RECONSIDERING RANKING-AND-SELECTION GUARANTEES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

David John Eckman

May 2019

RECONSIDERING RANKING-AND-SELECTION GUARANTEES

David John Eckman, Ph.D.

Cornell University 2019

This dissertation deals with the various statistical guarantees delivered by *ranking-and-selection* (R&S) procedures: a class of methods designed for the problem of selecting the best from among a finite number of simulated systems. Examples of such guarantees include ensuring that an optimal or near-optimal system is selected with high probability or that the expected performance gap between the selected system and the optimal system is below a specified threshold. We explore three fundamental issues concerning R&S guarantees that are of practical and theoretical interest to the simulation community.

First, we discuss the shortcomings of the popular indifference-zone-inspired guarantee on the probability of correct selection (PCS) and argue that delivering a guarantee on the probability of good selection (PGS) is a more justifiable goal. We present an overview of the PGS guarantee and examine numerous techniques for proving the PGS guarantee, including sufficient conditions under which R&S procedures that deliver the IZ-inspired PCS guarantee also deliver the PGS guarantee.

Second, we study Bayesian R&S guarantees, contrasting them with their frequentist counterparts and investigating the practical implications of this distinction. R&S procedures deliver Bayesian guarantees by terminating when a posterior quantity of interest—e.g., the posterior PCS or PGS—crosses some threshold. We develop several methods for improving the computational efficiency of checking this stopping rule when there are a large number of systems and

demonstrate their effectiveness compared to existing approaches.

Third, we study R&S guarantees for the setting in which a R&S procedure is run after a simulation-optimization search. We show that for searches that use the observed performance of explored systems to identify new systems, the simulation replications are conditionally dependent given the sequence of returned systems. We demonstrate that reusing replications taken during a search as input to a R&S procedure can result in an empirical PCS or PGS below the guaranteed threshold. Based on these negative findings, we call into question the guarantees of established R&S procedures that reuse search data.

BIOGRAPHICAL SKETCH

David John Eckman grew up in Fargo, North Dakota. He attended the University of Pittsburgh from the fall of 2009 to the spring of 2014 and received a Bachelor of Science in industrial engineering. As an undergraduate, he did three co-op rotations in the supply chain logistics department at Heinz North America. He met his wife at Pitt and they were later married in May 2017. In the fall of 2014, he began his Ph.D. in Operations Research at Cornell University. After completing his Ph.D., he will start a postdoctoral position at Northwestern University.

This dissertation is dedicated to my freshman calculus professor and friend, Chris Lennard, who showed me just how fun teaching and research can be.

ACKNOWLEDGEMENTS

There are so many people I wish to thank for their role in my Ph.D. journey.

I am thankful to my advisor, Prof. Shane Henderson, for his guidance and mentoring throughout my time at Cornell. I cannot express how much I have learned from him about being a better researcher, teacher, and colleague. I would also like to thank my other committee members, Profs. Peter Frazier and Eilyan Bitar, for their insights and dedication to seeing me through the program. I also owe a debt of thanks to the ORIE faculty and staff who have helped me too many times to count. And to the ORIE Ph.D. students I have befriended over the years: thank you making ORIE feel like a second home.

I would also like to thank those at Pitt who prepared me for this Ph.D. Thank you to Profs. Andrew Schaefer, Jeff Kharoufeh, and Chris Lennard for their invaluable mentoring and to Profs. Anna Balazs and Natasa Vidic for introducing me to simulation and fostering my passion for the subject.

As for my family, I am profoundly thankful to my parents for their steadfast support for my education over the years. I am also eternally grateful to my wife, Kaley, for her constant love and support and for putting up with the schedule that comes with having a Ph.D. student as a husband. I also wish to thank our dog, Kiah, for the countless walks and play sessions that led to research breakthroughs.

Finally, I wish to thank Barry Nelson, Jeff Hong, Michael Fu, Steve Chick, and Yijie Peng (among others) for their enlightening conversations.

This research has been supported by the National Science Foundation under grants DGE-1144153 and DGE-1650441. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In many problems of decision-making under uncertainty, the objective is to identify a solution that is optimal with respect to some performance measure of interest. For example, an ambulance fleet operator may wish to locate ambulance bases within a metropolitan area so as to minimize the expected call response time. One approach to solving this kind of optimization problem is to use stochastic simulation to evaluate the performance of a given feasible solution. More specifically, for a given arrangement of ambulance bases, a discrete-event simulation model can be used to simulate the fleet's operations for, say, a year's worth of call times and locations. Simulation can be an appealing method for problems such as this wherein the complexity of the physical system being studied does not admit an analytical expression for the objective function. Examples of approaches to decision-making under uncertainty that use simulation include multi-armed bandits (Bubeck and Cesa-Bianchi, 2012), stochastic programming (Birge and Louveaux, 2011), approximate dynamic programming (Powell, 2007), and simulation optimization (Fu, 2015; Fu and Henderson, 2017).

Compared to deterministic-optimization problems, simulation-optimization problems must deal with the added challenges of random error associated with evaluating a solution's performance and limited availability of unbiased gradient estimates. Moreover, finding a *global* optimal solution over a large feasible region is a hard task. A compromise is to restrict attention to a finite set of candidate solutions—henceforth referred to as systems or alternatives—and seek to select the best from among them. Candidate systems can be identified in various ways: enumeration when the feasible region is small and discrete; discretization

when the feasible region is continuous; or some form of search. Additionally, the number of candidate systems should be small enough so that each system can be simulated to some degree within the available computational budget. When the performances of different systems are regarded as unrelated to the systems' locations in the feasible region (i.e., no structural relationship is assumed), the resulting problem is referred to as *ranking and selection* (R&S); see Bechhofer et al. (1995), Kim and Nelson (2006b), and Chen et al. (2015) for introductory references.

Whereas early R&S problems dealt with physical experiments carried out on a small number of systems, advances in computing capabilities have made it tractable to solve R&S problems involving thousands of systems. The ubiquity of parallel computing environments has further pushed the limits of R&S procedures and motivated the design of efficient procedures that scale well with the number of systems (Luo et al., 2015; Ni et al., 2017; Hunter and Nelson, 2017). Commercial simulation software (e.g., Simio 2019) also features implementations of several R&S procedures, making them widely available to simulation practitioners.

A major strength of R&S procedures relative to other simulation-optimization algorithms is their ability to offer *finite-time* statistical guarantees. Many other simulation-optimization algorithms, on the other hand, only provide guarantees on their asymptotic performance, e.g., convergence to a local optimal solution (Broadie et al., 2011; Chang et al., 2013; Zabinsky, 2015). Upon termination, R&S procedures can instead deliver statistical guarantees on the performance of the returned system or subset of systems relative to the best. We focus on the setting in which the decision-maker specifies a desired guarantee

and runs a R&S procedure until it has taken a sufficient number of observations from each system to provide the guarantee.

The dominant paradigm in the R&S literature assumes that observations—the outputs of simulation replications—are normally distributed. This assumption can be approximately satisfied by batching individual replications into macroreplications and appealing to the Central Limit Theorem. For settings in which this assumption is not justified, other procedures that analyze the large-deviations behavior of the observations have been developed (Glynn and Juneja, 2004; Hunter and Pasupathy, 2010; Glynn and Juneja, 2018). While we mainly discuss R&S procedures that assume normality in this dissertation, most of the conclusions hold even when the normality assumption does not.

Statistical guarantees for R&S procedures come in two flavors: frequentist and Bayesian. Frequentist guarantees are made with respect to repeated runs of a given procedure on a fixed problem instance. Bayesian guarantees are instead made with respect to the decision-maker's posterior belief about the unknown problem instance given the collected observations and any prior information. These two kinds of guarantees are fundamentally different, both in terms of how they are interpreted and how procedures are designed to deliver them.

Under both the frequentist and Bayesian treatments, there are several common statistical guarantees for R&S procedures. For example, many R&S procedures guarantee that the probability of selecting one of the best systems—referred to as the probability of correct selection (PCS)—exceeds a given threshold $1 - \alpha$. Alternatively, some R&S procedures guarantee that the probability of selecting a system whose performance is within a given tolerance δ of the best—referred to as the probability of good selection (PGS)—exceeds $1 - \alpha$. Another

popular guarantee is to ensure that the expected opportunity cost (EOC)—the expected difference in performance between the selected system and the best—is less than some threshold β. The parameters appearing in these guarantees ($1-\alpha$, δ, and β) are chosen by the decision-maker to reflect his or her tolerance toward making a suboptimal selection.

The choice of guarantee is a primary consideration in the design and selection of R&S procedures. This dissertation provides a deeper understanding of these guarantees by examining three instances where aspects of them have been overlooked or overvalued. Specifically, we study how the choice of the guarantee affects the empirical performance—efficiency and accuracy—of an R&S procedure. Conversely, we analyze how the way in which an R&S procedure is run affects its guarantees and their interpretations. We believe that the results and discussion in this dissertation will be of general interest to simulation researchers seeking to develop new R&S procedures. Moreover, simulation software developers can benefit from considering these principles when choosing which procedures and guarantees to offer to users.

In Chapter 2, we argue that the PGS guarantee should overtake the popular indifference-zone-inspired PCS guarantee as the predominant design goal for frequentist R&S procedures. In Chapter 3, we contrast Bayesian and frequentist R&S guarantees and explore ways to improve the computational efficiency of checking stopping rules that deliver Bayesian guarantees. In Chapter 4, we show that the guarantees of R&S procedures that reuse search data can be undermined by the fact that data collected during a search are conditionally dependent—an observation that had until now been overlooked. We outline several open areas for future research in Chapter 5.

CHAPTER 2

FIXED-CONFIDENCE, FIXED-TOLERANCE GUARANTEES FOR RANKING-AND-SELECTION PROCEDURES

The majority of this chapter was submitted for publication at a journal (Eckman and Henderson, 2019a).

2.1 Introduction

An early development in R&S was the indifference-zone (IZ) formulation of Bechhofer (1954), and it has played a dominant role in the design of R&S procedures ever since; see, for example, the recent procedures of Frazier (2014) and Zhong and Hong (2017). Under the IZ formulation, procedures are designed to guarantee that the best system will be chosen with probability exceeding $1 - \alpha$ if its performance is at least δ better than those of the other systems, where both $1 - \alpha$ and δ are specified by the decision-maker. That is, IZ-inspired procedures guarantee that the probability of correct selection (PCS) is above a specified threshold whenever the best system is sufficiently better than the others. A serious shortcoming of this guarantee is that no statement is made about how a given procedure performs when there are close contenders to the best system. For such problem instances, it is reasonable to expect that the decision-maker would be equally satisfied with selecting any system whose performance is, in some sense, "good."

A more suitable goal is to guarantee that for *any* problem instance, a system with performance strictly within δ of the best will be chosen with probability exceeding $1 - \alpha$. In this case, the value of δ represents the decision-maker's

tolerance towards making a suboptimal decision. Compared to the IZ-inspired PCS guarantee, this guarantee on the probability of good selection (PGS) has received far less attention in the R&S literature and has often been treated as a secondary goal. We survey this neglected area of research and argue that the time is right for the PGS guarantee to displace the IZ-inspired PCS guarantee as the leading design goal for procedures delivering frequentist, fixed-confidence guarantees.

With this objective in mind, we explain the flaws of the IZ-inspired PCS guarantee and how it has come to be misunderstood. We also discuss how some procedures designed to deliver this guarantee may in fact be inefficient for problem instances in which there are several near-best systems. Efforts to show that certain procedures that deliver the IZ-inspired PCS guarantee also deliver the PGS guarantee have at times relied on erroneous proofs. Moreover, much of the past research on sufficient conditions under which the IZ-inspired PCS guarantee implies the PGS guarantee appears to have been overlooked by the simulation community. We clarify and extend these results and elucidate the key ideas behind other past proof techniques to present a unified treatment of the PGS guarantee.

In this chapter, we consider only R&S problems for which the systems have scalar performances. Although the IZ formulation has been extended to multi-objective R&S problems (Chen and Lee, 2009; Teng et al., 2010), we do not address it in this chapter because—for this class of problems—the definition of good selection remains unsettled (Branke et al., 2016; Hunter et al., 2019). We also do not consider R&S problems with stochastic constraints (Andradóttir and Kim, 2010; Hong et al., 2015). In addition, we focus on the setting of running an

R&S procedure until a fixed-confidence guarantee can be delivered. In contrast, many Bayesian R&S procedures are designed to be run until a fixed simulation budget has been exhausted (Chen et al., 2015).

Although we contend that the PGS guarantee is superior to the IZ-inspired PCS guarantee, it is not without its own shortcomings. Like the IZ-inspired PCS guarantee, the PGS guarantee offers no assurance about the performance of the selected system when it is a bad one. An even stronger goal that partially addresses this issue is guaranteeing that the expected linear loss—the expected difference in performance between the chosen system and the best system—does not exceed some threshold (Chick and Inoue, 2001b). Expected linear loss (also known as expected opportunity cost) may be a more pertinent metric for business and engineering decisions, but it can also be difficult to interpret and analyze. In relation to the PGS guarantee, expected linear loss can be used to bound PGS from below via Markov's inequality (Chick and Wu, 2005).

The R&S problem has also been studied from a Bayesian perspective in which the decision-maker's uncertainty about the unknown problem instance is described by a prior distribution (Berger and Deely, 1988; Gupta and Miescke, 1996). By taking observations, the decision-maker is able to update their beliefs in the form of a posterior distribution that can then be used to make inferences about selections. The *posterior* PGS of a given system is defined as the probability—under the posterior distribution—that the random problem instance is one for which that system is within δ of the best. An important advantage of the Bayesian formulation is that the posterior PGS of a system can be computed at any time and used in a stopping condition to deliver a Bayesian PGS guarantee. Frequentist and Bayesian PGS guarantees are fundamentally

different; we choose to focus on frequentist PGS guarantee in this chapter to accentuate its connection to the IZ-inspired PCS guarantee.

A related problem in which delivering fixed-confidence selection guarantees is a featured goal is the pure exploration problem for multi-armed bandits (Audibert et al., 2010). With a few exceptions (Chandrasekaran and Karp, 2014), the IZ formulation does not appear in the PCS guarantee studied in this setting (Garivier and Kaufmann, 2016). Furthermore, the PGS guarantee is referred to as the probably approximately correct (PAC) selection guarantee where "probably" refers to the fixed confidence, $1 - \alpha$, and "approximately correct" refers to the fixed tolerance, δ (Even-Dar et al., 2002, 2006). Research in this area has focused on the design and complexity analysis of efficient selection (Mannor and Tsitsiklis, 2004; Karnin et al., 2013) and subset-selection procedures (Kalyanakrishnan and Stone, 2010; Kalyanakrishnan et al., 2012; Zhou et al., 2014). Another variant of the PAC selection guarantee arises in machine learning and data mining for the problem of identifying a hypothesis with a low misclassification probability (Schuurmans and Greiner, 1995; Domingo et al., 2002).

Another approach for—among other things—classifying systems as good or bad is ordinal optimization (Lau and Ho, 1997; Ho et al., 2000). In the ordinal optimization paradigm, systems are classified based on the ordering of their performances, with either the top m systems or top m percent of systems being designated as good. This ordinal perspective of goodness, however, does not take into account potentially large differences in the performances of top systems. In this case, the decision-maker may not be satisfied with selecting a system that has a high ordering but a poor performance relative to the best. Instead, more control over the performance of a selected system can be achieved

by defining goodness as a cardinal property, i.e., regarding any system whose performance is within δ of the best as a good system, as we do in this chapter.

The remainder of this chapter is outlined as follows. In Section 2.2, we argue that the PGS guarantee is superior in many ways to the IZ-inspired PCS guarantee. In Sections 2.3 and 2.4, we present sufficient conditions under which selection and subset-selection procedures with the IZ-inspired PCS guarantee simultaneously deliver the PGS guarantee. In Section 2.5, we review other methods for proving the PGS guarantee and highlight several technical issues that arise. We discuss future research directions for the PGS guarantee in Section 2.6.

2.2 The IZ-Inspired PCS Guarantee versus the PGS Guarantee

Before mathematically defining the various fixed-confidence guarantees, we introduce some standard notations for R&S problems. Suppose there are k systems with performances $\mu_1, \ldots, \mu_k$ where, without loss of generality, we assume a higher performance is better. We refer to the vector $\mu = (\mu_1, \ldots, \mu_k)$ as the *configuration* of the systems' performances and use $[\cdot]$ to denote the indices of the systems when ordered by their performances, i.e., $\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}$. If some systems have tied performances, we will assume that the ordered indexing of the systems is arbitrary and fixed.

We define a selection procedure as one that determines how many observations should be taken from each system and then ultimately selects a single system as the best. The index of the selected system, denoted by K, is a random variable since the observations, and hence the estimators of systems' performances, are themselves random variables. Correct selection is then defined as

the event CS := $\{\mu_K = \mu_{[k]}\}$. Under this definition, when there are multiple systems with performances tied for the best, choosing any of the best systems is considered a correct selection. A fixed-confidence guarantee on the probability of correct selection for any configuration takes the form

$$\mathbb{P}_\mu(\text{CS}) \geq 1 - \alpha \quad \text{for all } \mu, \tag{Goal PCS}$$

where $1 - \alpha \in (1/k, 1)$ is the user-specified confidence and $\mathbb{P}_\mu$ is the probability measure induced through the combination of the selection procedure's sampling and the configuration of the systems' performances.

Without further assumptions, satisfying Goal PCS can be computationally expensive. Indeed, when the best system is only slightly better than the second-best system, a substantial amount of computational effort could be needed to distinguish between the two systems. From the decision-maker's perspective, it seems unreasonable to demand that a procedure makes a correct selection with high probability for *any* positive gap in performance. For this reason, selection procedures are rarely designed to deliver Goal PCS. The procedure of Fan et al. (2016) comes close; it guarantees that whenever there is a unique best system, the probability that it will be selected is at least $1 - \alpha$. The sampling complexity necessary to attain Goal PCS has also been studied in the multi-armed-bandit literature (Kaufmann et al., 2014).

In order to circumvent the issue with Goal PCS, Bechhofer (1954) proposed the indifference-zone formulation. The idea behind the IZ formulation is to specify a parameter $\delta > 0$ that divides the space of configurations into the preference zone $\text{PZ}(\delta) := \{\mu : \mu_{[k]} - \mu_{[k-1]} \geq \delta\}$ and the indifference zone $\text{IZ}(\delta) := \{\mu : \mu_{[k]} - \mu_{[k-1]} < \delta\}$. In the preference zone, the best system's performance is at least δ better than that of the second-best system, whereas in the

indifference zone, there are systems with performances within δ of the best. Under the IZ formulation, the PCS guarantee states that, only for configurations in the preference zone, the best system is selected with high probability:

$$\mathbb{P}_\mu(\text{CS}) \geq 1 - \alpha \quad \text{for all } \mu \in \text{PZ}(\delta). \tag{Goal PCS-PZ}$$

Ever since the conception of the indifference zone, Goal PCS-PZ has been a popular—though often misinterpreted—goal for selection procedures.

