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

AN EFFICIENT FULLY SEQUENTIAL SELECTION PROCEDURE GUARANTEEING PROBABLY APPROXIMATELY CORRECT SELECTION

Sijia Ma
Shane G. Henderson

Operations Research and Information Engineering
Cornell University
Ithaca, NY 14853, USA

ABSTRACT

Ranking and Selection (R&S) procedures are designed for selecting the best among a finite set of systems using stochastic simulation, guaranteeing the quality of the final selection. Instead of assuming a known lower bound on the difference between the best and others, we consider the probably approximately correct (PAC) selection formulation, which ensures a high quality solution with high probability for all configurations. In this paper, we present a new fully sequential selection procedure, called the Envelope Procedure (EP), which accommodates a variety of sampling rules that, together with a carefully defined termination condition, ensures a PAC selection. A particular sampling rule that achieves good efficiency is proposed. We compare the efficiency of the EP with some existing procedures in numerical experiments, and the results show that the EP saves considerable computational effort in many problem configurations.

1 INTRODUCTION

Stochastic simulation is often used in identification of the best among a finite set of alternative systems, where the best refers to the one with highest (or lowest) expected performance. The central problem is to decide how many samples to take from each system so that we have some guarantee on the quality of the selected system. When the number of systems is small enough that each system can be simulated, at least to some degree, such a problem is called a Ranking and Selection (R&S) problem. The goal in designing a R&S procedure is to allocate samples to systems wisely so that a good solution can be achieved while taking as few samples as possible. One may notice that the R&S problems are closely related to the best-arm problems or pure exploration problems in multi-armed bandit (MAB) literature. Though the algorithms designed for best-arm problems usually have different assumption on the simulation outputs, e.g., bounded (Even-Dar, Mannor, and Mansour 2002) or sub-Gaussian outputs (Jamieson et al. 2014), the essence of the problems are quite similar. In the work of Jamieson and Nowak (2014), the best-arm algorithms are divided into three general strategies: action elimination (AE), UCB and LUCB. Though the algorithm we propose in this paper belongs to none of these families, the philosophy in designing its stopping rule and sampling rule can be viewed as a variation of the mixture of them. Thus, we regard this work as a connection between R&S and best-arm problems.

There are two approaches for solving R&S problems in general: 1) given a desired frequentist guarantee on the quality of the selected system, the algorithm collects samples from systems until some stopping criterion for delivering the guarantee is met; 2) given a finite computational budget, the algorithm collects samples from systems in order to maximize some measure of the quality of the selection, e.g., the posterior probability of correct selection. The first approach is often called the frequentist approach, while the second approach is often referred to as the Bayesian approach. Both approaches can collect samples either deterministically or sequentially based on the statistics acquired so far.

We focus on the frequentist approach. A great deal of research in this area (Paulson 1964, Rinott 1978, Kim and Nelson 2001, Hong 2006, Frazier 2014) uses the indifference-zone (IZ) formulation, dating back to the seminal work of Bechhofer (1954). IZ procedures guarantee to select the unique best system with probability exceeding a given confidence level, assuming that the difference between the best and all others is sufficiently large. Another form of guarantee is called probably approximately correct (PAC) selection, which is also referred to as good selection (Nelson and Banerjee 2001, Ni et al. 2017). It guarantees, with high probability, to select a system whose performance is not too far away from that of the best system.

Bayesian approaches and related approaches based on approximation of the frequentist guarantee, though not explored in this paper, have been studied extensively in the literature, e.g., Chen et al. (2000), Chen and Lee (2011), Chick and Inoue (2001a), Chick and Inoue (2001b); see Chen et al. (2015). Procedures developed under these formulations typically require fewer samples than frequentist methods to deliver good solutions (Branke, Chick, and Schmidt 2007). However, they do not explicitly deliver an exact PCS or PAC guarantee.

In this paper, we develop a fully sequential procedure, that we call the envelope procedure (EP), which provides a PAC guarantee. The PAC guarantee of EP is based on the construction of the confidence intervals (CIs) of certain widths. The idea of linking CIs and R&S is also adopted in Lee and Nelson () and Lee and Nelson (2015), but they consider pairwise difference between two systems, while our procedure is based on events for each system one at a time, and the methods of constructing CIs are quite different. The EP is designed to preserve the PAC guarantee over a wide range of sampling rules, thereby affording considerable flexibility in design. Thus, the EP is really a family of procedures. We propose a particular sampling rule that is designed to achieve high efficiency. Numerical experiments demonstrate that, across various configurations, EP is more efficient than the KN procedure (Kim and Nelson 2001), broadly considered the state-of-the-art IZ R&S procedure. Comparing to BIZ (Frazier 2014), an IZ procedure that has a tight lower-bound on PCS for the worst-case configuration in continuous time, EP shows considerable improvement for configurations with spread-out expectations, especially for large-scale problems. We further compare thes procedures based on their empirical PAC instead of the input parameter for lower bound of PAC, and the EP is dramatically more efficient than the other two under this setting. The statistical validity of EP is also confirmed through numerical experiments.

The remainder of this paper is organized as follows. § 2 formally states the formulation of the problem. We discuss the Envelope Procedure in § 3, first introducing the procedure in § 3.1, then proving its statistical validity in § 3.2, followed by a specific choice of sampling rule in § 3.3 and discussion of certain input parameters in § 3.4. We then summarize the numerical results in § 4, followed by conclusions in § 5.

