

ADAPTIVE SIMULATION BUDGET ALLOCATION FOR DETERMINING THE BEST DESIGN

Qi Fan
Jiaqiao Hu

Department of Applied Mathematics and Statistics
State University of New York
Stony Brook, NY 11794, USA

ABSTRACT

We consider the problem of allocating a given simulation budget among a set of design alternatives in order to maximize the probability of correct selection. Prior work has focused on deriving static rules that predetermine the number of simulation replications to be allocated to each design. In contrast, we formulate the problem as a Markov decision process (MDP) and propose a dynamic myopic scheme to adaptively allocate simulation samples based on current estimates of the means and variances of the design alternatives. We provide numerical examples to illustrate the performance of the proposed dynamic allocation rule.

1 INTRODUCTION

We consider the problem of identifying the best design from a finite set of design alternatives. Each design is assumed to involve random uncertainty and requires stochastic simulation for performance estimation. When simulation is expensive and the number of design alternatives is relatively small, a well-known class of procedures for solving such problems is ranking and selection, where the goal is to determine the number of simulation runs to be allocated to each design in order to guarantee a pre-specified correct selection probability. Examples of ranking and selection methods include Rinott's two-stage indifference zone procedure (Rinott 1978), the expected value of information procedure (Chick and Inoue 2001), and the KN family of algorithms (Kim and Nelson 2006); some reviews and advances in this field can be found in e.g., Goldsman and Nelson (1998), Nelson et al. (2001), and Kim and Nelson (2007).

Chen (1995) approached the problem from a different perspective by determining the best allocation of a given simulation budget among the designs in order to maximize the probability of correct selection (PCS). In particular, Chen (1996) proposed to use a Bayesian approach to estimate the design performance measures based on prior sampling information and derived a lower bound for the correct selection probability. The idea was subsequently used in Chen, Chen, and Yücesan (2000) and Chen et al. (2000) to develop analytical allocation rules called Optimal Computing Budget Allocation (OCBA) that asymptotically optimize the lower bounds of the probability of correct selection. More recently, Chen et al. (2006) also investigated a dynamic allocation rule based on perfect information assumption, and suggested that a dynamic scheme could dramatically improve the performance of static allocation rules.

In this paper, we consider the setting of OCBA, i.e., maximizing PCS under a simulation budget constraint. However, unlike previous work, which has primarily focused on static rules that predetermine the number of simulation replications to be allocated to each design, we investigate a dynamic programming approach that adaptively allocates simulation samples based on current estimates of the means and variances of various designs. Our work can be viewed as an extension of that of Chen et al. (2006) based on perfect information. In particular, we model the simulation allocation process as a Markov decision process (MDP) with a terminal cost function. The state variable of the MDP model consists of the current sample mean of each design and the number of simulation replications allocated to each of them by assuming that

design variances are known. These variances can then be estimated by sample variances. Since analytically solving the MDP model is intractable, we further develop an upper bound to the optimal value function and propose a one-step lookahead index policy to myopically minimize the sum of the current one-stage cost and the upper bound of the optimal value function. Our preliminary numerical results indicate competitive performance of our approach with that of OCBA, especially when the simulation budget is small.

The rest of the paper is organized as follows: In Section 2, we define notations and describe the problem setting. In Section 3, we formulate the problem as an MDP, provide an upper bound to the optimal value function, and derive a myopic index policy for simulation allocation. Numerical examples are provided in Section 4 to illustrate the performance of our approach. Finally, we conclude the paper in Section 5.

2 PROBLEM SETTING

Consider the following optimization problem:

$$\min_{i \in \Theta} J_i \equiv \min_{i \in \Theta} E[L_i(\xi)]$$

where $\Theta = \{1, 2, \dots, k\}$ is a finite set of design indices and J_i is the true performance measure of design i . Note that J_i itself is the expectation of the sample performance $L_i(\xi)$, where the expectation is understood with respect to the distribution of the random variable ξ representing the stochastic uncertainty of the design. We assume that the expectation cannot be evaluated exactly; however, for a given simulation budget t , the performance measure J_i can be estimated by the sample mean

$$\bar{J}_i^t \equiv \frac{1}{N_i^t} \sum_{j=1}^{N_i^t} L_i(\xi_{ij}),$$

where N_i^t represents the number of replication runs allocated to design i and ξ_{ij} represents the j th realization of ξ (simulation sample path) from design i . Throughout this paper, we assume that the simulation outputs are independent of each other.

We begin by defining some notations.