In regards to good selection, the R&S literature has been inconsistent about whether a system with knife-edge performance $\mu_{[k]} - \delta$ is considered good or bad. Although this distinction may appear minor, it is important here that we consider the IZ formulation when defining good selection. Under the IZ formulation, the PCS guarantee holds whenever $\mu \in \text{PZ}(\delta)$, i.e., $\mu_{[k]} - \mu_{[k-1]} \geq \delta$. Thus the events of correct selection and good selection will agree over the entire preference zone if we define good selection as $\text{GS} := \{\mu_K > \mu_{[k]} - \delta\}$. Under this definition, only systems with performances *strictly* within δ of the best are considered good. A guarantee on the probability of good selection then has the form

$$\mathbb{P}_\mu(\text{GS}) \geq 1 - \alpha \quad \text{for all } \mu. \tag{Goal PGS}$$

Goal PCS implies Goal PGS since all correct systems are good, and Goal PGS implies Goal PCS-PZ since in the preference zone there is only one good system.

2.2.1 Why Goal PGS is Superior to Goal PCS-PZ

As a stand-alone guarantee, Goal PCS-PZ suffers from several flaws, the foremost being that it says nothing about a procedure's behavior when the configuration of systems' performances is in the indifference-zone. Does it deliver

Goal PGS? Does it even terminate in finite time almost surely? This shortcoming of Goal PCS-PZ is critical in practice, where the difference between the performances of the best and second-best systems is unknown and likely cannot be bounded from below with certainty. Furthermore, for problems with large numbers of systems, one might expect that the best system will not be well-separated from the others, suggesting that for reasonable values of δ, the configuration of systems' performances will be in the indifference zone. Likewise, in the case when an R&S procedure is used to "clean-up" after a simulation-optimization search (Boesel et al., 2003b), systems with similar performances are likely to be returned by the search. In this setting, IZ-inspired PCS guarantees are conditional on the *random* configuration of the returned systems' performances being in the preference zone, an event that the decision-maker cannot control or verify (Eckman and Henderson, 2018).

A related issue with Goal PCS-PZ is the presumption that the configuration is in the preference zone. According to Parnes and Srinivasan (1986), Goal PCS-PZ would only be useful if either (i) the decision-maker has prior knowledge that the configuration is almost certainly in the preference zone or (ii) in the event that the configuration is in the indifference zone, the error $\mu_{[k]} - \mu_K$ is unacceptably large with small probability. In the first case, the implicit assumption that the configuration lies in the preference zone suggests that a *Bayesian* selection procedure that can exploit this information—via a prior distribution on μ—may be preferable. In the second case, the decision-maker's interest in the linear loss function $\mu_{[k]} - \mu_K$ suggests that the expected opportunity cost may be a more relevant performance metric.

Another concern with Goal PCS-PZ is that the IZ parameter is commonly

misinterpreted. Under Goal PGS, δ represents the smallest difference in performance that is worth detecting; it classifies systems that, if selected, would or would not be acceptable to the decision-maker. Under Goal PCS-PZ, however, the IZ parameter is only of significance in stating that *if* the best system is at least δ better than the others, the decision-maker would only be satisfied with selecting the best system. In this way, the IZ parameter restricts the set of problems on which a selection procedure can be relied on to perform well. In addition, this role of the IZ parameter in Goal PCS-PZ can have adverse consequences. For example, a decision-maker may choose a small value for δ in an attempt to be more confident that the configuration lies in PZ(δ). Yet by choosing δ to be smaller than their tolerance, the decision-maker will end up with a more conservative selection procedure (Fan et al., 2016).

Our objective in discussing these flaws is not to argue that Goal PCS-PZ is without use. Indeed, in Section 2.3 we show that under various conditions it is equivalent to Goal PGS. We instead assert that Goal PCS-PZ is better suited as a tool for proving Goal PGS than as a stand-alone goal. Even so, we will later discuss how some sequential procedures designed to deliver Goal PCS-PZ can be inefficient for configurations in the indifference zone. One explanation for the persistence of Goal PCS-PZ is perhaps the relative mathematical ease of designing procedures to deliver this goal; a lower bound on the difference between the performances of the best and second-best systems is useful in proving a PCS guarantee. As we will see, some proofs of Goal PGS deal with technical challenges that are not present in the proofs of Goal PCS-PZ, such as needing to account for pairwise comparisons between good and bad systems and not just those involving the best system.

Although Goal PGS has none of the aforementioned issues with Goal PCS-PZ, it is not perfect. In particular, Goal PGS says nothing about what happens when a good system is not picked. Just how bad are the bad selections of procedures achieving Goal PGS? Intuitively, one might expect that in this event, a selection procedure would select slightly bad systems and only rarely select an extremely bad one. This argument, however, is not altogether different from the belief that procedures achieving Goal PCS-PZ still make good selections with high probability for configurations in the indifference zone. On the other hand, guarantees on the expected opportunity cost offer slightly more control on the performance of the selected system.

2.2.2 Distributional Assumptions

Before discussing how Goal PCS-PZ can be lifted to Goal PGS, we introduce some notation and distributional assumptions related to the observations of the systems' performances. Let X_{ij} denote the jth observation from System i for $i = 1, \ldots, k$. We assume that the vectors of observations $X_j = (X_{1j}, X_{2j}, \ldots, X_{kj})$, for $j = 1, 2, \ldots$, are drawn independently from some joint distribution F having marginal distributions F_i. Unless otherwise stated, we allow the observations $X_{1j}, X_{2j}, \ldots, X_{kj}$ to be dependent across systems, as would be the case if common random numbers were used.

The R&S and multi-armed-bandit communities differ in the assumptions they make about the marginal distributions F_i. Within the R&S community, a common assumption is that the observations are normally distributed and the performance measures μ_i are the corresponding means. This normality as-

sumption can often be approximately satisfied using batched means as a single observation and appealing to the Central Limit Theorem. The R&S problem has also been studied from a large-deviations perspective that does not rely on this assumption (Glynn and Juneja, 2004). For later research in this direction, see Broadie et al. (2007), Blanchet et al. (2008), Hunter and Pasupathy (2010), and Glynn and Juneja (2018). For multi-armed bandit problems, the distributions of observations are either assumed to have bounded support or to be sub-Gaussian with a known bound on the variance (Even-Dar et al., 2002, 2006).

To prove sufficient conditions under which Goal PCS-PZ implies Goal PGS, we require an identifiability assumption on the joint distribution of the observations.

Assumption 1. *The joint distribution F is identifiable with respect to the configuration μ, i.e., for any joint distributions F_μ and $F_{\mu'}$, $F_\mu = F_{\mu'}$ implies $\mu = \mu'$.*

Unlike other regularity conditions such as normality or bounded support, Assumption 1 does not control the large-deviations behavior of the observations. Therefore, under Assumption 1 alone, the sample sizes needed to detect differences in performance of δ cannot be predetermined. Rather than enabling the design of procedures achieving Goal PCS-PZ or Goal PGS, our purpose for Assumption 1 is to ensure that when we later manipulate the configuration, the probability measure $\mathbb{P}_\mu$ is unambiguously defined for each μ. An example of a joint distribution satisfying Assumption 1 is a location parameter family with respect to μ, e.g., the multivariate normal distribution.

Given the above setup, a selection procedure takes observations X_{ij} from all systems and calculates estimators Y_i of μ_i. As an illustration, for the standard R&S setting with normally distributed observations, the performances μ_i are the

means of the marginal distributions F_i, and the estimators Y_i are naturally the sample means. Other examples of estimators include generalized means and sample quantiles (Dudewicz and Dalal, 1975; Shin et al., 2016). With regards to the estimators, we make an additional assumption.

Assumption 2. *The estimators $Y_1, \ldots, Y_k$ have a joint probability density function.*

Assumption 2 is made as a matter of convenience, so that the event of ties among estimators occurs with probability zero. We expect that a careful accounting of ties will allow the results of Sections 2.3 and 2.4 to extend to the case where the estimators are discrete random variables. In support of this, we provide a proof of Goal PGS for the procedure of Sobel and Huyett (1957) for Bernoulli observations; see Appendix A.5.

For the conditions that follow in Section 2.3, it will be necessary for us to make an assumption about the rule for selecting the returned system.

Assumption 3. *A selection procedure selects the system with the highest estimator as the best, i.e., $K \in \arg\max_i Y_i$.*

Most frequentist selection procedures in the literature satisfy Assumption 3, whereas other selection rules have been studied for the Bayesian R&S problem (Peng et al., 2016). Under Assumptions 2 and 3, we can state that $\mathbb{P}_\mu(\text{Select } i) = \mathbb{P}_\mu(Y_i > Y_j \text{ for all } j \neq i)$, even though the event $\{\text{Select } i\}$ may include sample paths on which $Y_i = Y_j$ for at least some $j \neq i$ and System i is ultimately selected based on certain tie-breaking rules.

2.3 Goal PCS-PZ Can Imply Goal PGS

2.3.1 Counterexamples

As previously mentioned, Goal PGS implies Goal PCS-PZ. One might wonder if the converse holds: do all selection procedures that achieve Goal PCS-PZ also achieve Goal PGS? A supporting intuition is that for configurations in the indifference zone, the presence of good systems should make it more likely that one of them is selected. We show that this is not universally true by presenting two contrived selection procedures (Procedures 1 and 2) that achieve Goal PCS-PZ but not Goal PGS. Both procedures fail to deliver Goal PGS because they are designed to behave very differently when there appear to be multiple systems with performances that are close to the best. In both counterexamples, it is assumed that observations from System i are independent and identically distributed (i.i.d.) from normal distributions with means μ_i and known common variance σ^2 and that observations across systems are independent.

Procedure 1

Specify a confidence level $1-\alpha \in (1/k, 1)$ and an IZ parameter $\delta > 0$.
Choose a scalar $r > -\Phi^{-1}((2k)^{-1})$ arbitrarily.
Take $n = \lceil 2(h_B + r)^2\sigma^2\delta^{-2} \rceil$ observations from each system, where h_B is the constant of Bechhofer (1954).
Calculate the sample means $Y_i = n^{-1}\sum_{j=1}^{n} X_{ij}$ as the estimators of the systems' performances and denote the ordered estimators by $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(k)}$.
If $Y_{(k)} > Y_{(k-1)} + r\sigma\sqrt{2/n}$, select the system corresponding to $Y_{(k)}$ as the best.
Otherwise, select a system uniformly at random from those that do not correspond to $Y_{(k)}$ or $Y_{(k-1)}$.

Proposition 1. *Procedure 1 achieves Goal PCS-PZ, but for some values of α and k it does not achieve Goal PGS.*

The proof of Proposition 1 can be found in Appendix A.1.

Procedure 1 behaves bizarrely: If one estimator is clearly better than the others, the procedure selects the best-looking system. But if the top two estimators are close to each other, the procedure selects from among the worst-looking systems. The presence of good systems can therefore make it *less* likely that one of the good systems is selected. Procedure 1 does not satisfy Assumption 3, so we present a second counter-example procedure that does.

Procedure 2

Specify a confidence level $1 - \alpha \in (1/k, 1)$ and an IZ parameter $\delta > 0$.
Take

$$n_0 = \left\lceil \frac{16\sigma^2}{\delta^2} \left(\Phi^{-1} \left(\frac{1 - (1-\alpha)^{1/(2k)}}{2} \right) \right)^2 \right\rceil$$

observations from each system.
Calculate the sample means $Y_i = n_0^{-1} \sum_{j=1}^{n_0} X_{ij}$ as the estimators of the systems' performances and denote the ordered estimators by $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(k)}$.
If $Y_{(k)} - Y_{(k-1)} \geq \delta/2$, take $n_1 = \lceil 2h_B^2 \sigma^2 \delta^{-2} \rceil$ new observations from each system, where h_B is the constant of Bechhofer (1954) with a confidence of $\sqrt{1-\alpha}$. Select the system with the highest sample mean based on the n_1 observations.
Otherwise, take $n_1 = 1$ new observation from each system. Select the system with the highest sample mean based on the single observation.

Proposition 2. *Procedure 2 achieves Goal PCS-PZ, but for some values of α and k it does not achieve Goal PGS.*

The proof of Proposition 2 can be found in Appendix A.2.

Procedure 2 also behaves atypically: It uses an initial stage to infer whether the problem instance is in the preference zone or in the indifference zone. If it appears to be in the preference zone, a modified version of the procedure of Bechhofer (1954) is run in a second stage, but if it appears to be in the indiffer-

ence zone, the second stage takes only a single observation from each system before making a selection.

Despite the existence of contrived counter-examples, one might yet expect that "reasonable" selection procedures that achieve Goal PCS-PZ also achieve Goal PGS. Indeed, many existing procedures achieving Goal PCS-PZ have been shown to empirically deliver a PGS above $1 - \alpha$ when run on configurations in the indifference zone. Finding a counter-example is made all the more difficult by the fact that many procedures are designed using conservative bounds (e.g., Bonferroni's inequality) that further boost the empirical PGS. Although we are not aware of any selection procedure in the literature that achieves Goal PCS-PZ but *provably* does not achieve Goal PGS, we are concerned by the absence of proofs of Goal PGS. In the remainder of Section 2.3, we examine ways to rigorously prove that some procedures achieving Goal PCS-PZ also achieve Goal PGS.

2.3.2 Lifting Goal PCS-PZ

An effective approach for extending Goal PCS-PZ to Goal PGS is to relate the probability of good selection under any IZ configuration with that under a related PZ configuration. That is, for an arbitrary configuration $\mu \in \text{IZ}(\delta)$, one finds a configuration $\mu^* \in \text{PZ}(\delta)$ for which it can be shown that

$$\mathbb{P}_{\mu}(\text{GS}) \geq \mathbb{P}_{\mu^*}(\text{GS}), \tag{2.1}$$

where the notation $\mathbb{P}_{\mu}$ and $\mathbb{P}_{\mu^*}$ reflects the dependence of the probability measures on the configuration. Because $\mu^* \in \text{PZ}(\delta)$, it follows from Goal PCS-PZ that $\mathbb{P}_{\mu^*}(\text{GS}) = \mathbb{P}_{\mu^*}(\text{CS}) \geq 1 - \alpha$ and so $\mathbb{P}_{\mu}(\text{GS}) \geq 1 - \alpha$. Hence if it can be shown

that Inequality (2.1) holds for any arbitrary configuration $\mu \in \text{IZ}(\delta)$ and its corresponding configuration $\mu^* \in \text{PZ}(\delta)$, then Goal PCS-PZ implies Goal PGS.

Some care is needed in constructing the related configuration μ^*, as Inequality (2.1) should not be expected to hold for an arbitrary choice of μ^*. Intuitively, the configuration μ^* should closely resemble μ so that the probabilities of good selection can be easily compared. When manipulating systems' performances, it will be implicitly assumed that the ordered indices $[1], \ldots, [k]$ are with respect to the fixed configuration, μ, unless otherwise stated.

A simple choice for constructing μ^* is to only increase the performance of (one of) the best systems until it is exactly δ better than the second best, i.e., set $\mu^*_{[k]} = \mu_{[k-1]} + \delta$ and $\mu^*_{[i]} = \mu_{[i]}$ for all $i = 1, \ldots, k-1$ where $[k]$ is the index associated with (one of) the best systems in μ. While changing the performance of only one system would seem to simplify the analysis, it actually makes it harder to compare the PGS under the two configurations. This is because the PGS for a configuration $\mu^* \in \text{PZ}(\delta)$ is a function of the *differences* in performances between the best system and the bad systems, all of which are changed by shifting the best system's performance.

A better construction for μ^* is instead to decrease the performances of the good systems of μ while holding the performance of (one of) the best systems fixed. To formalize this idea, let $\mathcal{G} := \{i : \mu_i > \mu_{[k]} - \delta\}$ and $\mathcal{B} := \{i : \mu_i \leq \mu_{[k]} - \delta\}$ denote the sets of indices of the good and bad systems, respectively, for a configuration μ. The related configuration μ^* is then described by $\mu^*_i = \mu_i$ for $i \in \mathcal{B} \cup \{[k]\}$ and $\mu^*_i = \mu_{[k]} - \delta$ for $i \in \mathcal{G} \backslash \{[k]\}$. That is, μ^* is identical to μ except that the good systems of μ (other than the best) are now bad systems with knife-edge performance $\mu_{[k]} - \delta$. From this construction, $\mu^* \in \text{PZ}(\delta)$.

In Section 2.3.3, we show that under various conditions on selection procedures, Inequality (2.1) is satisfied for any configuration $\mu \in \text{IZ}(\delta)$ and this choice of μ^*.

2.3.3 Sufficient Conditions

Dating back to Fabian (1962), various independent efforts have been made to identify sufficient conditions under which selection procedures that achieve Goal PCS-PZ simultaneously achieve Goal PGS. Many of these past results come from the statistics literature, but it appears that they have until now received little attention within the simulation community. Our purpose in reproducing these results is to create a coherent presentation of their main ideas, call attention to some of the implicit assumptions that they have relied on, and apply them to existing selection procedures. In addition, we weaken some of the proposed conditions (e.g., Theorems 1 and 3) and establish analogous conditions for subset-selection procedures in Section 2.4.

We present two of the most general conditions in Theorems 1 and 2, both of which deal with probability statements about the ordering of estimators. For procedures achieving Goal PCS-PZ, each of the two conditions implies Inequality (2.1) and therefore Goal PGS. The sources and interpretations of the two conditions are discussed in greater detail following the statements of the theorems.

Theorem 1. *Let $\mathscr{R}$ be a selection procedure achieving Goal PCS-PZ. Then $\mathscr{R}$ also achieves Goal PGS if*

(C1) For all subsets $A \subseteq \{1, \ldots, k\}$ and for all pairs of configurations μ and $\tilde{\mu}$ such that

$\mu_i = \tilde{\mu}_i$ *for all* $i \in A$,

$$\mathbb{P}_\mu(Y_i > Y_j \text{ for all } j \in A\backslash\{i\}) = \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j \text{ for all } j \in A\backslash\{i\}) \quad \text{for all } i \in A,$$

where Y_i *and* $\tilde{Y}_i$ *denote the estimators of the performance of System* i *under configurations* μ *and* $\tilde{\mu}$*, respectively.*

Condition (C1) states that the probability that a system has the highest estimated performance among those in an arbitrary subset of systems does not depend on the true performances of systems not belonging to the subset. It generalizes two conditions presented by Guiard (1996) that we restate in Corollary 1 as Conditions (C2) and (C3).

