Proceedings of the 2017 Winter Simulation Conference
W. K. V. Chan, A. D'Ambrogio, G. Zacharewicz, N. Mustafee, G. Wainer, and E. Page, eds.

RANKING AND SELECTION UNDER INPUT UNCERTAINTY: A BUDGET ALLOCATION FORMULATION

Di Wu
Enlu Zhou

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

ABSTRACT

A widely acknowledged challenge in ranking and selection is how to allocate the simulation budget such that the probability of correction selection (PCS) is maximized. However, there is yet another challenge: when the input distributions are estimated using finite real-world data, simulation output is subject to input uncertainty and we may fail to identify the best system even using infinite simulation budget. We propose a new formulation that captures the tradeoff between collecting input data and running simulations. To solve the formulation, we develop an algorithm for two-stage allocation of finite budget. We use numerical experiment to demonstrate the performance of our algorithm.

1 INTRODUCTION

Simulation is a widely used tool for modeling and comparing large and complex systems. For instance, when the interest lies in the expected performance of a stochastic system, one can run multiple simulation replications and use sample mean to approximate the true mean. One driving force of stochastic simulation is a set of distributions that describe the stochasticity in the system being modeled. Such distributions can be regarded as input models to the simulation process, hence the name "input distributions". Under this setting, simulation output is subject to two different sources of uncertainty. On the one hand, the limited computing power only allows us to run a finite number of replications, leading to what we call the "stochastic uncertainty"; on the other hand, the input distributions are usually estimated using finite real-world data, and the estimates are affected by the so-called "input uncertainty".

When simulation is used to compare and select the best systems/designs according to their expected performance, it is often referred to as ranking & selection (R&S), or more generally as simulation optimization or Optimization via Simulation (OvS). Currently, most simulation literature assume full knowledge on input distributions, and stochastic uncertainty is the dominating factor. Since we can only run finite simulation replications, we cannot be absolutely certain about the order of the systems' performances, and the probability of the estimated best system being the true best one is called the "probability of correct selection" or PCS in short. There are mainly two types of problems to consider in terms of PCS: (i) how to attain a given PCS level using as few simulation runs as possible (see, e.g., Branke et al. (2005), Kim and Nelson (2007)); (ii) how to maximize the PCS under a finite computing budget (e.g., total runs) constraint.

The focus of this work is on the aforementioned second problem. When there is no input uncertainty, one of the most widely used algorithms is the Optimal Computing Budget Allocation (OCBA) method proposed by Chen et al. (2000). Its allocation strategy has an intuitive interpretation based on each system's signal-to-noise ratio, and it allows a sequential implementation to adaptively control the allocation process. Aside from the simulation community, the budget allocation problem also receives lots of attention from

those working on the multi-armed bandit problem and online learning. From a learning perspective, the problem of selecting the best "arm" is known as the "best-arm identification" or the "pure exploration problem". For example, Audibert and Bubeck (2010) proposes a Upper Confidence Bound (UCB) type algorithm for the case of bounded support, and Russo (2016) develops simple Bayesian algorithms which are shown to attain the best possible exponential rate of the PCS in some sense. Due to limited space, we refer the reader to Russo (2016) for a comprehensive and in-depth review on the history and recent development of this problem.

In practice, input uncertainty can have a significant impact on R&S. Specifically, when the input uncertainty is dominating, we cannot identify the true best system even using infinite computing budget. Some work has been done on indifference zone (IZ) ranking & selection under input uncertainty, recent progress includes but is not limited to Corlu and Biller (2013), Song et al. (2015), Song (2016). For the finite budget allocation problem, Gao et al. (2016) seems to be the only work taking input uncertainty into account, and they optimize the worst-case performance given a fixed finite number of input models. In this paper, we consider the problem of balancing input uncertainty and stochastic uncertainty in R&S. A new formulation is proposed to allow the flexibility of additional input data. To solve the problem, we develop a two-stage algorithm based on a nested asymptotic result. Finally, we use numerical examples to demonstrate the performance of our algorithm in terms of PCS.

2 PROBLEM FORMULATION

2.1 Traditional Formulation

Suppose we are given K stochastic systems $\mathscr{I} = \{1, \ldots, K\}$ and only one of them has the best expected performance. Consider a maximization problem, where the goal is to find

$$b := \arg\max_{i \in \mathscr{I}} \mathbb{E}_{F^c}[h_i(\xi)], \tag{1}$$

where $h_i : \mathbb{R}^m \to \mathbb{R}$ is system i's simulation output function which usually does not have a closed form, $\xi \in \mathbb{R}^m$ is a random vector following a distribution F^c. Since F^c must be specified to drive simulation, it is also called an "input distribution". In this work, we assume the same F^c across all systems, but this can be easily generalized. In traditional computing budget allocation settings, F^c is assumed given and we can estimate system i's performance via its sample mean,

$$\hat{H}_i := \frac{1}{M_i} \sum_{r=1}^{M_i} h_i(\xi_{ir}),$$

where ξ_{ir}'s are independent, identically distributed (i.i.d.) samples drawn from F^c, and M_i is the number of simulation runs/replications allocated to system i. Then, the system with the highest $\hat{H}_i$ is estimated as the best system, i.e.,

$$\hat{b} := \arg\max_{i \in \mathscr{I}} \hat{H}_i.$$

However, unless $M_i \to \infty$ for every i, the estimates $\hat{H}_i$'s are always subject to noise and $\hat{b}$ need not be b. Moreover, each run can be quite costly in terms of time or money. Under a finite computing budget (e.g., finite total runs), one practical goal is to maximize the probability of correctly selecting b . Such probability is usually referred to as the probability of correct selection (PCS), and a typical computing budget allocation problem looks like

