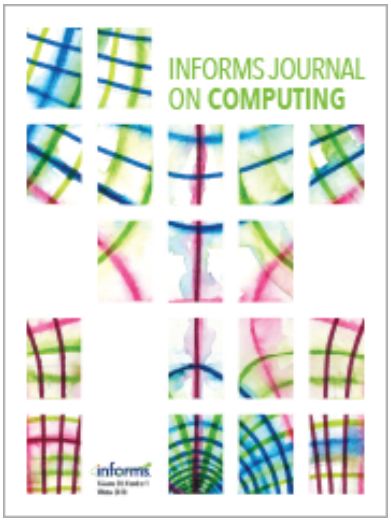


INFORMS Journal on Computing

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Posterior-Based Stopping Rules for Bayesian Ranking-and-Selection Procedures

David J. Eckman, Shane G. Henderson

To cite this article:
David J. Eckman, Shane G. Henderson (2022) Posterior-Based Stopping Rules for Bayesian Ranking-and-Selection Procedures. INFORMS Journal on Computing 34(3):1711-1728. https://doi.org/10.1287/ijoc.2021.1132

http://pubsonline.informs.org/journal/ijoc

INFORMS JOURNAL ON COMPUTING
Vol. 34, No. 3, May–June 2022, pp. 1711–1728
ISSN 1091-9856 (print), ISSN 1526-5528 (online)

Posterior-Based Stopping Rules for Bayesian Ranking-and-Selection Procedures

David J. Eckman,[a,*] Shane G. Henderson[b]

[a] Wm Michael Barnes '64 Department of Industrial and Systems Engineering, Texas A&M University, College Station, Texas 77843; [b] School of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14853
*Corresponding author

Contact: eckman@tamu.edu, https://orcid.org/0000-0002-6473-6434 (DJE); sgh9@cornell.edu, https://orcid.org/0000-0003-1004-4034 (SGH)

Received: October 16, 2020
Revised: March 20, 2021; August 4, 2021; August 24, 2021
Accepted: August 27, 2021
Published Online in Articles in Advance: January 19, 2022

https://doi.org/10.1287/ijoc.2021.1132

Abstract. Sequential ranking-and-selection procedures deliver Bayesian guarantees by repeatedly computing a posterior quantity of interest—for example, the posterior probability of good selection or the posterior expected opportunity cost—and terminating when this quantity crosses some threshold. Computing these posterior quantities entails nontrivial numerical computation. Thus, rather than exactly check such posterior-based stopping rules, it is common practice to use cheaply computable bounds on the posterior quantity of interest, for example, those based on Bonferroni's or Slepian's inequalities. The result is a conservative procedure that samples more simulation replications than are necessary. We explore how the time spent simulating these additional replications might be better spent computing the posterior quantity of interest via numerical integration, with the potential for terminating the procedure sooner. To this end, we develop several methods for improving the computational efficiency of exactly checking the stopping rules. Simulation experiments demonstrate that the proposed methods can, in some instances, significantly reduce a procedure's total sample size. We further show these savings can be attained with little added computational effort by making effective use of a Monte Carlo estimate of the posterior quantity of interest.

Summary of Contribution: The widespread use of commercial simulation software in industry has made ranking-and-selection (R&S) algorithms an accessible simulation-optimization tool for operations research practitioners. This paper addresses computational aspects of R&S procedures delivering finite-time Bayesian statistical guarantees, primarily the decision of when to terminate sampling. Checking stopping rules entails computing or approximating posterior quantities of interest perceived as being computationally intensive to evaluate. The main results of this paper show that these quantities can be efficiently computed via numerical integration and can yield substantial savings in sampling relative to the prevailing approach of using conservative bounds. In addition to enhancing the performance of Bayesian R&S procedures, the results have the potential to advance other research in this space, including the development of more efficient allocation rules.

History: Accepted by Bruno Tuffin, Area Editor for Simulation.
Funding: This work was supported by the National Science Foundation [Grants CMMI-1537394, CMMI-2035086, and DGE-1650441] and the Army Research Office [Grant W911NF-17-1-0094].
Supplemental Material: The online appendix is available at https://doi.org/10.1287/ijoc.2021.1132.

Keywords: ranking and selection • sequential procedures • Bayesian statistics

1. Introduction

Within the statistics and simulation communities, ranking and selection (R&S) refers to the problem of selecting the best from among a finite number of simulated alternatives. The R&S problem has been extensively studied from two differing statistical perspectives: frequentist and Bayesian (see Kim and Nelson (2006) and Chick (2006), respectively, for overviews). In both treatments, considerable attention has been given to the design of efficient procedures that offer finite-sample statistical guarantees (Kim and Nelson 2001, Branke et al. 2007, Chen et al. 2015, Hong et al. 2015).

Bayesian R&S procedures synthesize prior information and data obtained from simulating the alternatives to produce a posterior distribution on the unknown problem instance. This posterior distribution can then be used to evaluate selection decisions by computing criteria such as the (posterior) probability of correct selection (pPCS), probability of good selection (pPGS), and expected opportunity cost (pEOC) of the selected alternative. An R&S procedure straightforwardly delivers a Bayesian statistical guarantee by using the corresponding posterior quantity in a stopping rule. For example, to deliver a guarantee on the pEOC of the

selected alternative, it suffices for a procedure to terminate whenever that quantity drops below a specified threshold (Branke et al. 2007). The appeal of this approach is that repeatedly looking at the data does not invalidate a Bayesian guarantee, as it might for a frequentist one; consider the intricacies of continuously monitoring A/B tests (Deng et al. 2016, Dmitriev et al. 2017, Johari et al. 2017). Aside from the theoretical convenience of proving guarantees, Bayesian R&S procedures have become popular in recent years because of their demonstrated sampling efficiency relative to frequentist procedures (Branke et al. 2007). In particular, Bayesian procedures actively learn about the unknown problem instance and do not need to guard against pathological, worst-case problem instances, as do R&S procedures designed to deliver frequentist guarantees.

Bayesian R&S procedures are predominantly studied under a setting in which the sampling budget—the number of simulation replications that a procedure obtains—is fixed in advance (Chen et al. 2000; Chick and Inoue 2001a, b; Peng et al. 2018a). In this setting, procedures are designed to maximize the pPCS or pPGS or minimize the pEOC of the selected alternative upon exhausting the budget. This setup is pertinent to decision-making situations in which obtaining replications of the simulation model is time-consuming, to the point of being a limiting factor. We instead study the setting in which the decision maker specifies a desired statistical guarantee and runs an R&S procedure until the guarantee can be delivered (Branke et al. 2007). This setup is better suited to situations in which the decision maker can articulate his or her risk and tolerance toward making a suboptimal decision and desires this assurance.

Much research has also focused on the design of efficient rules for allocating replications among alternatives, with recent interest in rules that are, in a certain sense, asymptotically optimal (Peng et al. 2018b, Chen and Ryzhov 2019, Russo 2020). We instead investigate the underappreciated, but no less consequential, matter of how to check the stopping rule. Specifically, we harbor reservations about the common practice of using conservative bounds—typically derived via Bonferroni's or Slepian's inequality—on the posterior quantity of interest in lieu of an exact calculation (Branke et al. 2007, Chick et al. 2010, Görder and Kolonko 2019). Although these bounds are cheap to compute, we show that under fairly standard assumptions (normally distributed outputs, independent sampling, and independent prior beliefs), numerically integrating the posterior quantity is not as computationally onerous as may be suspected. We assess the quality of these bounds, both in terms of how well they approximate the posterior quantity of interest and their effects on extending the requisite sample sizes taken by Bayesian R&S procedures. Our numerical results indicate that while using such bounds for the pPGS leads to only a slight increase in a procedure's total sample size, using those for the pEOC can result in excessively inefficient procedures.

Even though these savings are there for the taking, so to speak, repeatedly computing the posterior quantity of interest poses a tradeoff between simulation time and computational time. We therefore examine the choice of how accurately to check a stopping rule and its impact on the overall run time of a procedure. We present several methods for exactly evaluating posterior quantities of interest and efficiently checking stopping rules. The potential reduction in a procedure's total sample size—relative to using bounds—can in fact be attained at little added computational cost by using a Monte Carlo estimate to "precheck" the stopping rule before applying our methods. For situations in which these combined methods are too computationally intensive, or in which our underlying assumptions are not satisfied, we suggest using the aforementioned Monte Carlo estimate to directly check the stopping rule, even though doing so sacrifices a rigorous Bayesian guarantee. Our numerical experiments indicate that the resulting deterioration in the guarantee is modest.

We believe that Bayesian R&S procedures have key statistical advantages over their frequentist counterparts, but that they lag in implementation because of their perceived computational difficulty. The primary goal of this paper is to allay some of these concerns, thereby increasing the accessibility of Bayesian methods in R&S. To achieve this goal, we make the following contributions:

- We review the difference between Bayesian and frequentist R&S guarantees and discuss the relative merits of the pPCS, pPGS, and pEOC guarantees. We highlight how the popular pPCS guarantee will tend to cause a procedure to incur long run-lengths for little practical gain.
- We review the Stopping Rule Principle which plays a central role in establishing the validity of stopping rules in Bayesian R&S yet does not seem to be well known or widely understood in the simulation research community.
- We develop several approaches to numerically compute integrals for the pPGS and pEOC and comment on their relative efficiencies. (The quantity pPCS is a special case of pPGS.)
- We provide empirical results showcasing the "slack" in the bounds that are typically used to determine stopping rules and explain why the pEOC bounds have more slack than the pPGS bounds.
- We develop structural results that can drastically reduce the computation needed to determine whether to stop; see Propositions 2 and 3. (The already well-known Proposition 5 also contributes to this goal.)

• We use numerical experiments to explore how the computational effort of numerical integration scales with the number of alternatives, k. For sufficiently large k, numerical integration becomes slow and may encounter numerical difficulties, thereby becoming a barrier to evaluating stopping rules in this way. Reasonable ranges for k over which numerical integration is effective depend on the computational cost of simulation replications, because if this is large, then the cost of numerical integration is small by comparison. Our numerical results indicate that the pPGS can be computed for k as large as 10,000 with reasonable computational effort and little numerical error, but as k approaches this value and goes beyond, one might prefer to check stopping criteria less frequently for computational expedience. For the pEOC, numerical integration is more challenging. A reasonable rule of thumb is that the pEOC can be efficiently computed for $k \leq 1{,}000$ and beyond that threshold requires nonnegligible computation. In that case, one can use Monte Carlo to determine when to stop, as mentioned later.

• We discuss how to use Monte Carlo to approximately determine when to stop by sampling directly from the posterior distribution and estimating the posterior quantity of interest. This is very fast and computationally inexpensive, even for very large k. This approach is natural but underused, perhaps because it is assumed that the sampling error would lead to premature stopping. We explain how posterior sampling can be effectively combined with numerical integration to avoid premature stopping, essentially by using posterior sampling as a "precheck" to determine whether it is worthwhile to numerically compute integrals.

• In the situation where k is so large as to make the computational cost of numerical integration prohibitive, one can still use posterior sampling to approximately check stopping rules. Our numerical results indicate that this strategy is unlikely to cause major problems with premature stopping.

Additionally, we prove a pair of new bounds—under an assumption of normality—that may be of independent interest:

• Proposition 1: a Slepian bound on the pPGS of alternatives with posterior means that are "good" relative to the best, for unknown sampling variances; and

• Equation (3): a Slepian bound on the pEOC of the alternative with the highest posterior mean.

All our results apply to the setting in which a decision maker is interested in comparing the mean performances of alternatives. In principle, one could develop selection rules and stopping criteria for other distributional quantities, such as quantiles, exploiting conjugate prior distributions to simplify calculations. The details of how to do this appear to us to be nontrivial and differ greatly from the presentation in this paper. Accordingly, we do not pursue that line of investigation here.