Proof. Fix an arbitrary configuration μ and define $\mathcal{G}$, $\mathcal{B}$, and μ^* accordingly. Let $Y_{[k]}$ denote the estimator associated with System $[k]$ where the index $[k]$ is with respect to the configuration μ. Then

$$\begin{aligned}
\mathbb{P}_\mu(\text{GS}) &\geq \mathbb{P}_\mu(Y_{[k]} > Y_i \text{ for all } i \in \mathcal{B}) \\
&= \mathbb{P}_{\mu^*}(Y^*_{[k]} > Y^*_i \text{ for all } i \in \mathcal{B}) \\
&\geq \mathbb{P}_{\mu^*}(Y^*_{[k]} > Y^*_i \text{ for all } i \neq [k]) \\
&= \mathbb{P}_{\mu^*}(\text{CS}) \\
&\geq 1 - \alpha.
\end{aligned}$$

The first inequality follows from the definition of good selection, while the first equality follows from Condition (C1), taking $A = \mathcal{B} \cup \{[k]\}$. The second inequality follows from including extra conditions and the last inequality follows from Goal PCS since $\mu^* \in \text{PZ}(\delta)$. □

Because Condition (C1) may be difficult to verify, we list four conditions in Corollary 1 that each imply Condition (C1), but may be easier to check. Proofs of the conditions of Corollary 1 and their relationships (see Figure 2.1) can be found in Appendices A.3 and A.4.

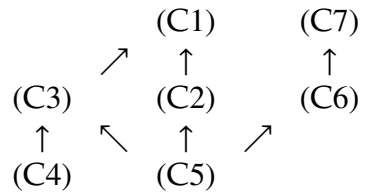

Figure 2.1: Relations of Conditions of Corollary 1 and Theorems 1, 2, and 3.

Corollary 1. *The following conditions each imply Condition (C1):*

(C2) Let B_1 and B_2 be disjoint subsets of $\{1, 2, \ldots, k\}$ and $\mathrm{IP} \subseteq B_1 \times B_2$ be a set of index pairs (i, j) with $i \in B_1$ and $j \in B_2$. For all (B_1, B_2, IP),

$$\mathbb{P}_\mu(Y_i > Y_j, \text{ for all } (i,j) \in \mathrm{IP}) \geq \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j, \text{ for all } (i,j) \in \mathrm{IP}),$$

for all pairs of configurations μ and $\tilde{\mu}$ satisfying $\mu_i - \mu_j \geq \tilde{\mu}_i - \tilde{\mu}_j$ for all $(i, j) \in \mathrm{IP}$.

(C3) For all subsets $A \subset \{1, \ldots, k\}$, the joint distribution of the estimators Y_i for $i \in A$ does not depend on μ_j for all $j \notin A$.

(C4) The estimators $Y_1, \ldots, Y_k$ are mutually independent.

(C5) The joint distribution of the estimators $Y_1, \ldots, Y_k$ is shift invariant, i.e., for any pair of configurations μ and $\tilde{\mu}$,

$$\begin{pmatrix} Y_1 - \mu_1 \\ Y_2 - \mu_2 \\ \vdots \\ Y_k - \mu_k \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} \tilde{Y}_1 - \tilde{\mu}_1 \\ \tilde{Y}_2 - \tilde{\mu}_2 \\ \vdots \\ \tilde{Y}_k - \tilde{\mu}_k \end{pmatrix}.$$

Condition (C2), which corresponds to class FND of Guiard (1996), has two main aspects. First, as the difference between the performances of systems increases, the probability that the systems with the higher true performances outperform the systems with the lower true performances does not decrease. Second, for the subset of systems whose performances remain unchanged relative to each other, the probability that any of the systems in this subset outperforms the others in the subset is unchanged. As shown in the proof of Corollary 1, it is because of this second property that Condition (C2) implies Condition (C1).

Condition (C3) corresponds to class F of Guiard (1996) and states that the joint distribution of the estimators of a subset of systems does not depend on the true performances of systems outside that subset. This statement stops short of asserting that the estimators are independent and thereby allows correlated observations across systems, as would be the case if common random numbers were used (Clark and Yang, 1986; Nelson and Matejcik, 1995). Furthermore, Condition (C3) can be applied to show that some selection procedures that use control-variate estimators achieve Goal PGS, e.g., Procedure 3 of Nelson and Staum (2006) and the WCS procedure of Tsai (2011).

Condition (C4), which corresponds to class FI of Guiard (1996), is satisfied by many early multi-stage selection procedures, e.g., the procedures of Dudewicz and Dalal (1975) and Rinott (1978). Condition (C4) can even be applied to some procedures for which the estimators are discrete random variables. An example of this is the procedure of Sobel and Huyett (1957) for selecting the Bernoulli population with the highest success probability; a proof of the procedure's PGS guarantee can be found in Appendix A.5.

Condition (C5) corresponds to class FS in Guiard (1996) and states that the

joint distribution of the estimators of the systems' performances is shift invariant with respect to the true performances. That is, when the true performances of the systems shift by a given amount, the joint distribution of their estimators shifts by the same amount. Condition (C5) does not follow automatically from an assumption that the joint distribution of the *observations* is shift invariant with respect to the configuration, as is the case when the observations are normally distributed and the estimators are the sample means.

Condition (C5) is a slight strengthening of the condition in Theorem 1 of Nelson and Matejcik (1995), which makes comparisons with the slippage configuration associated with System $[k]$ being the best, denoted by μ^{sc} where $\mu^{sc}_{[k]} = \mu_{[k]}$ and $\mu^{sc}_i = \mu^{sc}_{[k]} - \delta$ for all $i \neq [k]$:

$$\begin{pmatrix} Y_{[k]} \\ Y_{[k-1]} + (\mu_{[k]} - \mu_{[k-1]} - \delta) \\ \vdots \\ Y_{[1]} + (\mu_{[k]} - \mu_{[1]} - \delta) \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} Y^{sc}_{[k]} \\ Y^{sc}_{[k-1]} \\ \vdots \\ Y^{sc}_{[1]} \end{pmatrix}. \tag{2.2}$$

The left-hand side of Equation (2.2) features estimators under an arbitrary configuration μ and the right-hand side features estimators under the corresponding slippage configuration μ^{sc}. In contrast to Equation (2.2), Condition (C5) allows comparisons to be made between the joint distribution of the estimators under *any* two configurations, not just those for which the best systems have the same performance.

Prior to Guiard (1996), many of the conditions proposed for lifting Goal PCS-PZ to Goal PGS were unnecessarily restrictive. For example, Fabian (1962) proved that under a certain permutability assumption—one which implies Condition (C5)—procedures achieving Goal PCS-PZ also guarantee the stronger

probability statement

$$\mathbb{P}_\mu(\mu_K \geq \mu_{[k]} - D) \geq 1 - \alpha \quad \text{for all } \mu, \tag{2.3}$$

where $D = \max\{0, \delta - (Y_{(k)} - Y_{(k-1)})\}$ and $Y_{(i)}$ denotes the ith-lowest estimator. Equation (2.3) in turn implies Goal PGS, since $D \leq \delta$. Giani (1986) later extended the analysis to a more general selection goal.

Chiu (1974a) required that the probability that the best system outperforms all of the bad systems, i.e., $\mathbb{P}_\mu(Y_{[k]} > Y_i \text{ for all } i \in \mathcal{B})$, is increasing with respect to the differences $\mu_{[j+1]} - \mu_{[j]}$ for all $j = 1, \ldots, |\mathcal{B}|$, holding all other differences fixed. Yet the proof of Goal PGS in Chiu (1974a) also implicitly uses the assumption that $\mathbb{P}_\mu(Y_{[k]} > Y_i \text{ for all } i \in \mathcal{B})$ does not depend on the true means of the other systems, essentially Condition (C1). As a result, this monotonicity condition is unnecessary to prove that Goal PCS-PZ implies Goal PGS. The implicit assumption that Condition (C1) holds is also made in the proofs of Chiu (1974b) and Parnes and Srinivasan (1986).

Feigin and Weissman (1981) and Chen (1982) separately prove that Goal PCS-PZ implies Goal PGS for the case when the estimators $Y_1, \ldots, Y_k$ are mutually independent and are from a common family of stochastically increasing distributions that only differ in their location parameters. Specifically, Chen (1982) uses a monotonicity lemma developed independently by Alam and Rizvi (1966) and Mahamunulu (1967) (Lemmas 2.1 and 4.2 therein, respectively) to show that, for his procedure, PGS is minimized in the slippage configuration. The additional assumption of a stochastically increasing family of distribution functions for the estimators is unnecessary since having mutually independent estimators—Condition (C4)—is sufficient.

As we have shown, Conditions (C1)–(C5) can be used to prove Goal PGS for

many classical multi-stage selection procedures that determine necessary sample sizes for all systems and then select one as the best. On the other hand, Conditions (C1)–(C5) are unlikely to be satisfied for more elaborate selection procedures that sequentially eliminate (screen out) inferior systems, e.g., the procedures of Paulson (1964), Kim and Nelson (2001), and Frazier (2014). While these kinds of sequential selection procedures tend to be more efficient in terms of the number of observations taken, multi-stage procedures have the advantage of being easily parallelized since they require little communication across processors, thereby making them appealing for problems with thousands, or even millions, of systems (Ni et al., 2014, 2017).

Sequential selection procedures that iteratively eliminate systems from contention introduce two issues that make Conditions (C1)–(C5) harder to verify: the estimators of systems' performances are no longer well-defined and are highly dependent across systems. To resolve the first issue, we might set $Y_i = -\infty$ if System i is eliminated, to reflect the fact that System i will not be selected. Under this definition, however, multiple systems can have estimators of $-\infty$, thereby complicating the probability statements in Conditions (C1)–(C5). The second issue of dependent estimators cannot be remedied and immediately rules out Condition (C4), mutually independent estimators. In addition, shifting the performance of a given system can affect the number of samples taken from other systems and future elimination decisions, meaning Conditions (C3) and (C5) can also be ruled out. For similar reasons, Conditions (C1) and (C2) do not appear to have any better chances of holding for procedures of this kind.

Instead, Condition (C6), presented in Theorem 2, is more likely to be satisfied for sequential procedures that screen out systems. Theorem 2 is stated without

proof by Hayter (1994), so we provide one here.

Theorem 2. *Let $\mathscr{R}$ be a selection procedure achieving Goal PCS-PZ. Then $\mathscr{R}$ also achieves Goal PGS if*

(C6) For all systems $i = 1, \ldots, k$, $\mathbb{P}_\mu$(Select i) is nonincreasing in μ_j for every $j \neq i$.

Condition (C6) states that increasing the true performance of any system does not increase the probability that any other system is selected. It implies that $\mathbb{P}_\mu(\text{Select } i)$ is nondecreasing in μ_i—a monotonicity property that one might expect most selection procedures to satisfy. Unfortunately, directly verifying Condition (C6) or even formulating stronger conditions that imply it is difficult (Hayter, 1994). However, some multi-stage procedures, e.g., those of Bechhofer (1954), Dudewicz and Dalal (1975), and Rinott (1978), can be shown to satisfy Condition (C6) by using the fact that for these procedures, increasing the performance of a system does not affect the estimation of any other system's performance, i.e., Condition (C3).

Proof. Fix an arbitrary configuration μ and define $\mathcal{G}$ and $\mathcal{B}$ accordingly. If $|\mathcal{B}| = k - 1$, then there is only one good system in μ and so $\mu \in \text{PZ}(\delta)$, thus $\mathbb{P}_\mu(\text{GS}) = \mathbb{P}_\mu(\text{CS}) \geq 1 - \alpha$. Otherwise, define a configuration $\mu^{(1)}$ such that $\mu^{(1)}_{[|\mathcal{B}|+1]} = \mu_{[k]} - \delta$ and $\mu^{(1)}_i = \mu_i$ for all $i \neq [|\mathcal{B}| + 1]$, i.e., the worst good system is shifted down to $\mu_{[k]} - \delta$, thereby making it a bad system. By definition, selection procedures select a single system, thus

$$\mathbb{P}_\mu(\text{Select } [|\mathcal{B}|+1]) + \sum_{i \neq [|\mathcal{B}|+1]} \mathbb{P}_\mu(\text{Select } i) = 1 = \mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}|+1]) + \sum_{i \neq [|\mathcal{B}|+1]} \mathbb{P}_{\mu^{(1)}}(\text{Select } i). \tag{2.4}$$

By Condition (C6),

$$\mathbb{P}_\mu(\text{Select } i) \leq \mathbb{P}_{\mu^{(1)}}(\text{Select } i),$$

for all $i \neq [|\mathcal{B}|+1]$ because the performance of system $[|\mathcal{B}|+1]$ has been decreased. In particular, this holds for all bad systems. Together with Equation (2.4), we obtain

$$\mathbb{P}_{\mu}(\text{Select } [|\mathcal{B}|+1]) + \sum_{\substack{i:\mu_i \geq \mu_{[|\mathcal{B}|+1]} \\ i \neq [|\mathcal{B}|+1]}} \mathbb{P}_{\mu}(\text{Select } i) \geq \mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}|+1]) + \sum_{\substack{i:\mu_i \geq \mu_{[|\mathcal{B}|+1]} \\ i \neq [|\mathcal{B}|+1]}} \mathbb{P}_{\mu^{(1)}}(\text{Select } i).$$

The left-hand side is $\mathbb{P}_{\mu}(\text{GS})$ while the right-hand side is $\mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}| + 1]) + \mathbb{P}_{\mu^{(1)}}(\text{GS})$. Then since $\mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}| + 1]) \geq 0$, we have

$$\mathbb{P}_{\mu}(\text{GS}) \geq \mathbb{P}_{\mu^{(1)}}(\text{GS}).$$

This argument can be repeated to chain together inequalities of the form

$$\mathbb{P}_{\mu^{(\ell-1)}}(\text{GS}) \geq \mathbb{P}_{\mu^{(\ell)}}(\text{GS}),$$

for $\ell = 1, \ldots, |\mathcal{G}|-1$ where $\mu^{(0)} := \mu$ and we recursively define $\mu^{(\ell)}$ by $\mu^{(\ell)}_{[|\mathcal{B}|+\ell]} = \mu_{[k]} - \delta$ and $\mu^{(\ell)}_i = \mu^{(\ell-1)}_i$ for all $i \neq [|\mathcal{B}| + \ell]$; i.e., the worst ℓ good systems of μ have been made bad. From this definition, $\mu^{(|\mathcal{G}|-1)} = \mu^*$. Therefore the inequalities all together yield

$$\mathbb{P}_{\mu}(\text{GS}) \geq \mathbb{P}_{\mu^*}(\text{GS}) = \mathbb{P}_{\mu^*}(\text{CS}) \geq 1 - \alpha.$$

□

One might naturally expect Condition (C6) to hold for many selection procedures, including sequential ones that screen out systems. All else being equal, a given system's likelihood of being selected should suffer when one of its competitors is made stronger. Despite this appealing intuition, Condition (C6) does not universally hold for sequential procedures due to the complicated effect that changing the performances of systems can have on the selection decision. As a counterexample, consider the standard R&S setting in which the observations

are normally distributed and the performances are the means. Hayter (1994) describes the following two-stage selection procedure that simulates systems independently:

Procedure 3

Take n_0 i.i.d. samples for each system.
Eliminate all but the two systems with the highest sample means.
Take n_1 additional i.i.d. samples from each of the two surviving systems.
Select the surviving system with the highest overall sample mean.

To show that Procedure 3 can fail to satisfy Condition (C6), consider the case in which there are three systems with performances $\mu_1 < \mu_2 < \mu_3$, i.e., System 3 is the best. Following the argument given by Hayter (1994), we now demonstrate how, for fixed values of n_0 and n_1, increasing the performance of System 1 can actually *increase* the probability that System 2 is selected, violating Condition (C6).

For fixed $n_0 > 0$, consider the extreme cases of $n_1 = 0$ and $n_1 = \infty$. (While the procedure is not implementable for $n_1 = \infty$, it illustrates the case when the surviving systems are heavily sampled.) When $n_1 = 0$, there is no second stage and the procedure simply selects the system with the highest sample mean based on the first n_0 samples. Since systems are simulated independently, the probabilities of selecting Systems 2 and 3 will decrease as μ_1 increases.

On the other hand, when $n_1 = \infty$, the procedure always makes a correct selection from between whichever two systems survive screening. Therefore the probability that System 1 is selected is zero, while the probability that System 3 is selected is equal to the probability that System 3 survives screening. System 3 survives screening unless Systems 1 and 2 both advance to the second stage. Again, since systems are sampled independently, the event that Systems 1 and

2 both advance increases as μ_1 increases. This implies that the probability that System 3 is selected decreases as μ_1 increases. Thus the probability that System 2 is selected must increase as μ_1 increases. It is possible to then find a finite value of n_1 for which this relationship holds and thereby conclude that Condition (C6) does not hold for Procedure 3.

An example of a selection procedure that may violate Condition (C6) is that of Fairweather (1968), which closely resembles Procedure 3. The NSGS procedure of Nelson et al. (2001) might also fail to satisfy Condition (C6); selecting a high value of α_0 for the screening stage and a small value of α_1 for the selection stage would lead to large second-stage sample sizes, possibly mimicking the behavior of Procedure 3 for the case $n_1 = \infty$.

The proof of Condition (C6) indicates some ways that the condition can be weakened while still implying Goal PGS. First, instead of requiring that the probability of selecting each individual system is monotone with respect to the performances of other systems, it suffices that the probability of selecting from among a *subset* of systems is monotone. Second, this monotonicity condition only needs to hold for systems that are inferior to System i. Putting these ideas together, we present a more general condition with greater potential for holding for sequential selection procedures.

Theorem 3. *Let $\mathscr{R}$ be a selection procedure achieving Goal PCS-PZ. Then $\mathscr{R}$ also achieves Goal PGS if*

(C7) For all systems $i = 1, \ldots, k$, $\mathbb{P}_\mu$(Select some j for which $\mu_j < \mu_i$) is nonincreasing in μ_i.

Proof. The proof follows that of Theorem 2 with a few small changes.

Fix an arbitrary configuration μ and define $\mathcal{G}$, $\mathcal{B}$, and $\mu^{(1)}$ as in the proof of Theorem 2. Since selection procedures must select a single system,

$$\sum_{j:\mu_j<\mu_{[|\mathcal{B}|+1]}} \mathbb{P}_\mu(\text{Select } j) + \sum_{j:\mu_j\geq\mu_{[|\mathcal{B}|+1]}} \mathbb{P}_\mu(\text{Select } j) = 1$$

$$= \sum_{j:\mu_j<\mu_{[|\mathcal{B}|+1]}} \mathbb{P}_{\mu^{(1)}}(\text{Select } j) + \mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}|+1]) + \sum_{\substack{j:\mu_j\geq\mu_{[|\mathcal{B}|+1]} \\ j\neq|\mathcal{B}|+1]}} \mathbb{P}_{\mu^{(1)}}(\text{Select } j).$$

By Condition (C7),

$$\sum_{j:\mu_j<\mu_{[|\mathcal{B}|+1]}} \mathbb{P}_\mu(\text{Select } j) \leq \sum_{j:\mu_j<\mu_{[|\mathcal{B}|+1]}} \mathbb{P}_{\mu^{(1)}}(\text{Select } j),$$

because the performance of system $[|\mathcal{B}|+1]$ has been decreased. Thus

$$\sum_{j:\mu_j\geq\mu_{[|\mathcal{B}|+1]}} \mathbb{P}_\mu(\text{Select } j) \geq \mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}|+1]) + \sum_{\substack{j:\mu_j\geq\mu_{[|\mathcal{B}|+1]} \\ j\neq[|\mathcal{B}|+1]}} \mathbb{P}_{\mu^{(1)}}(\text{Select } j).$$

The left-hand side is $\mathbb{P}_\mu(\text{GS})$ while the right-hand side is $\mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}|+1]) + \mathbb{P}_{\mu^{(1)}}(\text{GS})$. Then since $\mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}|+1]) \geq 0$, we have

$$\mathbb{P}_\mu(\text{GS}) \geq \mathbb{P}_{\mu^{(1)}}(\text{GS}).$$

The rest of the proof follows from that of Theorem 2. □

Condition (C7) states that increasing the performance of a given system does not increase the probability that an inferior system is selected. This means that increasing the performance of a bad system so that it becomes a good system will not decrease the probability that a good system is selected. This condition resolves the issue with Procedure 3 where increasing the performance of a system increased the probability that a superior system was selected; this relationship is permitted under Condition (C7). The proof of Theorem 3 indicates that Condition (C7) can be further weakened to apply to only systems $i \in \mathcal{G}\backslash\{k\}$ instead of all $i = 1, \ldots, k$.

