf{x}}^t$.
2. Conditioned on $\mathscr{R}_t$ and $\bar{\mathbf{x}}^{\mathbf{t}}$, $\{Y_{i,\ell}^t - Y_{k,\ell}^t\}_{\ell=1}^{\infty}$ are independent and identically distributed (i.i.d.) for $i < k \in \mathscr{R}_t$.
3. The tolerance sequence $\{\varepsilon_t\}_{t=1}^N$ and $\{\varepsilon_t'\}_{t=1}^N$ satisfy $\varepsilon_t' > \varepsilon_t, t = 1, \ldots, N$.

Assumption 2.1 assumes normal simulation noises, which can be satisfied with batched simulation outputs. Assumption 2.2 assumes i.i.d. simulation outputs.

With Assumption 2, the confidence interval of $f(i,\vec{x}_i^t) - f(k,\vec{x}_k^t)$ with r samples is specified as $\bar{Y}_i - \bar{Y}_k \pm R(r;\cdot)/r$, where (i) $\bar{Y}_i$ and $\bar{Y}_k$, estimators of y_i^t and y_k^t, are sample average of r i.i.d. simulation outputs from system i and k, respectively; and (ii) $R(r;\cdot)$ is a non-negative function that decreases to 0 as r increases, which is computed as

$$R(r;v,\eta,z) := \max\left\{0, \frac{(r_0-1)\,\eta z}{v} - \frac{v}{2}r\right\}.$$

Here, $\eta = \frac{1}{2}\left((2\alpha_t')^{-\frac{2}{r_0-1}} - 1\right)$ is a function of the nominal error $\alpha_t' := \frac{\alpha}{2N|\mathscr{R}_t|(|\mathscr{R}_t|-1)}$ for the feasibility check and $z = S_{ik}^2$ is a one-time variance estimator for $Y_i^t - Y_k^t$ with r_0 samples. We refer the reader to Kim and Nelson (2006) for more details on the construction of the confidence interval. Algorithm 3 provides the full description of the **Pruning** step introduced in Section 3.

6 NUMERICAL STUDY

We consider a newsvendor example from Si and Zheng (2022). Suppose there are K products to be selected. For each product $1 \leq i \leq K$, there is an order cost c_i and a price p_i. Furthermore, the demand d_i for product i follows a Poisson distribution with mean λ_i. The decision maker wants to find the product as well as