$$\begin{aligned} \max_{M_1,\ldots,M_K} \quad & \text{PCS} \\ \text{s.t.} \quad & \sum_{i=1}^{K} M_i = T \\ & M_i \in \mathbb{Z}^+, \quad \forall i \in \mathscr{I}, \end{aligned} \tag{2}$$

where the total budget T is a positive integer and $\mathbb{Z}^+$ denotes the set of nonnegative integers. A key observation is that (2) only formulates a static problem: we are asked to determine $M_1,\ldots,M_K$ before running any simulation. Although this seems to be a heuristic approach, one can derive simple yet powerful algorithms from solving (2). For example, one of the most widely used algorithms, the OCBA method, is actually built on an approximate solution to (2). This suggests that solving a static problem may provide insight into developing simple but good heuristic dynamic algorithms.

2.2 A New Formulation Under Input Uncertainty

The traditional formulation always assumes full knowledge on F^c. But in practice F^c is rarely known exactly and often must be estimated using finite real-world data. Assuming that F^c has a known parametric form $F(\cdot;\theta^c)$ with some unknown parameter $\theta^c \in \Theta \subseteq \mathbb{R}^p$, we can define a function

$$H_i(\theta) := \mathbb{E}_{F(\cdot;\theta)}[h_i(\xi)]$$

to be the expected performance of system i under parameter θ. Further assume without loss of generality that $H_i(\theta^c) \neq H_j(\theta^c)$ for all $i \neq j \in \mathscr{I}$ so that the best system is unique. Suppose θ^c is estimated using maximum likelihood estimation (MLE), and $\hat{\theta}_N$ is the estimator of θ^c using N i.i.d. input data samples $\psi^n := (\xi_1,\xi_2,\ldots,\xi_N)$ from F^c. The estimated mean performance of system i is now

$$\hat{H}_i(\hat{\theta}_N) := \frac{1}{M_i}\sum_{r=1}^{M_i} h_i(\hat{\xi}_{ir}), \tag{3}$$

where $\hat{\xi}_{ir}$'s are i.i.d. samples drawn from $F(\cdot;\hat{\theta}_N)$, and the estimated best system is given by

$$\hat{b} := \arg\max_{i\in\mathscr{I}} \hat{H}_i(\hat{\theta}_N). \tag{4}$$

In this paper, we refer to the error in estimating the mean performance caused by finite simulation runs as "stochastic uncertainty", and the error in estimating F^c as "input uncertainty". To get a sense of how input uncertainty affects R&S, define

$$\mathscr{P} := \{\theta \in \Theta \mid \exists i \neq b \quad \text{s.t. } H_b(\theta) < H_i(\theta)\} \tag{5}$$

to be the set of parameters θ under which the best system is not b. We will refer to $\mathscr{P}$ as *the perturbation zone*. Since θ^c is estimated using finite input data, it is possible that $\hat{\theta}_N$ falls into $\mathscr{P}$. Should this happen, we will have $\arg\max_{i\in\mathscr{I}} H_i(\hat{\theta}_N) \neq \arg\max_{i\in\mathscr{I}} H_i(\theta^c)$, which means that the optimal system is perturbed, and the more we simulate, the less likely we will select the true best system b. Therefore, input uncertainty has an undeniable impact on R&S, and it cannot be controlled by increasing simulation runs. If we are not allowed to collect more input data, there is no guarantee that we will select b with a large probability. For this reason, we consider the following problem.

Suppose we are given a budget T, which can be used to collect input data and run simulations. The unit costs of input data and simulation run are c_1 and c_2, respectively. The budget can be thought of as time or money. For example, a company looking to launch a new product wants to collect information (data) about the potential market demand. Then, simulations are run to decide the optimal order quantities, inventory, etc.. Both collecting data and running simulations can be time-consuming, thus incurring an opportunity cost since its competitors may launch a similar product first and take up a market share. In this scenario, the decision is carried out in a two-stage style: first we determine N, the number of input data samples to collect; then using the input data, we compute the estimate $\hat{\theta}_N$ and use an adaptive algorithm to select the best system from $\mathscr{I}$ based on (3) and (4). Let

$$\mathscr{F}_r := \sigma\left(I_1, h^1_{I_1},\ldots,I_{r-1},h^{r-1}_{I_{r-1}}\right), \quad r = 2,3,\ldots,R$$

be the filtration generated by past decisions and observations up to (including) the $(r-1)$th run, where I_r is the system being simulated at the rth run and $h^r_{I_r}$ is the corresponding observation. The problem can be formulated as

$$\begin{aligned} \max_{N} \quad & \mathbb{E}\left[\max_{\{I_r\}_{r=1}^{R}} \mathbb{P}\left\{\hat{b}=b \mid \hat{\theta}_N\right\}\right] \\ \text{s.t.} \quad & c_1 N + c_2 R \le T \\ & N, R \in \mathbb{Z}^+ \\ & I_r \in \{1,\ldots,K\}, \quad r=1,2,\ldots,R \\ & I_r \in \mathscr{F}_r, \quad r=2,3,\ldots,R \end{aligned} \tag{6}$$

where R is the total number of simulation runs. The formulation (6) is interpreted as follows. After determining N, we collect N input data samples to compute $\hat{\theta}_N$. The remaining budget will be used to run simulations, where $\hat{\xi}_{ir}$'s are drawn independently from $F(\cdot;\hat{\theta}_N)$ and the total number of runs is upper bounded by $(T-c_1N)/c_2$. Which system to be simulated at the rth run must be adapted to $\mathscr{F}_r$. After R runs, we use (4) to output the estimated best system.