σ_i^2 : the true variance of design i , i.e., $\sigma_i^2 = \text{Var}(L_i(\xi))$, which can be estimated by its sample variance.

b_t : the index of the design that shows the current best sample performance after t simulation replications have been allocated, i.e., $\bar{J}_{b_t}^t \leq \min_i \bar{J}_i^t$.

s_t : the index of the design that shows the second best sample performance after t simulation runs have been allocated, i.e., $\bar{J}_{s_t}^t \leq \min_{i \neq b_t} \bar{J}_i^t$.

$\delta_{b_t, i}^t = \bar{J}_{b_t}^t - \bar{J}_i^t$: the difference between the sample performance of the current best and the i th designs.

$\sigma_{b_t, i}^t = \sqrt{\frac{\sigma_{b_t}^2}{N_{b_t}^t} + \frac{\sigma_i^2}{N_i^t}}$: the standard deviation of the random variable $\delta_{b_t, i}^t$.

Define the event of correct selection (CS) as the event that design b_t (i.e., the one with the current best sample performance) is actually the best design. For a given simulation budget, the goal is to find a way to maximize the probability of correct selection $P\{CS\}$.

We follow the Bayesian approach introduced in Chen (1996) and assume that the output performance measure $L_i(\xi)$ is normally distributed for each design. Let \tilde{J}_i be a random variable whose distribution is the posterior distribution of design i given the previous sampling information:

$$P\{\tilde{J}_i\} = P\{J_i | L_i(\xi_{ij}), j = 1, 2, \dots, N_i^t\},$$

for $i = 1, 2, \dots, k$. It can be shown that if no priori knowledge is given on the performance of each design, \tilde{J}_i has the normal distribution $\tilde{J}_i \sim N(\bar{J}_i, \frac{\sigma_i^2}{N_i^t})$. Thus, a lower bound to $P\{CS\}$ called approximate probability

of correct selection (APCS) can be obtained by applying Bonferroni's inequality.

$$\begin{aligned}
 P\{CS\} &= P\{J_{b_t} < J_i, \forall i \neq b_t | L_i(\xi_{ij}), j = 1, 2, \dots, N_i^t, i = 1, 2, \dots, k\} \\
 &= P\{\tilde{J}_{b_t} < \tilde{J}_i, \forall i \neq b_t\}, \\
 &= P\left\{\bigcap_{i \neq b_t} (\tilde{J}_{b_t} - \tilde{J}_i < 0)\right\} \\
 &\geq 1 - \sum_{i \neq b_t} P\{\tilde{J}_{b_t} - \tilde{J}_i > 0\} \\
 &= 1 - \sum_{i \neq b_t} \Phi\left(\frac{\delta_{b_t,i}^t}{\sigma_{b_t,i}^t}\right) \\
 &= APCS,
 \end{aligned}$$

where throughout this paper, we use Φ and ϕ to denote the c.d.f. and p.d.f. of the standard normal distribution.

Since $P\{CS\}$ is difficult to evaluate, whereas APCS can be computed analytically without resorting to additional simulation effort, Chen et al. (2000) proposed to use APCS as an approximation to the true probability of correct selection. For a given budget T , the goal is to find a simulation budget allocation that solves the following optimization problem.

$$\begin{aligned}
 &\min_{N_1^T, N_2^T, \dots, N_k^T} \sum_{i=1}^k \Phi\left(\frac{\delta_{b_T,i}^T}{\sigma_{b_T,i}^T}\right) \\
 &s.t. \sum_{i=1}^k N_i^T \leq T \quad \text{and} \quad N_i^T \geq 0.
 \end{aligned} \tag{1}$$

3 A DYNAMIC BUDGET ALLOCATION PROCEDURE

Motivated by the work of Hu et al. (2011) and that of Chen et al. (2006), we aim to solve the allocation problem by modeling the allocation process as an MDP model with a terminal cost function. So instead of allocating all T simulation replications at the beginning as in (1), we derive a dynamic policy that sequentially allocates the budget based on the estimated performance measure of all designs as well as the current sampling information.

3.1 Modeling the Allocation Process as an MDP

Given a total of T simulation samples, we start by assigning at step $t = 0$ a small number n_0 simulation replications to each of the k designs. Define the state variable $w_t = (\tilde{J}_i^t, N_i^t, i = 1, 2, \dots, k)^T$ as a vector containing the current performance estimates of all designs and the number of times each design has been sampled thus far, where $N_i^0 = n_0$ and $\tilde{J}_i^0 = \frac{1}{n_0} \sum_{j=1}^{n_0} L_i(\xi_{ij})$ for all i . Next consider a sequential allocation policy π that determines, at each step $t = 1, 2, \dots, T - n_0k$ based on w_t , whether one more replication run should be allocated to one of the k designs or the entire allocation process should be terminated. Let $\pi_t(w_t) \in \{0, 1, \dots, k\}$ be the action taken at step t , then

$$\pi_t(w_t) = \begin{cases} i & \text{allocate one simulate run to design } i, i = 1, 2, \dots, k \\ 0 & \text{stop the allocation process.} \end{cases} \tag{2}$$