Ultimately, Condition (C7) suffers from the same difficulties as Condition (C6), namely, deriving monotonicity relationships for the probabilities of selecting (subsets of) systems. It nonetheless may provide a way forward for checking whether new or existing procedures deliver Goal PGS.

2.3.4 Sequential Procedures

As pointed out in Section 2.3.3, lifting Goal PCS-PZ to Goal PGS for sequential selection procedures remains a challenge due to the complicated dependence among systems' estimators. One exception is the sequential (non-elimination) procedure $\mathscr{P}_B^*$ of Bechhofer et al. (1968) for the case in which observations are normally distributed with known, common variance. The $\mathscr{P}_B^*$ procedure iteratively takes one sample from each system and terminates as soon as the *posterior* probability of correct selection exceeds $1-\alpha$, selecting the system with the highest sample mean. For the $\mathscr{P}_B^*$ procedure, the posterior PCS is calculated as if the true problem instance was a permutation of the slippage configuration, making it different from the posterior PCS that typically appears in Bayesian R&S procedures (Branke et al., 2007). The proof that the $\mathscr{P}_B^*$ procedure achieves Goal PCS-PZ makes use of the fact that the posterior PCS in the slippage configuration is a lower bound on the posterior PCS for the unknown, true configuration. Sievers (1972) and Chiu (1977) proved that the $\mathscr{P}_B^*$ procedure also achieves Goal PGS by establishing that the posterior PCS in the slippage configuration is also a lower bound on the posterior PGS for the true configuration.

Aside from the $\mathscr{P}_B^*$ procedure of Bechhofer et al. (1968) and the Envelope procedure of Ma and Henderson (2017)—both of which do not eliminate

systems—few sequential selection procedures have been proven to achieve Goal PGS. For example, the proofs of Goal PGS presented for the KN procedure of Kim and Nelson (2001), the SSM procedure of Pichitlamken et al. (2006), and the CCSB procedure of Hong et al. (2015) are incorrect. These proofs all fail to properly account for the possibility that a good system can eliminate the best system early on and then be eliminated by a bad system. Furthermore, the proof of Goal PGS for the TSSD procedure of Osogami (2009) holds only for some parameter settings. Some sequential procedures instead asymptotically achieve Goal PGS, e.g., the bootstrapping procedure of Lee and Nelson (2019) and one of the indifference-zone-free procedures of Fan et al. (2016). Other recent sequential selection procedures, e.g., those of Kim and Dieker (2011), Frazier (2014), and Zhong and Hong (2017), achieve Goal PCS-PZ but lack proofs of Goal PGS. Alternatively, Kao and Lai (1980), Jennison et al. (1982) and Zhong and Hong (2018) show that some sequential selection procedures achieving Goal PCS-PZ, e.g., the procedure of Paulson (1964), can be modified to achieve Goal PGS. The proposed modifications, however, involve widening the continuation regions, resulting in procedures that are potentially excessively conservative.

Putting aside the question of whether the aforementioned procedures achieve Goal PGS, a crucial consideration is whether they would do so efficiently for configurations in the indifference zone. While the sample sizes of some multi-stage procedures are unaffected by the configuration of the systems' performances (Bechhofer, 1954; Dudewicz and Dalal, 1975; Rinott, 1978), many sequential selection procedures are designed to eliminate systems—or terminate—when the estimators of pairs or groups of systems become well-separated. This suggests that these procedures might require more observations to distinguish between systems with similar performances. For example, the

BIZ procedure of Frazier (2014) eliminates systems from contention and terminates when the posterior PCS for a surviving system exceeds a threshold. Similarly, the procedure of Kim and Dieker (2011) eliminates systems when a Brownian motion process observed in discrete time exits an elliptical continuation region. For configurations with multiple best systems, these two procedures might be expected to take, on average, more observations before these conditions for eliminating systems are met. Other sequential selection procedures that use pairwise comparisons to eliminate systems, e.g., the KN procedure of Kim and Nelson (2001) and the procedure of Zhong and Hong (2017), might also be expected to take more observations for configurations in the indifference zone; however, these two procedures exert some control on the maximum number of observations that are taken before terminating.

These observations suggest that designing a sequential selection procedure to efficiently deliver Goal PCS-PZ does not ensure that it will perform efficiently with respect to delivering Goal PGS. Furthermore, the empirical performance of the EP procedure of Ma and Henderson (2019) demonstrates how a sequential procedure designed for Goal PGS can be more efficient in the indifference zone than leading procedures designed for Goal PCS-PZ. This concern about the efficiency of sequential selection procedures is another reason why we advocate that Goal PGS supersede Goal PCS-PZ as the primary design criterion for frequentist R&S procedures.

2.4 Subset-Selection Procedures

In contrast to selection procedures, which choose a single system as the best, subset-selection procedures return a random subset of systems $I \subseteq \{1, \ldots, k\}$. Subset-selection procedures can be used to efficiently screen out inferior systems when the number of systems is large. Because subset-selection procedures were developed as an alternative to the indifference-zone formulation, they are often designed to give guarantees under any configuration. For example, the subset-selection procedure of Gupta (1965) for known common variance provides a guarantee of the form

$$\mathbb{P}_\mu(\text{CSS}) \geq 1 - \alpha \quad \text{for all } \mu, \tag{Goal PCSS}$$

where the event of correct subset selection is defined as $\text{CSS} := \{[k] \in I\}$. This definition of correct subset selection slightly differs from that of correct selection for selection procedures. When there is a unique best system, correct subset selection is the event that it is in the returned subset. When there are multiple best systems, however, correct subset selection is the event that a particular best system is in the returned subset. Goal PCSS therefore states that for *each* best system, the probability that it will be in the returned subset is at least $1 - \alpha$. This is stronger than guaranteeing that *at least one* of the best systems will be in the returned subset with probability at least $1 - \alpha$.

We can similarly define an IZ-inspired PCSS guarantee:

$$\mathbb{P}_\mu(\text{CSS}) \geq 1 - \alpha \quad \text{for all } \mu \in \text{PZ}(\delta). \tag{Goal PCSS-PZ}$$

Some subset-selection procedures achieve Goal PCSS-PZ, e.g., the Screen-to-the-Best procedure of Nelson et al. (2001). Because Goal PCSS-PZ is only with respect to configurations in the preference zone, subset-selection procedures de-

signed for this guarantee should be less conservative than those designed for Goal PCSS in the sense that they either return smaller subsets of systems or take fewer observations.

We define good subset selection as GSS := $\{\exists i \in I \text{ s.t. } \mu_i > \mu_{[k]} - \delta\}$, or equivalently GSS := $\{\mathcal{G} \cap I \neq \emptyset\}$, the event that at least one good system is in the returned subset. A guarantee on the probability of good subset selection is thus

$$\mathbb{P}_\mu(\text{GSS}) \geq 1 - \alpha \quad \text{for all } \mu. \tag{Goal PGSS}$$

Since the best system is always a good system, Goal PCSS implies Goal PGSS. The restricted subset-selection procedure of Sullivan and Wilson (1989) is one example of a procedure that achieves Goal PGSS. Good subset selection has alternatively been defined as the event that *all* of the good systems are in the returned subset (Lam, 1986; Wu and Yu, 2008) or the event that the returned subset contains only good systems (Desu, 1970; Santner, 1976). Others have studied the related problem of identifying the top m systems given a fixed sampling budget (Chen et al., 2008; Jia et al., 2013; Gao and Chen, 2015; Kaufmann et al., 2016; Zhang et al., 2016).

Subset-selection procedures are many and varied. For instance, some subset-selection procedures take a fixed number of observations, specified by the user, while others take a random number of observations over multiple stages. Subset-selection procedures also differ in their rules for determining the set of returned systems based on the estimators of the systems' performances. Some subset-selection procedures return a fixed number of systems, namely those with the m highest estimators, i.e., $I = \{i : Y_i \geq Y_{(k-m+1)}\}$ where $Y_{(j)}$ denotes the jth lowest estimator and m is specified by the user in advance (Mahamunulu, 1967; Desu and Sobel, 1968). Other subset selection-procedures return systems whose

estimators are above a certain threshold that depends on the highest estimator, i.e., $I = \{i : Y_i \geq Y_{(k)} - d\}$ for some constant $d \geq 0$ (Gupta, 1965; Dudewicz and Dalal, 1975).

Santner (1975) proposed a formulation known as *restricted* subset selection that integrates the fixed-number and fixed-threshold selection rules by defining the subset of returned systems as

$$I = \{i : Y_i \geq \max(Y_{(k-m+1)}, Y_{(k)} - d)\}.$$

The cases $d = \infty$ and $m = k$ reduce to fixed-number and fixed-threshold selection rules, respectively. The procedure of Sullivan and Wilson (1989) is an example of a restricted subset-selection procedure.

Other subset-selection procedures employ a selection rule that compares systems pairwise and retains the subset

$$I = \{i : Y_i \geq Y_j - W_{ij} \text{ for all } j \neq i\},$$

where $W_{ij} = W_{ji} \geq 0$ is a function of the observations that yielded Y_i and Y_j. In the Screen-to-the-Best procedure of Nelson et al. (2001), W_{ij} is a function of the sample variances of Systems i and j. Under this selection rule, we say that System j *eliminates* System i, denoted by $j \to i$, if $Y_i < Y_j - W_{ij}$.

When it comes to proving Goal PGSS for subset-selection procedures that use pairwise comparisons, an important property is transitive eliminations; i.e., if System i eliminates System j and System j eliminates System ℓ, then System i also eliminates System ℓ. Transitive eliminations imply that for any System $j \in I^c$, there exists a System $i \in I$ such that $i \to j$, a helpful result in proving Goal PGSS (Nelson et al., 2001). A sufficient condition for transitive eliminations is that $W_{ij} + W_{j\ell} \geq W_{i\ell}$ for all $i \neq j \neq \ell$. This triangle inequality, however,

is not automatically satisfied for subset-selection procedures that use pairwise comparisons; for example, it does not hold for the Screen-to-the-Best procedure when $\delta > 0$.

As in Section 2.3, we present sufficient conditions under which subset-selection procedures that achieve Goal PCSS-PZ also achieve Goal PGSS. Theorems 4 and 5 show how Conditions (C3) and (C6) can be adapted to subset selection; their proofs can be found in Appendices A.6 and A.7, respectively.

Theorem 4. *Let $\mathscr{S}$ be a subset-selection procedure achieving Goal PCSS-PZ.*

If $\mathscr{S}$ uses restricted subset-selection with $I = \{i : Y_i \geq \max\{Y_{(k-m+1)}, Y_{(k)} - d\}\}$, then $\mathscr{S}$ also achieves Goal PGSS if

(C8) For all subsets $A \subset \{1, \ldots, k\}$, the joint distribution of the estimators Y_i for $i \in A$ does not depend on μ_j for all $j \notin A$.

If $\mathscr{S}$ uses pairwise comparisons with $I = \{i : Y_i \geq Y_j - W_{ij}$ for all $j \neq i\}$ where $W_{ij} + W_{j\ell} \geq W_{i\ell}$ for all $i \neq j \neq \ell$, then $\mathscr{S}$ also achieves Goal PGSS if

(C9) For all subsets $A \subset \{1, \ldots, k\}$, the joint distribution of the terms Y_i for $i \in A$ and $W_{i\ell}$ for $i, \ell \in A$ does not depend on μ_j for all $j \notin A$.

Conditions (C8) and (C9) are satisfied for many subset-selection procedures because systems are commonly simulated independently, yielding independent estimators, e.g., the procedures of Mahamunulu (1967), Desu and Sobel (1968), and van der Laan (1992).

Condition (C6) can also be modified to work for subset-selection procedures in a way that does not require assumptions about the selection rule, as was the

case in Theorem 4.

Theorem 5. *Let $\mathscr{S}$ be a subset-selection procedure achieving Goal PCSS-PZ. Then $\mathscr{S}$ also achieves Goal PGSS if*

(C10) For every subset $A \subset \{1, \ldots, k\}$, $\mathbb{P}_\mu(A \cap I \neq \emptyset)$ is nondecreasing in μ_i for all $i \in A$ and nonincreasing in μ_j for all $j \notin A$ when all other components of μ are held fixed.

Condition (C10) is a generalization of Condition (C6) of Theorem 2 since for singleton subsets A and the choice of $I = \{i : Y_i > Y_j \text{ for all } j \neq i\}$, Condition (C10) reduces to Condition (C6). Condition (C6) alone is not sufficient to lift Goal PCSS-PZ to Goal PGSS because the proof of Theorem 2 relies on the fact that $\sum_{i=1}^{k} \mathbb{P}_\mu(\text{Select } i) = 1$, a statement that no longer holds when (Select i) is replaced by ($i \in I$).

A property related to Conditions (C10) and (C6) is that of strong monotonicity, as defined by Santner (1975). A subset-selection procedure is said to be *strongly monotone* in System i if $\mathbb{P}_\mu(i \in I)$ is nondecreasing in μ_i and nonincreasing in μ_j for all $j \neq i$ when all other components of μ are held constant. The condition that a subset-selection procedure is strongly monotone in all systems $i = 1, \ldots, k$ is equivalent to Condition (C10) with the added restriction that A is a singleton subset, and for the particular choice of $I := \{i : Y_i > Y_j \text{ for all } j \neq i\}$, it is equivalent to Condition (C6). That is, Condition (C6) is weaker than the condition of strong monotonicity in all systems, which in turn is weaker than Condition (C10).

Subset-selection procedures have also been used for the purpose of screening out inferior systems before running a selection procedure (Nelson et al.,

2001; Boesel et al., 2003b). If the observations used for the subset-selection stage are discarded and new observations are taken for the selection stage, then the results of Sections 2.3 and 2.4 can be used in tandem to prove an overall PGS guarantee. If instead the observations from the subset-selection stage are reused, the statistical analysis becomes more complicated; see, for example, the results of Nelson et al. (2001) for Goal PCS-PZ. How the results of Sections 2.3 and 2.4 can be combined in this setting to prove an overall PGS guarantee is left as a direction for future research.

2.5 Other Proof Methods

Extending Goal PCS-PZ is not the only way to prove Goal PGS for selection procedures. In this section, we review two other methods that have been used: multiple comparisons and concentration inequalities. Although these two approaches make use of fundamental ideas about good selection, they tend to be inherently conservative. Consequently, using these ideas in designing a selection procedure to achieve Goal PGS can result in an inefficient procedure.

2.5.1 Multiple Comparisons

In Section 2.3.3, we remarked that Nelson and Matejcik (1995) provide a shift-invariant assumption resembling Condition (C5) that implies Goal PGS. In their proof, it is first shown that for a selection procedure achieving Goal PCS-PZ and satisfying Equation (2.2), the following joint probability statement holds:

$$\mathbb{P}_\mu(Y_{[k]} - Y_i - (\mu_{[k]} - \mu_i) > -\delta, \text{ for all } i \neq [k]) \geq 1 - \alpha \quad \text{for all } \mu. \tag{2.5}$$

From Equation (2.2) and Goal PCS-PZ, for an arbitrary configuration μ,

$$\begin{aligned}
&\mathbb{P}_{\mu}(Y_{[k]} - Y_i - (\mu_{[k]} - \mu_i) > -\delta, \text{for all } i \neq [k]) \\
&\qquad = \mathbb{P}_{\mu^{sc}}(Y^{sc}_{[k]} - (Y^{sc}_i - \mu_{[k]} + \mu_i + \delta) - (\mu_{[k]} - \mu_i) > -\delta, \text{for all } i \neq [k]) \\
&\qquad = \mathbb{P}_{\mu^{sc}}(Y^{sc}_{[k]} - Y^{sc}_i > 0, \text{for all } i \neq [k]) \\
&\qquad \geq 1 - \alpha.
\end{aligned}$$

Goal PGS then follows directly from Equation (2.5):

$$\begin{aligned}
&\mathbb{P}_{\mu}(Y_{[k]} - Y_i - (\mu_{[k]} - \mu_i) > -\delta, \text{ for all } i \neq [k]) \\
&\qquad = \mathbb{P}_{\mu}(\mu_i > \mu_{[k]} - \delta + (Y_i - Y_{[k]}), \text{ for all } i \neq [k]) \\
&\qquad = \mathbb{P}_{\mu}(\mu_i > \mu_{[k]} - \delta + (Y_i - Y_{[k]}), \text{ for all } i = 1, \ldots, k) \\
&\qquad \leq \mathbb{P}_{\mu}(\mu_K > \mu_{[k]} - \delta + (Y_K - Y_{[k]})) \\
&\qquad \leq \mathbb{P}_{\mu}(\mu_K > \mu_{[k]} - \delta) \\
&\qquad = \mathbb{P}_{\mu}(\text{GS}),
\end{aligned}$$

where the first inequality comes from considering only the statement for $i = K$, and the second inequality comes from Assumption 3.

The use of Equation (2.5) in proving Goal PGS indicates that the fixed-confidence guarantee for good selection is related to the problem of obtaining fixed-width confidence intervals for the differences in performances between pairs of systems. Specifically, Equation (2.5) is a joint probability statement about the differences between each system's estimator and that of the best. It is closely related to the idea of constructing simultaneous confidence intervals for the differences between each system's performance and the best of the other systems, i.e., $\mu_i - \max_{j \neq i} \mu_j$, for $i = 1, \ldots, k$ (Hsu, 1984). In the statistics community, this kind of inference is referred to as multiple comparisons with the best (MCB); see Hsu (1996) and Hochberg and Tamhane (2009) for helpful references.

Matejcik and Nelson (1995) and Nelson and Matejcik (1995) show that some IZ-inspired selection procedures, e.g., those of Dudewicz and Dalal (1975), Rinott (1978), and Clark and Yang (1986), deliver Goal PCS-PZ and simultaneously allow MCB inference with the same confidence, thereby achieving Goal PGS. Nelson and Goldsman (2001) and Ni et al. (2017) also use MCB to prove Goal PGS for some of their procedures. The selection procedures of Yang and Nelson (1991) and Nelson and Staum (2006) that use control variates also achieve Goal PGS as a consequence of MCB. Nelson and Banerjee (2001) use similar multiple comparisons statements to obtain a lower confidence bound on PGS after a selection procedure has been run.