Formulation (6) characterizes a tradeoff between input uncertainty and stochastic uncertainty: if N is too small, then $\hat{\theta}_N$ is likely to fall into the perturbation zone and simulation does not help; if N is too large, there will not be enough budget to perform simulation, and the PCS may also be very low. Although (6) is a well-defined problem, it is nearly intractable even if we know the closed forms of h_i's and the true parameter θ^c. There are two major difficulties: (i) we cannot evaluate the inner-layer optimal value, otherwise the problem reduces to a two-stage stochastic program with recourse; (ii) the inner-layer optimization is essentially a dynamic program of which an optimal policy is still an open question. To develop a heuristic algorithm, we consider the following static problem.

$$\begin{aligned} \max_{N,M_1,\ldots,M_K} \quad & \text{PCS} \\ \text{s.t.} \quad & c_1 N + c_2 \sum_{i=1}^{K} M_i = T \\ & N \in \mathbb{Z}^+, \quad M_i \in \mathbb{Z}^+, \forall i \in \mathscr{I} \end{aligned} \tag{7}$$

Formulation (7) can be viewed as an approximation to (6). Similar to formulation (2), formulation (7) requires us to predetermine N and M_i's prior to data collection and simulation. Given full knowledge of h_i's and θ^c, (7) is an integer program with a nonlinear objective and some linear constraints. Later we will see that in a special case, the integrality constraints can be dropped and the problem becomes a standard nonlinear optimization problem. At the expense of generality, (7) has better tractability than (6). Hopefully, by solving (7), we can derive a simple and effective algorithm to solve the dynamic problem (6).

3 ALGORITHM

3.1 Asymptotic Normality Under Input Uncertainty

In order to solve (7), we need to know more about the objective function, or how the PCS depends on N and M_i's. By definition, we have

$$\text{PCS} = \mathbb{P}\left\{\bigcap_{i\neq b, i\in\mathscr{I}} \left\{\hat{H}_b(\hat{\theta}_N) > \hat{H}_i(\hat{\theta}_N)\right\}\right\},$$

where the estimate $\hat{H}_i$ implicitly depends on the number of runs M_i allocated to system i. Generally speaking, the PCS usually does not allow a closed-form expression except for some very special cases (e.g.,

two systems with normal performance). We follow the standard approach in the literature and approximate the PCS using Bonferroni's inequality:

$$\text{PCS} \geq 1 - \sum_{i \neq b, i \in \mathscr{I}} \mathbb{P}\left\{\hat{H}_b(\hat{\theta}_N) \leq \hat{H}_i(\hat{\theta}_N)\right\}. \tag{8}$$

The next step is to approximate each probability in the summation. In the derivation of OCBA under no input uncertainty, the idea is basically to approximate the distribution of $\hat{H}_b(\theta^c) - \hat{H}_i(\theta^c)$ using the central limit theorem (CLT). With input uncertainty, however, we cannot directly apply CLT, and we need to derive asymptotic results tailored to the nested structure. To begin with, we will assume that the asymptotic normality of MLE (see, e.g., (Lehmann and Casella 2006)) holds throughout this paper, i.e.,

$$\sqrt{N}(\hat{\theta}_N - \theta^c) \Rightarrow \mathscr{N}(0, [I(\theta^c)]^{-1}) \quad \text{as } N \to \infty,$$

where "$\Rightarrow$" means convergence in distribution, $\mathscr{N}$ stands for a normal distribution, and $I(\theta^c)$ is the Fisher information that a sample from $F(\cdot;\theta^c)$ carries about the true parameter θ^c. For a given system $i \in \mathscr{I}$, assuming that H_i is differentiable at θ^c, by the delta theorem (see, e.g., Casella and Berger (2002)),

$$\sqrt{N}[H_i(\hat{\theta}_N) - H_i(\theta^c)] \Rightarrow \mathscr{N}(0, \sigma^2_{H_i}), \tag{9}$$

where $\sigma^2_{H_i} := \nabla_\theta H_i(\theta^c)^\intercal [I(\theta^c)]^{-1} \nabla_\theta H_i(\theta^c)$. When H_i cannot be evaluated exactly and is approximated by sample mean as in (3), a straightforward asymptotic result is given by the following iterated limits.

$$\lim_{M_i \to \infty} \sqrt{N}[\hat{H}_i(\hat{\theta}_N) - H_i(\theta^c)] \Rightarrow \mathscr{N}(0, \sigma^2_{H_i}) \quad \text{as } N \to \infty,$$

where the limit of $M_i \to \infty$ is based on strong law of large numbers (SLLN). However, this is hardly useful because in practice we can never send M_i to infinity. Let us derive an asymptotic result that allows more flexibility for the growth of N and M_i. For a system $i \in \mathscr{I}$, define

$$\sigma_i^2(\theta) := \text{Var}_{F(\cdot;\theta)}[h_i(\xi)]$$

to be h_i's variance under parameter θ. Let "a.s." be short for "almost surely". To make the setting more rigorous, we make the following assumptions.

Assumption 1

(1) $\hat{\theta}_N \to \theta^c$ a.s. as $N \to \infty$ (strong consistency of MLE).
(2) For every system $i \in \mathscr{I}$, $\sigma_i^2(\theta)$ is a continuous function of θ.
(3) For every system $i \in \mathscr{I}$, H_i is differentiable at θ^c.

Note that Assumption 1(2) implies that $\sigma_i^2(\hat{\theta}_N) \to \sigma_i^2(\theta^c)$ a.s. by continuous mapping theorem. The following theorem will be useful in developing an algorithm to solve (7).