For every allocation policy π described above, it can be seen that $\{w_t\}$ is a Markov chain with the following state transition dynamics:

$$\bar{J}_i^{t+1} = \frac{N_i^t \bar{J}_i^t + Y_i I_{\{a=i\}}}{N_i^t + I_{\{a=i\}}} \quad \text{for } i = 1, 2, \dots, k \quad (3)$$

$$N_i^{t+1} = N_i^t + I_{\{a=i\}} \quad \text{for } i = 1, 2, \dots, k, \quad (4)$$

where I is the indicator function, a is the action taken at step t under π , and Y_i is the simulation output performance measure of design i after the additional allocation. Based on (3) and (4), the updating formula for the current best sample mean is given by:

$$\bar{J}_{b_{t+1}}^{t+1} = \begin{cases} \bar{J}_{s_t} - \left(\bar{J}_{s_t} - \frac{N_{b_t}^t \bar{J}_{b_t}^t + Y_{b_t}}{N_{b_t}^t + I_{\{a=b_t\}}} \right)^+ & \text{if } a = b_t \\ \bar{J}_{b_t} - \left(\bar{J}_{b_t} - \frac{N_a^t \bar{J}_a^t + Y_a I_{\{a=i\}}}{N_a^t + I_{\{a=i\}}} \right)^+ & \text{if } a \neq b_t, \end{cases} \quad (5)$$

where $Z^+ = \max\{0, Z\}$.

Let $w = (\bar{J}_i, N_i, i = 1, \dots, k)^T$ be a given state and define $b = \arg \min_i \bar{J}_i$, $\delta_{b,i} = \bar{J}_b - \bar{J}_i$, and $\sigma_{b,i} = \sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_i^2}{N_i}}$. By associating the state action pair (w, a) with the following one-stage cost function:

$$R_t(w, a) = \begin{cases} 0 & \text{if } a \neq 0 \\ \sum_{i=1}^k \Phi\left(\frac{\delta_{b,i}}{\sigma_{b,i}}\right) & \text{if } a = 0 \end{cases}$$

and $R_s(w, a) = 0$ for all $s \geq t + 1$ whenever $\pi_t(w) = 0$, we obtain a $(T - n_0k + 1)$ -horizon MDP with the total cost

$$V^\pi(w) = E \left[\sum_{t=0}^{T-n_0k} R_t(w_t, \pi_t(w_t)) \mid w_0 = w \right],$$

where the expectation is taken respect to the probability measure induced by π . For a given initial state $w_0 = w$, the objective is to find an optimal simulation allocation policy π^* to minimize the total cost accumulated before the allocation process terminates.

3.2 A Myopic Index Policy

Since obtaining the exact optimal policy for the MDP model is intractable, we derive a myopic index policy using one-step lookahead optimization. The following result provides an upper bound to the optimal value function.

Theorem 1 Let V_t be the optimal cost-to-go function at stage t of the MDP defined in Section 3.1. For every $t = 0, 1, 2, \dots, T - n_0k$ and $w = (\bar{J}_i, N_i, i = 1, \dots, k)^T$, we have

$$V_t(w) \leq \sum_{i=1}^k \Phi\left(\frac{\delta_{b,i}}{\sigma_{b,i}}\right). \quad (6)$$

Proof. At the final stage $t = T - n_0k$, all T replications have been exhausted, so the only option is to stop the allocation process. Therefore, we must have

$$\pi_{T-n_0k}^*(w) = 0 \text{ and } V_{T-n_0k}(w) = \sum_{i=1}^k \Phi\left(\frac{\delta_{b,i}}{\sigma_{b,i}}\right).$$

It follows that when $t = T - n_0k - 1$,

$$V_{T-n_0k-1}(w) = \min_a E_a [R_{T-n_0k-1}(w, a) + V_{T-n_0k}(w^a)]$$

where $a \in \{0, 1, 2, \dots, k\}$ and $w' = (\bar{J}'_i, N'_i, i = 1, 2, \dots, k)^T$ is the next state generated according to the transition dynamics (3) and (4) when action a is taken, in particular, $\bar{J}'_i = \frac{N_i \bar{J}_i + Y_i I_{\{a=i\}}}{N_i + I_{\{a=i\}}}$ and $N'_i = N_i + I_{\{a=i\}}$. Therefore,