Although MCB statements have been used to prove Goal PGS for many selection procedures, this approach has several limitations. First, if the objective is to design a procedure that achieves Goal PGS, working with MCB statements will result in conservative—and therefore less efficient—procedures. Second, ensuring Equation (2.5) holds with a prespecified confidence is hard to achieve for procedures that take observations sequentially. A recent development in this area is using bootstrapping to estimate the probabilities of multiple comparison events and stopping when the estimated probability exceeds $1-\alpha$ (Lee and Nelson, 2019). This bootstrapping approach sacrifices any kind of finite-sample-size guarantee, but can deliver different asymptotic versions of Goal PGS.

MCB statements can also be used for subset selection. For example, returning the subset of systems whose MCB upper confidence bounds exceed zero will ensure Goal PCSS, and hence Goal PGSS (Hsu, 1984; Kim and Nelson, 2006b).

2.5.2 Concentration Inequalities

In the multi-armed bandit community, it is usually assumed that the marginal distributions F_i have bounded support or are sub-Gaussian with known scale, i.e., a known upper bound on the variance. These regularity conditions control the large-deviations behavior of the estimators and therefore allow the use of concentration inequalities. A standard approach for proving Goal PGS under these assumptions is as follows: First, use concentration inequalities to bound the probabilities that the estimators Y_i differ from their true parameter values μ_i by a fixed amount. Next, obtain a bound on the probability that a given bad system outperforms the best system. Finally, use a union bound to obtain an upper bound on the probability of making a bad selection.

As an illustration, consider the standard multi-armed bandit setting where the observations of each system take values in the interval $[0, 1]$, the performance of System i is $\mu_i = \mathbb{E}[X_{ij}]$, and systems are simulated independently. For this problem, Even-Dar et al. (2006) propose a "Naive" algorithm achieving Goal PGS that takes $n = (2/\delta^2)\ln(2k/\alpha)$ observations from each system and selects the system with the highest sample mean, $Y_i = n^{-1}\sum_{j=1}^{n} X_{ij}$.

For a bad system, i, to be selected instead of the best system, $[k]$, then either $Y_i > \mu_i + \delta/2$ or $Y_{[k]} < \mu_{[k]} - \delta/2$, or both. Therefore

$$\begin{aligned}
\mathbb{P}_\mu(Y_i > Y_{[k]}) &\le \mathbb{P}_\mu(Y_i > \mu_i + \delta/2 \text{ or } Y_{[k]} < \mu_{[k]} - \delta/2) \\
&\le \mathbb{P}_\mu(Y_i > \mu_i + \delta/2) + \mathbb{P}_\mu(Y_{[k]} < \mu_{[k]} - \delta/2) \\
&\le 2\exp(-2n(\delta/2)^2),
\end{aligned}$$

where the last inequality is the result of applying Hoeffding's inequality twice. From the choice of n, $\mathbb{P}_\mu(Y_i > Y_{[k]}) \le \alpha/k$. Another union bound shows that the

"Naive" algorithm achieves Goal PGS:

$$1 - \mathbb{P}_\mu(\text{GS}) \leq \mathbb{P}_\mu\left(\cup_{i \neq [k]}\{Y_i > Y_{[k]}\}\right) \leq \sum_{i \neq [k]} \mathbb{P}_\mu(Y_i > Y_{[k]}) \leq (k-1)\frac{\alpha}{k} \leq \alpha.$$

Aside from Hoeffding's inequality, other concentration inequalities such as Chernoff's bound can be used in the same way if the marginal distributions are sub-Gaussian with known scale. While concentration inequalities are useful in the above proof of Goal PGS, this approach still requires the use of the conservative Bonferroni inequality to lift statements about pairwise comparisons to one about good selection. Some multi-armed bandit algorithms for the full-exploration problem eliminate systems in stages, such as the Successive Elimination and Median Elimination algorithms of Even-Dar et al. (2006). For these algorithms, concentration inequalities are used similarly to analyze pairwise comparisons and then combined with other conservative inequalities to bound the probability of making a bad selection.

In the R&S literature, the standard assumption that observations are normally distributed does not by itself allow the use of concentration inequalities, unless the variances are known or there are known upper bounds on the variances. Still, the approach of forming confidence bands around each system's estimator can be leveraged to yield a proof of Goal PGS, as is done for the Envelope procedure of Ma and Henderson (2017). The Envelope procedure designs upper and lower confidence limits for the performances of each system in such a way that with probability exceeding $1 - \alpha$, the upper confidence limit of the true best system stays above its performance and, simultaneously, the lower confidence limits of the other systems stay below their performances throughout the entire procedure. The procedure obtains observations from systems and updates the confidence limits over time. Once the lower confidence limit of the

estimated best system exceeds the highest upper confidence limit of the other systems minus δ, the procedure terminates. Selecting the system with the best estimated performance thereby guarantees that a good selection is made with probability exceeding $1 - \alpha$.

2.6 Conclusion and Future Work

In this chapter, we give a comprehensive overview of fixed-confidence, fixed-tolerance guarantees, with the objective of reorienting the simulation community towards designing R&S procedures with such guarantees. We point out several flaws of the more-popular IZ-inspired PCS guarantee and clarify sufficient conditions under which it is equivalent to the PGS guarantee. Some of the sufficient conditions for selection procedures are then adapted to work for subset-selection procedures. We also survey past results from the R&S and multi-armed-bandit literature to present a variety of approaches for proving the PGS guarantee.

A strength of the multi-armed-bandit literature is its analysis of the sampling complexity needed to deliver the PGS guarantee. Section 2.3.4 comes close to discussing this matter, but very little has been done in the R&S literature. Nearly matching complexity bounds have been developed by Ma and Henderson (2019), building on the results of Jennison et al. (1982) and Mannor and Tsitsiklis (2004), but presumably much more can be done.

It remains an open question whether some of the state-of-the-art selection procedures designed under the IZ formulation also deliver the PGS guarantee, possibly as a consequence of verifying Conditions (C6) or (C7). On the other

hand, rigorously proving or finding empirical evidence that some of these procedures do not deliver the PGS guarantee is also intriguing. Moreover, there is an opportunity for designing procedures that sequentially eliminate systems while delivering the PGS guarantee more efficiently than existing IZ-inspired procedures. Another approach that merits further investigation is using bootstrapping to obtain asymptotic guarantees.

CHAPTER 3

BAYESIAN RANKING-AND-SELECTION GUARANTEES

The majority of this chapter is from a manuscript under preparation (Eckman and Henderson, 2019b).

3.1 Introduction

The R&S problem has been extensively studied from two statistical perspectives—frequentist and Bayesian—that differ conceptually in what is regarded as random or fixed; see Kim and Nelson (2006b) and Chick (2006b) for overviews of these two areas. From the frequentist perspective, the underlying problem instance is *fixed* and the alternative selected by the procedure is *random*, the result of the observations drawn from the sampling distribution corresponding to the problem instance. From the Bayesian perspective, the problem instance is *random* and the alternative selected by the procedure is *fixed*, the result of a deterministic decision made based on the realized observations. Both treatments of the problem have received considerable attention in terms of designing efficient procedures that offer finite-time statistical guarantees (Kim and Nelson, 2001; Branke et al., 2007; Chen et al., 2015; Hong et al., 2015).

Three popular criteria for procedures designed under either framework are the probability of correct selection (PCS), the probability of good selection (PGS), and the expected opportunity cost (EOC). Under the frequentist framework, the probability measure associated with these criteria is defined with respect to repeated runs of the procedure on a fixed problem instance. Under the Bayesian framework, it is defined with respect to the posterior distribution on the prob-

lem instance given the collected observations and any prior information. Much of the analysis to date on Bayesian R&S procedures considers the setting in which the simulation budget—the maximum number of replications that can be taken—is fixed in advance (Chen et al., 2000; Chick and Inoue, 2001a,b; Peng et al., 2018). In this setting, procedures are designed to maximize the posterior PCS or PGS or minimize the posterior EOC after exhausting the budget. We instead consider the setting in which the decision-maker specifies a desired guarantee and then runs a procedure that takes observations until it can be delivered (Branke et al., 2007).

The frequentist treatment of the R&S problem dates back to the seminal work of Bechhofer (1954) and Gupta (1965) on selection and subset-selection procedures, respectively. Under the frequentist framework, R&S procedures are designed to deliver statistical guarantees that hold for a broad class of problem instances: either all problem instances (Ni et al., 2017; Ma and Henderson, 2019) or those in the preference zone, i.e., those for which the performances of the best and second-best alternatives differ by more than some specified amount (Kim and Nelson, 2001; Frazier, 2014; Zhong and Hong, 2017). Frequentist procedures are inherently conservative because of their need to guard against worst-case problem instances, such as the slippage configuration, which are unlikely to arise in practice (Frazier, 2010).

The Bayesian paradigm offers another way of handling uncertainty in simulation modeling and analysis (Andradottir and Bier, 2000; Chick, 2006a; Merrick, 2009). The Bayesian treatment of the R&S problem dates back to the work of Berger and Deely (1988) and Gupta and Kim (1980) on selection and subset-selection procedures, respectively. Bayesian procedures use the collected obser-

vations and any prior information to construct a posterior distribution on the unknown problem instance that is then used to evaluate the quality of a selection given the evidence. In this way, the statistical guarantees of Bayesian R&S procedures are for a fixed selection with respect to problem instances weighted according to the posterior distribution. Bayesian guarantees do not hold for *all* fixed problem instances; they instead hold in an average-case sense over problem instances drawn from the prior distribution, as explained in Section 3.2.2.

A straightforward way of ensuring that a procedure delivers a Bayesian guarantee is to use a posterior quantity of interest as a stopping rule. For example, to deliver a Bayesian EOC guarantee, it suffices for a procedure to terminate as soon as the posterior EOC of the best-looking alternative 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 guarantee. Aside from the theoretical convenience of proving their statistical guarantees, Bayesian R&S procedures have become popular in recent years because of their efficiency relative to frequentist procedures (Branke et al., 2007).

Although the frequentist and Bayesian treatments of the R&S problem are conceptually different, some research has studied their intersection. For example, the $\mathcal{P}^*_B$ procedure of Bechhofer et al. (1968) and the BIZ procedure of Frazier (2014) deliver a frequentist PCS guarantee via a Bayesian analysis of the posterior PCS where the posterior distribution of the problem instance is over the space of slippage configurations. In addition, Kadane et al. (1996) and Inoue and Chick (1998) compare the posterior PCS to frequentist p-values and Deng et al. (2016) compare Bayesian and frequentist A/B testing.

Using posterior-based stopping rules to deliver Bayesian R&S guarantees gives rise to several concerns that we address in this chapter. First, we explore in greater detail the conceptual and empirical differences between frequentist and Bayesian guarantees. We discuss how this distinction might factor into a decision-maker's choice of guarantee, depending on the situation in which R&S is used. Second, we investigate the decision of how often and how accurately to check stopping rules and its impact on the overall efficiency of a procedure. We present several methods for exactly evaluating the posterior quantity of interest and efficiently checking stopping rules, especially when the number of alternatives is large. Numerical experiments give a sense of the reduction in the total number of simulation replications relative to methods in the literature that use cheaply computable bounds.

The remainder of this chapter is outlined as follows. Section 3.2 introduces the mathematical notation and distributional assumptions, while defining PCS, PGS, and EOC under the frequentist and Bayesian frameworks. Section 3.3 justifies the use of posterior PCS, PGS, or EOC as stopping rules and studies the empirical performance of Bayesian procedures with such guarantees. Sections 3.4 and 3.5 highlight the computational challenges of checking the stopping conditions and presents several ways to improve the overall efficiency of Bayesian procedures. The proposed improvements are evaluated via simulation experiments in Section 3.6. Section 3.7 summarizes our findings and lays out directions for future research.

3.2 Frequentist and Bayesian R&S Guarantees

In this section, we formally define frequentist and Bayesian R&S guarantees for the purposes of comparing and contrasting them.

3.2.1 Frequentist R&S Guarantees

Suppose there are k alternatives under consideration and that the performance of Alternative i is denoted by w_i, for $i = 1, \ldots, k$. We refer to the vector of performances $\boldsymbol{w} = (w_1, \ldots, w_k)$ as the problem instance or configuration. From the frequentist perspective, the problem instance is fixed, but unknown to the decision-maker and describes a unique probability distribution from which observations are drawn. We use the notation $w_{[i]}$ to refer to the ith smallest performance where ties in indexing are broken arbitrarily; i.e., the ordered performances satisfy the relationship $w_{[1]} \leq \cdots \leq w_{[k]}$. Without loss of generality, we assume that larger performance values are better, hence Alternative $[k]$ is (one of) the best.

Let D denote the index of the alternative chosen by a R&S procedure. From the frequentist perspective, D is a random variable since it depends on the random observations sampled from the fixed problem instance. Three popular criteria by which frequentist R&S procedures are measured are the following:

- Probability of Correct Selection:

$$\text{PCS} := \mathbb{P}_{\boldsymbol{w}}(w_D = w_{[k]}).$$

- Probability of Good Selection:

$$\text{PGS} := \mathbb{P}_{\boldsymbol{w}}(w_D \geq w_{[k]} - \delta) \text{ for a fixed } \delta > 0.$$

- Expected Opportunity Cost:

$$\text{EOC} := \mathbb{E}_{\boldsymbol{w}}[w_{[k]} - w_D].$$

The subscript $\boldsymbol{w}$ in the three criteria indicates that the probability or expectation is taken with respect to the probability distribution associated with $\boldsymbol{w}$ from which random observations are drawn. In other words, the probabilities and expectations are defined in the frequentist sense of repeated runs of the selection procedure on a fixed problem instance. These criteria are mainly used to design selection procedures that deliver frequentist guarantees or to compare the empirical performance of R&S procedures.

PCS is the probability that the procedure selects the best or one of the best alternatives, and PGS is the probability that a δ-optimal alternative is selected. This definition of PGS differs slightly from that introduced in Chapter 2 when we examined the indifference-zone formulation. We redefine PGS in this chapter so that we can treat PCS as a special case where $\delta = 0$. Expected opportunity cost—also referred to as expected linear loss—is the expected difference between the performance of the best alternative and the selected alternative; i.e., the expected optimality gap of the procedure.

Many frequentist R&S procedures deliver statistical guarantees featuring these criteria:

- Indifference-zone (IZ) PCS Guarantee:

$$\mathbb{P}_{\boldsymbol{w}}(w_D = w_{[k]}) \geq 1 - \alpha \text{ for all } \boldsymbol{w} \text{ such that } w_{[k]} - w_{[k-1]} \geq \delta_{\text{IZ}}.$$

If $w_{[k]} - w_{[k-1]} < \delta_{\text{IZ}}$, no guarantee is provided.

- PGS Guarantee:

$$\mathbb{P}_{\boldsymbol{w}}(w_D \geq w_{[k]} - \delta) \geq 1 - \alpha \text{ for all } \boldsymbol{w}.$$

- EOC Guarantee:

$$\mathbb{E}_{\boldsymbol{w}}[w_{[k]} - w_D] \leq \beta \text{ for all } \boldsymbol{w}.$$

For these guarantees, the decision-maker specifies the values of $1 - \alpha$, δ_{IZ}, δ and β in advance. In the IZ PCS and PGS guarantees, the threshold $1 - \alpha$ reflects the decision-maker's desired degree of confidence in making a correct or good selection. The role of the IZ parameter, δ_{IZ}, is to specify the problem instances on which a procedure can be relied upon to select the best alternative with high probability, whereas the good-selection parameter, δ, represents the decision-maker's tolerance towards making a suboptimal decision. 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).

To deliver these frequentist guarantees, R&S procedures must guard against the hardest problem instances, making them inherently conservative. That is, for easier problem instances, the PCS or PGS may greatly exceed $1 - \alpha$, or the EOC may fall well below the threshold of β, indicating that the procedure has taken more observations than necessary.

3.2.2 Bayesian R&S Guarantees

In the Bayesian treatment of the R&S problem, the performances of the alternatives are viewed as random variables rather than unknown constants. To reflect this difference, we denote the performance of Alternative i by W_i for $i = 1, \ldots, k$. The random problem instance is denoted by the vector $\boldsymbol{W} = (W_1, \ldots, W_k)$ and the performances are ordered as $W_{[1]} \leq \cdots \leq W_{[k]}$.

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 observations 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. The posterior distribution can be used to define Bayesian analogs to the aforementioned frequentist criteria:

- Posterior Probability of Correct Selection of Alternative i:

$$\text{pPCS}_i := \mathbb{P}(W_i = W_{[k]} \mid \mathcal{E}).$$

- Posterior Probability of Good Selection of Alternative i:

$$\text{pPGS}_i := \mathbb{P}(W_i \geq W_{[k]} - \delta \mid \mathcal{E}).$$

- Posterior Expected Opportunity Cost of Alternative i:

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

The probabilities and expectations that appear in the Bayesian criteria are with respect to the posterior distribution given the evidence (simulated data)—denoted by $\mathcal{E}$—and the prior distribution. Furthermore these posterior quantities are defined in terms of selecting a given alternative, as designated by the subscript index. To better distinguish the frequentist and Bayesian criteria, we append a lowercase "p" to the abbreviations for *posterior* quantities. While this choice of notation is unconventional, we feel that it avoids unnecessary confusion. In contrast to their frequentist counterparts, the Bayesian criteria are properties of the evidence instead of properties of the procedure and thus can be calculated within a procedure; we discuss ways to calculate, estimate, and bound them in Sections 3.4 and 3.5.

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. Lastly, 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, D, is determined by the (fixed) observed data and the prior distribution. We will assume that Bayesian R&S procedures do not use randomized selection rules, i.e., 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 admit R&S guarantees with respect to the selected alternative:

- pPCS Guarantee:

$$\text{pPCS}_D \geq 1 - \alpha.$$

- pPGS Guarantee:

$$\text{pPGS}_D \geq 1 - \alpha.$$

- pEOC Guarantee:

$$\text{pEOC}_D \leq \beta.$$

Instead of referring to a fixed problem instance, Bayesian R&S guarantees involve an expectation over the unknown problem instance. More specifically, these Bayesian guarantees state that given the evidence, the posterior probability that the selected alternative is correct or δ-optimal is at least $1 - \alpha$, or that the posterior expected optimality gap of the selected alternative is below β. To perhaps belabor a point that will be important later, a Bayesian guarantee will *not* deliver a frequentist guarantee that holds for all problem instances.

Even though Bayesian guarantees are not stated with respect to a fixed problem instance, they can still be interpreted in a frequentist sense. For *repeated runs* of the following two-step process, the selected alternative will satisfy a comparable guarantee.

1. Generate a random problem instance from the prior distribution.
2. On the problem instance, run a given R&S procedure with a Bayesian guarantee.