The remainder of this paper is outlined as follows. Section 2 introduces the mathematical notation and distributional assumptions, while defining the pPCS, pPGS, and pEOC under the Bayesian framework. Sections 3 and 4 highlight the computational challenges associated with checking the pPGS and pEOC stopping rules, respectively, and present formulae for efficiently computing the posterior quantities of interest. The proposed improvements are evaluated via simulation experiments in Section 5. In Section 6, we demonstrate further efficiency gains from using a Monte Carlo estimate of the posterior quantity of interest to "precheck" the stopping rule. Section 7 summarizes our findings and lays out directions for future research.

2. Posterior-Based Guarantees and Stopping Rules

2.1. Bayesian R&S Guarantees

Suppose there are k alternatives under consideration and that the (expected) performance of Alternative i is denoted by a scalar W_i, for $i = 1, \ldots, k$. We refer to the vector of performances $\boldsymbol{W} = (W_1, \ldots, W_k)$ as the (random) problem instance or configuration and use the notation $W_{[i]}$ to refer to the ith smallest performance where ties in indexing are broken arbitrarily; that is, the ordered performances satisfy the relationship $W_{[1]} \leq \cdots \leq W_{[k]}$. Without loss of generality, we assume that larger performances are better; hence, Alternative $[k]$ is (one of) the best. In the Bayesian treatment, the identity of the unknown (best) alternative is not fixed, but rather depends on the realization of the performances.

The decision maker assumes a prior distribution over the space of problem instances based on previously gathered data or the opinions of subject matter experts. In the absence of such information, the prior distribution can instead reflect a general uncertainty about the problem instance (i.e., a noninformative prior). After taking replications from the alternatives, a Bayesian R&S procedure applies Bayes' rule to obtain a posterior distribution on the problem instance. The posterior distribution reflects the decision maker's remaining uncertainty about the performances of the alternatives after observing the data and incorporating any prior beliefs. It can be used to define different Bayesian decision criteria with respect to any fixed Alternative i:

• Posterior PCS of Alternative i:
$$\text{pPCS}_i := \mathbb{P}(W_i = W_{[k]} \mid \mathcal{E}),$$
• Posterior PGS of Alternative i:
$$\text{pPGS}_i := \mathbb{P}(W_i \geq W_{[k]} - \delta \mid \mathcal{E}), \text{ and}$$
• Posterior EOC of Alternative i:
$$\text{pEOC}_i := \mathbb{E}[W_{[k]} - W_i \mid \mathcal{E}].$$

The previous probabilities and expectations are with respect to the posterior distribution given the evidence (observed simulation outputs)—denoted by $\mathcal{E}$—and the prior distribution. In contrast to their frequentist counterparts, these Bayesian criteria are functions of the evidence and thus can be calculated within a procedure; we discuss ways to bound, calculate, and estimate them in this paper.

For a given Alternative i, pPCS_i is the probability under the posterior distribution that the random problem instance is one for which Alternative i is (one of) the best. Similarly, pPGS_i is the posterior probability that the random problem instance is one for which Alternative i is δ-optimal. From these definitions, pPCS_i corresponds to pPGS_i for the case $\delta = 0$. Last, pEOC_i is the expected optimality gap associated with selecting Alternative i over all problem instances weighted according to the posterior distribution.

Under the Bayesian framework, the index of the selected alternative, denoted by d, is determined by the (fixed) observed data and the prior distribution. We assume for simplicity that Bayesian R&S procedures do not use randomized selection rules; that is, given the observed data, d is deterministic. In the event of ties in posterior quantities, we will assume that there is a ranking of the alternatives' indices—fixed a priori—that is used to break ties. The three Bayesian criteria lend themselves to guarantees with respect to the selected alternative:

- pPCS Guarantee: $\text{pPCS}_d \geq 1 - \alpha$,
- pPGS Guarantee: $\text{pPGS}_d \geq 1 - \alpha$, and
- pEOC Guarantee: $\text{pEOC}_d \leq \beta$.

For these guarantees, the decision maker specifies the values of $1 - \alpha$ and δ, or of β, in advance. The threshold $1 - \alpha$ reflects the decision maker's desired degree of confidence in making a correct or good selection. The values of δ and β have clear interpretations in terms of the largest or average difference in performance to which the decision maker is indifferent. A reasonable choice of β is the good-selection parameter, δ, times the allowable probability of making a bad selection, α (Chen et al. 2015).

2.2. Issues with the pPCS Guarantee

The pPCS guarantee stipulates that the decision maker insists on selecting the best alternative with high probability and will not be satisfied with selecting a suboptimal alternative, no matter how close its performance is to the best. By this reasoning, extremely small differences in performances are worth detecting, even at great computational expense. Consequently, when the difference between the performances of the best and second-best alternatives is small, a Bayesian R&S procedure will take many replications before the pPCS of any alternative rises above $1 - \alpha$. It is hard to justify expending so much computational effort to detect differences that are of less-than-practical significance. This insistence on finding the optimal solution to the R&S problem also ignores the fact that there is inherently some degree of model error associated with the simulation model. In contrast, the pPGS guarantee takes a more lenient approach, allowing the decision maker to specify a tolerance in performance to which he or she is indifferent.

Related concerns with the pPCS guarantee arise in the event that multiple alternatives are tied for the best. Although it would be convenient to assume that practical problems do not have multiple alternatives with tied performances, some do, if not properly formulated. For example, the prototypical buffer-allocation problem of Pichitlamken et al. (2006) has multiple optimal solutions because of symmetries in the tandem-queuing system; see Ni et al. (2017) for a detailed description.

The Bayesian resolution to this issue is that for continuous posterior distributions, for example, multivariate normal, the probability that the realized performances of two or more alternatives are tied is zero. In the situation where two or more alternatives are tied for the best, the posterior distributions will concentrate around the true (tied) means but not coalesce on a point mass in finite time. Computationally there is no difference with the situation where multiple alternatives are *nearly* tied for the best, so here too the pPGS guarantee does not lead to excessive sampling. See section 3.3.2 of Eckman (2019) for further discussion.

If it was known in advance that certain alternatives had the same performances, perhaps because of symmetry, that information *could* be incorporated into the prior distribution. The drawback of this approach, however, is the loss of conjugacy; as a consequence, updating the posterior distribution becomes computationally intensive. Moreover, establishing that two alternatives have the same performance or that a problem has a unique optimal solution is nontrivial, especially when there are many alternatives. Given these challenges, and the fact that ties are rather neatly handled when one uses continuous posterior distributions, we do not pursue such an approach.

Notwithstanding these concerns, many Bayesian R&S procedures designed for the fixed-budget setting use pPCS to allocate replications across alternatives (Chen et al. 2000) and as an overall performance criterion (Branke et al. 2007, Peng et al. 2016, Russo 2020). This raises an important question that we do not address in this paper: Does allocating replications based on pPCS have a deleterious effect on the quality of the selected alternative, especially on problem instances with multiple near-optimal alternatives?

2.3. Stopping Rule Principle

In Bayesian statistics, the Stopping Rule Principle states that given observed evidence, inference about

an unknown parameter of interest should not depend on the rule used to terminate an experiment (Berger 1993). In other words, an experimenter can ignore the stopping rule when carrying out a statistical analysis after an experiment. The sequential tests supported by the Stopping Rule Principle have the upside of taking only as many samples as necessary, in contrast to fixed-sample-size tests. The Stopping Rule Principle is a natural consequence of the Likelihood Principle, a foundational argument in Bayesian statistics claiming that all experimental information about the unknown parameter of interest is reflected in the likelihood function. Perhaps in light of its simplifying appeal, the Stopping Rule Principle is the subject of some controversy among statisticians; we direct the interested reader to an illuminating discussion in section 7.7 of Berger (1993).

The Stopping Rule Principle has several remarkable consequences. First, the rule used to terminate a procedure does not affect the calculation of any posterior quantity, meaning that the pPCS, pPGS, and pEOC can appear in stopping rules (Chick and Inoue 2001b, Chick 2006, Chen et al. 2015):

- pPCS Stopping Rule: Terminate when $\text{pPCS}_i \geq 1-\alpha$ for *any* $i = 1,\ldots,k$ and select Alternative i;
- pPGS Stopping Rule: Terminate when $\text{pPGS}_i \geq 1-\alpha$ for *any* $i = 1,\ldots,k$ and select Alternative i; and
- pEOC Stopping Rule: Terminate when $\text{pEOC}_i \leq \beta$ for *any* $i = 1,\ldots,k$ and select Alternative i.

Put succinctly, the aforementioned Bayesian guarantees are attained by terminating a sequential R&S procedure when the posterior quantity of interest crosses some threshold.

Chick et al. (2010) refer to these kinds of stopping rules as *adaptive* stopping rules because they depend on the replications collected, as opposed to the rule of stopping when a fixed budget has been exhausted. We choose to call them *posterior-based* stopping rules to emphasize that they involve quantities calculated from the posterior distribution of the problem instance. We cannot understate the significance of being able to compute (and recompute) posterior quantities in the stopping rules and still deliver Bayesian guarantees. Unless special care is taken, repeatedly looking at the data in this way can invalidate frequentist guarantees; sequential analysis methods are a notable exception (Wald 1973).

A second important consequence of the Stopping Rule Principle is that because Bayesian R&S guarantees follow from the stopping rule, the user has complete flexibility in allocating replications across alternatives. These stopping rules can therefore be used in conjunction with popular allocation rules: for example, value-of-information (VIP), optimal computing budget allocation (OCBA), Thompson sampling (TS), and knowledge-gradient (KG). Allocation rules are also allowed to use posterior quantities; for example, OCBA and TS rules use the pPCS of alternatives (Chen et al. 2000, Russo 2020) and some VIP rules use the pEOC of alternatives, or approximations thereof (Chick and Inoue 2001b).

2.4. Distributional Assumptions

Thus far, we defined Bayesian criteria and guarantees without imposing any distributional assumptions on the simulated outputs of the alternatives' performances or the decision maker's beliefs. We now make several standard assumptions so that we can derive specific results for checking stopping rules involving the pPCS, pPGS, and pEOC.

Let X_{ij} denote the jth observation from Alternative i and define $\overrightarrow{X}_i = \{X_{i1}, X_{i2}, \ldots\}$ for $i = 1,\ldots,k$.

Assumption 1. *For each $i = 1,\ldots,k$ and $j \geq 1$, X_{ij} is normally distributed with mean w_i and variance σ_i^2.*

Assumption 1 is commonly made in the simulation community and can be partially justified by batching replications because the batch means will be approximately normally distributed. The R&S problem has also been studied for Bernoulli-distributed outputs (Even-Dar et al. 2006, Russo 2020) and from a large-deviations perspective (Glynn and Juneja 2004, Hunter and Pasupathy 2010, Glynn and Juneja 2018).

Assumption 2. *For each $i = 1,\ldots,k$, the sequence $\overrightarrow{X}_i$ consists of independent outputs.*

Assumption 2 implies that the outputs are exchangeable—a standard assumption for the derivation of the posterior distribution.

Assumption 3. *The sequences $\overrightarrow{X}_1, \overrightarrow{X}_2, \ldots, \overrightarrow{X}_k$ are independent.*

Assumption 3 rules out the use of common random numbers (CRN) in generating replications. Although CRN are helpful in comparing the performances of alternatives, their use complicates the statistical analysis of Bayesian R&S procedures because of two notable challenges concerning the number of replications taken from each alternative. First, when an allocation rule produces a monotone missing pattern with unequal sample sizes, updating the posterior distribution requires careful attention and accounting (Görder and Kolonko 2019). Second, under the reference prior, at least k replications must be taken from each alternative to ensure that the sample covariance matrix is invertible (Chick and Inoue 2001a).