2 PROBLEM FORMULATION

Suppose we have k alternative systems. Let X_{ij} indicate the jth sample from System i. We assume that $(X_{ij} : 1 \le i \le k, j \ge 1)$ are independent and normally distributed, and that X_{ij} has mean μ_i and variance σ_i^2 for all $1 \le i \le k$ and $j \ge 1$. Here we assume that $\sigma_i^2, i = 1,2,\ldots,k$ are known, but can be unequal. Our goal is to identify the index i with the highest mean μ_i by observing samples sequentially. Without loss of generality, we assume that the true means of the systems are indexed so that $\mu_1 \le \mu_2 \le \cdots \le \mu_k$, although we do not know or exploit this ordering except to simplify the presentation.

The quality of the solution provided by R&S procedures can be measured by two forms of probabilistic guarantee. The dominant form of guarantee is the *indifference-zone* (IZ) guarantee, which is given in the form of the *probability of correct selection* (PCS). That is, the probability of selecting the best system k is no less than $1-\alpha$ for some confidence level $0 < 1-\alpha < 1$, provided that the difference between the expectation of the best system and that of the second best system is at least δ. Formally, let I^* be the system selected as best, then an IZ guarantee takes the form

$$\text{PCS} = \mathbb{P}(I^* = k) \ge 1-\alpha, \quad \text{if } \mu_k - \mu_{k-1} \ge \delta.$$

Here the assumption on the problem configuration $\mu_k - \mu_{k-1} \geq \delta$ is often referred as the indifference-zone assumption, and the constant δ is called the indifference-zone parameter.

The other form of guarantee is called a *probably approximately correct* (PAC) selection, which ensures that the *probability of being approximately correct* (also abbreviated PAC), i.e., selecting a system whose expectation is within δ of the best, is no less than $1-\alpha$. This can be stated as

$$\text{PAC} = \mathbb{P}(\mu_{I^*} \geq \mu_k - \delta) \geq 1 - \alpha.$$

We prefer this form of guarantee since it requires no assumption on the problem configuration, and it is not hard to see that all procedures that provide a PAC guarantee also provide a PCS guarantee, while the converse does not necessarily hold.

A fully sequential R&S procedure usually has three key parts: a sampling rule, a screening rule and a stopping rule. The sampling rule defines how to allocate the computational budget to each system in each iteration. The computational budget allocation is either pre-determined or determined based on statistics collected and computed in the screening stage. The whole procedure terminates when the stopping rule is satisfied.

3 THE ENVELOPE PROCEDURE

In this section we introduce the *Envelope Procedure* (EP) for the known-variance case. We first describe the procedure in § 3.1, where the sampling rule is not specified since it can incorporate any valid sampling rule. Then we prove the statistical validity of the procedure in § 3.2. A specific choice of sampling rule is discussed in § 3.3, followed by discussion about certain input parameters in § 3.4.

3.1 The Procedure

The EP is a fully sequential procedure, meaning that it iteratively allocates samples to each system and screens the systems until the stopping rule is met. We denote the total number of samples of System i up to iteration $r \geq 0$ as $n_i(r)$ and the number of samples to take from System i in iteration r as $m_i(r)$, so $n_i(r+1) = n_i(r) + m_i(r)$, where $n_i(0) = n_0$, $i = 1,2,\ldots,k$ is the number of samples to take from each system in the initial stage. In each iteration, $B = \sum_{i=1}^{k} m_i(r)$ samples will be drawn from systems in total. We let $\bar{X}_i(n) = \frac{1}{n}\sum_{j=1}^{n} X_{ij}$ denote the sample mean of the first n samples from System i. The sample size of each system in the initial stage n_0, together with the confidence level $1/k < 1-\alpha < 1$, the indifference-zone parameter $\delta > 0$ and the computational budget of each iteration B are given as inputs.

Envelope Procedure

1. Setup. Let $a = 1-(1-\alpha)^{1/k}$. Compute N and η, such that

$$\mathbb{P}\left(W_n \leq \eta\sqrt{n},\ \forall n = 1,2,\ldots,N\right) \geq 1-a, \tag{1}$$

   where W_n is a standard Brownian motion observed at discrete time n, and

$$\frac{2\eta \max_i \sigma_i}{\sqrt{N}} \leq \delta. \tag{2}$$

   Here N is the maximum number of samples that can be taken from a system. The procedure is guaranteed to stop before any $n_i(r)$ exceeds N. We show later that it is possible to choose N and η in this manner.
2. Initialization. Initialize the iterator $r = 0$. Obtain n_0 samples X_{ij}, $j = 1,2,\ldots,n_0$ from each system $i = 1,2,\ldots,k$. Compute sample means $\bar{X}_i(n(0))$ and let

$$i^* = \arg\max_{i=1,2,\ldots,k} \bar{X}_i(n(0))$$

be the index of the system with highest sample mean, breaking ties arbitrarily.

3. Stopping Rule. If

$$\bar{X}_{i^*}(n_{i^*}(r)) - \frac{\eta\sigma_{i^*}}{\sqrt{n_{i^*}(r)}} \geq \max_{j\neq i^*} \bar{X}_j(n_j(r)) + \frac{\eta\sigma_j}{\sqrt{n_j(r)}} - \delta, \tag{3}$$

go to Step 5. Otherwise go to Step 4.