$$\begin{aligned} V_{T-n_0k-1}(w) &= \min_a E_a [R_{T-n_0k-1}(w, a) + V_{T-n_0k}(w')] \\ &= \min \left\{ \sum_{i=1}^k \Phi \left(\frac{\delta_{b,i}}{\sigma_{b,i}} \right), \min_{a \neq 0} E_a \left[\sum_{i=1}^k \Phi \left(\frac{\delta_{b',i}}{\sigma_{b',i}} \right) \right] \right\} \\ &\leq \sum_{i=1}^k \Phi \left(\frac{\delta_{b,i}}{\sigma_{b,i}} \right), \end{aligned}$$

where $b' = \arg \min_i \bar{J}'_i$, $\delta_{b',i} = \bar{J}_{b'} - \bar{J}'_i$, and $\sigma_{b',i} = \sqrt{\frac{\sigma_{b'}^2}{N_{b'}} + \frac{\sigma_i^2}{N_i}}$.

Now proceed by induction and assume that $V_{t+1}(w) \leq \sum_{i=1}^k \Phi \left(\frac{\delta_{b,i}}{\sigma_{b,i}} \right)$ for all w . Then

$$\begin{aligned} V_t(w) &= \min_a E_a [R_t(w, a) + V_{t+1}(w')] \\ &\leq \min \left\{ \sum_{i=1}^k \Phi \left(\frac{\delta_{b,i}}{\sigma_{b,i}} \right), \min_{a \neq 0} E_a \left[\sum_{i=1}^k \Phi \left(\frac{\delta_{b',i}}{\sigma_{b',i}} \right) \right] \right\} \\ &\leq \sum_{i=1}^k \Phi \left(\frac{\delta_{b,i}}{\sigma_{b,i}} \right). \end{aligned}$$

This completes the proof of the theorem. □

Motivated by Theorem 1, we propose a simple (stationary) greedy policy that minimizes the sum of the current one-stage cost function and the upper bound of the optimal cost-to-go function at each step:

$$\hat{\pi}(w) = \begin{cases} 0 & \text{if } \sum_{i=1}^k \Phi \left(\frac{\delta_{b,i}}{\sigma_{b,i}} \right) \leq \min_{a \neq 0} E_a \left[\sum_{i=1}^k \Phi \left(\frac{\delta_{b',i}}{\sigma_{b',i}} \right) \right] \\ \arg \min_{a \neq 0} E_a \left[\sum_{i=1}^k \Phi \left(\frac{\delta_{b',i}}{\sigma_{b',i}} \right) \right] & \text{otherwise.} \end{cases} \quad (7)$$

By connecting $\hat{\pi}$ to (1), it is not difficult to see that if one more simulation sample is needed, such a policy myopically allocates the next sample in such a way so that the APCS in the next step is maximized after the additional allocation. Since $\hat{\pi}$ is an index policy, we can create an index for each action a based on the current state w . Denote by $index(a)$ as the index of action a and let $w' = (\bar{J}'_i, N'_i, i = 1, \dots, k)^T$ be the sampled next state when action a is taken, we have three different cases based on the transition dynamics (3), (4), and (5):

case 1) If $a = 0$,

$$index(0) = \sum_{i=1}^k \Phi \left(\frac{\bar{J}_b - \bar{J}_i}{\sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_i^2}{N_i}}} \right), \quad (8)$$

case 2) If $a = b$,

$$index(b) = E_b \left[\sum_{l=1, l \neq b}^k \Phi \left(\frac{\bar{J}_s - \left(\bar{J}_s - \frac{N_b \bar{J}_b + Y_b}{N_b + 1} \right)^+ - \bar{J}_l}{\sqrt{\frac{\sigma_{b'}^2}{N_{b'}} + \frac{\sigma_l^2}{N_l}}} \right) + \Phi \left(\frac{\bar{J}_s - \left(\bar{J}_s - \frac{N_b \bar{J}_b + Y_b}{N_b + 1} \right)^+ - \frac{N_b \bar{J}_b + Y_b}{N_b + 1}}{\sqrt{\frac{\sigma_{b'}^2}{N_{b'}} + \frac{\sigma_b^2}{N_b}}} \right) \right], \quad (9)$$

case 3) If $a \neq b$ and $a \neq 0$,

$$index(a) = E_a \left[\sum_{l=1, l \neq a}^k \Phi \left(\frac{\bar{J}_b - \left(\bar{J}_b - \frac{N_a \bar{J}_a + Y_a}{N_a + 1} \right)^+ - \bar{J}_l}{\sqrt{\frac{\sigma_{b'}^2}{N_{b'}} + \frac{\sigma_l^2}{N_l}}} \right) + \Phi \left(\frac{\bar{J}_b - \left(\bar{J}_b - \frac{N_a \bar{J}_a + Y_a}{N_a + 1} \right)^+ - \frac{N_a \bar{J}_a + Y_a}{N_a + 1}}{\sqrt{\frac{\sigma_{b'}^2}{N_{b'}} + \frac{\sigma_a^2}{N_a}}} \right) \right], \quad (10)$$

where $b' = \arg \min_i \bar{J}_i$ and $\sigma_{b'}^2$ is the variance of design b' .