Theorem 1 Let Assumption 1 hold. If there exist constants $\alpha_i > 0$ such that $\lim_{N\to\infty} N/M_i = \alpha_i$ for all $i \in \mathscr{I}$, then for any two systems i and j,

$$\sqrt{N}\left\{[\hat{H}_i(\hat{\theta}_N) - \hat{H}_j(\hat{\theta}_N)] - [H_i(\theta^c) - H_j(\theta^c)]\right\} \Rightarrow \mathscr{N}(0, \tilde{\sigma}^2_{ij}) \quad \text{as } N \to \infty,$$

where $\tilde{\sigma}^2_{ij} := \sigma^2_{H_{ij}} + \alpha_i \sigma_i^2(\theta^c) + \alpha_j \sigma_j^2(\theta^c)$ and

$$\sigma^2_{H_{ij}} := [\nabla H_i(\theta^c) - \nabla H_j(\theta^c)]^\intercal [I(\theta^c)]^{-1} [\nabla H_i(\theta^c) - \nabla H_j(\theta^c)].$$

Proof. Decompose the difference as follows.

$$\begin{aligned}
&\sqrt{N}\left\{[\hat{H}_i(\hat{\theta}_N)-\hat{H}_j(\hat{\theta}_N)]-[H_i(\theta^c)-H_j(\theta^c)]\right\}\\
=&\underbrace{\sqrt{N}\left\{[\hat{H}_i(\hat{\theta}_N)-\hat{H}_j(\hat{\theta}_N)]-[H_i(\hat{\theta}_N)-H_j(\hat{\theta}_N)]\right\}}_{X_N}+\underbrace{\sqrt{N}[\left\{H_i(\hat{\theta}_N)-H_j(\hat{\theta}_N)\right]-[H_i(\theta^c)-H_j(\theta^c)]\right\}}_{Y_N}.
\end{aligned}$$

Observe that Y_N is $\sigma(\hat{\theta}_N)$-measurable and $Y_N \Rightarrow \mathcal{N}(0,\sigma^2_{H_{ij}})$. Let $\mathbf{i}$ denote the imaginary number $\sqrt{-1}$. We study the characteristic function (ch.f.) of X_n+Y_n. By Theorem 3.3.8 and Theorem 3.4.2 from Durrett (2010),

$$\begin{aligned}
&\mathbb{E}[e^{\mathbf{i}t(X_N+Y_N)}]\\
&=\mathbb{E}\left\{e^{\mathbf{i}tY_N}\cdot\mathbb{E}[e^{\mathbf{i}tX_N}\mid\hat{\theta}_N]\right\}\\
&=\mathbb{E}\left\{e^{\mathbf{i}tY_N}\cdot\mathbb{E}\left[\exp\left(\mathbf{i}t\sqrt{N}[\hat{H}_i(\hat{\theta}_N)-H_i(\hat{\theta}_N)]\right)\middle|\hat{\theta}_N\right]\cdot\mathbb{E}\left[\exp\left(\mathbf{i}t\sqrt{N}[-\hat{H}_j(\hat{\theta}_N)+H_j(\hat{\theta}_N)]\right)\middle|\hat{\theta}_N\right]\right\}\\
&=\mathbb{E}\left\{e^{\mathbf{i}tY_N}\cdot\left(1-\sigma_i^2(\hat{\theta}_N)\frac{t^2}{2M_i}\frac{N}{M_i}+o\left(\frac{1}{M_i}\right)\right)^{M_i}\cdot\left(1-\sigma_j^2(\hat{\theta}_N)\frac{t^2}{2M_j}\frac{N}{M_j}+o\left(\frac{1}{M_j}\right)\right)^{M_j}\right\}\\
&\to\exp\left(-\frac{\sigma^2_{H_{ij}}t^2}{2}\right)\exp\left(-\frac{\alpha_i\sigma_i^2t^2}{2}\right)\exp\left(-\frac{\alpha_j\sigma_j^2t^2}{2}\right),
\end{aligned}$$

where the last step follows from the fact that $|e^{\mathbf{i}t}|\le 1$ and the dominated convergence theorem from complex analysis. Since (X_N+Y_N)'s ch.f. converges to the ch.f. of $\mathcal{N}(0,\tilde{\sigma}^2_{ij})$, the proof is complete. □

Theorem 1 characterizes the asymptotic normality of the difference between systems i and j's estimated mean performance. The intuition behind theorem 1 is that as N grows large, $\hat{\theta}_N$ approximately equals θ^c, and $\hat{\xi}_{ik}$'s can be roughly treated as i.i.d. random variables drawn from F^c. Thus, the correlation between $\hat{\xi}_{ik}$'s and $\hat{\theta}_N$ diminishes as $N\to\infty$, leading to a sum of limiting variances in $\tilde{\sigma}^2_{ij}$. Another way to interpret this result is that for N large,

$$\hat{H}_i(\hat{\theta}_N)-\hat{H}_j(\hat{\theta}_N)\overset{\mathscr{D}}{\approx}\mathcal{N}\left(H_i(\theta^c)-H_j(\theta^c),\tilde{\sigma}^2_{ij}/N\right),$$

where $\overset{\mathscr{D}}{\approx}$ means "approximately distributed as" and

$$\frac{\tilde{\sigma}^2_{ij}}{N}=\frac{\sigma^2_{H_{ij}}}{N}+\frac{\sigma_i^2(\theta^c)}{M_i}+\frac{\sigma_j^2(\theta^c)}{M_j}.$$

Here, $\sigma^2_{H_{ij}}$ characterizes the relative sensitivity of i and j's true performance to input uncertainty (error in estimating θ^c), while $\sigma_i^2(\theta^c)$ and $\sigma_j^2(\theta^c)$ capture the amount of stochastic uncertainty (error in estimating H) in simulation. Clearly, increasing N,M_i and M_j controls the variance caused by each individual source of uncertainty.