4. Sampling and Screening. Compute $m_i(r)$ for $i = 1,2,\ldots,k$ according to the sampling rule, and take $m_i(r)$ samples from each system i. Let $r = r+1$ and update $n_i(r)$ and $\bar{X}_i(n_i(r))$. Compute i^*, the index of the system with highest sample mean. Go to Step 3.
5. Termination. Stop and select system $I^* = i^*$ as the best.

We motivate the EP as follows. First, for each iteration r, we interpret

$$\mathrm{UCL}_i(n_i(r)) = \bar{X}_i(n_i(r)) + \frac{\eta\sigma_i}{\sqrt{n_i(r)}}$$

$$\mathrm{LCL}_i(n_i(r)) = \bar{X}_i(n_i(r)) - \frac{\eta\sigma_i}{\sqrt{n_i(r)}}$$

as the upper and lower confidence limit of μ_i, respectively, and the range between them as its confidence interval. We call this range between the UCL and LCL the "envelope", which is why this procedure is called the Envelope Procedure. The width of the confidence interval $2\eta\sigma_i n_i(r)^{-1/2}$ monotonically decreases as more samples are drawn. We carefully choose parameters η and N so that with high probability,

$$\mathrm{LCL}_i(n) \leq \mu_i \text{ for } i = 1,2,\ldots,k-1, \text{ and}$$

$$\mathrm{UCL}_k(n) \geq \mu_k,$$

for all $n = 1,2,\ldots,N$. Since N is chosen so that the procedure will surely terminate before the sample size of any system exceeds N, the inequalities for $\mathrm{UCL}_i(n)$ and $\mathrm{LCL}_i(n)$ both hold throughout the whole procedure with high probability. We demonstrate that such η and N exist, and show how to compute them, in § 3.4.

In each iteration, we allocate samples to systems according to the sampling rule, take samples and update the upper and lower confidence limits. Once we see that the lower confidence limit $\mathrm{LCL}_{i^*}(n_{i^*}(r))$ of System i^*, the one with the highest sample mean, is no less than the highest $\mathrm{UCL}_j(n_j(r)) - \delta$, $j \neq i^*$, we select System i^* and stop. Since the inequalities for the upper and lower confidence limits should hold with high probability when the procedure terminates, we expect that the selected system i^* is probably a good system. We prove this result in § 3.2.

The behavior of the EP is illustrated in Figure 1. The example has $k = 3$ systems. The sample mean $\bar{X}_i(n_i(r))$ and upper and lower confidence limits $\mathrm{UCL}_i(n_i(r))$ and $\mathrm{LCL}_i(n_i(r))$ for each system i are plotted versus the iterator r. In this example we use the sampling rule discussed in § 3.3. Samples are drawn from systems iteratively, and the total sample size varies across systems. Due to the sampling rule, the green system has fewer samples and hence wider confidence interval, since its mean is relatively lower than those of the red one and blue one. As $n_i(r)$ increases for each system, the sample means are converging to their true means and the confidence intervals are shrinking. Finally, when the difference between the lower confidence limit of System i^* (the red one) and the highest upper confidence limit of the others is no more than δ, the procedure stops and we take i^* as the output. Here the final upper confidence limits of the inferior systems (green and blue) are overlapping (the second dotted line from the top) since they are very close, which is due to our choice of the sampling rule.

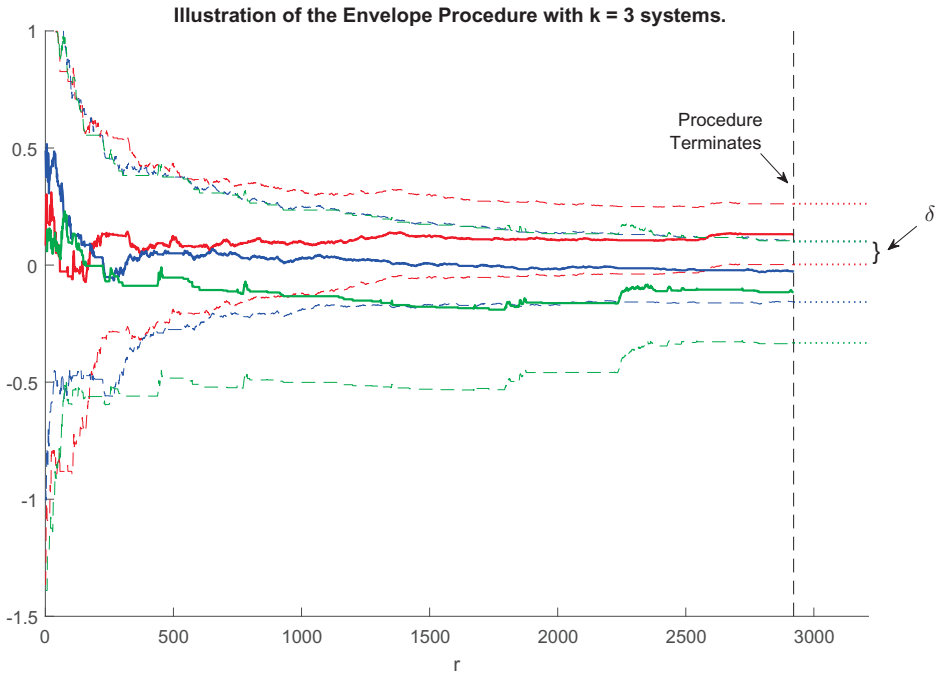

Figure 1: Illustration of EP. Systems are distinguished by different colors. The solid lines are $\bar{X}_i(n)$'s and the dashed lines are $\mathrm{UCL}_i(n)$'s and $\mathrm{LCL}_i(n)$'s throughout the procedure. The dotted lines are the final upper and lower confidence bands when the procedure stops. The black dashed line indicates termination of the procedure when the stopping rule is met.