Let $B_s = \{ \frac{N_b \bar{J}_b + Y_b}{N_b + 1} \leq \bar{J}_s \}$. The $(\cdot)^+$ operator can be removed from (9) by conditioning on event B_s :

$$\begin{aligned} index(b) &= E_b \left[\sum_{l \neq b} \Phi \left(\frac{\frac{N_b \bar{J}_b + Y_b}{N_b + 1} - \bar{J}_l}{\sqrt{\frac{\sigma_b^2}{N_b + 1} + \frac{\sigma_l^2}{N_l}}} \right) + \frac{1}{2} \middle| B_s \right] \times P(B_s) \\ &+ E_b \left[\sum_{l \neq b} \Phi \left(\frac{\bar{J}_s - \bar{J}_l}{\sqrt{\frac{\sigma_s^2}{N_s} + \frac{\sigma_l^2}{N_l}}} \right) + \Phi \left(\frac{\bar{J}_s - \frac{N_b \bar{J}_b + Y_b}{N_b + 1}}{\sqrt{\frac{\sigma_s^2}{N_s} + \frac{\sigma_b^2}{N_b + 1}}} \right) \middle| B_s^c \right] \times P(B_s^c). \end{aligned} \quad (11)$$

Similarly, by conditioning on $B_b = \{ \frac{N_b \bar{J}_b + Y_b}{N_b + 1} \leq \bar{J}_b \}$, the index in case 3) can be obtained as

$$\begin{aligned} index(a) &= E_a \left[\sum_{l \neq a} \Phi \left(\frac{\frac{N_a \bar{J}_a + Y_a}{N_a + 1} - \bar{J}_l}{\sqrt{\frac{\sigma_a^2}{N_a + 1} + \frac{\sigma_l^2}{N_l}}} \right) + \frac{1}{2} \middle| B_b \right] \times P(B_b) \\ &+ E_a \left[\sum_{l \neq a} \Phi \left(\frac{\bar{J}_b - \bar{J}_l}{\sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_l^2}{N_l}}} \right) + \Phi \left(\frac{\bar{J}_b - \frac{N_a \bar{J}_a + Y_a}{N_a + 1}}{\sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_a^2}{N_a + 1}}} \right) \middle| B_b^c \right] \times P(B_b^c). \end{aligned} \quad (12)$$

3.3 A Dynamic Budget Allocation Algorithm

Note that when $a = 0$ the index in case 1) can be calculated analytically, whereas calculating the performance indices in (11) and (12) require evaluating the expectations with respect to the design distributions. One natural approach to evaluate/estimate these expectations is to use Taylor expansion. Taking (11) as an example, we can treat each respective term as a function of the sample mean $\frac{N_b \bar{J}_b + Y_b}{N_b + 1}$ and perform a first order Taylor expansion of the term around \bar{J}_b . In addition, by replacing the true mean of the current best design b with its sample mean \bar{J}_b , $P(B_s)$ can be approximated by $\Phi \left(\frac{(N_b + 1)(\bar{J}_s - \bar{J}_b)}{\sigma_b} \right)$. Thus when $a = b$, we can approximate the index of action b by the following analytical formula:

$$\begin{aligned} \widehat{index(b)} &= \left[\frac{1}{2} + \sum_{l \neq b} \Phi \left(\frac{\bar{J}_b - \bar{J}_l}{\sqrt{\frac{\sigma_b^2}{N_b + 1} + \frac{\sigma_l^2}{N_l}}} \right) \right] \times \Phi \left(\frac{(N_b + 1)(\bar{J}_s - \bar{J}_b)}{\sigma_b} \right) \\ &- \sum_{l \neq b} \phi \left(\frac{\bar{J}_b - \bar{J}_l}{\sqrt{\frac{\sigma_b^2}{N_b + 1} + \frac{\sigma_l^2}{N_l}}} \right) \frac{\sigma_b}{(N_b + 1) \sqrt{\frac{\sigma_b^2}{N_b + 1} + \frac{\sigma_l^2}{N_l}}} \times \phi \left(\frac{(N_b + 1)(\bar{J}_s - \bar{J}_b)}{\sigma_b} \right) \\ &+ \left[\Phi \left(\frac{\bar{J}_s - \bar{J}_b}{\sqrt{\frac{\sigma_s^2}{N_s} + \frac{\sigma_b^2}{N_b + 1}}} \right) + \sum_{l \neq b} \Phi \left(\frac{\bar{J}_s - \bar{J}_l}{\sqrt{\frac{\sigma_s^2}{N_s} + \frac{\sigma_l^2}{N_l}}} \right) \right] \times \left(1 - \Phi \left(\frac{(N_b + 1)(\bar{J}_s - \bar{J}_b)}{\sigma_b} \right) \right) \\ &- \phi \left(\frac{\bar{J}_s - \bar{J}_b}{\sqrt{\frac{\sigma_s^2}{N_s} + \frac{\sigma_b^2}{N_b + 1}}} \right) \frac{\sigma_b}{(N_b + 1) \sqrt{\frac{\sigma_s^2}{N_s} + \frac{\sigma_b^2}{N_b + 1}}} \times \phi \left(\frac{(N_b + 1)(\bar{J}_s - \bar{J}_b)}{\sigma_b} \right). \end{aligned} \quad (13)$$

Similarly, when $a \neq b$ and $a \neq 0$, $index(a)$ can be approximated by

$$\begin{aligned}
 \widehat{index}(a) &= \left[\frac{1}{2} + \sum_{l \neq a} \Phi \left(\frac{\bar{J}_a - \bar{J}_l}{\sqrt{\frac{\sigma_a^2}{N_{a+1}} + \frac{\sigma_l^2}{N_l}}} \right) \right] \times \Phi \left(\frac{(N_a + 1)(\bar{J}_b - \bar{J}_a)}{\sigma_a} \right) \\
 &- \sum_{l \neq a} \phi \left(\frac{\bar{J}_a - \bar{J}_l}{\sqrt{\frac{\sigma_a^2}{N_{a+1}} + \frac{\sigma_l^2}{N_l}}} \right) \frac{\sigma_a}{(N_a + 1)\sqrt{\frac{\sigma_a^2}{N_{a+1}} + \frac{\sigma_l^2}{N_l}}} \times \phi \left(\frac{(N_a + 1)(\bar{J}_b - \bar{J}_a)}{\sigma_a} \right) \\
 &+ \left[\Phi \left(\frac{\bar{J}_b - \bar{J}_a}{\sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_a^2}{N_{a+1}}}} \right) + \sum_{l \neq a} \Phi \left(\frac{\bar{J}_b - \bar{J}_l}{\sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_l^2}{N_l}}} \right) \right] \times \left(1 - \Phi \left(\frac{(N_a + 1)(\bar{J}_b - \bar{J}_a)}{\sigma_a} \right) \right) \\
 &- \phi \left(\frac{\bar{J}_b - \bar{J}_a}{\sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_a^2}{N_{a+1}}}} \right) \frac{\sigma_a}{(N_a + 1)\sqrt{\frac{\sigma_b^2}{N_b} + \frac{\sigma_a^2}{N_{a+1}}}} \times \phi \left(\frac{(N_a + 1)(\bar{J}_b - \bar{J}_a)}{\sigma_a} \right). \tag{14}
 \end{aligned}$$

Finally, by replacing true variances with sample variances, we propose the following algorithm for simulation budget allocation.

Dynamic Simulation Budget Allocation (DSBA)

- Step 0:** Perform n_0 simulation replications for all designs. Calculate the sample mean and sample variance for each design.
- Step 1:** For each action $a \in \{0, \dots, k\}$, compute the index of action a according to (8), (13), and (14).
- Step 2:** Select the action a^* with the smallest index value. If $a^* = 0$, then stop the allocation process; else if $a^* = i$, perform one simulation replication for design i , update the sample mean and sample variance of design i . Increase the number of simulation replications to design i by 1 and go back to Step 1 until the given budget is exhausted.

4 NUMERICAL RESULTS

In this section, we test the proposed DSBA algorithm and compare its performance with that of OCBA on some simple examples. OCBA was derived based on analytically solving the static optimization problem (1). It has been shown in Chen et al. (2000) that the asymptotically optimal solution to the problem as $T \rightarrow \infty$ satisfies the following conditions.

- (1) $\frac{N_i}{N_j} = \left(\frac{\sigma_i/\delta_{b,i}}{\sigma_j/\delta_{b,j}} \right)^2$
- (2) $N_b = \sigma_b \sqrt{\sum_{i=1, i \neq b}^k \frac{N_i^2}{\sigma_i^2}}$,

where N_i is the number of samples allocated to design i , $\delta_{b,i} = \bar{J}_b - \bar{J}_i$, $\bar{J}_b \leq \min_i \bar{J}_i$, and σ_i is the standard deviation of the performance measure for design i , which can be estimated by sample variance. OCBA sequentially allocates a given budget T by splitting it into batches of size Δ . Then at each step, the current computing budget is increased by Δ and a budget allocation is calculated using conditions (1) and (2) based on the updated computing budget. This allocation is then used to determine the number of additional simulation runs need to be allocated to each design. The process continues until all budget has been consumed.