3.2 Approximate Solution

With theorem 1, we can derive an approximate solution to problem (7). For ease of presentation, we write $\sigma_i^2(\theta^c)$ as σ_i^2 for short. Let

$$\delta_{bi}:=H_b(\theta^c)-H_i(\theta^c),\quad \sigma^2_{bi}:=\frac{\sigma^2_{H_{bi}}}{N}+\frac{\sigma_b^2}{M_b}+\frac{\sigma_i^2}{M_i}.$$

Recall that we have the following lower bound on the PCS.

$$\text{PCS} \geq 1 - \sum_{i \neq b, i \in \mathscr{I}} \mathbb{P}\left\{\hat{H}_b(\hat{\theta}_N) - \hat{H}_i(\hat{\theta}_N) \leq 0\right\}.$$

From theorem 1 we know that for N large, $\hat{H}_b(\hat{\theta}_N) - \hat{H}_i(\hat{\theta}_N) \overset{\mathscr{D}}{\approx} \mathscr{N}(\delta_{bi}, \sigma_{bi}^2)$. Using this fact, we have the following approximate PCS.

$$\text{APCS} := 1 - \sum_{i \neq b, i \in \mathscr{I}} \int_{-\infty}^{-\frac{\delta_{bi}}{\sigma_{bi}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

The APCS here is almost the same as that in OCBA except that σ_{bi} now depends on $\sigma_{H_{bi}}^2$ and the input data size N. Since σ_{bi}^2 is a continuous function of N and M_i's, we may ignore the minor technicality that N and M_i's are integers. Following the derivation of OCBA, we drop the nonnegativity constraints for the moment. So the problem becomes

$$\begin{aligned} &\max_{N, M_1, \ldots, M_K} \text{APCS} \\ &\text{s.t.} \quad c_1 N + c_2 \sum_{i=1}^{K} M_i = T. \end{aligned} \tag{10}$$

The Lagrangian function is given by

$$F := \text{APCS} + \lambda \left(c_1 N + c_2 \sum_{i=1}^{K} M_i - T \right).$$

Then, from the Karush-Kuhn-Tucker (KKT) conditions we have

$$\frac{\partial F}{\partial N} = c_1 \lambda - \sum_{i \neq b} \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_{bi}^2}{2\sigma_{bi}^2} \right) \frac{\delta_{bi} \sigma_{H_{bi}}^2}{\sigma_{bi}^3 N^2} = 0, \tag{11}$$

$$\frac{\partial F}{\partial M_i} = c_2 \lambda - \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_{bi}^2}{2\sigma_{bi}^2} \right) \frac{\delta_{bi} \sigma_i^2}{\sigma_{bi}^3 {M_i}^2} = 0 \quad \forall i \neq b, \tag{12}$$

$$\frac{\partial F}{\partial M_b} = c_2 \lambda - \sum_{i \neq b} \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_{bi}^2}{2\sigma_{bi}^2} \right) \frac{\delta_{bi} \sigma_b^2}{\sigma_{bi}^3 {M_b}^2} = 0, \tag{13}$$

$$N + \sum_{i=1}^{K} M_i = T. \tag{14}$$

Note that from (12) we have

$$\frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{\delta_{bi}^2}{2\sigma_{bi}^2} \right) \frac{\delta_{bi}}{\sigma_{bi}^3} = c_2 \lambda \frac{M_i^2}{\sigma_i^2},$$

and plugging this into (11) yields

$$N = \sqrt{\frac{c_2}{c_1} \sum_{i \neq b} \frac{M_i^2 \sigma_{H_{bi}}^2}{\sigma_i^2}}. \tag{15}$$

Once M_i's are determined, (15) can be interpreted in a very intuitive way. First, how big N is depends on the cost ratio c_2/c_1: the more expensive the input data is compared with simulation, the fewer input data

samples we should collect. Second, N is related to the weighted sum of $M_i^2/\sigma_i^2, i \neq \mathscr{I}$, where the weights are the asymptotic variances $\sigma^2_{H_{bi}}$. From definition we know that $\sigma^2_{H_{bi}}$ depends on the gradient difference $\nabla H_b(\theta^c) - \nabla H_i(\theta^c)$, which is the relative sensitivity information.

To understand the sensitivity information, consider an extreme case where there exists constants C_i such that $H_i(\theta) = C_i + \theta$. As θ changes, the systems' mean performances simply shift by the same amount and their order is always preserved. Therefore, we do not need to collect any input data: simply plug in any $\theta \in \Theta$ would work. Notice that in this example, $\nabla H_i(\theta^c)$ is the same for all systems, so $\sigma^2_{H_{bi}} = 0$ and (15) tells us $N = 0$. Similarly, a bigger $\sigma^2_{H_{bi}}$ means that system i has a larger impact on N, because the difference between H_b and H_i is very sensitive to the estimation error of $\hat{\theta}_N$. The importance of using sensitivity information is also observed in Song (2016).