3.2 Probabilistic Guarantee

The statistical guarantee of the EP relies on the following assumption on the distribution of the simulation output, which is common in the R&S literature.

Theorem 1 Suppose that for each system $i = 1,2,\ldots,k$, the simulation outputs $X_{ij}, j = 1,2,\ldots$ are i.i.d. samples from distribution $\mathcal{N}(\mu_i, \sigma_i^2)$, and are independent for different i. Then the EP terminates in finite time and selects a system I^* that satisfies $\mu_{I^*} \geq \mu_k - \delta$ with probability at least $1-\alpha$, i.e., the EP provides a PAC selection.

We discuss the key insights for proving Theorem 1, and defer the full proof to a journal version of this paper.

Sketch of Proof. First, by (2), the stopping rule (3) is guaranteed to be satisfied before any $n_i(r)$ exceeding N. Second, consider the event that the sample means $\bar{X}_i$ never cross the upper (for $i = 1,2,\ldots,k-1$) or lower (for $i = k$) confidence limits before the procedure terminates. By (1), the probabilities of those events are no less than $1-a$, so the probability of their intersection, denoted as A, is no less than $1-\alpha$, by the independence assumption.

Then we consider a sample path $\omega \in A$. The stopping rule (3) and the way we construct A jointly ensure that the final selection of ω is a good selection. Therefore, A implies good selection, so the PAC of EP is lower bounded by $\mathbb{P}(A) \geq 1-\alpha$. □

3.3 The Sampling Rule

The statistical validity of the EP does not depend on a specific sampling rule. However, the choice of sampling rule affects the efficiency of the procedure, i.e., the total number of samples to draw before the procedure terminates. Since we do not eliminate systems, the sampling rule determines how much effort is wasted on inferior systems. The philosophy is that we want to focus computational effort on those systems that are more likely to be the best system, or one of its strong competitors.

Here we propose a sampling rule that allocates the computing budget in a greedy fashion. In each step it tries to shrink the absolute value of the (negative) gap between the lower confidence bound of the system with highest sample mean and the highest upper confidence bound of all the other systems. In other words, it attempts to minimize the difference between the right-hand side and left-hand side of the inequality (3), ignoring the change in the sample means.

Formally, in each iteration r we solve the optimization problem

$$\begin{aligned}\underset{m_1(r),m_2(r),\cdots,m_k(r)}{\arg\min} & \left\{\max_{j\neq i^*}\bar{X}_j(n_j(r))+\frac{\eta\sigma_j}{\sqrt{n_j(r)+m_j(r)}}-\left[\bar{X}_{i^*}(n_{i^*}(r))-\frac{\eta\sigma_{i^*}}{\sqrt{n_{i^*}(r)+m_{i^*}(r)}}\right]\right\},\\ \text{s.t.}&\sum_i m_i(r)=B,\end{aligned}$$

which is equivalent to

$$\begin{aligned}\underset{m_{i^*}(r)}{\arg\min} & \left\{\frac{\eta\sigma_{i^*}}{\sqrt{n_{i^*}(r)+m_{i^*}(r)}}+\underset{m_j(r),j\neq i^*}{\arg\min}\left\{\max_{j\neq i^*}\bar{X}_j(n_j(r))+\frac{\eta\sigma_j}{\sqrt{n_j(r)+m_j(r)}}\right\}\right\},\\ \text{s.t.}&\sum_i m_i(r)=B\end{aligned}$$

For convenience we drop the iterator r. Let $U_i=\bar{X}_i+\eta\sigma_i(n_i+m_i)^{-1/2}$ for $i=1,2,\ldots k$ be the shorthand of UCL_i, the upper confidence limit of System i. We enumerate m_{i^*} and solve the inner optimization problem. Finally take m_{i^*} that minimizes the outer objective function and the corresponding m_j, $j\neq i$ as our solution; see Algorithm 1.

Algorithm 1 Greedy sampling rule

Let $U_i\leftarrow\bar{X}_i+\eta\sigma_i/\sqrt{n_i}$ for each i, $i^*\leftarrow\arg\max_i\bar{X}_i$, $j^*\leftarrow\arg\max_{j\neq i^*}U_j$.
Let $m_j\leftarrow 0$ for $j\neq i^*$, $m_{i^*}\leftarrow B$, $b\leftarrow 1$.
Let $minV\leftarrow\frac{\eta\sigma_{i^*}}{\sqrt{n_{i^*}+B}}+U_{j^*}$, $minM_i=m_i$ for each i.
while $m_{i^*}>0$ **do**
    Let $m_{i^*}\leftarrow m_{i^*}-b$, $m_{j^*}\leftarrow m_{j^*}+b$, $U_{j^*}\leftarrow\bar{X}_{j^*}+\eta\sigma_{j^*}/\sqrt{n_{j^*}+m_{j^*}}$.
    Let $j^*\leftarrow\arg\max_{j\neq i^*}U_j$, $V\leftarrow\eta\sigma_{i^*}/\sqrt{n_{i^*}+m_{i^*}}+U_{j^*}$.
    if $V<minV$ **then**
        Let $minV\leftarrow V$, $minM_i=m_i$ for each i.
Let $m_i\leftarrow minM_i$ for each i.