Assumption 4. *The decision maker has independent prior beliefs about the performances of alternatives, $W_1,\ldots,W_k$.*

Assumption 4 is another standard assumption in the literature, although the setting of correlated beliefs

and the use of CRN have also been studied (Frazier et al. 2009, Xie et al. 2016). Assumptions 3 and 4 are typically made for analytical convenience as together, they imply that the posterior distribution of $\boldsymbol{W}$ is the product of the marginal posterior distributions of W_i for $i = 1, \ldots, k$. We later exploit this property when performing numerical integration. Enforcing independent beliefs in the prior distribution entails discarding any available structural information about the optimization problem, for example, convexity or symmetry. In doing so, the decision maker sacrifices prior knowledge about the relationships among alternatives for computational convenience.

Our purpose for making Assumptions 1–4 is to provide theoretical results that lay the groundwork for our more general arguments about checking posterior-based stopping rules. If the assumptions do not all hold, many of the methods presented in this paper will not directly apply. For instance, the results dealing with Slepian's bound and the quasi-Pareto relationship of pPGS (Proposition 3) almost certainly do not hold if Assumption 1 is not satisfied. Furthermore, without Assumptions 3 and 4, the integral equations we provide for exactly calculating the pPGS and pEOC no longer hold. In such cases, using a Monte Carlo estimator to precheck a stopping rule, as developed in Section 6, remains an option. We expect that many of our main conclusions will still hold when these distributional assumptions are relaxed, for example, the potential for a sizable loss in efficiency from using conservative bounds on posterior quantities to check stopping rules.

Given Assumptions 1–4, we now mathematically describe the marginal posterior distributions of W_i when using the conjugate (normal-gamma) reference prior for the mean and precision of Alternative i. After observing $x_{i1}, \ldots, x_{in_i}$ from Alternative i, the *marginal* posterior distribution of the performance of Alternative i is given by

$$W_i \sim t_{n_i-1}(\bar{x}_i, s_i^2/n_i) \equiv t_{\nu_i}(\mu_i, \rho_i^2),$$

where s_i^2 is the sample variance of $x_{i1}, \ldots, x_{in_i}$, and $t_\nu(\mu, \rho^2)$ denotes the distribution of a three-parameter t-distributed random variable $Z = \mu + \rho T_\nu$, where T_ν is a t-distributed random variable with ν degrees of freedom (Chick and Inoue 2001b). When the variances σ_i^2 are known, the *marginal* posterior distribution for the performance of Alternative i is

$$W_i \sim \mathcal{N}(\bar{x}_i, \sigma_i^2/n_i) \equiv \mathcal{N}(\mu_i, \rho_i^2).$$

We will refer to μ_i and ρ_i^2 as the posterior mean and variance of the performance of Alternative i and use subscript $(\cdot)$ to denote the ordered indices of the posterior means; that is, $\mu_{(1)} \leq \mu_{(2)} \leq \cdots \leq \mu_{(k)}$ and Alternative (k) is the best-looking alternative. More complete derivations of the marginal posterior distributions and their parameters can be found in DeGroot (2004), Branke et al. (2007), and Eckman (2019). As can be seen from the previous formulae, using conjugate prior distributions greatly simplifies the task of updating the posterior distribution. When conjugate prior distributions are not used, the pPCS, pPGS, and pEOC can still be estimated via Markov chain Monte Carlo.

3. Checking the pPCS and pPGS Stopping Rules

3.1. Computational Considerations

When running a Bayesian R&S procedure, two critical decisions are how frequently and how accurately to check the stopping rule, as these choices affect the run time. At one extreme, one could allocate replications one at a time and precisely calculate the posterior quantity of interest of *every* alternative after *every* replication. Although this approach would ensure that the procedure takes no unnecessary replications, it would likely be expensive in terms of the computational overhead. At the other extreme, one could allocate replications in large batches and calculate a cheap bound for the posterior quantity of interest of, say, only the best-looking alternative. This approach would be cheaper in terms of overhead but would necessitate more simulation time. Within the Bayesian framework, these decisions can even be made adaptively, based on the data collected thus far. In studying these tradeoffs, we assume that the time required to run a simulation replication is long enough—perhaps on the order of seconds or tens of seconds—to justify expending some nontrivial amount of time checking the stopping rule.

The pPCS and pPGS stopping rules give rise to computational challenges similar to those of the pEOC stopping rule, but differing in important ways. We choose to consider them separately. Throughout this section and the next, we maintain Assumptions 1–4 and assume that the conjugate prior is used. Hereafter, we focus our discussion on the pPGS stopping rule and treat the pPCS stopping rule as a special case.

At first glance, computing pPGS_i appears to involve evaluating a k-dimensional integral of the joint posterior distribution of $W_1, \ldots, W_k$ over a polyhedron described by the $k-1$ inequalities $W_i \geq W_j - \delta$ for $j \neq i$. It can alternatively be expressed as a $(k-1)$-dimensional integral with respect to the positively correlated random variables $W_j - W_i$ for $j \neq i$. When the number of alternatives is small (fewer than 25, say), these integrals can be evaluated numerically using quadrature, for example, MATLAB's `mvncdf` function if the true variances are known. As the number of alternatives

increases, this approach quickly becomes computationally impracticable.

Checking the pPGS stopping rule entails frequently computing the pPGS of one or more alternatives; therefore, for problems with even a modest number of alternatives, other methods for evaluating pPGS_i are necessary. We compare two: cheaply computable lower bounds and an equivalent one-dimensional integral.

3.2. Cheap Lower Bounds

One approach that avoids calculating pPGS_i is to instead compute a cheap lower bound and terminate the procedure when it exceeds $1-\alpha$—doing so will still yield the desired Bayesian guarantee. This approach lowers the computational cost of checking the stopping rule, but any slack in the bound will likely cause the procedure to take additional replications relative to the approach of exactly calculating pPGS_i.

One such lower bound follows from Bonferroni's inequality: for any Alternative i,

$$\begin{aligned}\text{pPGS}_i &= \mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid \mathcal{E}) \\ &\geq 1 - \sum_{j \neq i} \mathbb{P}(W_i < W_j - \delta \mid \mathcal{E}) \equiv \text{pPGS}_i^{\text{Bonf}}.\end{aligned}$$

When the true variances are unknown, the term $\mathbb{P}(W_i < W_j - \delta \mid \mathcal{E})$ involves the cumulative distribution function (cdf) of the difference of two t-distributed random variables with possibly different degrees of freedom. Chick and Inoue (2001b) apply the approximation of Welch (1938) to this term to simplify the computation of $\text{pPGS}_i^{\text{Bonf}}$.

Another cheap bound on pPGS_i can be derived from Slepian's inequality (Slepian 1962). Although Slepian's inequality concerns *normal* random variables—which $W_1, \ldots, W_k$ would be if the true variances were known—applying the result when the variances are unknown is less straightforward. For this setting, Branke et al. (2007) proposed a Slepian-type bound on the pPGS of the alternative with the highest posterior mean, Alternative (k), yet offered limited justification. We rigorously show in Proposition 1 that with the help of an inequality from Tamhane and Bechhofer (1977) (see Lemma 2 in Appendix A in the online appendix), Slepian's inequality can indeed be applied to the case of unknown variances. Proposition 1 also generalizes the Slepian-type bound of Branke et al. (2007) by extending the bounds on pPGS_i to alternatives with δ-optimal posterior means; see Appendix A in the online appendix for a proof.

Proposition 1. *Under Assumptions 1–4, for any Alternative i with posterior mean μ_i satisfying $\mu_i \geq \mu_{(k)} - \delta$,*

$$\begin{aligned}\text{pPGS}_i &= \mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid \mathcal{E}) \\ &\geq \prod_{j \neq i} \mathbb{P}(W_i \geq W_j - \delta \mid \mathcal{E}) \equiv \text{pPGS}_i^{\text{Slep}}.\end{aligned}$$

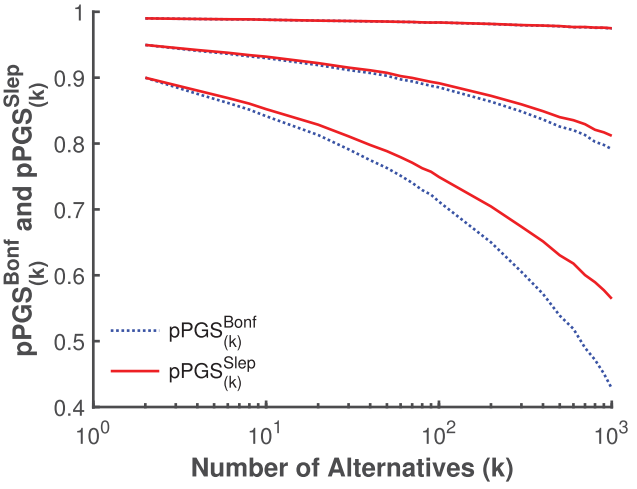
Whereas the expression for pPGS_i deals with the maximum of $k-1$ positively correlated random variables $W_j - W_i$ for $j \neq i$, the expression for $\text{pPGS}_i^{\text{Slep}}$ resembles the maximum of $k-1$ *independent* random variables. In this way, Slepian's inequality ignores the positive correlations and treats the terms in the pPGS_i expression as independent, thereby replacing a joint probability statement with a product of marginal ones. Peng et al. (2018a) claim that these ignored correlations are unimportant in a high-confidence setting, when $1-\alpha$ is close to one but are more likely to make a difference in a low-confidence setting. Our experimental results below bear this out.

Unlike the $\text{pPGS}^{\text{Slep}}$ bound, which requires a multivariate normal posterior distribution, the $\text{pPGS}^{\text{Bonf}}$ bound holds even when Assumptions 1–4 do not. In particular, the bound is still valid when $W_1, \ldots, W_k$ are not independent under the posterior distribution, for example, when CRN are used. Although pPGS_i is harder to compute in such instances, the $\text{pPGS}^{\text{Bonf}}$ bound remains one of few resources available.

Various Bayesian allocation rules use the $\text{pPGS}^{\text{Bonf}}$ (Chen et al. 2000) or $\text{pPGS}^{\text{Slep}}$ (Chick and Inoue 2001b) bounds (with $\delta = 0$). Other procedures have used these bounds as stopping rules; for example, the procedure of Görder and Kolonko (2019) for CRN terminates when $\text{pPGS}_{(k)}^{\text{Bonf}}$ first exceeds $1-\alpha$, and the procedures of Branke et al. (2005, 2007) terminate when $\text{pPGS}_{(k)}^{\text{Slep}}$ first exceeds $1-\alpha$.