We consider the following examples in our computational experiment.

Example 1 : This is a special example where the best design has a zero variance and the rest two designs have the same performance:

$$X_{1j} \sim N(0, 0^2), X_{2j} \sim N(0.4, 3^2), X_{3j} \sim N(0.4, 3^2)$$

Example 2 : There are five design alternatives with the best design being deterministic:

$$X_{1j} \sim N(0, 0^2), X_{2j} \sim N(0.4, 1.5^2), X_{3j} \sim N(0.4, 3^2), X_{4j} \sim N(1, 3^2), X_{5j} \sim N(2, 3^2)$$

Example 3 : This example is an extension of the previous one with the deterministic design removed:

$$X_{1j} \sim N(0, 1.5^2), X_{2j} \sim N(0.6, 3^2), X_{3j} \sim N(1, 3^2), X_{4j} \sim N(2, 3^2)$$

Example 4 : This is an example with three alternatives, all of which are random:

$$X_{1j} \sim N(1, 1^2), X_{2j} \sim N(1.5, 3^2), X_{3j} \sim N(1.5, 3^2)$$

In our experiment, the initial number of replications n_0 is set to 10 for both DSBA and OCBA. Figure 1 shows the performance of both algorithms for each of the four respective test cases, where the true $P\{CS\}$ in each case is estimated by the proportional of times the best design is found by an algorithm out of 10,000 independent experiments. The figure indicates competitive performance of DSBA with that of OCBA in all test cases. In particular, DSBA outperforms OCBA when the simulation budget is small, whereas OCBA shows slightly better performance when the budget is increased, especially in the last case. Our conjecture is that this is due to the asymptotic optimality of OCBA, whereas DSBA is myopic in nature.

5 CONCLUSION

In this paper, we have introduced a dynamic simulation budget allocation procedure for determining the best design from a set of finite design alternatives. The idea is to use a myopic one-step lookahead policy to approximately solve an underlying MDP characterizing the budget allocation process. Such a policy gives rise to a stationary index rule that adaptively determines at each step which design should be simulated next in order to myopically maximize the approximate probability of correction after the additional allocation. Our preliminary numerical results indicate that our approach may provide competitive performance with that of OCBA, especially when the computing budget is small.

ACKNOWLEDGMENTS

This work was supported in part by the Air Force Office of Scientific Research under Grant FA95501010340 and by the National Science Foundation under Grant CMMI-1130761.

REFERENCES

- Chen, C.-H. 1995. "An Effective Approach to Smartly Allocate Computing Budget for Discrete Event Simulation". In *Proceedings of the 34th IEEE Conference on Decision and Control*, 2598–2605. Piscataway, NJ: IEEE.
- Chen, C.-H. 1996. "A Low Bound for the Correct Subset-Selection Probability and Its Application to Discrete-Event System Simulation". *IEEE Transaction on Automatic Control* 41:1227–1231.
- Chen, C.-H., D. He, and M. Fu. 2006. "Efficient Dynamic Simulation Allocation in Ordinal Optimization". *IEEE Transaction on Automatic Control* 51:2005–2009.
- 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 System: Theory and Applications* 10:251–270.
- Chen, H.-C., C.-H. Chen, and E. Yücesan. 2000. "Computing Efforts Allocation for Ordinal Optimization and Discrete Event Simulation". *IEEE Transaction on Automatic Control* 45:960–964.
- Chick, S. E., and K. Inoue. 2001. "New Two-Stage and Sequential Procedures for Selecting the Best Simulated System". *Operations Research* 49:732–743.
- Goldsman, D., and B. L. Nelson. 1998. "Comparing systems via simulation". In *Handbook of simulation*, edited by J. Banks, 273–306. New York: John Wiley.