Initially we set $m_{i^*}=B$ and $m_j=0$ for each $j\neq i^*$, and initialize the minimum value of the objective function *minV* and the corresponding allocation *minM* to be the current value of the objective function and current allocation, respectively. Then we iteratively update the allocation. In each iteration, we take $b=1$ sample(s) allocated to i^* and instead allocate them to j^*, the system with highest U_j, so that its confidence bound is shrunk and hence the inner objective function decreases since it only depends on U_{j^*}. Notice that

the identity of j^* may change after that. Then we recalculate V, the value of the objective function, and update *minV* and *minM* if $V < minV$. Finally, we set our allocation $\boldsymbol{m} = (m_1, m_2, \ldots, m_k)$ to be *minM*, the allocation that minimizes the objective function.

In our iterations, we keep finding j^*, the minimizer of U_j, allocating samples to it and updating U_{j^*}. In order to do that efficiently, we construct and maintain a min heap H for U_j. Each time we extract the node with the minimum value of U_j, which is j^*, update its value and push it back to H. The time complexity of these operations is $\mathcal{O}(\log k)$, so the total running time is $\mathcal{O}(B \log k/b)$.

We can set the step length b to a larger value than 1 in an attempt to increase the efficiency of the algorithm. Notice that the value of B/b determines the maximum number of systems from which we can draw samples in each iteration. In practice we suggest $b = B/10$ or $b = B/100$.

3.4 On Computing the Parameters η and N

3.4.1 Existence of η and N

The parameter η determines the width of the confidence interval, and hence how quickly the stopping rule is met. In order for the procedure to be valid, we need to choose η together with N, so that both (1) and (2) are satisfied at the same time. It is not immediately clear that one can satisfy both conditions. We first establish this, and then give a method for computing N and η by simulation. To prove the existence, we need the following lemma.

Lemma 1 Let $(W_n, n = 1, 2, \ldots)$ be a random walk with independent increments X_n with support $\mathbb{R}$, and $F(\cdot)$ be the cumulative distribution function of X_n. then

$$\mathbb{P}\left(W_n \leq z \mid W_1 \leq y_1, W_2 \leq y_2, \ldots, W_{n-1} \leq y_{n-1}\right) \geq \mathbb{P}(W_n \leq z),\ \forall z, y_1, y_2, \ldots, y_{n-1}. \tag{4}$$

We provide a sketch of the proof of the above lemma and the following proposition, referring to the journal version of this paper for the full proof.

Sketch of Proof. We prove it by induction. The base case automatically holds. Second, suppose that (4) holds for $n-1$. Let V_{n-1} be the random variable that follows the conditional distribution $G_{n-1}(x) = \mathbb{P}(W_{n-1} \leq x | W_1 \leq y_1, W_2 \leq y_2, \ldots, W_{n-1} \leq y_{n-1})$. Using the induction hypothesis, we can prove that V_{n-1} is smaller than W_{n-1} in the usual stochastic order (denoted by $V_{n-1} \leq_{\text{st}} W_{n-1}$). Next, since $W_n = W_{n+1} + X_n$, $\mathbb{P}\left(W_n \leq z \mid W_1 \leq y_1, \ldots, W_{n-1} \leq y_{n-1}\right)$ and $\mathbb{P}(W_n \leq z)$ can be written as $\mathbb{E}\left[F\left(z - V_{n-1}\right) \mid V_{n-1}\right]$ and $\mathbb{E}\left[F\left(z - W_{n-1}\right) \mid W_{n-1}\right]$, respectively. Then we can compare these two expectations by using Theorem 1.A.3 of Shaked and Shanthikumar (2007) on stochastic order for monotonic functions and (1.A.7) of Shaked and Shanthikumar (2007) on stochastic order of expectations, to complete the induction step and hence the proof. □

Now we prove the existence of parameters N and η that satisfy both conditions.

Proposition 1 For any $a \in (0, 1]$, $c > 0$, $\exists\ (\eta, N)$, such that

$$\mathbb{P}\left(W_n \leq \eta\sqrt{n},\ \forall n = 1, 2, \ldots, N\right) \geq 1 - a, \text{ and} \tag{5}$$

$$c\eta \leq \sqrt{N}. \tag{6}$$

Sketch of Proof. First, by lemma 1, $\mathbb{P}\left(W_n \leq \eta\sqrt{n},\ \forall n = 1, 2, \ldots, N\right) \geq \Phi(\eta)^N$. Setting $\eta = \sqrt{N}/c$ to make (6) hold, and using a well-known tail inequality for the standard normal distribution $1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}$ (Feller 1968, Section 7.1), in order to prove (6), it suffices to show that $\frac{c}{\sqrt{2\pi N}} e^{-\frac{N}{2c}} \leq 1 - (1-a)^{\frac{1}{N}}$. Then we can prove that this inequality holds for large enough N by considering the limit of the ratio between the left-hand side and the Taylor expansion of the right-hand side, which confirms the existence of such N and corresponding η.

□

In Proposition 1, (5) and (6) are equivalent to (1) and (2) by setting $c = 2\max_i \sigma_i/\delta$. Hence, the existence of η and N is assured.