Algorithm 3 Fully Sequential Pruning

Input: Remaining set $\mathscr{R}_t$; current solution $\vec{x}_k^t$ for $k \in \mathscr{R}_t$; tolerance $\varepsilon_t, \varepsilon_t'$; initial number of simulations r_0.
Initialize: Set $q_t \leftarrow \frac{\varepsilon_t+\varepsilon_t'}{2}$ and $\tau_t \leftarrow \frac{\varepsilon_t'-\varepsilon_t}{2}$; set $\alpha_t' \leftarrow \frac{\alpha}{2N|\mathscr{R}_t|(|\mathscr{R}_t|-1)}$; compute $\eta \leftarrow \frac{1}{2}\left((2\alpha_t')^{-\frac{2}{r_0-1}}-1\right)$; set ON $\leftarrow \mathscr{R}_t$; and set $\text{STOP}_{ik,1} \leftarrow$ False and $\text{STOP}_{ik,2} \leftarrow$ False, $\forall i < k \in \mathscr{R}_t$.
For each $k \in \mathscr{R}_t$, generate r_0 i.i.d. samples $(Y_{k,\ell})_{\ell=1}^{r_0}$ of Y_k^t and compute $\bar{Y}_k \leftarrow \frac{1}{r_0}\sum_{\ell=1}^{r_0} Y_{k,\ell}$; compute $S_{ik}^2 \leftarrow \frac{1}{r_0-1}\sum_{\ell=1}^{r_0}\left(Y_{i,\ell}-Y_{k,\ell}-\bar{Y}_i-\bar{Y}_k\right)^2$, $\forall i < k \in \mathscr{R}_t$; and set $r \leftarrow r_0$.
while $|\text{ON}| \geq 2$ **do**
  for $i < k \in$ ON **do**
    $Z_{ik} \leftarrow R(r; \tau_t, \eta, S_{ik}^2)$
    if $\text{STOP}_{ik,1} =$ False and $\bar{Y}_i - \bar{Y}_k - \frac{Z_{ik}}{r} \geq q_t$ **then**
      $\text{STOP}_{ik,1} =$ True, $\mathscr{R}_t = \mathscr{R}_t \setminus \{i\}$, ON $=$ ON $\setminus \{i\}$.
    else if $\text{STOP}_{ik,1} =$ False and $\bar{Y}_i - \bar{Y}_k + \frac{Z_{ik}}{r} \leq q_t$ **then**
      $\text{STOP}_{ik,1} =$ True.
    end if
    if $\text{STOP}_{ik,2} =$ False and $\bar{Y}_i - \bar{Y}_k + \frac{Z_{ik}}{r} \leq -q_t$ **then**
      $\text{STOP}_{ik,2} =$ True, $\mathscr{R}_t = \mathscr{R}_t \setminus \{k\}$, ON $=$ ON $\setminus \{k\}$.
    else if $\text{STOP}_{ik,2} =$ False and $\bar{Y}_i - \bar{Y}_k - \frac{Z_{ik}}{r} \geq -q_t$ **then**
      $\text{STOP}_{ik,2} =$ True.
    end if
  end for
  for $k \in$ ON **do**
    if $\text{STOP}_{ik,1} = \text{STOP}_{ik,2} =$ True, $\forall i < k \in$ ON and $\text{STOP}_{ki,1} = \text{STOP}_{ki,2} =$ True, $\forall i > k \in$ ON **then**
      ON $=$ ON $\setminus \{k\}$.
    end if
  end for
  $r \leftarrow r+1$. Generate 1 more sample $Y_{k,r}$ of Y_k and update $\bar{Y}_k, \forall k \in$ ON.
end while
Output: Set $\mathscr{R}_t$.

its corresponding optimal ordering amount, which obtains the largest expected profit. In particular, with ordering amount x_i and a realization of demand d_i, the profit is

$$F(i, x_i, d_i) = p_i \min\{d_i, x_i\} - c_i x_i.$$

The corresponding expected profit for product i is then

$$f(i, x_i) := p_i \mathbb{E}[\min\{d_i, x_i\}] - c_i x_i.$$

The goal is to find the ε-optimal system (product)-decision (ordering amount) solution. We set $K = 10$, $c_i = 1 + 0.2(i-1)$, $p_i = 2 + 0.8(i-1)$, and $\lambda_i = 6 - 0.5(i-1)$ for $i = 1, 2, \ldots, 10$. The true optimal product is $k^* = 5$ with optimal ordering amount $x_5^* = 4$ and optimal expected profit $f(5, x_5^*) = 8.277$. Meanwhile, the second optimal product is the 4th product with optimal expected profit $f(4, x_4^*) = 8.083$. We set the tolerance $\varepsilon = 0.1$ such that product 5 is the unique ε-optimal product. We also set the nominal error to be $\alpha = 0.1$ in running the algorithm.

For comparison baselines, to our best knowledge, only Si and Zheng (2022) and this paper consider the generalized R&S with an optimized continuous decision variable. Moreover, the method by Si and Zheng (2022) can only be applied when a computing budget (total number of SGD iterations) is given

ahead while our method determines this computing budget with the given tolerance ε as well as the nominal error α. For these reasons, it is difficult to compare our method with Si and Zheng (2022). Therefore, we only test our method with different choices of $N = 1, 2, \ldots, 6$ for the number of stages. For the sequence of tolerance values, we set $\varepsilon_t = \frac{2}{5} 2^{N-t} \varepsilon$ and $\varepsilon'_t = \frac{3}{5} 2^{N-t} \varepsilon, t = 1, \ldots, N$. Note that $N = 1$ corresponds to the direct approach in Section 3.1. For more implementation details, we set the sub-Gaussian parameter $\sigma_{G,i} = \frac{1}{4} p_i$ and $M_i = p_i - c_i$. The domain of possible ordering amounts for each product is set to $[0, 10]$, thus $D_i = 5\sqrt{2}$. The initial solution for all products is set to $\vec{x}_k^0 = 0$ for $1 \leq k \leq K$.