We examine the potential slack in the $\text{pPGS}^{\text{Bonf}}$ and $\text{pPGS}^{\text{Slep}}$ bounds, by which we mean the gap between the approximations and the exact pPGS. Consider a slippage configuration of posterior means, that is, $\mu_j = \mu_{(k)} - \Delta$ for all $j \neq (k)$ for some $\Delta > 0$, with a common posterior variance ρ^2. For this configuration, $\text{pPGS}_{(k)} = 1-\alpha$ if $\Delta = h_B\sqrt{2\rho^2} - \delta$ (provided δ is small enough so that Δ is positive), where h_B is a constant depending on $1-\alpha$ and k (Kim and Nelson 2006). To be precise, h_B is the $1-\alpha$ quantile of the maximum of a $(k-1)$-dimensional multivariate normal vector with zero means, unit variances, and common pairwise correlations of $1/2$. It can be checked that for this configuration, $\text{pPGS}_{(k)}^{\text{Bonf}} = 1-(k-1)\Phi(-h_B)$ and $\text{pPGS}_{(k)}^{\text{Slep}} = \Phi(h_B)^{k-1}$, irrespective of ρ^2 and δ. Figure 1 shows these bounds when $\text{pPGS}_{(k)} = 1-\alpha = 0.90, 0.95$, and 0.99 for this configuration for different numbers of alternatives. Several trends are evident: First, both $\text{pPGS}_{(k)}^{\text{Bonf}}$ and $\text{pPGS}_{(k)}^{\text{Slep}}$ become less tight as the number of alternatives increases. Second, the slack in both bounds is greater—in an absolute sense—for values of $1-\alpha$

Figure 1. (Color online) Values of $\text{pPGS}^{\text{Bonf}}_{(k)}$ and $\text{pPGS}^{\text{Slep}}_{(k)}$ in a Slippage Configuration of Posterior Means in Which $\text{pPGS}_{(k)} = 1-\alpha$ for $1-\alpha = 0.90, 0.95$, and 0.99



farther from one; for $1-\alpha = 0.99$, the bounds are roughly equivalent. Third, $\text{pPGS}^{\text{Bonf}}_{(k)}$ appears to be a looser bound than $\text{pPGS}^{\text{Slep}}_{(k)}$, with the difference in the bounds growing with the number of alternatives.

3.3. Numerical Integration

Given the potential slack in the $\text{pPGS}^{\text{Bonf}}$ and $\text{pPGS}^{\text{Slep}}$ bounds, we return to the challenge of calculating pPGS_i in a way that is cheap and scales well with the number of alternatives. The approach we take relies on the observation of Peng et al. (2016) and Russo (2020) that pPGS_i can be expressed as a one-dimensional integral by conditioning on the performance of Alternative i:

$$\begin{aligned}\text{pPGS}_i &= \mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid \mathcal{E}) \\ &= \mathbb{E}\left[\mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid W_i, \mathcal{E}) \mid \mathcal{E}\right] \\ &= \mathbb{E}\left[\prod_{j\neq i} \mathbb{P}(W_j \leq W_i + \delta \mid W_i, \mathcal{E}) \mid \mathcal{E}\right] \\ &= \int_{-\infty}^{\infty}\left[\prod_{j\neq i} F_{W_j|\mathcal{E}}(w+\delta)\right] f_{W_i|\mathcal{E}}(w)\, dw, \end{aligned} \tag{1}$$

where $f_{W_i|\mathcal{E}}(\cdot)$ and $F_{W_j|\mathcal{E}}(\cdot)$ are the posterior marginal probability density function (pdf) of W_i and cdf of W_j, respectively. The third equality in Equation (1) makes use of the product form of the posterior joint distribution under Assumptions 3 and 4. The integrand in Equation (1) is the product of $k-1$ cdfs and a pdf, all of which under our assumptions are either for normal or t-distributed random variables. Equation (1) thereby avoids the need to approximate the difference of two t-distributed random variables with different degrees of freedom, as was the case for the $\text{pPGS}^{\text{Bonf}}$ and $\text{pPGS}^{\text{Slep}}$ bounds.

To give a sense of the computational time associated with numerically integrating Equation (1), we use MATLAB's integral function for values of k ranging from 10 to 100,000. MATLAB's integral function performs global adaptive quadrature with a 7-point Gauss formula and 15-point Kronrod extension (Shampine 2008). All experiments were run on a high-performance computing cluster using eight cores on a compute node with 256 GB of RAM. Source code for all experiments is available at https://github.com/daveckman/posterior-stopping-rules. For each value of k, we generate 100 random posterior distributions according to $\mu_i \sim \mathcal{N}(0,25)$ and $\rho_i^2 \sim \text{ChiSquared}(4)$, independent for all $i = 1,\ldots,k$, and compute $\text{pPGS}_{(k)}$, without loss of generality, for both known and unknown variances. Table 1 reports the average times as well the 10th and 90th percentiles. Times for unknown variances are about 6 to 20 times slower than those for known variances, because of the need to evaluate gamma and incomplete-beta functions when evaluating the pdfs and cdfs of t-distributed random variables. The results suggest that, even for fairly large numbers of alternatives, pPGS_i can be quickly computed.

Remark 1. Under the default tolerance settings for MATLAB's integral function, the absolute error between the returned value and the true value of an integral is estimated to be less than the larger of 10^{-10} or 10^{-6} times the returned value. For our pPGS_i calculations, this translates to a bound of 10^{-6} on the absolute error. For our pEOC_i calculations in Section 4, which feature sums of integrals or double integrals, a very conservative bound on the absolute error is 10^{-3}. The numerical integration results are therefore sufficiently accurate for our purpose of checking stopping rules.

3.4. Which Alternatives to Evaluate

In cases where repeatedly computing the pPGS of all alternatives is computationally prohibitive, it is worthwhile to reduce the number of alternatives for which the pPGS must be evaluated to check the stopping rule. This motivates two important questions: Which alternatives can have the highest pPGS? Can the highest pPGS of these alternatives exceed $1-\alpha$?

To answer these questions, we first make use of a simple upper bound on the pPGS of alternatives whose posterior means are at least δ below the highest posterior mean.

Proposition 2. *Under Assumptions 1 and 2, for any Alternative i with posterior mean μ_i satisfying $\mu_i < \mu_{(k)} - \delta$, $\text{pPGS}_i < 1/2$.*

Table 1. Average and 10th/90th Percentiles of Time (Seconds) to Numerically Integrate Equation (1) with Known/Unknown Variances for Different Numbers of Alternatives (k) Based on 100 Randomly Generated Problem Instances

k	10	100	1,000	10,000	100,000
Average (known)	0.006	0.007	0.031	0.14	0.60
Percentiles (known)	[0.004, 0.007]	[0.007, 0.008]	[0.025, 0.037]	[0.13, 0.16]	[0.53, 0.69]
Average (unknown)	0.036	0.062	0.17	1.52	12.5
Percentiles (unknown)	[0.029, 0.043]	[0.049, 0.071]	[0.16, 0.18]	[1.41, 1.65]	[11.3, 14.0]

Proof of Proposition 2. For any Alternative i satisfying $\mu_i < \mu_{(k)} - \delta$,

$$\begin{aligned}\text{pPGS}_i &= \mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid \mathcal{E}) \\ &\leq \mathbb{P}(W_i \geq W_{(k)} - \delta \mid \mathcal{E}) = \mathbb{P}(W_{(k)} - W_i \leq \delta \mid \mathcal{E}) \\ &< 1/2,\end{aligned}$$

where the last inequality comes from the fact that $W_{(k)} - W_i$ is symmetrically distributed with mean $\mu_{(k)} - \mu_i > \delta$. □

Proposition 2 implies that only alternatives whose posterior means are within δ of $\mu_{(k)}$ can satisfy the pPGS stopping rule when $1 - \alpha \geq 1/2$. Hence the fact that Proposition 1 defines the $\text{pPGS}^{\text{Slep}}$ bound for only these alternatives is not necessarily a limitation. Similarly, only the alternative with the highest posterior mean can satisfy the pPCS stopping rule in this high-confidence setting. Proposition 2 holds even when Assumptions 3 or 4 do not, because $W_{(k)} - W_i$ is still symmetrically distributed under the conjugate prior assumption.

We further reduce the number of alternatives for which we need to calculate the pPGS by identifying sufficient conditions under which the pPGS of an alternative is greater than that of another. The motivating idea is to use the posterior means and variances to form a partial order on the set of alternatives. Thus, there is no need to calculate the pPGS of any dominated alternative because any alternative that dominates it has a higher pPGS.

Peng et al. (2016) show that if the posterior means of all alternatives are equal, the alternative with the highest posterior variance has the highest pPCS; see proposition A.2 therein. When trying to select the alternative with the highest pPCS, the authors also suggest that one should favor alternatives that have high posterior means *and* variances. Proposition 3 formalizes this assertion and extends it to the pPGS criterion; its proof appears in Appendix B in the online appendix.

Proposition 3. *Under Assumptions 1–4, suppose that the true variances are known. For any pair of Alternatives i and j with posterior means μ_i and μ_j and posterior variances ρ_i^2 and ρ_j^2 satisfying $\mu_i \leq \mu_j - \delta/2$ and $\rho_i^2 \leq \rho_j^2$, $\text{pPGS}_i < \text{pPGS}_j$.*

The $\delta/2$ term in Proposition 3 is tight in a sense made precise in Proposition 4; its proof appears in Appendix C in the online appendix.

Proposition 4. *Under Assumptions 1–4, for any $k \geq 3$ and any pair of Alternatives i and j with posterior means μ_i and μ_j for which $\mu_i = \mu_j - \delta/2 + \gamma$ where $\gamma > 0$, there exist posterior means μ_ℓ for $\ell \neq i, j$ and posterior variances ρ_ℓ^2 for $\ell = 1, \ldots, k$ such that $\rho_i^2 \leq \rho_j^2$ and $\text{pPGS}_i > \text{pPGS}_j$.*

When the true variances are unknown, the result of Proposition 3 is unlikely to hold in general since the t distributions arising from unequal sample sizes do not satisfy the stochastic dominance relationship used in the proof. Proposition 3, however, can still be used as a heuristic for determining for which alternatives i to compute pPGS_i. If a procedure fails to compute the pPGS of the alternative with the highest pPGS, it might miss an opportunity to terminate when the pPGS stopping rule is first satisfied, but the procedure's Bayesian guarantee would not be invalidated. Moreover, because the alternative with the highest posterior mean would always remain in consideration, this approach would entail taking no more observations than the approach of tracking a lower bound on $\text{pPGS}_{(k)}$.

We exploit Propositions 2 and 3 to develop a framework for exactly checking the pPGS stopping rule—outlined in Procedure 1—and demonstrate its potential savings in Section 5.

Procedure 1

1. Take n_0 initial replications from each alternative.
2. Let $\mathcal{S} = \{1, \ldots, k\}$.
3. For each Alternative $i \in \mathcal{S}$, remove Alternative i from $\mathcal{S}$ if $\mu_i < \mu_{(k)} - \delta$.
4. For each Alternative $i \in \mathcal{S}$, remove Alternative i from $\mathcal{S}$ if $\mu_i \leq \mu_j - \delta/2$ and $\rho_i^2 \leq \rho_j^2$ for some $j \in \mathcal{S}$.
5. For each Alternative $i \in \mathcal{S}$, calculate pPGS_i as in Equation (1).
6. If $\text{pPGS}_i > 1 - \alpha$ for some Alternative $i \in \mathcal{S}$, stop. Otherwise take additional replications and return to Step 2.

4. Checking the pEOC Stopping Rule

Like pPGS_i, pEOC_i can be computed as a k-dimensional integral. For even a modest number of alternatives, computing pEOC_i in this way is too time consuming. We present a sequence of ideas—similar to those in Section 3—for efficiently checking the pEOC stopping rule: identifying the alternative with the lowest pEOC and examining different approaches for evaluating its pEOC.

In contrast to the analysis of the pPGS stopping rule, determining which alternative has the lowest pEOC is straightforward. Recall that (k) is the index of the alternative with the highest posterior mean. Proposition 5 states a well-known result regarding the pEOC of Alternative (k).

Proposition 5. *Alternative* $(k) \in \arg\min_{1\le i\le k} \text{pEOC}_i$.

Proof of Proposition 5. For an arbitrary Alternative i,

$$\text{pEOC}_i = \mathbb{E}[W_{[k]} - W_i \mid \mathcal{E}] = \mathbb{E}[W_{[k]} \mid \mathcal{E}] - \mathbb{E}[W_i \mid \mathcal{E}],$$

where $[k]$ is the index of the alternative with the highest performance. Because the term $\mathbb{E}[W_{[k]} \mid \mathcal{E}]$ is independent of i, the alternative with the highest posterior mean, Alternative (k), has the lowest pEOC. □

Proposition 5 implies that to check the pEOC stopping rule, a procedure needs to compute the pEOC of only Alternative (k). It holds regardless of the form of the posterior distribution, for example, even when Assumptions 1–4 are not satisfied.