For example, if a Bayesian R&S procedure with the pPGS guarantee is run in Step 2, then for repeated runs of the two-step process, the selected alternative

will be δ-optimal with probability exceeding $1-\alpha$. This frequentist interpretation follows from the law of total expectation:

$$\mathbb{P}(W_D \geq W_{[k]} - \delta) = \mathbb{E}[\mathbb{P}(W_D \geq W_{[k]} - \delta \mid \mathcal{E}_\tau)] \geq \mathbb{E}[1-\alpha] = 1-\alpha,$$

where $\mathcal{E}_\tau$ is the evidence collected by the time at which the procedure terminates, τ, and $\mathbb{E}$ denotes the expectation under the probability measure in which the problem instance is drawn from the prior distribution and observations are sampled from the probability distribution associated with that problem instance. Similar conclusions hold if Bayesian R&S procedures delivering the pPCS and pEOC guarantees are used in Step 2.

A similar derivation appears in the proofs of the $\mathcal{P}^*_B$ and BIZ procedures of Bechhofer et al. (1968) and Frazier (2014). Both procedures deliver frequentist IZ PCS guarantees, but do so by tracking the pPCS given that the true problem instance is in the slippage configuration, i.e., there is a unique best alternative whose performance is exactly δ_{IZ} better than all of the others. Under a uniform prior distribution on the set of k slippage configurations, pPCS_i is exactly the posterior probability of the slippage configuration for which Alternative i is the best. Further details for how this special choice of prior distribution leads to the IZ PCS guarantee can be found in Sections 3.1 and 3.2 of Bechhofer et al. (1968) and the proofs of Lemma 8 and Theorem 1 of Frazier (2014).

This analysis shows that if the prior distribution correctly describes the process that generates the problem instance, a Bayesian R&S procedure will deliver an average-case guarantee. A pertinent question that we do not address in this chapter is: How should we value Bayesian R&S guarantees when the prior is poorly chosen? Section 3.7 of Berger (1993) offers some insight on this important issue.

3.2.3 Frequentist and Bayesian R&S Guarantees in Different Situations

In this section, we examine how the conceptual differences between frequentist and Bayesian R&S guarantees might lead a decision-maker to favor one type of guarantee over the other, depending on the situation. We consider three situations in which R&S procedures may be used.

R&S for a one-time, expensive decision. Consider a decision with expensive implications, e.g., opening a new fulfillment center or hospital ward, that will presumably be made only once. The significant financial cost associated with the decision justifies an investment of time and resources to design and run a simulation model to solve the optimization problem. After the simulation model has been developed, verified, and calibrated, and the feasible alternatives have been identified, the problem instance for the resulting R&S problem can be viewed as fixed, but unknown.

Given the important ramifications of this one-time decision, it would be natural for a decision-maker to be especially averse to making a costly mistake. In this case, a decision-maker might be reassured by a frequentist guarantee that holds for *all* problem instances, even if the hardest problem instances are very unlikely. The decision-maker may also be more willing to incur the longer run times of frequentist R&S procedures relative to Bayesian R&S procedures.

R&S for repeated decisions. Suppose that the decision-maker wishes to solve a sequence of related problems using R&S. In this setting, a common simulation model is used for the problems, but its input parameters reflect the most up-to-date conditions of the physical system being modeled. If the conditions of

the physical system evolve according to some exogenous process, the sequence of problem instances can be viewed as a stochastic process. This observation suggests that Bayesian guarantees may be better-suited for this application of R&S. However, it is important to consider how the decision-maker specifies the prior distribution.

If the prior distribution incorporates the decision-maker's knowledge of the exogenous process and the relationship between the input parameters and the problem instance, the frequentist interpretation of Bayesian R&S guarantees may align well. For instance, for a long sequence of decisions made using a Bayesian R&S procedure with the pPGS guarantee and a common prior distribution, one would expect that about $(1-\alpha)\times 100\%$ of them would be δ-optimal. Using a frequentist R&S procedure with the PGS guarantee would instead ensure that each decision is δ-optimal with probability exceeding $1-\alpha$, independent across decisions. A common prior distribution, however, would not represent the uncertainty about a *specific* problem instance, given the current conditions of the system. Alternatively, the prior distribution could be updated over time based on observations taken while solving past R&S problems. In this way, Bayesian R&S procedures offer a convenient means of learning the relationship between the input parameters and the problem instance.

R&S after search. Suppose that a simulation-optimization search is run to identify a set of promising alternatives and then a R&S procedure is used to "clean up" afterwards (Boesel et al., 2003b). Since the resulting problem instance is random, the frequentist interpretation of Bayesian guarantees is appealing, provided that the prior distribution accounts for the process that generates the problem instance—in this case, the simulation-optimization search. Coming

up with an *informative* prior about the random problem instance returned by a search is nontrivial; the prior would have to combine information about the objective function and how the search operates. Frequentist R&S guarantees are another option, but it is very unlikely that a search will return the kind of worst-case problem instances that they are designed to guard against. Instead, one would expect a search to return a problem instance with several good alternatives.

Another related issue is the reuse of observations collected during the search, either in the first stage of a frequentist R&S procedure or in the prior distribution of a Bayesian R&S procedure. In either case, reusing search data is problematic, since search observations are conditionally dependent given the set of returned alternatives (Eckman and Henderson, 2018).

3.2.4 Issues with Posterior PCS

Returning to the Bayesian R&S guarantees introduced in Section 3.2.2, we bring attention to two serious issues with the pPCS guarantee that the pPGS guarantee avoids. Eckman and Henderson (2019a) have previously argued, for different reasons, that the PGS guarantee is superior to the IZ PCS guarantee.

The first concern is that for the pPCS guarantee, 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 more observations (on average) 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 there there is inherently some degree of error associated with the simulation model. On the other hand, 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.

The second concern with the pPCS guarantee stems from the case in which multiple alternatives are tied for the best. While it would be convenient to assume that practical problems do not have multiple alternatives with tied performances, some do. For example, the prototypical buffer-allocation problem of Pichitlamken et al. (2006) has multiple optimal solutions due to symmetries in the tandem-queuing system; see Ni et al. (2017) for a detailed description. In addition, min-max problems that arise in distributionally robust optimization can have multiple solutions with the same performance due to the inner maximum. For problem instances in which there are multiple optimal solutions, simulating until the pPCS of one of the alternatives exceeds $1 - \alpha$ will take a long time. A procedure that delivers the pPCS guarantee will therefore be chasing noise and will only terminate when a large-deviations event occurs.

The Bayesian resolution to this issue is that for continuous posterior distributions (e.g., multivariate normal), the probability that the performances of two or more alternatives are tied is zero. And if it were known in advance that certain alternatives had the same performances, perhaps due to symmetry, that information could be incorporated into the prior distribution. The cost of this

approach, however, is the loss of conjugacy—updating the posterior distribution would become computationally intensive. Moreover, determining that two alternatives have the same performance or that a problem has a unique optimal solution is nontrivial, especially when there are many alternatives.

Despite these concerns with the pPCS guarantee, many Bayesian R&S procedures designed for a fixed-budget setting use pPCS as the overall performance criteria (Branke et al., 2007; Peng et al., 2016; Russo, 2016). Some Bayesian R&S procedures also use the pPCS criteria to allocate simulation replications across alternatives (Chen et al., 2000). An interesting, open question is how the sampling efficiency of Bayesian R&S procedures is affected by the use of pPCS in their allocation rules, especially on problem instances with tied alternatives.

3.2.5 Assumptions

Thus far we have defined frequentist and Bayesian criteria without imposing any distributional assumptions on the observations of the alternative's 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 pPCS, pPGS, and pEOC. If these assumptions do not hold, we still expect that many of our main conclusions will, e.g., the potential benefits of numerically integrating posterior quantities of interest, even though the proposed methods may not.

Let X_{ij} denote the jth observation from Alternative i.

Assumption 4. *For any $i = 1, \ldots, k$ and any $j \geq 1$, X_{ij} is normally distributed with mean w_i and variance σ_i^2.*

Assumption 5. *For any $i = 1, \ldots, k$, the observations $X_{i1}, X_{i2}, \ldots$ are independent.*

Assumption 6. *For any $j \geq 1$, the observations $X_{1j}, \ldots, X_{kj}$ are independent.*

Assumption 7. *The decision-maker has independent prior beliefs about the performances of alternatives.*

Assumption 4 is commonly made in the simulation community and can be partially justified by batching observations as the batched means will be approximately normally distributed. The R&S problem has also been studied for Bernoulli-distributed observations, e.g., Even-Dar et al. (2006) and Russo (2016), and from a large-deviations perspective (Glynn and Juneja, 2004; Hunter and Pasupathy, 2010; Glynn and Juneja, 2018).

Assumption 5 is easily satisfied when observations are generated via stochastic simulation. In addition, this independence assumption implies that the observations are exchangeable, a necessary assumption for the upcoming derivation of the posterior distribution.

Assumption 6 rules out the use of common random numbers (CRN) in generating observations. While CRN are helpful in comparing the performances of alternatives, their use in a R&S procedure complicates the statistical analysis under both the frequentist (Nelson and Matejcik, 1995) and Bayesian (Chick and Inoue, 2001a; Fu et al., 2004; Gorder and Kolonko, 2019) paradigms. In the Bayesian treatment, CRN pose two notable challenges concerning the number of observations taken from each alternative. First, updating the posterior distribution when sample sizes are unequal requires careful attention and accounting; see Gorder and Kolonko (2019). Second, under the reference prior, at least k observations must be taken from each alternative to ensure that the sample

covariance matrix is invertible (Chick and Inoue, 2001a).

Assumption 7 is another standard assumption in the literature, although the setting of correlated beliefs has also been well studied (Frazier et al., 2009; Xie et al., 2016). Assumptions 6 and 7 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 $i = 1, \ldots, k$. However, enforcing independent beliefs in the prior distribution requires discarding any available structural information about the optimization problem, e.g., convexity or symmetry. In doing so, the decision-maker sacrifices prior knowledge about the relationships among alternatives for computational convenience.

Given Assumptions 4–7, we now mathematically describe the conjugate prior and posterior distributions. We will find it easier to work with the precisions $\lambda_i = 1/\sigma_i^2$ for $i = 1, \ldots, k$. The (joint) conjugate prior distribution for the mean and precision of Alternative i is normal-gamma: $\Lambda_i \sim \mathcal{G}(\alpha_i^0, \beta_i^0)$ and $W_i \mid \lambda_i \sim \mathcal{N}(\mu_i^0, 1/(\nu_i^0 \lambda_i))$ where the gamma distribution $\mathcal{G}(\alpha, \beta)$ has mean α/β and variance α/β^2. (This α parameter should not be confused with the one that appears in the pPCS and pPGS guarantees.) The prior hyperparameters $\alpha_i^0, \beta_i^0, \mu_i^0$, and ν_i^0 are specified by the decision-maker.

After observing samples $x_{i1}, \ldots, x_{in_i}$ from Alternative i, the posterior distribution of W_i and Λ_i is normal-gamma given by:

$$\Lambda_i \sim \mathcal{G}\left(\alpha_i^0 + \frac{n_i}{2}, \beta_i^0 + \frac{1}{2}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2 + \frac{n_i \nu_i^0}{\nu_i^0 + n_i}\frac{(\bar{x} - \mu_i^0)^2}{2}\right) \quad \text{and}$$

$$W_i|\lambda_i \sim \mathcal{N}\left(\frac{\nu_i^0 \mu_i^0 + n_i \bar{x}_i}{\nu_i^0 + n_i}, \frac{1}{(\nu_i^0 + n_i)\lambda_i}\right),$$

where $\bar{x}_i = n_i^{-1} \sum_{i=1}^{n_i} x_{ij}$ (Degroot, 2004). The *marginal* posterior distribution of the

performance of Alternative i is given by

$$W_i \sim t_{2\alpha_i^0+n_i}\left(\frac{\nu_i^0\mu_i^0+n_i\bar{x}_i}{\nu_i^0+n_i}, \frac{\beta_i^0+\frac{1}{2}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2+\frac{n_i\nu_i^0}{\nu_i^0+n_i}\frac{(\bar{x}_i-\mu_i^0)^2}{2}}{\left(\alpha_i^0+\frac{n_i}{2}\right)\left(\nu_i^0+n_i\right)}\right) =: t_{\nu_i}(\mu_i, \hat{\sigma}_i^2),$$

where $t_\nu(\mu, \sigma^2)$ denotes the distribution of a three-parameter Student-t random variable $Z = \mu + \sigma T_\nu$ where T_ν is a Student-t random variable with ν degrees of freedom (Chick and Inoue, 2001b). We will refer to μ_i and $\hat{\sigma}_i^2$ 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, i.e., $\mu_{(1)} \leq \mu_{(2)} \leq \cdots \leq \mu_{(k)}$ and Alternative (k) is the best-looking alternative.

The reference prior is specified by the prior hyperparameters $\mu_i^0 = \cdot$, $\nu_i^0 = 0$, $\alpha_i^0 = -1/2$ and $\beta_0^i = 0$ for all $i = 1, \ldots, k$. Under the reference prior, the marginal posterior distribution for the performance of Alternative i becomes $W_i \sim t_{n_i-1}(\bar{x}_i, s_i^2/n_i)$ where s_i^2 is the sample variance of the n_i observations taken from Alternative i. When the variances σ_i^2 are known, the conjugate prior distribution for the performance of Alternative i is $W_i \sim \mathcal{N}(\mu_i^0, 1/\lambda_i^0)$ where λ_i^0 is a prior hyperparameter, and the posterior distribution is

$$W_i \sim \mathcal{N}\left(\frac{\sigma_i^2/\lambda_i^0}{\sigma_i^2+n/\lambda_i^0}\left(\mu_i^0\lambda_i^0+\frac{n_i\bar{x}_i}{\sigma_i^2}\right), \frac{\sigma_i^2/\lambda_i^0}{\sigma_i^2+n_i/\lambda_i^0}\right).$$

Under the reference prior ($\mu_i^0 = \cdot, \lambda_i^0 = 0$), the posterior distribution is $W_i \sim \mathcal{N}(\bar{x}_i, \sigma_i^2/n_i)$.

As can be seen from the above formulae, using conjugate prior distributions greatly simplifies the task of updating the posterior distribution. When conjugate prior distributions are not used, the posterior distribution can still be computed numerically; pPCS, pPGS, and pEOC can then be calculated via Markov Chain Monte Carlo.

3.3 Posterior-Based Stopping Rules

In this section, we discuss how R&S procedures that take observations sequentially can deliver the Bayesian guarantees described in Section 3.2.2. We first introduce an important principle from Bayesian statistics which shows that these guarantees hold after terminating a procedure when the posterior quantity of interest crosses some threshold.

3.3.1 Stopping Rule Principle

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. While the Stopping Rule Principle has convenient practical ramifications, the idea remains disputed by frequentist statisticians; see, for example, the sequence of psychology articles: Yu et al. (2014), Sanborn and Hills (2014), Rouder (2014), and Sanborn et al. (2014). A/B testing is another field in which the frequentist and Bayesian guarantees of continuously monitored tests are actively studied (Johari et al., 2015; Deng, 2015; Deng et al., 2016; Dmitriev et al., 2017; Johari et al., 2017).

Because an experimenter's choice of stopping rule can influence his or her conclusions, the Stopping Rule Principle can appear counter-intuitive (Berger,

1993). We bring up a well-studied example that illustrates the seeming paradox of the Stopping Rule Principle. Consider an experimenter who samples independent random variables $X_1, X_2, \ldots$ that are each $\mathcal{N}(\theta, 1)$ and wishes to report a credible interval for the unknown parameter θ. Suppose that the experimenter stops sampling the first time, n, at which $|\sum_{j=1}^{n} X_j| \geq 2\sqrt{n}$ and—based on a noninformative prior—reports the 95% credible interval of $(\overline{X}_n - 1.96/\sqrt{n}, \overline{X}_n + 1.96/\sqrt{n})$.

For this choice of stopping rule, the experiment will terminate in finite time almost surely, even when $\theta = 0$ (by the Law of the Iterated Logarithm). However, the credible interval will have a coverage probability of zero at the point $\theta = 0$, even if the true parameter value was $\theta = 0$. The Bayesian resolution to this paradox is that the value of $\theta = 0$ is very unlikely to produce the data observed at the time of stopping, and if the experimenter were to put a positive prior probability on $\theta = 0$, the coverage issues would disappear. We will see later in Section 3.3.2 that a similar situation arises for Bayesian R&S procedures that use pPCS in a stopping rule.

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

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

select Alternative i.

Chick et al. (2010) refer to these kinds of stopping rules as *adaptive* stopping rules because they depend on the observations 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. The significance of being able to compute (and recompute) posterior quantities in the stopping rules and still deliver Bayesian guarantees cannot be understated. 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 since Bayesian R&S guarantees follow from the stopping rule, the user has complete flexibility in allocating simulation replications across alternatives. These stopping rules can therefore be used in conjunction with popular allocation rules: e.g., 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; e.g., OCBA and TS rules use the pPCS of alternatives (Chen et al., 2000; Russo, 2016) and some VIP rules use the pEOC of alternatives, or approximations thereof (Chick and Inoue, 2001b).

3.3.2 Visualization

To illustrate posterior-based stopping rules, we consider a selection procedure that iteratively takes one observation from each alternative until it can deliver

a given Bayesian R&S guarantee. For the purposes of this example, we will assume that there are $k = 2$ alternatives with a common known variance σ^2 and that the reference prior is used. Under Assumptions 4–7, the posterior distribution of the performance of Alternative i—after taking n observations from each alternative—is given by $W_i \sim \mathcal{N}(\bar{x}_i, \sigma^2/n)$ for $i = 1, 2$ and

$$\text{pPCS}_i = \Phi\left(\frac{\bar{x}_i - \bar{x}_j}{\sqrt{2\sigma^2/n}}\right), \tag{3.1}$$

$$\text{pPGS}_i = \Phi\left(\frac{\bar{x}_i - \bar{x}_j + \delta}{\sqrt{2\sigma^2/n}}\right), \quad \text{and} \tag{3.2}$$

$$\text{pEOC}_i = \sqrt{\frac{2\sigma^2}{n}}\phi\left(\frac{\bar{x}_i - \bar{x}_j}{\sqrt{2\sigma^2/n}}\right) - (\bar{x}_i - \bar{x}_j)\Phi\left(\frac{\bar{x}_j - \bar{x}_i}{\sqrt{2\sigma^2/n}}\right), \tag{3.3}$$

where $j \neq i$ and $\phi(\cdot)$ and $\Phi(\cdot)$ are the probability density function (pdf) and cumulative distribution function (cdf) of a standard normal random variable (Branke et al., 2005). Equation (3.3) follows from Equation (3.7) (see Section 3.5) for the case $\nu_{ij} = \infty$.

When the problem instance is viewed as fixed and the sample means are viewed as random variables, the difference of the partial sums, $\sum_{j=1}^{n}\left(X_{1j} - X_{2j}\right) = n(\bar{X}_1 - \bar{X}_2)$, is a Brownian motion with drift $w_1 - w_2$ and diffusion $2\sigma^2$ observed at discrete points in time, $n = 1, 2, \ldots$. The posterior-based stopping rules of Section 3.3.1 are therefore tantamount to terminating the procedure when this difference-of-partial-sums process first exits a continuation region.