3.4.2 Computing η and N

In the EP, only η is used during the procedure, while N is an auxiliary parameter for computing η. Hence we only consider (1) and use naive Monte Carlo to calculate η for different values of k and N by simulating the random walk with normally distributed increments. Here we fix $\alpha = 0.05$. As seen in Table 1, for fixed k (or equivalently, fixed a), as N increases by orders of magnitude, the corresponding η only increases by a small amount. For example, when $k = 1000$, η only increases by 1% when N increases from 10^4 to 10^5, which only causes an increase in the total number of samples by 1%. Practically speaking, then, we can pick a large enough N so that (2) is always satisfied, and then set η to be the corresponding value. On the other hand, as we will see in § 4, the estimated PAC of EP is always much larger than needed, so in practice even though the value of η we pick is a little smaller than needed in the inequality (1), we expect that the procedure is still able to deliver the pre-specified PAC. Therefore, even though in our experiments we pick large enough N so that the theoretical validity of EP is guaranteed, in practice, as long as the value of η and N are not much smaller than required by (1) and (2), we expect that the EP should still work well and deliver the pre-specified PAC.

Table 1: η for different values of k and N.

k	$N = 10^3$	$N = 10^4$	$N = 10^5$
10	3.58	3.69	3.77
100	4.20	4.30	4.37
1000	4.73	4.83	4.88
10000	5.22	5.28	5.35

Based on our simulation results, we give the empirical formula

$$\eta \approx 3.015 + 0.711 \log_{10} k - 0.035 (\log_{10} k)^2$$

for computing η for a given k. This should prove sufficient in practice provided that the maximum number of samples needed for one system, N, is not too large, e.g., $N \leq 10^9$.

4 NUMERICAL EXPERIMENTS

In this section, we summarize the results of our numerical experiments to demonstrate the performance of the EP on standard test problems, and compare it to two leading R&S procedures: the KN procedure (Kim and Nelson 2001) and the BIZ procedure (Frazier 2014). The KN procedure improves the efficiency over previous IZ procedures in a number of configurations, and is somewhat of a yardstick for selection procedures. The BIZ procedure is a Bayes-inspired procedure, but it delivers a pre-specified frequentist PCS and has a tight lower bound on worst-case PCS under the IZ setting. BIZ has been seen to require fewer samples than KN on a variety of problems (Frazier 2014) and is regarded as a state-of-the-art procedure. Since we focus on known-variance cases, we modify KN from its original version so that it exploits the fact that the variances are known. All free parameters are set to their default values in the original papers. We use the sampling rule described in § 3.3.

4.1 Statistical Validity and Efficiency

We use 4 classes of test problem configurations: 1) Slippage Configuration (SC), where $\mu_1 = \mu_2 = \ldots = \mu_{k-1} = 0$ and $\mu_k = \delta$. This is typically the most difficult configuration since the difference of the expectations between the best (and the only good system) and all other systems is exactly δ; 2) Monotone Increasing

Means (MIM), where $\mu_i = \delta(i-1)$; 3) Random problem instances 1 (RPI1), where $\mu_i \sim \mathcal{N}(0, 4\delta^2)$, and are i.i.d.; and 4) Random problem instances 2 (RPI2), where $\mu_i \sim \mathcal{N}(0, 25\delta^2)$, and are i.i.d. The variances for all configurations are Chi-squared distributed: $\sigma_i^2 \sim \chi^2(4)$, and are i.i.d. We set $\delta = 0.1$ and $\alpha = 0.05$ for all configurations. The samples X_{ij} are normally distributed with mean μ_i and variance σ_i^2 and are independent, as assumed in our problem formulation. For the size of the problems, we choose $k = 100, 1000, 10000$ for small, medium and large-scale problems, respectively.

Regarding statistical validity, we report the empirical PAC in Table 2, which is equal to PCS for SC and MIM problems. For all of the configurations the estimated PAC of EP is 100%, which means that the EP never errs in selecting an approximately correct system as the final output. By way of comparison, the PAC of KN is about 98% in SC, and is 100% in all other configurations. The PAC of BIZ is lowest, and it almost achieves the desired 95% accuracy in SC. These results indicate that EP is reliable, but this over-delivery of the PAC guarantee suggests that EP takes more samples than are needed. Hence, even though EP is more efficient than KN and BIZ in many of our instances, there is room for further improvement.

Table 2: PAC of different procedures.

Configuration	k	KN	BIZ	EP
SC	100	0.981	0.961	1.000
	1000	0.976	0.955	1.000
	10000	0.985	0.959	1.000
MIM	100	1.000	0.998	1.000
	1000	1.000	0.987	1.000
	10000	1.000	0.976	1.000
RPI1	100	1.000	1.000	1.000
	1000	1.000	0.999	1.000
	10000	1.000	0.999	1.000
RPI2	100	1.000	0.998	1.000
	1000	1.000	1.000	1.000
	10000	1.000	1.000	1.000

We compare the efficiency of EP with KN and BIZ based on $n = \sum_{i=1}^{k} n_i$, the total number of samples needed for the procedures to terminate. For each configuration, we run 100 experiments, and give the average ratio between the sample size of EP and that of KN, and the average ratio between the sample size of EP and BIZ. We also give the average sample size of each procedure for each configuration. For each class of configurations, the variances (and expectations for RPI1 and RPI2) are randomly generated, and so the average ratio perhaps better captures the difference between the efficiencies of the different procedures. The average sample sizes are given to show the magnitude of the problem instances. The results are shown in Table 3.

For the Slippage Configuration, the EP is about 20% more efficient than KN in medium and large scaled problems. BIZ is nearly tight (tight when viewed in continuous time and with common variance) for this configuration, so it is not surprising that BIZ is more efficient than EP in this case.