4.1. Cheap Upper Bounds

Computing a cheap upper bound on $\text{pEOC}_{(k)}$ and terminating when it drops below β will ensure that a procedure delivers the pEOC guarantee. Chick and Inoue (2001b) provide one such bound, though their name for it is somewhat of a misnomer:

$$\text{pEOC}_i = \mathbb{E}\left[\max_{j\ne i} (W_j - W_i)^+ \mid \mathcal{E}\right] \le \mathbb{E}\left[\sum_{j\ne i} (W_j - W_i)^+ \mid \mathcal{E}\right],$$

$$= \sum_{j\ne i} \mathbb{E}\left[(W_j - W_i)^+ \mid \mathcal{E}\right] \equiv \text{pEOC}_i^{\text{Bonf}}.$$

When the true variances are unknown, the terms summed in $\text{pEOC}_i^{\text{Bonf}}$ can be approximated (Branke et al. 2005). The $\text{pEOC}^{\text{Bonf}}$ bound holds even when Assumptions 1–4 do not, meaning that it can be applied when CRN are used.

We present another upper bound on $\text{pEOC}_{(k)}$—one derived from Slepian's inequality—that is, to the best of our knowledge, the first of its kind. Our approach makes use of the fact that the expected value of the nonnegative random variable $W_{[k]} - W_i$ can be written as an integral over its complementary cdf:

$$\begin{aligned}\text{pEOC}_i &= \mathbb{E}[W_{[k]} - W_i \mid \mathcal{E}] \\ &= \int_0^\infty \mathbb{P}(W_{[k]} - W_i > \delta \mid \mathcal{E})\, d\delta \\ &= \int_0^\infty \mathbb{P}(W_i < W_j - \delta \text{ for some } j \ne i \mid \mathcal{E})\, d\delta \\ &= \int_0^\infty \left[1 - \mathbb{P}(W_i \ge W_j - \delta \text{ for all } j \ne i \mid \mathcal{E})\right] d\delta \\ &= \int_0^\infty \left[1 - \text{pPGS}_i\right] d\delta, \end{aligned} \tag{2}$$

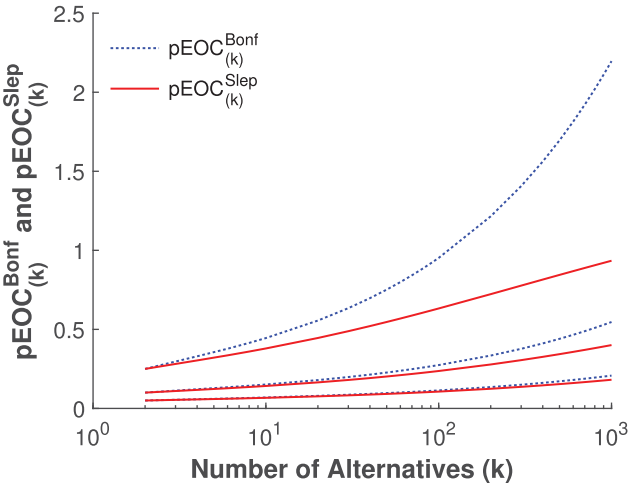
where pPGS_i is implicitly a function of δ. Proposition 1 implies that $\text{pPGS}_{(k)} \ge \text{pPGS}_{(k)}^{\text{Slep}}$ for all $\delta \ge 0$, hence for Alternative (k),

$$\begin{aligned}\text{pEOC}_{(k)} &\le \int_0^\infty \left[1 - \text{pPGS}_{(k)}^{\text{Slep}}\right] d\delta \\ &= \int_0^\infty \left[1 - \prod_{j\ne(k)} \mathbb{P}(W_{(k)} \ge W_j - \delta \mid \mathcal{E})\right] d\delta \\ &\equiv \text{pEOC}_{(k)}^{\text{Slep}}. \end{aligned} \tag{3}$$

The integrand in $\text{pEOC}_{(k)}^{\text{Slep}}$ features a product of $k-1$ cdfs and should therefore take roughly as long to numerically integrate as Equation (1). One minor difference is that the cdfs in $\text{pEOC}_{(k)}^{\text{Slep}}$ deal with the difference of two normal or t-distributed random variables, necessitating some form of analytical or numerical approximation when the true variances are unknown and the sample sizes are unequal.

To illustrate the potential slack in the $\text{pEOC}_{(k)}^{\text{Bonf}}$ and $\text{pEOC}_{(k)}^{\text{Slep}}$ bounds, we again take the approach of evaluating them for slippage configurations of posterior means with a common posterior variance ($\rho^2 = 1$). We use a line search to identify the spacing of posterior means for which $\text{pEOC}_{(k)} = \beta$ for the settings of $\beta = 0.05$, 0.1, and 0.25 and compute $\text{pEOC}_{(k)}^{\text{Bonf}}$ and $\text{pEOC}_{(k)}^{\text{Slep}}$ for the posterior distributions. Figure 2 shows that the tightness of both bounds deteriorates as the number of alternatives increases, with the $\text{pEOC}_{(k)}^{\text{Bonf}}$ bound faring worse. The quality of the bounds also appears to suffer for larger values of β. In many instances, the absolute error of these bounds is several times larger than the true value of $\text{pEOC}_{(k)}$, suggesting that the use of these conservative bounds as surrogates for $\text{pEOC}_{(k)}$ may cause a procedure to take significantly more replications than are necessary to deliver the pEOC guarantee. This conjecture is borne out in the experimental results of Section 5.

Figure 2. (Color online) Values of $\text{pEOC}_{(k)}^{\text{Bonf}}$ and $\text{pEOC}_{(k)}^{\text{Slep}}$ in a Slippage Configuration of Posterior Means in Which $\text{pEOC}_{(k)} = \beta$ for $\beta = 0.05, 0.10$, and 0.25



4.2. Numerical Integration

We turn our attention to the task of computing $\text{pEOC}_{(k)}$ without resorting to evaluating a k-dimensional integral. The proof of Proposition 5 indicates that one way is to calculate $\mathbb{E}[W_{[k]} \mid \mathcal{E}]$ and then subtract the posterior mean $\mathbb{E}[W_i \mid \mathcal{E}] = \mu_i$. Under the posterior distribution, the marginal pdf of $W_{[k]}$ is given by $f_{W_{[k]}|\mathcal{E}}(w) = \sum_{i=1}^{k} f_{W_i|\mathcal{E}}(w) \prod_{j \neq i} F_{W_j|\mathcal{E}}(w)$. One can thus calculate

$$\begin{aligned}\mathbb{E}[W_{[k]} \mid \mathcal{E}] &= \int_{-\infty}^{\infty} w \sum_{i=1}^{k} f_{W_i|\mathcal{E}}(w) \prod_{j \neq i} F_{W_j|\mathcal{E}}(w)\, dw \\ &= \sum_{i=1}^{k} \int_{-\infty}^{\infty} w \left[\prod_{j \neq i} F_{W_j|\mathcal{E}}(w) \right] f_{W_i|\mathcal{E}}(w)\, dw. \end{aligned} \tag{4}$$

Each of the k integrals in Equation (4) resembles that of Equation (1) with $\delta = 0$ and an extra factor of w in the integrand. We evaluate Equation (4) for the same experimental setup of random problem instances as in Section 3.3 and report the timing results in Table 2. Unsurprisingly, the computational times are roughly equivalent to k times the computational times in Table 1. For more than a modest number of alternatives, numerically integrating Equation (4) becomes too intensive for the purposes of checking the pEOC stopping rule.

Another approach to computing pEOC_i is to substitute Equation (1) for pPGS_i into Equation (2), yielding a two-dimensional integral:

$$\begin{aligned}\text{pEOC}_i &= \int_0^{\infty} \left[1 - \int_{-\infty}^{\infty} \left[\prod_{j \neq i} F_{W_j|\mathcal{E}}(w + \delta) \right] f_{W_i|\mathcal{E}}(w)\, dw \right] d\delta \\ &= \int_0^{\infty} \int_{-\infty}^{\infty} \left[1 - \prod_{j \neq i} F_{W_j|\mathcal{E}}(w + \delta) \right] f_{W_i|\mathcal{E}}(w)\, dw\, d\delta. \end{aligned} \tag{5}$$

Table 2 shows that the time required to evaluate Equation (5) scales better with k than does evaluating Equation (4), but for small numbers of alternatives it is more intensive. This may be because in Equation (5), the complexity of only the integrand increases with k, whereas for Equation (4), the number of integrals to evaluate also increases linearly in k. However, the operation of computing the latter can be more easily parallelized. Based on these timings, one might choose to evaluate Equation (4) for problems with $k \leq 100$ and evaluate Equation (5) for problems with $k > 100$. Although pEOC_i is several orders of magnitude more expensive to compute than pPGS_i, we show in Section 6 that the number of times a procedure needs to compute pEOC_i can be greatly reduced by using a Monte Carlo estimate.

5. Experimental Results

In this section, we investigate the tradeoff between the time spent checking the stopping rule and the total number of simulation replications taken by a procedure. Simulation experiments give a sense of the potential reduction in the number of replications from using the proposed methods for exactly checking the pPGS and pEOC stopping rules relative to using bounds on these quantities. Overall savings in a procedure's run time can be worked out by comparing the computational times reported in Sections 3 and 4—along with the frequency with which the stopping rule

Table 2. Average and 10th/90th Percentiles of Time (Seconds) to Numerically Integrate Equations (4) and (5) for Known/Unknown Variances and Different Numbers of Alternatives (k) Based on 100 Randomly Generated Problem Instances

	Equation (4)			Equation (5)		
k	10	100	1,000	10	100	1,000
Average (known)	0.07	0.90	23.0	0.99	1.66	5.45
Percentiles (known)	[0.06, 0.08]	[0.90, 1.30]	[19.5, 23.7]	[0.85, 1.23]	[1.48, 1.86]	[4.56, 6.29]
Average (unknown)	0.24	3.8	140	4.88	7.22	22.3
Percentiles (unknown)	[0.20, 0.28]	[3.6, 4.1]	[132, 147]	[4.03, 5.83]	[6.44, 8.35]	[20.0, 25.1]

is checked—to the average time required to simulate a replication and the number of replications taken.

We test the proposed methods on three allocation rules: EA, TS, and OCBA. Our goal in these experiments is not to compare the relative efficiency of allocation rules, but rather to gain an understanding of the potential savings from exactly checking the stopping rule for various well-known allocation rules. The EA rule allocates replications in batches of size k, with one additional replication taken from each alternative. Although EA can be inefficient—it continues to sample alternatives that are clearly inferior—it provides a baseline for potential savings. We run the Thompson sampling and OCBA rules fully sequentially, that is, allocating one replication at a time, and check the stopping rule after every replication. The TS rule randomly allocates the next replication to Alternative i with probability pPCS_i for $i = 1, \ldots, k$ (Russo 2020). This is achieved by generating a random problem instance from the posterior distribution and allocating the next replication to the alternative with the highest performance. We implement fully sequential versions of the OCBA rule as presented in Branke et al. (2007). Specifically, for the pPGS stopping rule, the OCBA rule allocates the next replication to the alternative that would yield the highest value of $\text{pPGS}_{(k)}^{\text{Slep}}$ if an extra observation were taken from it and its posterior mean were unchanged. Likewise, for the pEOC stopping rule, the OCBA rule allocates the next replication to the alternative that would yield the lowest value of $\text{pEOC}_{(k)}^{\text{Bonf}}$.