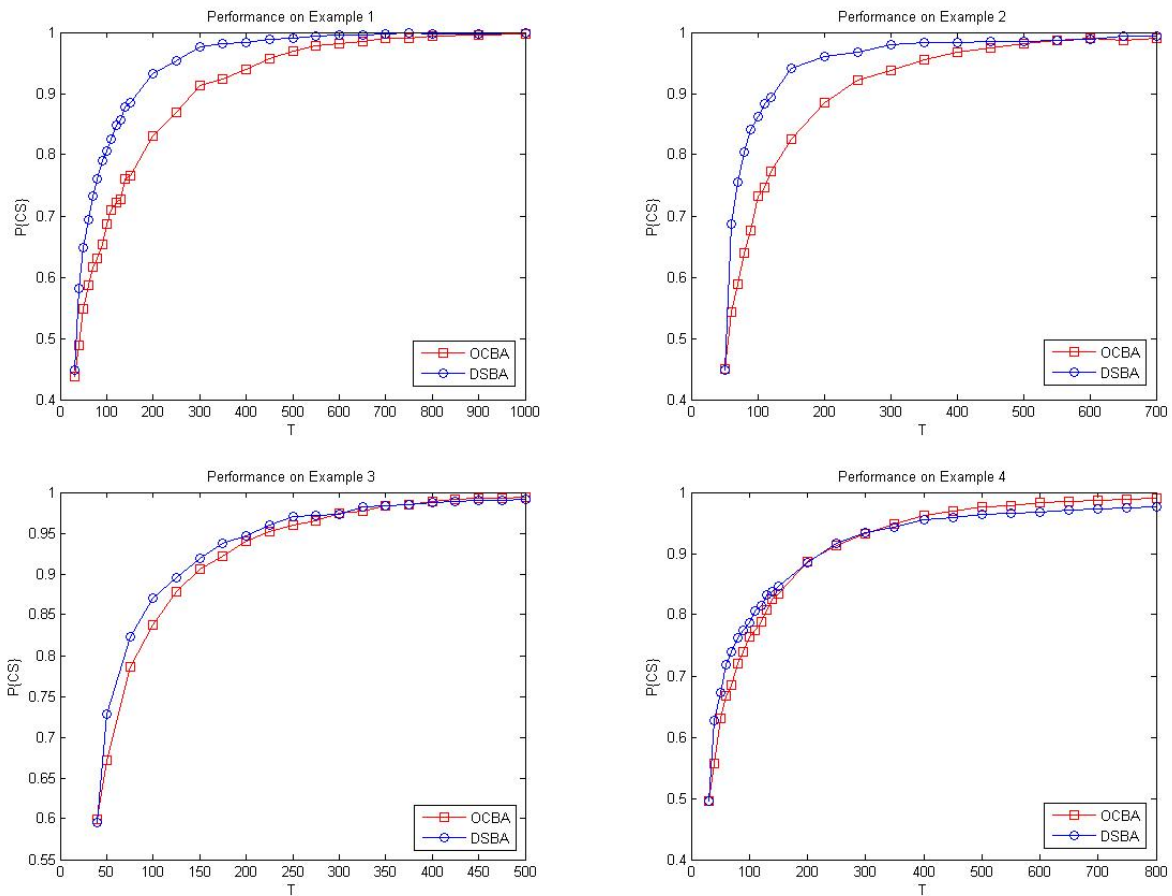


Figure 1: Comparison of OCBA and DSBA.

Hu, J., H. S. Chang, M. C. Fu, and S. I. Marcus. 2011. “Dynamic Sample Budget Allocation in Model-Based Optimization”. *Journal of Global Optimization* 50:575–569.

Kim, S.-H., and B. L. Nelson. 2006. “Selecting the Best System”. In *Handbooks in Operations Research and Management Science: Simulation*, edited by S. G. Henderson and B. L. Nelson, Chapter 17, 501–534. Oxford, UK: Elsevier Science.

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. Henderson, B. Biler, M.-H. Hsieh, J. Shortle, J. Tew, and R. Barton, 162–172. Piscataway, NJ: IEEE.

Nelson, B. L., J. Swann, D. Goldsman, and W. Song. 2001. “Simple Procedures for Selecting the Best Simulated System when the Number of Alternatives Is Large”. *Operations Research* 49:950–963.

Rinott, Y. 1978. “On Two-Stage Selection Procedures and Related Probability Inequalities”. *Communications in Statistics - Theory and Methods* 7:799–811.

AUTHOR BIOGRAPHIES

Qi Fan is a Ph.D. student in the Department of Applied Mathematics and Statistic at the State University of New York, Stony Brook. He received the B.S. degree in mathematics from Zhejiang University, China in 2011. His research interests include Markov decision processes, optimization and simulation. His e-mail address is qfan@ams.sunysb.edu.

JIAQIAO HU is an Associate Professor in the Department of Applied Mathematics and Statistics at the State University of New York, Stony Brook. He received the B.S. degree in automation from Shanghai Jiao Tong University, the M.S. degree in applied mathematics from the University of Maryland, Baltimore County, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park. His research interests include Markov decision processes, applied probability, and simulation optimization. His email address is jqhu@ams.sunysb.edu.