We run 100 macro-replications for each N. For each macro-replication run, we say one trial of an algorithm is a **success** if it outputs an ε-optimal system-decision solution, which is the optimal product along with its ε-optimal ordering amount. In the following table, we list the average total number of iterations of SGD (denoted as SGD), the average total number of simulation outputs (denoted as simulation) and the average success rate (denoted as success rate, which is the number of successes divided by the number of macro-replications).

Table 1: Performance comparison for different choices of N.

N	N=1	N=2	N=3	N=4	N=5	N=6
SGD	9.03×10^7	2.9×10^7	1.53×10^7	1.22×10^7	1.17×10^7	1.18×10^7
simulation	2.06×10^6	1.75×10^6	1.71×10^6	1.86×10^6	1.87×10^6	1.93×10^6
success rate	1.0	1.0	1.0	1.0	1.0	1.0

As Table 1 indicates, in terms of the total number of SGD iterations, Algorithm 1 runs the least number of iterations of SGD, 1.17×10^7, when $N = 5$, which is 87% less than that by Algorithm 1 when $N = 1$. The average number of simulation outputs for the feasibility check ranges from $[1.71, 2.06] \times 10^6$, which does not vary much among different choices of N. The reason why $N > 1$ reduces the total number of SGD iterations can be explained by Table 2, where we list, for each N, the average number of (sub-optimal) systems removed by Algorithm 1 with different error tolerances (i.e., $\varepsilon_t + \varepsilon'_t = 2^{N-t}\varepsilon, t = 1, \ldots, N$). As we can see, when $N = 1$, all systems are solved to ε-optimal, which requires a large number of SGD iterations. As N increases, some designs are removed with a large tolerance, which requires less number of SGD iterations and help reduce the total number of SGD iterations required by Algorithm 1. Regarding the

Table 2: Average number of systems removed with different error tolerances.

tolerance	6.4	3.2	1.6	0.8	0.4	0.2
N=1						9.0
N=2					9.0	0.0
N=3				8.0	1.0	0.0
N=4			7.14	0.86	1.0	0.0
N=5		6.0	1.27	0.73	1.0	0.0
N=6	4.06	1.94	1.29	0.71	1.0	0.0

average success rate, Algorithm 1 consistently achieves the ε-optimal system-decision solution across 100 macro-replications for various selections of N. However, the average success rate surpasses the targeted confidence level of 90%, indicating potential conservativeness in the sample efficiency of Algorithm 1. One

major factor contributing to this conservativeness is the concentration result in Lemma 1, which induces an unnecessarily large number of SGD iterations to ensure the specified error tolerance and confidence level.

7 CONCLUSION AND FUTURE WORK

In this paper, we have introduced a multi-stage optimization-pruning framework for addressing the generalized ranking and selection problem where system performance depends on a continuous decision variable. The proposed framework alternates between optimizing the decision variable for each system and comparing the performances of different systems. This framework is designed to enhance computational efficiency by identifying and pruning sub-optimal systems with low accuracy early on, thereby avoiding redundant computational efforts associated with optimizing these systems.

Further refinements and extensions of our framework could be explored as future work. This includes investigating alternative optimization techniques to mitigate conservativeness in the **Optimization** step, incorporating additional constraints or higher-dimensional decision variables, and considering the distributional uncertainty of underlying random variables.

REFERENCES

Cakmak, S., Y. Wang, S. Gao, and E. Zhou. 2024. "Contextual Ranking and Selection with Gaussian Processes and Optimal Computing Budget Allocation". *ACM Transactions on Modeling and Computer Simulation* 34(2):1–24.

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.