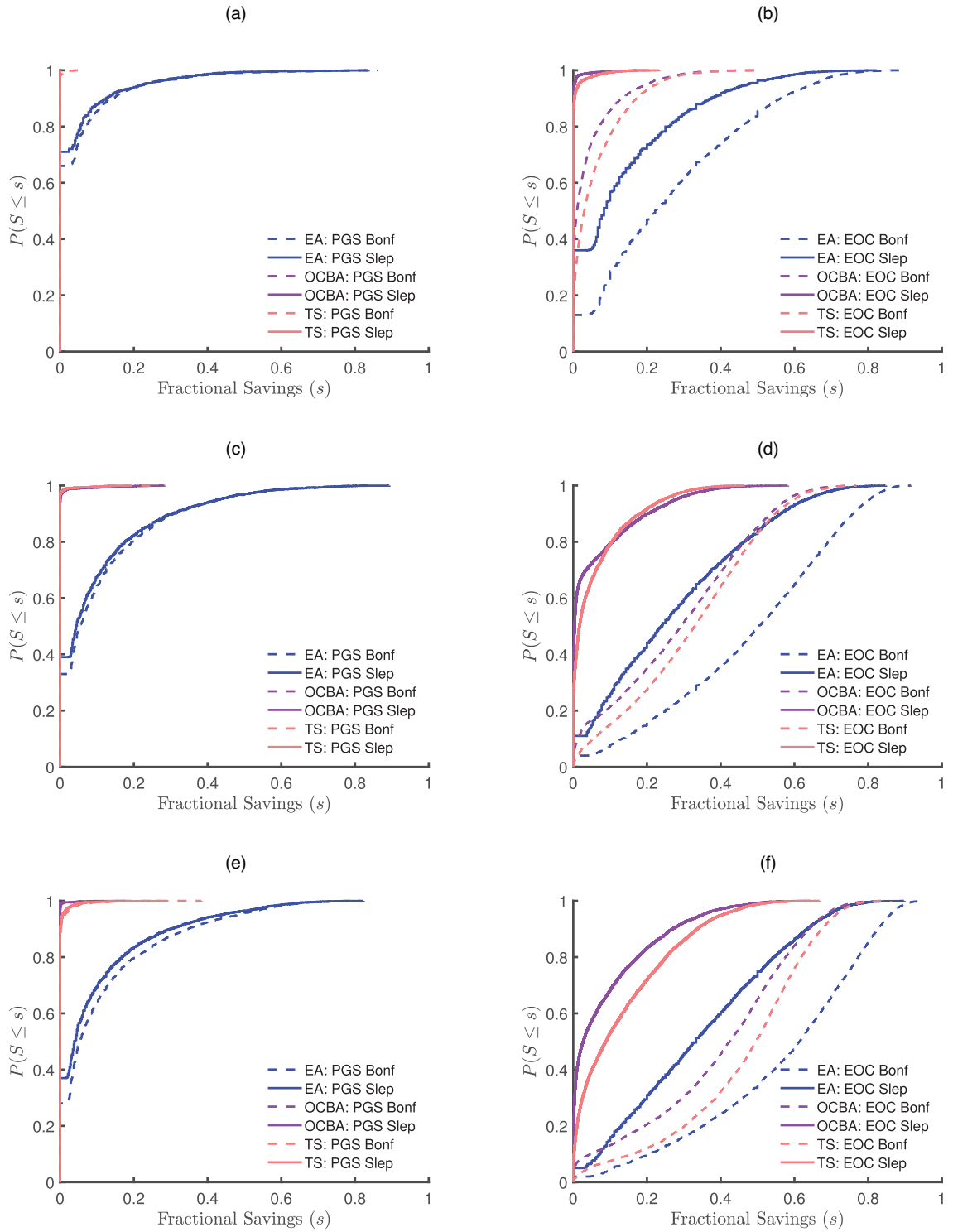
Let N_e be the (random) number of replications taken by a procedure using the proposed methods for exactly checking a given stopping rule and let N_b be the number of replications taken by a procedure that terminates when a bound for the posterior quantity first crosses the specified threshold. For both the pPGS and pEOC stopping rules, we implement the latter approach with the Bonferroni-type and Slepian-type bounds, for a total of four variations: (1) terminate when $\text{pPGS}_{(k)}^{\text{Bonf}} \geq 1 - \alpha$, (2) terminate when $\text{pPGS}_{(k)}^{\text{Slep}} \geq 1 - \alpha$, (3) terminate when $\text{pEOC}_{(k)}^{\text{Bonf}} \leq \beta$, and (4) terminate when $\text{pEOC}_{(k)}^{\text{Slep}} \leq \beta$. Branke et al. (2007) test the second and third variations in their experiments. For any given allocation rule and stopping rule, $N_e \leq N_b$ almost surely because of the slack in the bounds. We focus on the *relative* difference in the number of replications taken by a procedure when using the aforementioned approaches, defining the fractional savings from exactly checking the stopping rule as $S \equiv (N_b - N_e)/N_b$.

We use a conditional Monte Carlo method known as *splitting* to speed up the experiments (Asmussen and Glynn 2007). On each macroreplication, we record the outputs obtained up until the exact stopping condition is first met and then generate q realizations of the remainder of the procedure that each run until a given bound on the posterior quantity of Alternative (k) crosses the specified threshold. In this way, one macroreplication yields q observations of S for a given allocation rule and bound. While this method is not guaranteed to reduce the variance of the Monte Carlo estimator of $\mathbb{E}[S]$, it allows us to quickly generate identically distributed (albeit dependent) observations of the fractional savings. Multiple independent macroreplications then yield the desired estimator as a sample average over the macroreplication results. Because the allocation rules are carried out independent of how the stopping rule is checked, we further exploit the splitting method to simultaneously test pairs of bounds. More precisely, for a given allocation rule and posterior quantity of interest, each macroreplication is split into q realizations that terminate based on the corresponding Bonferroni-type bound *and* q realizations that terminate based on the corresponding Slepian-type bound. As a consequence, some of the observations of S for, say, OCBA with the $\text{pPGS}_{(k)}^{\text{Bonf}}$ and $\text{pPGS}_{(k)}^{\text{Slep}}$ bounds are dependent.

We test the procedures on random problem instances generated in three settings, denoted by RPI–1, RPI–2, and RPI–3. In each setting, $-\mu_1, \ldots, -\mu_k$ are independent and identically distributed (i.i.d.) from a Weibull distribution with scale parameter a and shape parameter b; that is, the density of $-\mu_i$ is given by $f(x; a, b) = (b/a)(x/a)^{b-1} e^{-(x/a)^b}$. We fix $b = 2$ and vary a across settings, taking $a = 4$ in RPI–1, $a = 1.5$ in RPI–2, and $a = 1$ in RPI–3. The Weibull distribution is chosen to produce problem instances for which most of the alternatives have performances clustered below that of the best alternative, with a tail of alternatives having very poor performances. A smaller scale parameter, a, causes the performances of the alternatives to become more clustered, hence random problem instances in RPI–3 have more good alternatives than those in RPI–1. In all settings, the sampling variances σ_i^2, $i = 1, \ldots, k$, are i.i.d. from a chi-squared distribution with four degrees of freedom, independent of μ_i.

In all experiments, we randomly generate problem instances with $k = 100$ alternatives. We choose an initial sample size of $n_0 = 5$ for each alternative and values of $\delta = 1$, $1 - \alpha = 0.90$ and $\beta = 0.50$. For this choice of δ, the proportions of good alternatives in RPI–1, RPI–2, and RPI–3 are approximately 6%, 36%, and 63%, respectively. Problem instances like those in RPI–1 may arise when randomly sampling alternatives over a large feasible region, whereas problem instances like those in RPI–2 and RPI–3 are more likely

Figure 3. (Color online) ecdfs of Fractional Savings for the pPGS and pEOC Stopping Rules Relative to Bonferroni and Slepian Bounds EA, TS, and OCBA Allocation Rules in RPI–1, RPI–2, and RPI–3



Notes. The height of each curve at any value of s is accurate to within ± 0.1 with 95% confidence. (a) pPGS in RPI–1. (b) pEOC in RPI–1. (c) pPGS in RPI–2. (d) pEOC in RPI–2. (e) pPGS in RPI–3. (f) pEOC in RPI–3.

to arise when the alternatives are feasible solutions visited by a simulation-optimization search (Boesel et al. 2003).

For each pairing of an allocation rule and bound, we generate 5,000 observations from $m = 100$ macroreplications with $q = 50$ splits. Figure 3 shows the

empirical cdfs (ecdfs) of the fractional savings from exactly checking the pPGS and pEOC stopping rules. Curves that are further from the upper-left corner indicate greater fractional savings. Average fractional savings and error estimates are reported in Table 3.

For all allocation rules, posterior quantities, and bounds, the fractional savings are greatest in RPI–3 and smallest in RPI–1. Our explanation for this is that when the true performances are clustered closer together, the posterior means are more likely to be clustered together, which in turn leads to more slack in the Bonferroni- and Slepian-type bounds. In all cases, the potential savings when using the EA rule are also greater than those when using more efficient allocation rules. This can be partially explained by the fact that failing to stop as soon as possible incurs a cost of at least k additional replications under EA.

For the EA rule, the upper tails of the fractional savings relative to the $\text{pPGS}^{\text{Bonf}}_{(k)}$ and $\text{pPGS}^{\text{Slep}}_{(k)}$ bounds exceed 10% for a nontrivial proportion of the runs in all settings. In contrast, the average fractional savings for the TS and OCBA rules are minimal, though in RPI–2 and RPI–3, there is the potential for fractional savings of up to 20%. Overall, the results suggest that when using efficient allocation rules, there is limited upside to exactly checking the pPGS stopping rule in realistic problem instances. The results also support the conjecture of Peng et al. (2018a) that in high-confidence settings, the $\text{pPGS}^{\text{Slep}}$ bound will closely align with the pPGS.

The fractional savings for the pEOC stopping rule with the $\text{pEOC}^{\text{Bonf}}_{(k)}$ and $\text{pEOC}^{\text{Slep}}_{(k)}$ bounds are much greater than those for the pPGS bounds. This can be explained by the looseness of the $\text{pEOC}^{\text{Bonf}}_{(k)}$ and $\text{pEOC}^{\text{Slep}}_{(k)}$ bounds, especially when the posterior means are clustered together. Compared with the $\text{pEOC}^{\text{Bonf}}_{(k)}$ bound that appears in the literature, exactly checking the stopping rule in RPI–2 and RPI–3 yields average fractional savings of about 25–45% for the TS and OCBA allocation rules, with the potential for two- and three-fold savings. Compared with the $\text{pEOC}^{\text{Slep}}_{(k)}$ bound we introduce, exactly checking the stopping rule in RPI–2 and RPI–3 yields average fractional savings of about 5–15% for the TS and OCBA allocation rules.

6. Monte Carlo Precheck

The results in Section 5 demonstrate how exactly checking the stopping rule can reduce the number of replications taken until a procedure terminates. These savings come at a cost, however, since frequently evaluating the posterior quantity of interest can amount to a nontrivial computational time, even with the proposed methods for accelerating these operations. Can one achieve the best of both worlds: smaller sample sizes with little extra computational effort?

We provide an affirmative answer by taking advantage of the observation that when the posterior quantity of interest is far from its intended threshold, it is not worth computing. Specifically, we explore the use of a cheap Monte Carlo estimate of the posterior quantity to "precheck" the stopping rule. By this we mean that the posterior quantity is exactly evaluated and compared with the threshold *only* when a Monte Carlo estimate of it crosses the threshold. In this way, the relatively expensive operation of exactly computing the posterior quantity is performed less frequently, yet the statistical validity of the procedure's Bayesian guarantee is maintained.