For $k = 2$, the alternative with the higher posterior mean—in this case the higher sample mean—will have the higher pPCS and pPGS and the lower pEOC. Thus the stopping rules can be expressed in terms of $|n(\bar{X}_1 - \bar{X}_2)|$, implying that the boundaries of the procedure's continuation region will be symmetric.

Analytically inverting Equations (3.1) and (3.2) for $\text{pPCS}_i = 1 - \alpha$ and $\text{pPGS}_i =$

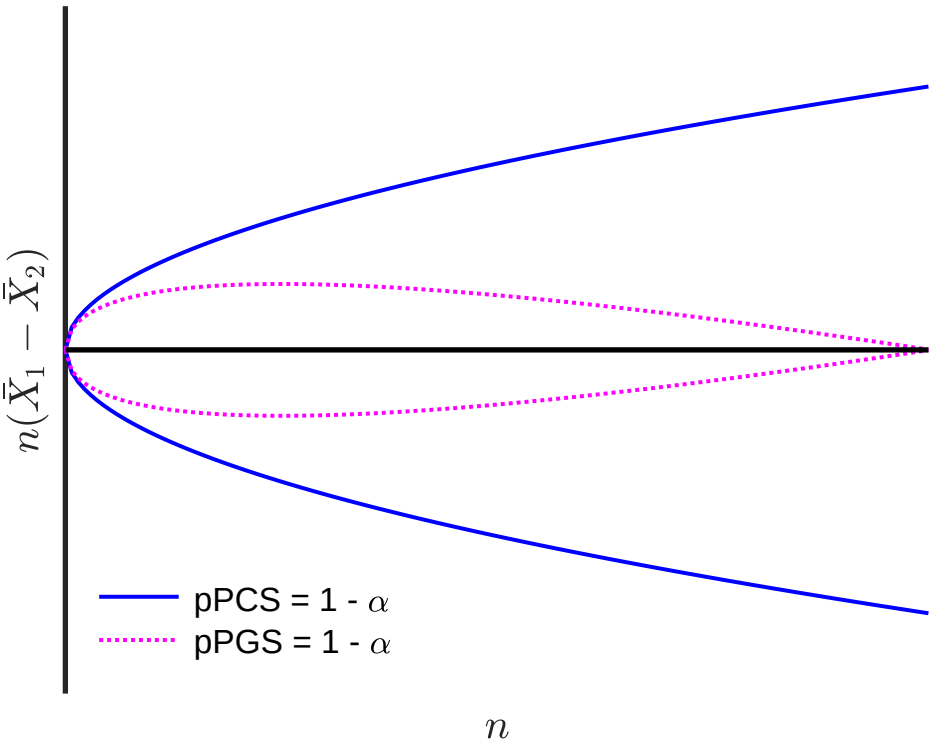

Figure 3.1: Continuation regions for the pPCS and pPGS stopping rules for equal allocation across two alternatives with common known variance and the reference prior.

$1 - \alpha$ gives the following descriptions of the pPCS and pPGS stopping rules for this equal allocation scheme:

1. pPCS stopping rule: Terminate when $|n(\bar{X}_1 - \bar{X}_2)| \geq \sqrt{2n\sigma^2}\Phi^{-1}(1-\alpha)$.
2. pPGS stopping rule: Terminate when $|n(\bar{X}_1 - \bar{X}_2)| \geq \sqrt{2n\sigma^2}\Phi^{-1}(1-\alpha) - \delta n$.

As depicted in Figure 3.1, the continuation region for the pPCS stopping rule gives rise to an open procedure (i.e., one with unbounded run-lengths), whereas the continuation region for the pPGS stopping rule gives rise to a closed procedure. The formula for the pPGS stopping rule also indicates that increasing

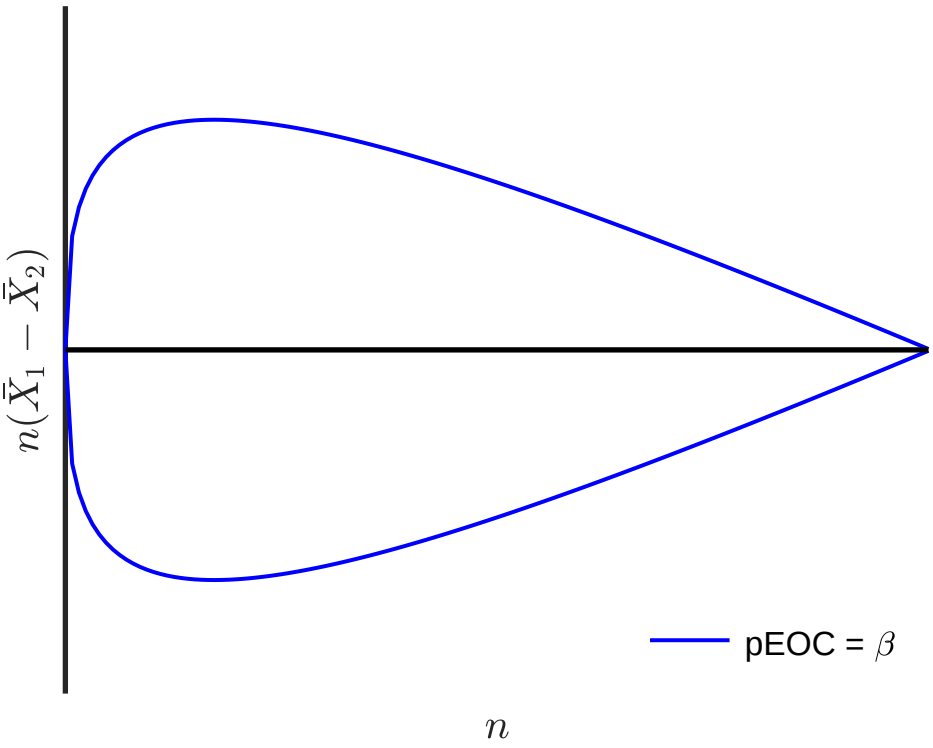

Figure 3.2: Continuation region for the pEOC stopping rule for equal allocation across two alternatives with common known variance and the reference prior.

δ narrows the continuation region, causing the procedure to terminate earlier. The continuation region for the pPGS stopping rule closely resembles that of the frequentist procedure of Zhong and Hong (2017)—at least for the case $k = 2$—except that for their procedure, the α term is replaced with an α/n term that "spends" the probability of making an incorrect selection over time. In contrast, for $k = 2$, the continuation region of the frequentist BIZ and $\mathcal{P}^*_B$ procedures of Frazier (2014) and Bechhofer et al. (1968) is described by a constant boundary: terminate when

$$|n(\bar{X}_1 - \bar{X}_2)| \geq \frac{\sigma^2}{\delta_{\text{IZ}}} \ln\left(\frac{1-\alpha}{\alpha}\right).$$

Equation (3.3) for pEOC cannot be analytically inverted when set equal to β;

instead, we numerically invert it. Figure 3.2 shows that the continuation region for the pEOC stopping rule gives rise to a closed procedure.

3.3.3 Empirical Performance

We explore how Bayesian R&S procedures that use posterior-based stopping rules perform with respect to the frequentist criteria of PCS, PGS, and EOC for *fixed* problem instances. Extensive experiments of this kind have been carried out to compare the performance and efficiency of Bayesian and frequentist R&S procedures for different parameter settings (Branke et al., 2005, 2007). The purpose of our own limited experiments is to illustrate how the empirical performance of these procedures varies depending on the problem instance.

We continue to use the setup assumed in Section 3.3.2: two alternatives with common known variance, with the added assumption that $\sigma^2 = 1$. We implement the equal allocation procedure that uses the reference prior and terminates when a given posterior-based stopping rule is first satisfied. Upon stopping, the procedure selects whichever alternative satisfies the stopping rule or the one with the higher posterior mean if both alternatives satisfy the stopping rule. For each stopping rule, we run 10,000 macroreplications of the procedure on fixed problem instances with differences in performances ($w_1 - w_2$) ranging from 0.05 to 1.

We first test the pPGS stopping rule with $1 - \alpha = 0.95$ and $\delta = 0$, 0.05, 0.10, and 0.25—the case of $\delta = 0$ corresponds to the pPCS stopping rule. The empirical PGS is calculated as the proportion of runs for which the procedure makes a δ-optimal selection *relative* to the fixed problem instance.

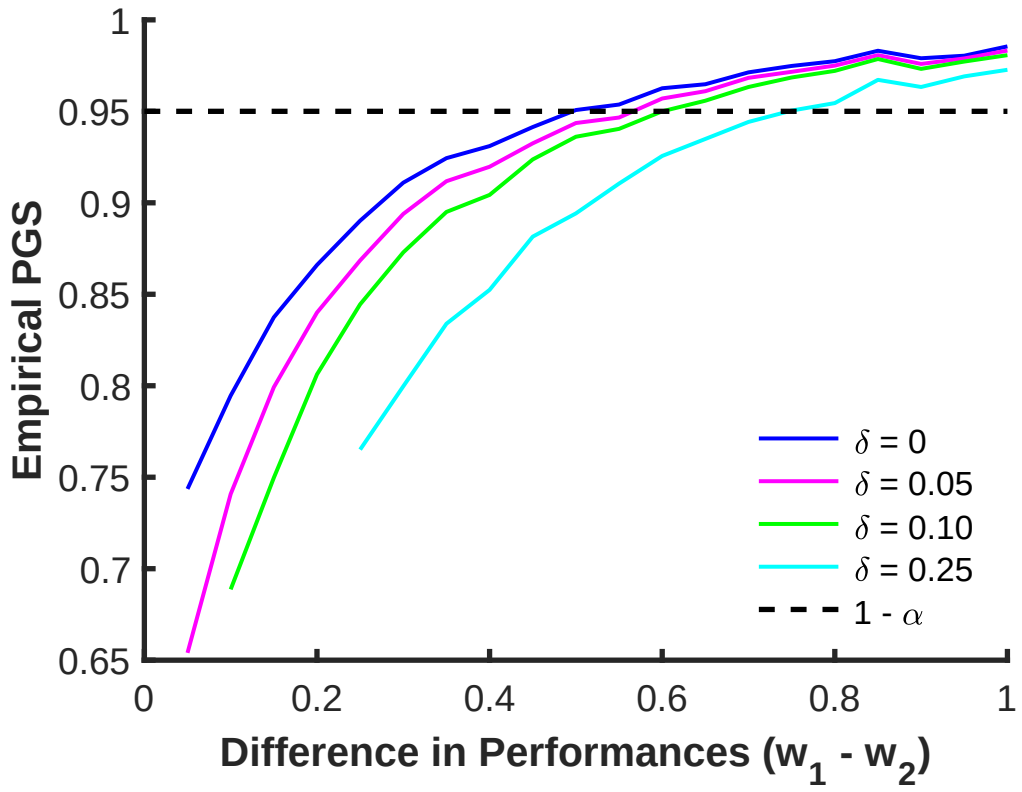

Figure 3.3: Empirical PGS for the pPGS stopping rules for equal allocation across two alternatives with common known variance and the reference prior.

The curves in Figure 3.3 show that when the true difference in performances is slightly greater than δ, the empirical PGS is less than $1 - \alpha$, whereas when the true difference in performances is large, the empirical PGS is above $1 - \alpha$. In other words, for problem instances in which there is only one good alternative, and it is hard to distinguish between the two alternatives, the procedure is less likely to make a good selection. While it is unsurprising that a Bayesian procedure performs worse for these kinds of hard problem instances (frequentist procedures tend to behave in the same way), it is noteworthy that for a range of problem instances, the empirical PGS is below $1 - \alpha$. For a procedure with a frequentist PGS guarantee, on the other hand, the empirical PGS curve would sit above $1 - \alpha$ for all problem instances. Figure 3.3 demonstrates how a Bayesian

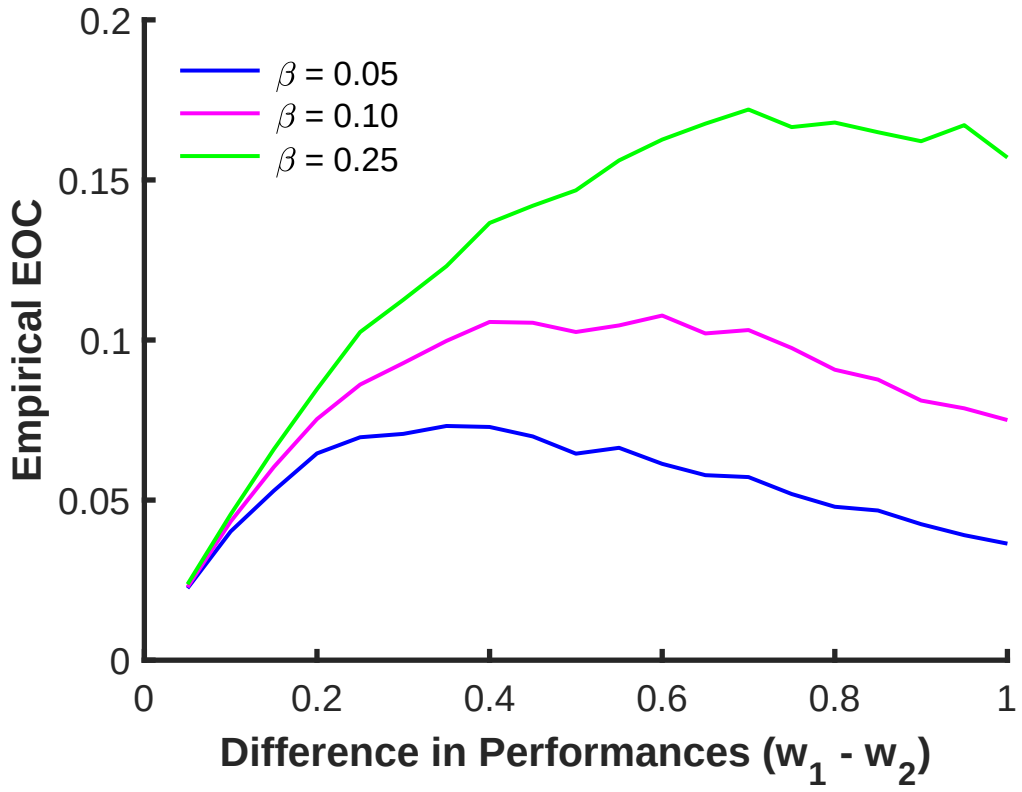

Figure 3.4: Empirical EOC for the pEOC stopping rules for equal allocation across two alternatives with common known variance and the reference prior.

R&S guarantee does not hold for all problem instances. Instead, the average empirical PGS of a Bayesian R&S procedure—where the average is taken over the prior distribution—will be above $1 - \alpha$.

Figure 3.3 also shows how as δ increases, the empirical PGS curve shifts downward for problem instances in which the difference in performances is at least δ. This is explained by the fact that increasing δ causes the procedure to stop earlier, making the procedure less likely to make a good selection when there is only one good alternative. The same behavior can be observed for the empirical PGS of frequentist R&S procedures with PGS guarantees since larger values of δ tend to correspond to taking fewer observations.

We also test the equal allocation procedure using the pEOC stopping rule for $\beta = 0.05, 0.10$, and 0.25. For a given problem instance, the empirical EOC was calculated as the average optimality gap between the selected alternative and the best. Figure 3.4 shows that for procedures using the pEOC stopping rule with $\beta = 0.05$ and $\beta = 0.10$, the empirical EOC exceeds β for a range of problem instances and is below β when the difference in performances is relatively small or large. The empirical EOC curve has this shape because for problem instances in which it is hard to make a correct selection (i.e., those with small differences in performances), making an incorrect selection is not heavily penalized under the linear loss function. Even though the procedure has the longest run-lengths in these problem instances, its empirical EOC is low. On the other hand, for problem instances in which it is easy to make a correct selection, an incorrect selection is rare.

3.3.4 Relating Posterior PCS to Frequentist PCS

For the equal allocation procedure using the reference prior, the pPCS of the alternative with the highest posterior mean has a special relationship to the PCS of the procedure. In the following discussion, we generalize to the case where there are k alternatives and assume that there is a unique best alternative.

Consider a procedure that takes n observations from each alternative and selects the alternative with the highest sample mean. The frequentist PCS of this procedure is equal to the probability that Alternative $[k]$ has the highest sample mean. For Alternative i, its sample mean, $\bar{X}_i$, is normally distributed

with mean w_i and variance σ^2/n, hence

$$\begin{aligned}
\text{PCS} &= \mathbb{P}_{\boldsymbol{w}}(\bar{X}_{[k]} > \bar{X}_i \text{ for all } i \neq [k]) \\
&= \mathbb{E}_{\boldsymbol{w}}\left[\mathbb{P}_{\boldsymbol{w}}(\bar{X}_{[k]} > \bar{X}_i \text{ for all } i \neq [k] \mid \bar{X}_{[k]})\right] \\
&= \mathbb{E}_{\boldsymbol{w}}\left[\prod_{i \neq [k]} \mathbb{P}_{\boldsymbol{w}}(\bar{X}_i \leq \bar{X}_{[k]} \mid \bar{X}_{[k]})\right] \\
&= \int_{-\infty}^{\infty}\left[\prod_{i \neq [k]} \Phi\left(z + \frac{w_{[k]} - w_i}{\sigma/\sqrt{n}}\right)\right]\phi(z)\,dz. \qquad (3.4)
\end{aligned}$$

Recall that under the reference prior, the posterior distribution of W_i is $\mathcal{N}(\bar{x}_i, \sigma^2/n)$. Given n observations from each alternative, the pPCS of the alternative with the highest posterior mean, Alternative (k), is

$$\begin{aligned}
\text{pPCS}_{(k)} &= \mathbb{P}(W_{(k)} > W_i \text{ for all } i \neq (k) \mid \mathcal{E}) \\
&= \mathbb{E}\left[\mathbb{P}(W_{(k)} > W_i \text{ for all } i \neq (k) \mid W_{(k)}, \mathcal{E}) \mid \mathcal{E}\right] \\
&= \mathbb{E}\left[\prod_{i \neq (k)} \mathbb{P}(W_i < W_{(k)} \mid W_{(k)}, \mathcal{E}) \mid \mathcal{E}\right] \\
&= \int_{-\infty}^{\infty}\left[\prod_{i \neq (k)} \Phi\left(z + \frac{\bar{x}_{(k)} - \bar{x}_i}{\sigma/\sqrt{n}}\right)\right]\phi(z)dz. \qquad (3.5)
\end{aligned}$$

The only difference between Equations (3.4) and (3.5) is that $w_{[k]} - w_i$ is replaced by $\bar{x}_{(k)} - \bar{x}_i$. This can be explained by the fact that the posterior distribution for the problem instance and the distribution of the sample means resemble each other: both are multivariate normal distributions with the same diagonal covariance matrix, but one is centered at $(w_1, w_2, \ldots, w_k)$ while the other is centered at $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)$.