In practice we rarely see problems as difficult as the SC. Usually the means of systems are more spread out, as in our other instances. For the MIM configurations, EP outperforms KN, especially for medium-scale and large-scale problems, where EP uses only about half of the samples used by KN. EP and BIZ are comparable, with EP having a slight edge over BIZ for medium and large-scale instances.

For RPI1, EP shows great improvement in efficiency over both KN and BIZ. For RPI1, EP only uses 72%, 40% and 29% of those used for KN for different scaled problems. Comparing with BIZ, EP is dramatically more efficient for medium-scale and large-scale problems. For RPI2, the improvement is even greater. Especially for large-scale problems, EP saves about 85% and 77% of the samples required

by KN and BIZ respectively, which is a welcome improvement. Moreover it appears that EP saves more computational effort as the size of the problem increases, suggesting that EP is more efficient than other leading procedures in large-scale problems.

Table 3: Average sample size n ($\times 10^5$) and average ratio of sample sizes of different procedures. All numbers are rounded to provide an approximate indication of statistical precision.

Configuration	k	KN	BIZ	EP	mean(EP / KN)	mean(EP / BIZ)
SC	100	3.18	2.18	3.4	1.1	1.6
	1000	40.8	24.1	33.4	0.83	1.40
	10000	502	255	385	0.78	1.51
MIM	100	0.239	0.152	0.175	0.72	1.16
	1000	0.427	0.241	0.239	0.55	0.99
	10000	0.694	0.384	0.379	0.54	0.98
RPI1	100	0.99	0.70	0.74	0.72	1.04
	1000	8.53	5.49	3.42	0.40	0.62
	10000	81.7	51.4	23.6	0.29	0.46
RPI2	100	0.484	0.319	0.268	0.51	0.78
	1000	3.74	2.36	0.86	0.22	0.35
	10000	35	21.9	5.10	0.15	0.23

4.2 Empirical PAC vs. Average Sample Size

As seen above, the procedures over-deliver PAC in our experiments. The target PAC, $1-\alpha$, that we provide as an input parameter is a lower bound on their true PAC, and the lower bound can be very loose. When we compare the accuracy of the procedures, we may prefer to compare their true PAC instead of the input parameter $1-\alpha$, e.g., when we want to find a good system with limited computational budget. To that end, we conduct some experiments to compare the efficiency of the EP and that of other procedures based on their true PAC.

We consider the slippage configuration with common variance, where $\mu_1 = \mu_2 = \ldots = \mu_{k-1} = 0$, $\mu_k = \delta = 0.1$, $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_k^2 = 4$, and $k = 100$. To compare the efficiency of the procedures with different PAC, we vary the PAC by changing the value of the input α, which is equivalent to changing the value of some parameters in the procedure, e.g., η in KN in the original paper (Kim and Nelson 2001) and η in EP. These two η's have different meanings. In order to differentiate them, we call them η_{KN} and η_{EP}, respectively. By changing the value of those parameters, essentially we are changing the total sample size, n. Intuitively, higher n should bring higher PAC. To achieve the desired range of PAC, for KN we set $\eta_{\text{KN}} \in [4.12, 10.81]$, for EP we set $\eta_{\text{EP}} \in [2.22, 4.6]$, and for BIZ, which is nearly tight for this configuration, we set $1-\alpha \in [0.8, 0.999]$. The empirical PAC and corresponding average sample size is obtained using 10,000 independent replications at each point.

The experimental results are shown in Figure 2. We see that all procedures achieve higher PAC as computational effort increases. Among the three procedures, EP outperforms the other two consistently. In particular, to achieve PAC $= 0.95$, on average KN requires 3.06×10^5 samples, BIZ requires 2.19×10^5 samples, while EP only requires 1.50×10^5 samples, which is 51% less than KN and 32% less than BIZ. From another perspective, with the same computational cost, EP attains higher empirical PAC. In particular, with 2×10^5 samples, the empirical PAC of KN, BIZ and EP is $0.827, 0.935$ and 0.987, respectively.

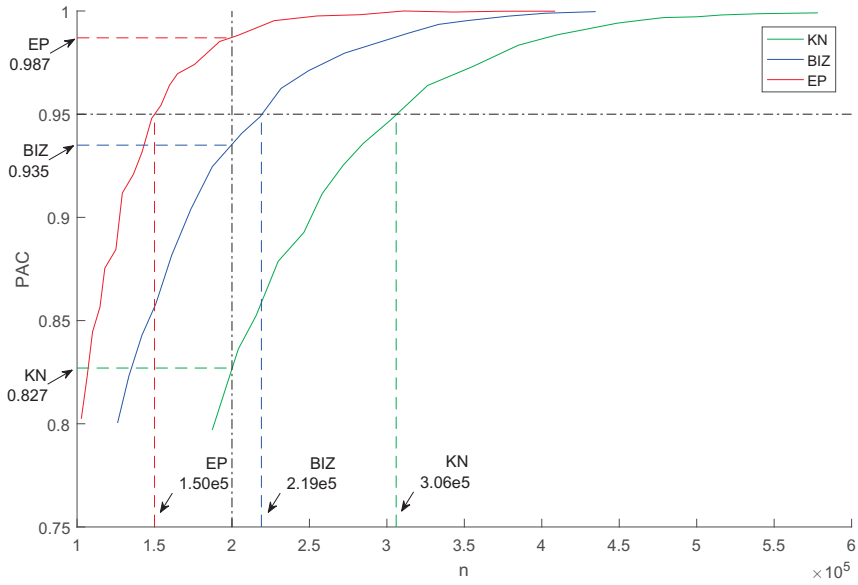

Figure 2: Empirical PAC vs. total sample size n for different procedures in SC with common variance.