Monte Carlo estimators of pPGS_i and $\text{pEOC}_{(k)}$ can be attained by generating r independent random problem instances $\boldsymbol{W}^{(1)}, \ldots, \boldsymbol{W}^{(r)}$ from the posterior distribution and computing

$$\widehat{\text{pPGS}}_i \equiv \frac{1}{r}\sum_{l=1}^{r} \mathbf{1}\left\{W_i^{(l)} \geq W_{[k]}^{(l)} - \delta\right\} \tag{6}$$

and

$$\widehat{\text{pEOC}}_{(k)} \equiv \frac{1}{r}\sum_{l=1}^{r}\left(W_{[k]}^{(l)} - W_{(k)}^{(l)}\right), \tag{7}$$

where the index $[k]$ in Equations (6) and (7) is defined with respect to each $\boldsymbol{W}^{(l)}$. The variances of these Monte Carlo estimators can be further reduced by using conditional Monte Carlo schemes that exploit properties of elliptical distributions (Ahn and Kim 2018, Jian and Henderson 2020), though our empirical results suggest that these crude estimators are adequate for at least modest-sized problems. The finer precision offered by such techniques should be weighed against the added computational cost of implementing them.

Procedure 2 outlines the Monte Carlo precheck method for the pEOC stopping rule.

Procedure 2

1. Take n_0 initial replications from each alternative.
2. Calculate $\widehat{\text{pEOC}}_{(k)}$ as in Equation (7).
3. If $\widehat{\text{pEOC}}_{(k)} < \beta$, calculate $\text{pEOC}_{(k)}$ as in Equation (4) or (5). Otherwise take additional replications and return to Step 2.
4. If $\text{pEOC}_{(k)} < \beta$, stop. Otherwise take additional replications and return to Step 2.

To illustrate the effectiveness of the Monte Carlo precheck method, we implement Procedure 2 with $r = 10{,}000$ for the TS allocation rule and RPI–2 using the experimental setup from Section 5. We evaluate four methods for checking the stopping rule: using the

Table 3. Average Fractional Savings (with Error Estimates) Relative to Using Bounds for the EA, TS, and OCBA Allocation Rules

		Bound			
Problem instance	Allocation rule	$\text{pPGS}_{(k)}^{\text{Bonf}}$	$\text{pPGS}_{(k)}^{\text{Slep}}$	$\text{pEOC}_{(k)}^{\text{Bonf}}$	$\text{pEOC}_{(k)}^{\text{Slep}}$
RPI–1	EA	0.04 ± 0.02	0.04 ± 0.01	0.26 ± 0.03	0.14 ± 0.02
	TS	0.00 ± 0.00	0.00 ± 0.00	0.04 ± 0.01	0.00 ± 0.00
	OCBA	0.00 ± 0.00	0.00 ± 0.00	0.06 ± 0.01	0.00 ± 0.00
RPI–2	EA	0.11 ± 0.02	0.10 ± 0.02	0.47 ± 0.02	0.27 ± 0.02
	TS	0.00 ± 0.00	0.00 ± 0.00	0.29 ± 0.02	0.05 ± 0.01
	OCBA	0.00 ± 0.00	0.00 ± 0.00	0.32 ± 0.02	0.06 ± 0.01
RPI–3	EA	0.12 ± 0.02	0.10 ± 0.02	0.56 ± 0.02	0.34 ± 0.02
	TS	0.00 ± 0.00	0.00 ± 0.00	0.39 ± 0.02	0.09 ± 0.02
	OCBA	0.00 ± 0.00	0.00 ± 0.00	0.46 ± 0.02	0.13 ± 0.01

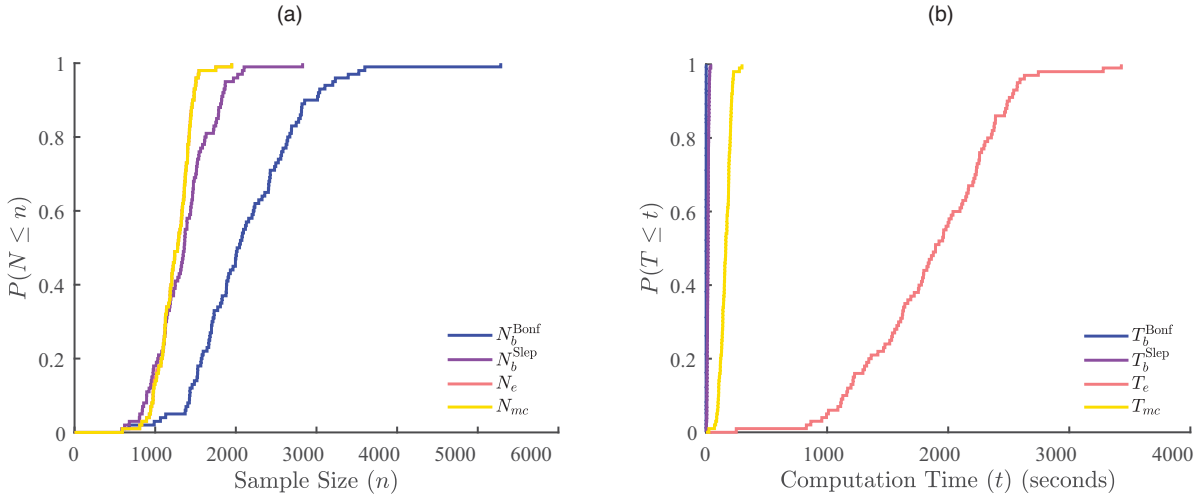
Note. Estimates of 0.00 ± 0.00 denote small average fractional savings that round down to zero.

$\text{pEOC}_{(k)}^{\text{Bonf}}$ bound, using the $\text{pEOC}_{(k)}^{\text{Slep}}$ bound, exactly calculating $\text{pEOC}_{(k)}$ via Equation (4), and Monte Carlo prechecking. For each method, we track the total number of replications taken, N, as well as the total computational time spent determining whether to terminate, denoted by T. Let N_{mc} and T_{mc} denote these quantities for the Monte Carlo precheck method. On each of $M = 100$ macroreplications, we couple the procedures by generating the replications determined by the TS allocation rule, simultaneously collecting sample size and timing data for all four procedures and terminating when the last procedure stops.

Because of the fixed initial sample size of $n_0 = 5$ for each alternative, 500 is a lower bound on N_b, N_e, and N_{mc}, as is evident in their ecdfs shown in Figure 4(a). The ecdf of N_{mc} is virtually indistinguishable from that of N_e, demonstrating that the Monte Carlo precheck method captures nearly all of the sample-size savings relative to using the $\text{pEOC}_{(k)}^{\text{Bonf}}$ and $\text{pEOC}_{(k)}^{\text{Slep}}$ bounds. (The unexpected crossing of the ecdf curves for N_b^{Slep} and N_e is likely because of the imprecision of the Welch approximation in the $\text{pEOC}_{(k)}^{\text{Slep}}$ bound.) Furthermore, the computational times associated with the Monte Carlo precheck method are roughly 1/10th those of exactly checking the stopping rule, as illustrated in Figure 4(b). Together, these plots demonstrate how the Monte Carlo precheck method can further reduce the overall run time of a procedure. Although using bounds remains the cheapest approach in terms of the computational time spent checking whether to terminate, T, in many if not most practical settings, we expect this to be outweighed by the excessive sample sizes depicted in Figure 4(a).

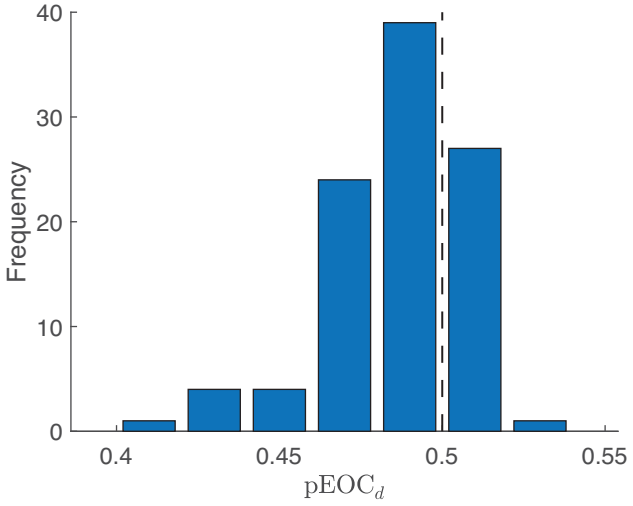
These results further suggest that using a Monte Carlo estimator to *directly* check the stopping rule may yield a procedure with sample sizes comparable to N_e *and* computational times comparable to T_b^{Bonf} and T_b^{Slep}. However, terminating a procedure based solely on

Figure 4. (Color online) ecdfs of the Total Sample Size and Time Spent Checking the pEOC Stopping Rule with the TS Allocation for RPI–2



Notes. The height of each curve at any value of n or t is accurate to within ±0.1 with 95% confidence. (a) Sample size. (b) Computation time.

Figure 5. (Color online) Histogram of the Posterior EOC of the Selected Alternative (pEOC_d) upon Termination When Using the Monte Carlo Estimator $\widehat{\text{pEOC}}_{(k)}$ to Directly Check the Stopping Rule



Note. The vertical dashed line indicates the nominal guarantee of $\text{pEOC}_d \le \beta = 0.5$.

when a Monte Carlo estimate crosses the specified threshold would invalidate the Bayesian guarantee. (Terminating a procedure based on Monte Carlo integration of Equations (1), (4), or (5), although appealing when the number of alternatives is large, would likewise lead to invalidated guarantees.) This raises a critical question: how might one weigh this added computational efficiency against the loss of a rigorous Bayesian guarantee?

Before answering, we stress that this is not purely a hypothetical question. When Assumption 1—normally distributed observations—is not justified, Slepian-type bounds are generally unavailable. Moreover, when one uses common random numbers or a prior distribution with correlated beliefs, Assumptions 3 and 4, respectively, are not satisfied and the posterior distribution no longer has a product form. Consequently, the k-dimensional integrals for pPGS_i and pEOC_i cannot be reduced to one- and two-dimensional integrals. In short, the $\text{pPGS}^{\text{Bonf}}_{(k)}$ and $\text{pEOC}^{\text{Bonf}}_{(k)}$ bounds are among the few computationally tractable means available for checking stopping rules while maintaining a procedure's statistical validity. From Figure 4, it is evident that using a Monte Carlo estimator to directly check the pEOC stopping rule would yield an appreciable reduction in a procedure's run time relative to using the $\text{pEOC}^{\text{Bonf}}_{(k)}$ bound.

In such cases, we advocate for using a Monte Carlo estimator to directly check a stopping rule, believing that the substantial savings in a procedure's run time more than compensate for a slight deterioration in its Bayesian guarantee. In support of this contention, we track a fifth procedure within the previous experiment—one that uses the Monte Carlo estimator $\widehat{\text{pEOC}}_{(k)}$ to directly check the stopping rule. Figure 5 shows a histogram of the pEOC of the selected alternative upon termination. Indeed, on 72% of the macroreplications, the pEOC guarantee is attained, and when it is not, departures from the nominal $\beta = 0.5$ are minimal. Alternatively, using a Monte Carlo upper confidence bound on $\text{pEOC}_{(k)}$ to directly check the stopping rule could make it more likely that $\text{pEOC}_d < 0.5$, although with slightly larger total sample sizes.

For those with reservations about sacrificing a rigorous Bayesian guarantee, we point out that even in the case where Assumption 1 holds, the Welch approximations commonly used in $\text{pPGS}^{\text{Bonf}}_{(k)}$, $\text{pPGS}^{\text{Slep}}_{(k)}$, and $\text{pEOC}^{\text{Bonf}}_{(k)}$ introduce an error that could compromise the guarantee, and yet they receive no such scrutiny.