Corlu, C. G. and B. Biller. 2015. "Subset Selection for Simulations Accounting for Input Uncertainty". In *2015 Winter Simulation Conference (WSC)*, 437–446 https://doi.org/10.1109/WSC.2015.7408185.

Du, J., S. Gao, and C.-H. Chen. 2024. "A Contextual Ranking and Selection Method for Personalized Medicine". *Manufacturing & Service Operations Management* 26(1):167–181.

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

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

Hunter, S. R. and R. Pasupathy. 2013. "Optimal Sampling Laws for Stochastically Cnstrained Simulation Optimization on Fnite Sets". *INFORMS Journal on Computing* 25(3):527–542.

Kim, S. and B. Nelson. 2006. "Selecting the Best System". In *Simulation, Handbooks in Operations Research and Management Science*, edited by S. Henderson and B. Nelson, 501Í C534.

Kim, T. and E. Song. 2022. "Optimizing Input Data Acquisition for Ranking and Selection: A View Through the Most Probable Best". In *2022 Winter Simulation Conference (WSC)*, 2258–2269 https://doi.org/10.1109/WSC57314.2022.10015453.

Lan, G. 2020. *First-order and Stochastic Optimization Methods for Machine Learning*. Springer Nature.

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 (TOMACS)* 25(1):1–26.

Shen, H., L. J. Hong, and X. Zhang. 2021. "Ranking and Selection with Covariates for Personalized Decision Making". *INFORMS Journal on Computing* 33(4):1500–1519.

Si, N. and Z. Zheng. 2022. "Selecting the Best Optimizing System". *arXiv preprint arXiv:2201.03065*.

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

Wang, Y. and E. Zhou. 2023. "Input Data Collection Versus Simulation: Simultaneous Resource Allocation". In *2023 Winter Simulation Conference (WSC)*, 3657–3668 https://doi.org/10.1109/WSC60868.2023.10408130.

Yuhao Wang and Enlu Zhou 2024. "Optimal Computing Budget Allocation for Data-driven Ranking and Selection".

Wu, D., Y. Wang, and E. Zhou. 2022. "Data-Driven Ranking and Selection Under Input Uncertainty". *Operations Research* 72(2):781–795.

Wu, D. and E. Zhou. 2017. "Ranking and Selection Under Input Uncertainty: A Budget Allocation Formulation". In *2017 Winter Simulation Conference (WSC)*, 2245–2256 https://doi.org/10.1109/WSC.2017.8247956.

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.

Xiao, H. and S. Gao. 2018. "Simulation Budget Allocation for Selecting the Top-m Designs with Input Uncertainty". *IEEE Transactions on Automatic Control* 63(9):3127–3134.

Xu, J., Z. Zheng, and P. W. Glynn. 2020. "Joint Resource Allocation for Input Data Collection and Simulation". In *2020 Winter Simulation Conference (WSC)*, 2126–2137 https://doi.org/10.1109/WSC48552.2020.9384050.

Zhou, Y., S. Andradóttir, S.-H. Kim, and C. Park. 2022. "Finding Feasible Systems for Subjective Constraints Using Recycled Observations". *INFORMS Journal on Computing* 34(6):3080–3095.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support by the Air Force Office of Scientific Research under Grant FA9550-22-1-0244, the National Science Foundation under Grant NSF-DMS2053489 and the NSF AI Institute for Advances in Optimization under Grant NSF-2112533.

AUTHOR BIOGRAPHIES

YUHAO WANG is a Ph.D. student at the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology. He received his B.S. degree from the Department of Mathematics at Nanjing University, China, in 2021. His research interests include simulation and stochastic optimization and reinforcement learning. His email address is yuhaowang@gatech.edu.

SEONG-HEE KIM is a Professor in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. She received her Ph.D. in Industrial Engineering and Management Sciences from Northwestern University in 2001. Her email address is skim@isye.gatech.edu and her website is https://www2.isye.gatech.edu/∼skim/.

ENLU ZHOU is a 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 simulation optimization, stochastic optimization, and stochastic control. Her email address is enlu.zhou@isye.gatech.edu, and her web page is https://www.enluzhou.gatech.edu/.