The last step is to solve for the M_i's. Unfortunately, we cannot find simple analytical expressions for them. Instead, we will resort to the well-known noise-to-signal allocation rule result from OCBA:

$$M_b = \sigma_b \sqrt{\sum_{i \neq b} \frac{M_i^2}{\sigma_i^2}}, \qquad \frac{M_i}{M_j} = \frac{\sigma_i^2/\delta_{bi}^2}{\sigma_j^2/\delta_{bj}^2}, \quad i \neq j \neq b. \tag{16}$$

In other words, we will use (16) as an approximate solution to (10). The reasons are threefold: (i) the problem's dimension scales up as the number of systems increases; (ii) during implementation, we only have noisy estimates of $\sigma^2_{H_{bi}}, \delta_{bi}$, etc., so there is no need to solve (10) exactly; (iii) keep in mind that we actually want to solve a dynamic problem, and the OCBA allocation rule is known to possess good and robust dynamic performance.

Let $\alpha_0 := N/T$ and $\alpha_i := M_i/T$. We summarize the approximate solution as follows.

$$\begin{aligned} &\alpha_0 = \sqrt{\frac{c_2}{c_1} \sum_{i \neq b} \frac{\alpha_i^2 \sigma^2_{H_{bi}}}{\sigma_i^2}}, && \alpha_b = \sigma_b \sqrt{\sum_{i \neq b} \frac{\alpha_i^2}{\sigma_i^2}} \\ &\frac{\alpha_i}{\alpha_j} = \frac{\sigma_i^2/\delta_{bi}^2}{\sigma_j^2/\delta_{bj}^2}, \quad i \neq j \neq b, && c_1\alpha_0 + c_2 \sum_{i=1}^{K} \alpha_i = 1. \end{aligned} \tag{17}$$

In (17), the fractions of N and M_i's of T are independent of T, so this approximate solution is a fixed allocation scheme. Nevertheless, we do not follow this fixed allocation in implementation. Instead, we only use (17) to determine N, whereas M_i's are computed using an adaptive algorithm. To justify this, observe that the PCS can be written as

$$\text{PCS} = \mathbb{P}\{\hat{b} = b \mid \hat{\theta}_N \in \mathscr{P}\} \cdot \mathbb{P}\{\hat{\theta}_N \in \mathscr{P}\} + \mathbb{P}\{\hat{b} = b \mid \hat{\theta}_N \notin \mathscr{P}\} \cdot \mathbb{P}\{\hat{\theta}_N \notin \mathscr{P}\},$$

where $\mathscr{P}$ is the perturbation zone defined in (5). When N is fixed, $\mathbb{P}\{\hat{\theta}_N \notin \mathscr{P}\}$ is fixed and we can only control the conditional PCS. If we are unlucky and $\hat{\theta}_N \in \mathscr{P}$, no allocation rule can guarantee a good $\mathbb{P}\{\hat{b} = b \mid \hat{\theta}_N \in \mathscr{P}\}$. Thus, our only hope is to maximize $\mathbb{P}\{\hat{b} = b \mid \hat{\theta}_N \notin \mathscr{P}\}$. But if $\hat{\theta}_N \notin \mathscr{P}$, the problem reduces to the traditional computing budget allocation problem with no input uncertainty, where OCBA can be used to effectively solve the problem. From this perspective, the key to solving (6) is to choose a reasonable N. After that, any state-of-the-art allocation procedure can be called as a submodule.

3.3 Algorithm Design

Several aspects need to be taken into account when it comes to designing an implementable algorithm. On the one hand, some intermediate quantities in (17) are unknown and must be estimated; on the other hand, these estimates are subject to error, and we want to correct the error as the total budget increases. This requires us to design a sequential allocation scheme with online updating, and it turns out that some difficulties arise in the presence of input uncertainty. Let us first look at the estimation part.

Algorithm 1 OCBA

Input: $\hat{\theta}_N, M_0, \Delta, B$.
Initialization: Run M_0 replications for all systems. Compute $\hat{H}_i(\hat{\theta}_N)$ and $\hat{\sigma}_i^2(\hat{\theta}_N)$ via (18). Set $M_i \leftarrow M_0$ for all i, and $B' \leftarrow KM_0 + \Delta$.
while $\sum_{i \in \mathscr{I}} M_i \leq B$ **do**
    Compute $\beta_1, \ldots, \beta_K$ using (16).
    for $i = 1, \ldots, K$ **do**
        Run $\max\{0, \lfloor \beta_i B' \rfloor - M_i\}$ replications for system i.
        $M_i \leftarrow \max\{M_i, \lfloor \beta_i B' \rfloor\}$.
        Update $\hat{H}_i(\hat{\theta}_N)$ and $\hat{\sigma}_i^2(\hat{\theta}_N)$.
    end for
    $B' \leftarrow B' + \Delta$.
end while
Output: $\hat{b} := \arg\max_{i \in \mathscr{I}} \hat{H}_i(\hat{\theta}_N)$.

The intermediate quantities to be estimated are (i) the parameter θ^c; (ii) the mean performance $H_i(\theta^c)$; (iii) the performance variance σ_i^2; (iv) the sensitivity information $\nabla H_i(\theta^c)$. If no input data has been collected, then first use a small budget to collect N_0 (e.g., 20) input data samples. Then, MLE can be used to get an estimate $\hat{\theta}_N$, and we can directly use the input data to run N_0 simulation replications for each system to get $h_i(\xi_r)$'s. Now $H_i(\theta^c)$'s and $\sigma_i^2(\theta^c)$'s are estimated via

$$\hat{H}_i(\theta^c) = \frac{1}{N_0} \sum_{r=1}^{N_0} h_i(\xi_r), \quad \hat{\sigma}_i^2 = \frac{1}{N_0 - 1} \sum_{r=1}^{N_0} [h_i(\xi_r) - \hat{H}_i(\theta^c)]^2. \tag{18}$$