7. Conclusion

We study R&S procedures that deliver Bayesian guarantees by tracking a posterior quantity of interest—such as the pPGS or pEOC—and terminating when it crosses a threshold. For the pPGS stopping rule, we devise several methods for restricting attention to a small set of alternatives that could satisfy the stopping rule. We also derive a new Slepian-type bound on the pEOC of the alternative with the highest posterior mean. We investigate ways to exactly compute the pPGS and pEOC of an alternative and demonstrate the looseness of cheap bounds on these quantities for problems with large numbers of alternatives. Numerical experiments indicate that implementing these exact methods can translate into savings in the number of replications a procedure takes. Savings for the pPGS stopping rule are limited, whereas those for the pEOC stopping rule can be substantial, particularly in problem instances in which the performances of alternatives are similar. In addition, potential savings relative to using Bonferroni-type bounds are greater than those relative to using Slepian-type bounds. An interesting, open question is how using the exact values (as opposed to bounds) of the pPGS and pEOC of alternatives *within* allocation rules could impact their sampling efficiency.

We also examine how using a Monte Carlo estimator to precheck a stopping rule can yield appreciable computational savings while preserving a procedure's Bayesian guarantee. Using a Monte Carlo estimator to directly check a stopping rule, without calculating the exact posterior quantity, can further reduce the computational time, while slightly weakening the Bayesian guarantee. This is an appealing approach for situations

in which the posterior quantity cannot be conveniently computed via numerical integration, for example, when sampling or prior beliefs are correlated. Although these Monte Carlo calculations did not represent a significant computational hurdle in our experiments, further efficiencies may be possible through the use of variance reduction schemes or other strategies such as quasi-Monte Carlo or randomized quasi-Monte Carlo.

It is hard to provide absolute recommendations among the various methods because their relative efficiencies depend on the costs of simulating replications and checking the stopping rules. These costs can vary from macroreplication to macroreplication and across computing platforms. Using a Monte Carlo estimator to directly check the stopping rule appears to have little loss with regards to premature stopping, scales well with the number of alternatives, and applies under the most general conditions, requiring only the ability to sample from the posterior distribution. For users who want the assurance of a rigorous Bayesian guarantee, the Monte Carlo precheck method performs well on problems with up to at least 100 alternatives. For larger problems, it may be advantageous to check the stopping rule less frequently, for example, by allocating replications in batches.

Our work also motivates further research on the scalability of Bayesian selection procedures for problems with large numbers of alternatives. Important concerns in this setting are the computational costs of integrating or estimating posterior quantities of interest and implementing allocation rules. Although frequentist procedures have been developed to tackle R&S problems with up to a million alternatives (Ni et al. 2017, Pei et al. 2018), the limits of Bayesian selection procedures remain underexplored. Moreover, the standard guarantees on pPGS_d and pEOC_d may be ill-suited objectives for large-scale problems.

R&S problems with many alternatives are well suited to parallel computation (Luo and Hong 2011, Luo et al. 2015, Ni et al. 2017). How might parallel computing be exploited in the Bayesian setting beyond the obvious use in running simulation replications of alternatives in parallel? It would be straightforward to use parallel computing for the proposed Monte Carlo precheck, but that step is not computationally intensive, so the advantage in doing so is not clear. But what of numerical integration schemes? The adaptive quadrature integration methods we use here adaptively choose points at which to compute the marginal posterior densities and cdfs. These selections depend heavily on the part of the support of $f_{W_i|\mathcal{E}}$ where the density takes nontrivial values—as in Equations (1), (4), and (5)—and thus depend on the Alternative i under consideration. Therefore, an effective use of parallel computing to accelerate numerical integration is not straightforward. However, it is conceivable that some scheme could be engineered to precompute the densities and cdfs on some grid of values using multiple processors and then share those values across other processors dedicated to evaluating metrics like the pPGS and pEOC of alternatives of interest. Such a scheme would require careful coordination across different alternatives but might yield important speedups in numerical evaluation. We leave this topic for future research.

Acknowledgments

The authors thank Steve Chick, Peter Frazier, and Yijie Peng for helpful conversations.

References

Ahn D, Kim KK (2018) Efficient simulation for expectations over the union of half-spaces. *ACM Trans. Modeling Comput. Simulation* 28(3):1–20.

Asmussen S, Glynn PW (2007) *Stochastic Simulation: Algorithms and Analysis*, vol. 57 (Springer Science & Business Media, New York).

Berger JO (1993) *Statistical Decision Theory and Bayesian Analysis* (Springer-Verlag, New York).

Boesel J, Nelson BL, Kim SH (2003) Using ranking and selection to "clean up" after simulation optimization. *Oper. Res.* 51(5): 814–825.

Branke J, Chick SE, Schmidt C (2005) New developments in ranking and selection: An empirical comparison of the three main approaches. Kuhl ME, Steiger NM, Armstrong FB, Joines JA, eds. *Proc. 2005 Winter Simulation Conf.* (Institute of Electrical and Electronics Engineers, Piscataway, NJ), 708–717.

Branke J, Chick SE, Schmidt C (2007) Selecting a selection procedure. *Management Sci.* 53(12):1916–1932.

Chen Y, Ryzhov I (2019) Balancing optimal large deviations in sequential selection. Technical report, University of Maryland, Baltimore).

Chen CH, Chick SE, Lee LH, Pujowidianto NA (2015) Ranking and selection: Efficient simulation budget allocation. Fu MC, ed. *Handbook of Simulation Optimization* (Springer, New York), 45–80.

Chen CH, Lin J, Yücesan E, Chick SE (2000) Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems* 10(3):251–270.

Chick SE (2006) Subjective probability and Bayesian methodology. Henderson SG, Nelson BL, eds. *Simulation*, vol. 13 of Handbooks in Operations Research and Management Science (Elsevier, North Holland), 225–257.

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

Chick SE, Inoue K (2001b) New two-stage and sequential procedures for selecting the best simulated system. *Oper. Res.* 49(5): 732–743.

Chick SE, Branke J, Schmidt C (2010) Sequential sampling to myopically maximize the expected value of information. *INFORMS J. Comput.* 22(1):71–80.

DeGroot MH (2004) *Optimal Statistical Decisions* (John Wiley & Sons, Hoboken, NJ).

Deng A, Lu J, Chen S (2016) Continuous monitoring of A/B tests without pain: Optional stopping in Bayesian testing. *Proc. IEEE Internat. Conf. on Data Science and Advanced Analytics* (IEEE, New York), 243–252.

Dmitriev P, Gupta S, Dong Woo K, Vaz G (2017) A dirty dozen: Twelve common metric interpretation pitfalls in online controlled

experiments. Matwin S, Yu S, Farooq F, eds. *Proc. 23rd ACM SIGKDD Internat. Conf. on Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York), 1427–1436.

Eckman DJ (2019) Reconsidering ranking-and-selection guarantees. PhD thesis, Cornell University, Ithaca, NY.

Even-Dar E, Mannor S, Mansour Y (2006) Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *J. Machine Learning Res.* 7(June):1079–1105.

Frazier P, Powell W, Dayanik S (2009) The knowledge-gradient policy for correlated normal beliefs. *INFORMS J. Comput.* 21(4):599–613.

Glynn P, Juneja S (2004) A large deviations perspective on ordinal optimization. Ingalls RG, Rossetti MD, Smith JS, Peters BA, eds. *Proc. 2004 Winter Simulation Conf.* (IEEE, New York), 577–585.

Glynn P, Juneja S (2018) Selecting the best system and multi-armed bandits. Preprint, submitted September 10, 2018, https://arxiv.org/abs/1507.04564v3.

Görder B, Kolonko M (2019) Ranking and selection: A new sequential Bayesian procedure for use with common random numbers. *ACM Trans. Modeling Comput. Simulation* 29(1):1–24.

Hong LJ, Nelson BL, Xu J (2015) Discrete optimization via simulation. Fu MC, ed. *Handbook of Simulation Optimization* (Springer, New York), 9–44.

Hunter SR, Pasupathy R (2010) Large-deviation sampling laws for constrained simulation optimization on finite sets. Johansson B, Jain S, Montoya-Torres J, Hugan J, Yücesan E, eds. *Proc. 2010 Winter Simulation Conf.* (IEEE, New York), 995–1002.

Jian N, Henderson SG (2020) Estimating the probability that a function observed with noise is convex. *INFORMS J. Comput.* 32(2):376–389.

Johari R, Pekelis L, Koomen P, Walsh D (2017) Peeking at A/B tests: Why it matters, and what to do about it. Matwin S, Yu S, Farooq F, eds. *Proc. 23rd ACM SIGKDD Internat. Conf. on Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York), 1517–1525.

Kim SH, Nelson BL (2001) A fully sequential procedure for indifference-zone selection in simulation. *ACM Trans. Modeling Comput. Simulation* 11(3):251–273.

Kim SH, Nelson BL (2006) Selecting the best system. Henderson SG, Nelson BL, eds. *Simulation*, vol. 13 of Handbooks in Operations Research and Management Science (Elsevier, North Holland), 501–534.

Luo J, Hong LJ (2011) Large-scale ranking and selection using cloud computing. Jain S, Creasey RR, Himmelspach J, White KP, Fu M, eds. *Proc. 011 Winter Simulation Conf.* (IEEE, New York), 4051–4061.

Luo J, Hong LJ, Nelson BL, Wu Y (2015) Fully sequential procedures for large-scale ranking-and-selection problems in parallel computing environments. *Oper. Res.* 63(5):1177–1194.

Ni E, Ciocan D, Henderson S, Hunter S (2017) Efficient ranking and selection in parallel computing environments. *Oper. Res.* 65(3):821–836.

Pei L, Nelson BL, Hunter S (2018) A new framework for parallel ranking & selection using an adaptive standard. Rabe M, Juan AA, Mustafee N, Skoogh A, Jain S, Johansson B, eds. *Proc. 2018 Winter Simulation Conf.* (IEEE, New York), 2201–2212.

Peng Y, Chen CH, Fu MC, Hu JQ (2016) Dynamic sampling allocation and design selection. *INFORMS J. Comput.* 28(2):195–208.

Peng Y, Chen CH, Fu MC, Hu JQ (2018a) Gradient-based myopic allocation policy: An efficient sampling procedure in a low-confidence scenario. *IEEE Trans. Automated Control* 63(9):3091–3097.

Peng Y, Chong EK, Chen CH, Fu MC (2018b) Ranking and selection as stochastic control. *IEEE Trans. Automated Control* 63(8):2359–2373.

Pichitlamken J, Nelson BL, Hong LJ (2006) A sequential procedure for neighborhood selection-of-the-best in optimization via simulation. *Eur. J. Oper. Res.* 173(1):283–298.

Russo D (2020) Simple Bayesian algorithms for best-arm identification. *Oper. Res.* 68(6):1625–1647.

Shampine LF (2008) Vectorized adaptive quadrature in MATLAB. *J. Comput. Appl. Math.* 211(2):131–140.

Slepian D (1962) The one-sided barrier problem for Gaussian noise. *Bell Systems Tech. J.* 41(2):463–501.

Tamhane AC, Bechhofer RE (1977) A two-stage minimax procedure with screening for selecting the largest normal mean. *Comm. Statist. Theory Methods* 6(11):1003–1033.

Wald A (1973) *Sequential Analysis* (Courier Corporation, North Chelmsford, MA).

Welch BL (1938) The significance of the difference between two means when the population variances are unequal. *Biometrika* 29(3/4):350–362.

Xie J, Frazier PI, Chick SE (2016) Bayesian optimization via simulation with pairwise sampling and correlated prior beliefs. *Oper. Res.* 64(2):542–559.