5 CONCLUSIONS

In this paper, we propose EP, a new frequentist R&S sequential procedure that delivers a PAC statistical guarantee. The EP allows a variety of sampling rules, so it is actually a family of procedures. We also design a particular sampling rule to make the EP more efficient. Instead of eliminating inferior systems completely, the EP with this sampling rule saves computational effort by allocating different numbers of samples to systems based on their performance so far. Numerical experiments show that EP is very reliable, and is more efficient than the state-of-the-art procedures KN and BIZ on a variety of problems, especially for some large-scale problems. Comparisons of the estimated PAC versus sampling effort further demonstrate the empirical efficiency of EP.

The results also suggest possibilities for further improving the efficiency of EP. The EP over-delivers the PAC, which is a common trait of R&S procedures. The results in § 4.2 suggest considerable room for further improvement if this problem can be alleviated. There are also other possibilities, e.g., to seek a better sampling rule or a tighter envelope region than the current one determined by the value of η.

ACKNOWLEDGMENTS

This research was supported, in part, by National Science Foundation grant CMMI 1537394 and Army Research Office grant W911NF-17-1-0094.

REFERENCES

Bechhofer, R. E. 1954. "A single-sample multiple decision procedure for ranking means of normal populations with known variances". *The Annals of Mathematical Statistics*:16–39.

Branke, J., S. E. Chick, and C. Schmidt. 2007. "Selecting a selection procedure". *Management Science* 53 (12): 1916–1932.

Chen, C.-H., S. E. Chick, L. H. Lee, and N. A. Pujowidianto. 2015. "Ranking and selection: efficient simulation budget allocation". In *Handbook of Simulation Optimization*, 45–80. Springer.

Chen, C.-h., and L. H. Lee. 2011. *Stochastic simulation optimization: an optimal computing budget allocation*, Volume 1. World scientific.

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.

Chick, S. E., and K. Inoue. 2001a. "New procedures to select the best simulated system using common random numbers". *Management Science* 47 (8): 1133–1149.

Chick, S. E., and K. Inoue. 2001b. "New two-stage and sequential procedures for selecting the best simulated system". *Operations Research* 49 (5): 732–743.

Even-Dar, E., S. Mannor, and Y. Mansour. 2002. "PAC bounds for multi-armed bandit and Markov decision processes". In *International Conference on Computational Learning Theory*, 255–270. Springer.

Feller, W. 1968. *An introduction to probability theory and its applications: volume I*, Volume 3. John Wiley & Sons New York.

Frazier, P. I. 2014. "A fully sequential elimination procedure for indifference-zone ranking and selection with tight bounds on probability of correct selection". *Operations Research* 62 (4): 926–942.

Hong, L. J. 2006. "Fully sequential indifference-zone selection procedures with variance-dependent sampling". *Naval Research Logistics (NRL)* 53 (5): 464–476.

Jamieson, K., M. Malloy, R. Nowak, and S. Bubeck. 2014. "lilucb: An optimal exploration algorithm for multi-armed bandits". In *Conference on Learning Theory*, 423–439.

Jamieson, K., and R. Nowak. 2014. "Best-arm identification algorithms for multi-armed bandits in the fixed confidence setting". In *Information Sciences and Systems (CISS), 2014 48th Annual Conference on*, 1–6. IEEE.

Kim, S.-H., and B. L. Nelson. 2001. "A fully sequential procedure for indifference-zone selection in simulation". *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 11 (3): 251–273.

Lee, S., and B. L. Nelson. "Bootstrap ranking & selection revisited". In *Proceedings of the 2014 Winter Simulation Conference*, edited by A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, 3857–3868. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lee, S., and B. L. Nelson. 2015. "Computational improvements in bootstrap ranking & selection procedures via multiple comparison with the best". 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, 3758–3767. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Nelson, B. L., and S. Banerjee. 2001. "Selecting a good system: Procedures and inference". *IIE Transactions* 33 (3): 149–166.

Ni, E. C., D. F. Ciocan, S. G. Henderson, and S. R. Hunter. 2017. "Efficient ranking and selection in parallel computing environments". *Operations Research. To appear. arXiv preprint arXiv:1506.04986*.

Paulson, E. 1964. "A sequential procedure for selecting the population with the largest mean from k normal populations". *The Annals of Mathematical Statistics*:174–180.

Rinott, Y. 1978. "On two-stage selection procedures and related probability-inequalities". *Communications in Statistics-Theory and methods* 7 (8): 799–811.

Shaked, M., and G. Shanthikumar. 2007. *Stochastic orders*. Springer Science & Business Media.

AUTHOR BIOGRAPHIES

SIJIA MA is a Ph.D. student in the School of Operations Research and Information Engineering at Cornell University. He received B.S. in Information & Computing Science from Xi'an Jiaotong University in 2012, and received M.S. in Scientific Computing from New York University in 2014. His research interests include stochastic simulation, optimization and machine learning. His e-mail address is sm2462@cornell.edu.

SHANE G. HENDERSON is a professor in the School of Operations Research and Information Engineering at Cornell University. His research interests include discrete-event simulation and simulation optimization, and he has worked for some time with emergency services. He co-edited the Proceedings of the 2007 Winter Simulation Conference. His web page is https://people.orie.cornell.edu/shane/.