There are many ways to estimate the gradient $\nabla H_i(\theta^c)$ (see e.g., Fu (2006) for a review). Since we are considering a parametric case, we assume that $F(\cdot;\theta)$ has a density $f(\cdot;\theta)$ to which we can apply the likelihood ratio (LR) method (or the score function method), i.e., use

$$\frac{1}{N_0} \sum_{r=1}^{N_0} h_i(\xi_r) \frac{\nabla_\theta f(\xi_r; \theta^c)}{f(\xi_r; \theta^c)} \tag{19}$$

as an unbiased and consistent estimator of $\nabla H(\theta^c)$. In practice, θ^c is replaced by $\hat{\theta}_N$ in (19). We may also apply LR if F has a probability mass function (p.m.f.) which is differentiable w.r.t. θ (e.g., Bernoulli distribution). The benefit of using LR is that once we have the simulation outputs $h_i(\xi_r)$, it takes almost no extra time to compute the weighted average (19). The downside, however, is that the LR estimator may have a large variance due to the change of measure. Since the focus of this work is not gradient estimation, we refer interested readers to Rubinstein and Shapiro (1993) for more extensive discussions.

Before we present a two-stage algorithm, let us briefly review the OCBA algorithm in our context. The pseudo code of OCBA is given in algorithm 1. Apart from the aforementioned quantities, we also need to specify Δ, the increment of the current budget B' at each iteration. Here the budget B refers to the total number of runs, and $\beta_i = M_i/(\sum_i M_i)$. To initialize the algorithm, we use M_0 replications to get a rough estimate of the mean and variance of each system's performance. Then we enter the loop, where the current budget B' gets increased by Δ at every iteration. The variable M_i represents the budget that has been allocated to system i. We only run more replications if $\lfloor \beta_i B' \rfloor > M_i$, because even if $\lfloor \beta_i B' \rfloor < M_i$, we cannot take back those $M_i - \lfloor \beta_i B' \rfloor$ replications. New simulation outputs are used to update the $\hat{H}_i(\hat{\theta}_N)$ and $\hat{\sigma}_i^2(\hat{\theta}_N)$. After the total budget is exhausted, the system with the highest sample mean is selected as the best system.

Now we are ready to present a two-stage algorithm, OCBAIU, which stands for "OCBA under Input Uncertainty". The details of OCBAIU are provided in algorithm 2. OCBAIU starts off by collecting a small

Algorithm 2 OCBAIU

1: **Input**: $c_1, c_2, N_0, M_0, \Delta, T$.
2: **Initialization:** Collect N_0 input data samples and compute $\hat{\theta}_N$ using MLE. Run M_0 replications for all systems using the input data. Compute $\hat{H}_i(\theta^c)$ and $\hat{\sigma}_i^2$ via (18), and compute the gradients via (19).
3: Use (17) to determine α_0.
4: **if** $\lfloor \alpha_0 T \rfloor > N_0$ **then**
5: Collect additional $\lfloor \alpha_0 T \rfloor - N_0$ input data samples.
6: Update $\hat{\theta}_N$. Set $N \leftarrow \lfloor \alpha_0 T \rfloor$.
7: **end if**
8: $B \leftarrow \lfloor (T - c_1 N - c_2 K N_0)/c_2 \rfloor$. Call algorithm 1.
9: **Output:** $\hat{b} := \arg\max_{i \in \mathscr{I}} \hat{H}_i(\hat{\theta}_N)$.

number of input data samples, which are fed directly into simulation for estimating intermediate quantities. After N is determined, OCBA is used to solve the second-stage problem, i.e., finding $\arg\max_i H_i(\hat{\theta}_N)$.

4 NUMERICAL STUDY

The best way to demonstrate OCBAIU's performance is to show how the PCS changes as the total budget T increases. For illustrative purpose, we apply our algorithm to a simple quadratic case,

$$h_i(\xi) = -(i-\xi)^2, \quad \mathscr{I} = \{-3,-2,-1,\ldots,5,6\}.$$

In words, we have ten systems corresponding to i ranging from -3 to 6. We assume that the input data ξ follows an exponential distribution with rate $\theta^c = 0.5$. Then H_i has a closed form

$$H_i(\theta) = -\mathbb{E}[(i-\xi)^2] = -\frac{2}{\theta^2} + \frac{2i}{\theta} - i^2.$$

The optimal design is $i = \mathbb{E}\xi = 2$, namely, $b = 6$. The LR estimator of $H_i(\theta^c)$ is given by

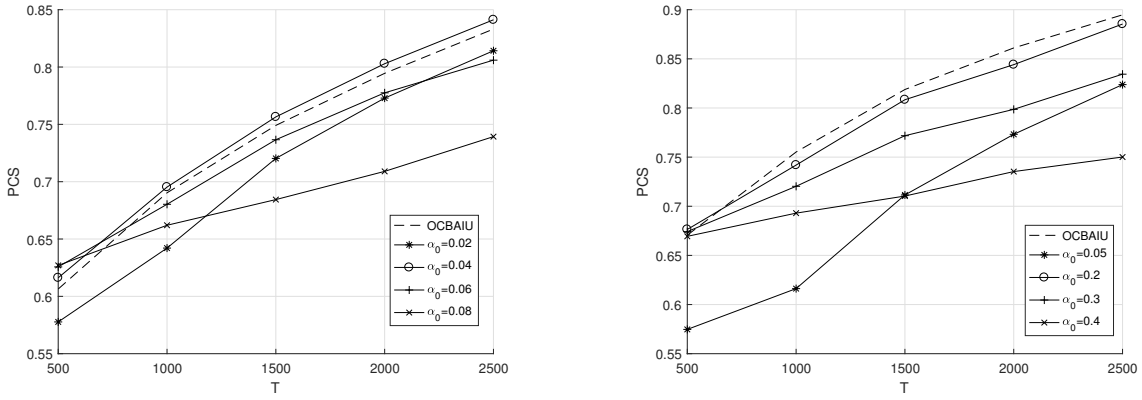
$$\widehat{\nabla H_i(\theta^c)} = \frac{1}{M_i} \sum_{r=1}^{M_i} h_i(\hat{\xi}_{ir}) \left(\frac{1}{\hat{\theta}_N} - \hat{\xi}_{ir} \right).$$

To study how the cost parameters c_1 and c_2 affects OCBAIU's allocation, we study two cases: (i) $c_1 = 10, c_2 = 1$ and (ii) $c_1 = 2, c_2 = 1$. Let the total budget T increase from 500 to 2500 with an increment of 500. For a given budget, the PCS is estimated using 10,000 independent replications. We set $N_0 = M_0 = 20$ and $\Delta = 50$. To examine how well OCBAIU determines α_0, we also compare OCBAIU's allocation with α_0 fixed at different levels. The resulting PCS curves are shown in Figure 1.

In the first case, OCBAIU outperforms $\alpha_0 = 0.02, 0.06$ and 0.08 while it is slightly worse than $\alpha = 0.04$. In the second case, OCBAIU is better than all fixed values of α_0. Also, the α_0 computed by OCBAIU increases from around 0.04 to 0.2 as the cost of input data decreases from 10 to 2. This agrees with our intuition that more input data should be bought if it is cheaper, and it demonstrates the adaptiveness of our algorithm. In particular, if a decision maker does not make use of OCBAIU and simply allocates half (or any arbitrary portion) of the budget to collecting input data, then α_0 will be 0.05 and 0.25 in these two cases, both of which will lead to an inferior PCS than OCBAIU.

Admittedly, the performance of OCBAIU is not totally satisfactory, mostly due to the fact that OCBAIU is a "semi-heursitic" algorithm based on many approximations. It is possible to find an even better α_0 by enumerating more possibilities. However, in practice running the PCS curves is extremely expensive, and brute-force enumeration is infeasible. Furthermore, note that the formulation we consider is a complex two-stage stochastic program which does not have a well-studied solution. The purpose of this paper is to draw attention to this problem and provide one efficient method to solve it. Two ways to improve the algorithm might be the following.

1. Explore the structure of PCS to see if a better approximation is possible. A deeper understanding of the objective function is always helpful.
2. Improve the algorithm for the case of no input uncertainty. This is equivalent to improving the solution quality of the second-stage problem.



(a) Case i: $c_1 = 10, c_2 = 1$.

(b) Case ii: $c_1 = 2, c_2 = 1$.

Figure 1: PCS curves.

5 CONCLUSION AND FUTURE WORK

This paper considers a new formulation for simulation budget allocation under input uncertainty. By studying the asymptotics of a system's performance, we apply the framework of OCBA and develop an algorithm, OCBAIU, which can be viewed as an extension of OCBA to balance the effort in learning the input model and doing optimization. Nevertheless, to further enhance the two-stage algorithm towards a multi-stage sequential allocation procedure, we must perform online updating in an effective way so that the PCS converges fast. Another open question is whether the PCS still has exponential decay in the presence of input uncertainty. These are some of the fundamental questions we hope to answer in our future work.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation under Grant CAREER CMMI-1453934, and Air Force Office of Scientific Research under Grant YIP FA-9550-14-1-0059.

REFERENCES

Audibert, J.-Y., and S. Bubeck. 2010. "Best arm identification in multi-armed bandits". In *COLT-23th Conference on Learning Theory-2010*, 13–p.

Branke, J., S. E. Chick, and C. Schmidt. 2005. "New developments in ranking and selection: an empirical comparison of the three main approaches". In *Proceedings of the 2005 conference on Winter simulation*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 708–717. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Casella, G., and R. L. Berger. 2002. *Statistical inference*, Volume 2. Duxbury Pacific Grove, CA.

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. 2013. "A subset selection procedure under input parameter uncertainty". In *Simulation Conference (WSC), 2013 Winter*, edited by R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, 463–473. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Durrett, R. 2010. *Probability: theory and examples*. Cambridge university press.

Fu, M. C. 2006. "Gradient estimation". *Handbooks in operations research and management science* 13:575–616.

Gao, S., H. Xiao, E. Zhou, and W. Chen. 2016. "Optimal computing budget allocation with input uncertainty". 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, 839–846. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Kim, S.-H., and B. L. Nelson. 2007. "Recent advances in ranking and selection". In *Proceedings of the 2007 Winter Simulation Conference*, edited by S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 162–172. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lehmann, E. L., and G. Casella. 2006. *Theory of point estimation*. Springer Science & Business Media.

Rubinstein, R. Y., and A. Shapiro. 1993. *Discrete event systems: sensitivity analysis and stochastic optimization by the score function method*. John Wiley & Sons Inc.

Russo, D. 2016. "Simple bayesian algorithms for best arm identification". *arXiv preprint arXiv:1602.08448*.

Song, E. 2016. "Input-output uncertainty comparisons for optimization via simulation". 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, 3666–3667. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

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.

AUTHOR BIOGRAPHIES

DI WU is a Ph.D. student in the H. Milton Stewart School of Industrial and Systems Engineering at Georgia Institute of Technology. He received his B.E. in electrical engineering from East China University of Science and Technology, China, in 2011, and 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, with applications towards financial engineering. Her email address is enlu.zhou@isye.gatech.edu and her web page is http://enluzhou.gatech.edu/.