This observation is closely related to the problem of estimating the PCS of a selection procedure *after* it has been run (Venter, 1989; Bofinger, 1990; Sohn and Kang, 1992). For the purposes of this discussion, we put aside concerns

about the validity or interpretability of such estimates of the PCS; see Bofinger (1994) for more on this matter. Olkin et al. (1976, 1982) proposed an estimator for the frequentist PCS that amounts to plugging the sample means into Equation (3.4); i.e., the estimator is exactly $\text{pPCS}_{(k)}$ as in Equation (3.5). Asymptotically, this estimator coincides with the frequentist PCS, but Faltin and McCulloch (1983) found it to be seriously flawed for finite sample sizes. In particular, they observed that the estimator is biased high when the performances of the top alternatives are close together and biased low when their performances are well-separated. When the differences in true performances are small, the sample means are likely to be more spread out than the true performances, making the problem instance look easier than it is. Plugging in the differences between the best sample mean and the others will therefore lead to a high estimated PCS.

This "means-spreading" phenomenon has also been observed for the problem of estimating the run-time of a given R&S procedure (Ma and Henderson, 2018). A naive approach is to take a preliminary round of simulation replications from each alternative and then run the R&S procedure using synthetic observations drawn from a multivariate normal distribution corresponding to the observed sample means and variances. Ma and Henderson (2018) observed that the average run-time computed using this approach is biased low because the sample means tend to be more spread out than the true means. Thus for especially hard problem instances, e.g., the slippage configuration, the problem instance looks like one for which the R&S would take fewer observations before terminating.

3.4 Checking the pPCS and pPGS Stopping Rules

In this section and the next, we explore ways to improve the computational efficiency of R&S procedures that deliver Bayesian guarantees via a posterior-based stopping rule. As mentioned previously, simulation replications can be allocated in any way without affecting Bayesian guarantees. Instead of studying how best to allocate simulation replications, we concentrate on how best to check the stopping rule. In particular, we focus on the computational cost of checking the stopping rule and its impact on the sampling efficiency of a procedure.

For R&S procedures that deliver Bayesian guarantees, two critical decisions are how frequently to check the stopping rule and how much computational effort to spend checking it. Together, these choices determine how much time it takes to run a procedure. At one extreme, one could allocate simulation replications one at a time and precisely calculate the posterior quantity of interest of *every* alternative after *every* replication. Though this approach would ensure that the procedure does not take any unnecessary simulation replications, it would be expensive in terms of the computational overhead. At the other extreme, simulation replications could be allocated in large batches and a cheap bound could be calculated for the posterior quantity of interest of (say) only the best-looking alternative. This approach will be cheaper in terms of overhead, but will amount to more simulation time. In studying these trade-offs, we will assume that the time required to run a simulation replication is long enough—perhaps on the order of tens of seconds—to justify expending some nontrivial amount of time checking the stopping rule.

The pPCS and pPGS stopping rules give rise to different computational challenges than the pEOC stopping rule, so we choose to consider them separately. Throughout this section and the next, we maintain Assumptions 4–7 and assume that a conjugate prior is used.

Recall that for Alternative i,

$$\text{pPCS}_i = \mathbb{P}(W_i \geq W_j \text{ for all } j \neq i \mid \mathcal{E}) \quad \text{and} \quad \text{pPGS}_i = \mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid \mathcal{E}).$$

In this section, we focus our discussion on the pPGS and treat the pPCS as a special case ($\delta = 0$).

Computing pPGS_i involves 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 $j \neq i$. Alternatively, pPGS_i can 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), both integrals could be evaluated numerically using quadrature, e.g., MATLAB's `mvncdf` function if the variances are known. As the number of alternatives increases, this approach quickly becomes computationally intractable.

Checking the pPGS stopping rule requires frequently computing the pPGS of one or more alternatives. For problems with even a modest number of alternatives, other methods for evaluating the pPGS of alternatives are necessary. We explore two: cheaply computable lower bounds and numerical integration of an equivalent one-dimensional integral.

3.4.1 Which Alternatives to Evaluate

Given the computational expense of computing the pPGS of an alternative, it is worthwhile to find ways 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 rule out alternatives that cannot satisfy the pPGS stopping rule by coming up with a simple upper bound on the pPGS of alternatives whose posterior means are not within δ of the best. Recall that the posterior mean of Alternative i is denoted by μ_i.

Proposition 3. *If* $\mu_i < \mu_{(k)} - \delta$, *then* $\text{pPGS}_i < 1/2$.

Proof. For an 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 3 implies that only alternatives whose posterior means are within δ of the best can satisfy the pPGS stopping rule when $1 - \alpha \geq 1/2$. Similarly, in this high-confidence setting, only the alternative with the highest posterior mean can satisfy the pPCS stopping rule. Proposition 3 holds even when Assumptions 6 or 7 do not, since $W_{(k)} - W_i$ is still symmetrically distributed.

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 of alternatives to construct a *quasi*-Pareto frontier. There is then no need to calculate the pPGS of dominated alternatives because some other alternative has a higher pPGS—if a dominated alternative satisfies the pPGS stopping rule, so does the alternative that dominates it.

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, it is also suggested (without proof) that one should favor alternatives that have high posterior means *and* variances. Proposition 4 formalizes this assertion and extends it to the pPGS criterion. Its proof can be found in Appendix B.1.

Proposition 4. *Suppose that the variances are known. For any pair of Alternatives i and j satisfying $\mu_i \leq \mu_j - \delta/2$ and $\hat{\sigma}_i^2 \leq \hat{\sigma}_j^2$ (where at least one inequality is strict),* $\text{pPGS}_i < \text{pPGS}_j$.

When the variances are known, combining Propositions 3 and 4 provides a method for reducing the number of alternatives for which the pPGS needs to be evaluated in order to check the pPGS stopping rule:

1. Eliminate from consideration all alternatives whose posterior means are at least δ below that of Alternative (k).
2. Among the alternatives still in consideration, construct the quasi-Pareto frontier and eliminate from consideration any dominated alternatives.

3. For the alternatives still in consideration, evaluate their pPGS.

When the variances are unknown, this method may still be used as a heuristic for eliminating alternatives from consideration. If the alternative with the highest pPGS is accidentally eliminated from consideration, the procedure may fail to identify when the pPGS stopping rule is first satisfied. The procedure's Bayesian guarantee, however, will not be invalidated. Moreover, since the alternative with the highest posterior mean will always remain in consideration, this method will entail taking no more observations than the approach of tracking a lower bound on the pPGS of only the alternative with the highest posterior mean (Branke et al., 2007).

3.4.2 Cheap Lower Bounds

One approach that avoids calculating the pPGS of an alternative is to instead compute a cheap lower bound for it. Terminating a procedure when this lower bound exceeds $1 - \alpha$ will still yield a Bayesian guarantee. This approach has the appeal of cutting down on the overhead, but any slack in the bound will likely cause the procedure to take additional replications relative to the approach of exactly calculating the pPGS.

One example of a cheap lower bound on the pPGS follows from Bonferroni's inequality:

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

When the variances are unknown, the terms $\mathbb{P}(W_i < W_j - \delta \mid \mathcal{E})$ for $j \neq i$ involve the cdf of the difference of two Student t-distributed random variables

with possibly different degrees of freedom. The approximation of Welch (1938) can be used to simplify this computation; see Chick and Inoue (2001b) for more details. The $\text{pPGS}^{\text{Bonf}}$ bound holds even when Assumptions 4–7 do not. In particular, the bound is still valid when $W_1, \ldots, W_k$ are not independent under the posterior distribution, i.e., when the procedure uses CRN.

Another cheap bound on the pPGS can be derived from Slepian's inequality (Slepian, 1962). Although Slepian's inequality is stated in terms of independent normal random variables—which would be the case for $W_1, \ldots, W_k$ if the variances were known—we show in Proposition 5 that it can be applied to alternatives with δ-optimal posterior means when the variances are unknown. Branke et al. (2007) have used Slepian's bound on the alternative with the highest posterior mean, but we provide a rigorous proof of our more general result in Appendix B.2.

Proposition 5. *For an Alternative i satisfying $\mu_i \geq \mu_{(k)} - \delta$,*

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

The restriction in Proposition 5 to alternatives whose posterior means are within δ of the best is not a limitation, since Proposition 3 implies that when $1 - \alpha > 1/2$, these are the only alternatives for which we need to evaluate the pPGS.

Whereas the expression for the pPGS deals with the maximum of the $k - 1$ positively correlated random variables $W_j - W_i$ for $j \neq i$, the expression for the $\text{pPCS}^{\text{Slep}}$ bound 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 expression as independent, thereby replacing a joint proba-

bility statement with a product of marginal ones. Peng et al. (2018) claim that the correlations of $W_j - W_i$ for $j \neq i$ that are ignored by Slepian's inequality are unimportant in a high-confidence setting (when $1 - \alpha$ is close to one) but can make a difference in a low-confidence setting.

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

As an illustration of the potential slack in the $\text{pPGS}^{\text{Bonf}}$ and $\text{pPGS}^{\text{Slep}}$ bounds, we consider a slippage configuration of posterior means, i.e., $\mu_{(k)} = \mu_j + \Delta$ for all $j \neq (k)$ for some $\Delta > 0$, with a common posterior variance $\hat{\sigma}^2$. For this configuration, $\text{pPCS}_{(k)} = 1 - \alpha$ if $\Delta = h_B \sqrt{2\hat{\sigma}^2} - \delta$ (provided δ is small enough so that Δ is positive), where $h_B(1-\alpha, k)$ 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 $1/2$; see Bechhofer (1954) and Kim and Nelson (2006b). On the other hand, it can be checked that $\text{pPGS}_{(k)}^{\text{Bonf}} = 1 - (k-1)\Phi(-h_B)$ and $\text{pPGS}_{(k)}^{\text{Slep}} = \Phi(h_B)^{k-1}$.

Figure 3.5 shows $\text{pPGS}_{(k)}^{\text{Bonf}}$ and $\text{pPGS}_{(k)}^{\text{Slep}}$ when $\text{pPGS}_{(k)} = 1 - \alpha = 0.90, 0.95$, and 0.99 for the slippage configuration of posterior means for different numbers of alternatives. Several trends are evident in Figure 3.5. 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 $1 - \alpha$ farther from one. Third, $\text{pPGS}_{(k)}^{\text{Bonf}}$ appears to be a looser bound than $\text{pPGS}_{(k)}^{\text{Slep}}$,

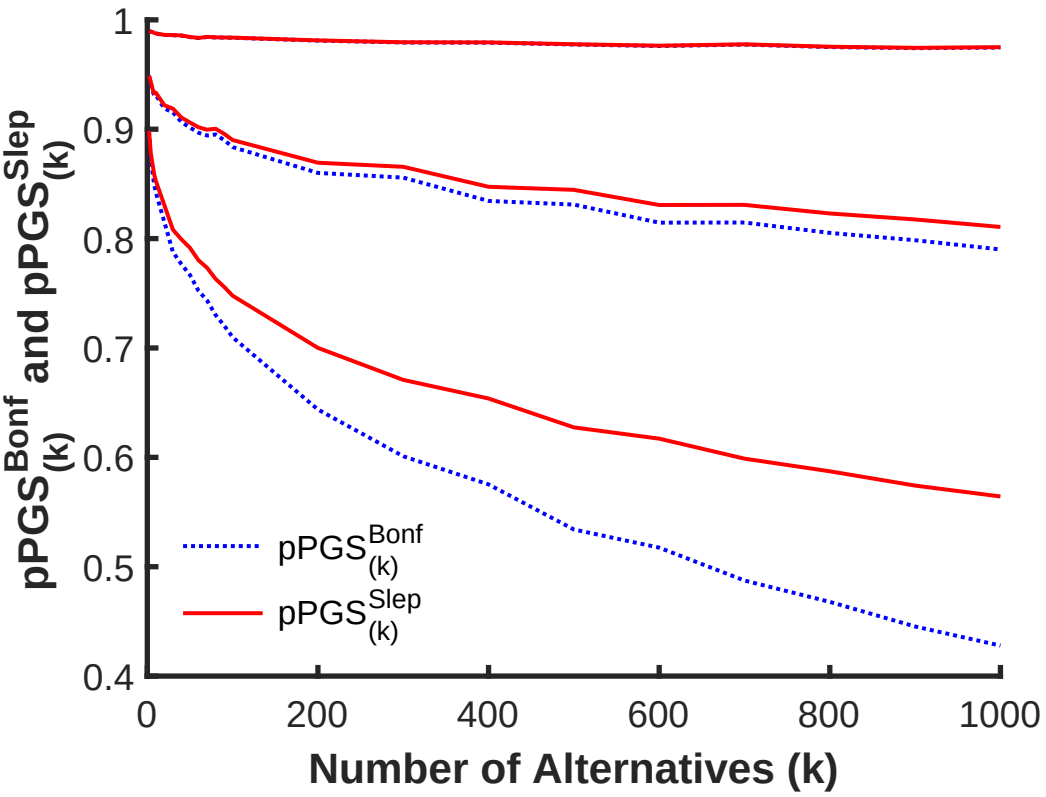

Figure 3.5: $\text{pPGS}^{\text{Bonf}}_{(k)}$ and $\text{pPGS}^{\text{Slep}}_{(k)}$ in a slippage configuration of posterior means in which $\text{pPGS}_{(k)} = 1 - \alpha$ for $1 - \alpha = 0.90$, 0.95, and 0.99.

with the difference in the bounds growing with the number of alternatives; for $1 - \alpha = 0.99$, the bounds are roughly equivalent.

3.4.3 Numerical Integration

Given the potential slack in the $\text{pPGS}^{\text{Bonf}}$ and $\text{pPGS}^{\text{Slep}}$ bounds depicted in Figure 3.5, we return to the challenge of calculating the pPGS in a way that is cheap to compute and scales well with the number of alternatives. The approach we take relies on the observation of Peng et al. (2016) that pPGS_i can be expressed as

a one-dimensional integral by conditioning on the performance of Alternative i:

$$\begin{aligned}
\text{pPGS}_i &= \mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid \mathcal{E}) \\
&= \mathbb{E}\left[\mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid W_i, \mathcal{E}) \mid \mathcal{E}\right] \\
&= \mathbb{E}\left[\prod_{j \neq i} \mathbb{P}(W_j \leq W_i + \delta \mid W_i, \mathcal{E}) \mid \mathcal{E}\right] \\
&= \int_{-\infty}^{\infty} \left[\prod_{j \neq i} F_{W_j|\mathcal{E}}(w + \delta)\right] f_{W_i|\mathcal{E}}(w) dw, \qquad (3.6)
\end{aligned}$$

where $f_{W_i|\mathcal{E}}(\cdot)$ and $F_{W_i|\mathcal{E}}(\cdot)$ are the marginal posterior pdf and cdf of W_i.

The integrand in Equation (3.6) is the product of $k-1$ cdfs and a pdf, all of which are either for normal or Student t-distributed random variables. Equation (3.6) thereby avoids the need to approximate the difference of two Student 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 evaluating Equation (3.6), we use MATLAB's `integral` function to perform the numerical integration for values of k ranging from 10 to 100000. For each value of k, we generate 10000 random posterior distributions according to $\mu_i \sim \mathcal{N}(0, 25)$ and $\hat{\sigma}_i^2 \sim \text{ChiSquared}(4)$, independent for all $i = 1, \ldots, k$, and compute the pPGS of Alternative 1. Table 3.1 reports the average times (according to MATLAB's `tic` and `toc` functions).

Table 3.1: Computational time to numerically integrate Equation (3.6) for the pPGS for 10000 random problem instances for different numbers of alternatives. All times are accurate to within ±0.002 seconds.

Number of Alternatives	10	100	1000	10000	100000
Average Time (seconds)	0.012	0.013	0.018	0.069	0.478

The average times in Table 3.1 suggest that even for fairly large numbers

of alternatives, the pPGS can be quickly computed. When the number of alternatives is small (<100), it may not be worthwhile to implement the methods proposed in Section 3.4.1 for reducing the number of alternatives for which the pPGS is calculated. Depending on the time required to simulate one replication, numerical integration may be too slow for very large numbers of alternatives (>100000), as each evaluation of the pPGS takes at least half a second and will need to be performed many times on many alternatives. For such large problems, approximating the pPGS using Monte Carlo integration of Equation (3.6) may be a viable alternative. We defer our discussion of Monte Carlo methods to Section 3.7. In Section 3.6, we demonstrate the potential savings in the number of observations taken from numerically integrating Equation (3.6) in conjunction with the method presented in Section 3.4.1.

3.5 Checking the pEOC Stopping Rule

Recall that the pEOC of Alternative i is defined as $\text{pEOC}_i = \mathbb{E}[W_{[k]} - W_i \mid \mathcal{E}]$. Computing pEOC_i involves evaluating a k-dimensional integral of the function $W_{[k]} - W_i$ times the joint posterior pdf of $\boldsymbol{W}$ over a union of half-spaces given by $\cup_{j \neq i}\{W_j - W_i \geq 0\}$. Repeatedly computing pEOC as a k-dimensional integral is likely too time-intensive, even for small numbers of alternatives. We present a sequence of ideas—similar to those in Section 3.4—for efficiently checking the pEOC stopping rule: identifying which alternatives can have the lowest pEOC and examining different approaches for evaluating that pEOC.

In contrast to the analysis of the pPGS stopping rule, determining which alternative has the lowest pEOC is straightforward. Proposition 6 formalizes a

well-known result.

Proposition 6.

$$\text{pEOC}_{(k)} = \min_{1 \le i \le k} \text{pEOC}_i.$$

Proof. For an arbitrary Alternative i,

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

Since the term $\mathbb{E}[W_{[k]} \mid \mathcal{E}]$ is independent of i, the alternative with the highest posterior mean, Alternative (k), will have the lowest pEOC. □

Proposition 6 implies that to check the pEOC stopping rule, a procedure needs to compute the pEOC for only the alternative with the highest posterior mean. The result of Proposition 6 holds regardless of the form of the posterior distribution, i.e., even when Assumptions 4–7 are not satisfied.

3.5.1 Cheap Upper Bounds

As was the case with the pPGS, computing a cheap upper bound on the pEOC of an alternative and terminating when it drops below β will result in a procedure delivering the pEOC guarantee. Branke et al. (2005) provide one such bound, though their name for it is somewhat of a misnomer:

$$\begin{aligned}\text{pEOC}_i = \mathbb{E}[\max_{j \ne i} (W_j - W_i)^+ \mid \mathcal{E}] &\le \mathbb{E}\left[\sum_{j \ne i} (W_j - W_i)^+ \mid \mathcal{E}\right] \\ &= \sum_{j \ne i} \mathbb{E}[(W_j - W_i)^+ \mid \mathcal{E}] =: \text{pEOC}_i^{\text{Bonf}}.\end{aligned}$$

The $\text{pEOC}^{\text{Bonf}}$ bounds holds even when Assumptions 4–7 do not, meaning that it can be applied when CRN are used. Branke et al. (2005) derive approximations of the terms summed in $\text{pEOC}_i^{\text{Bonf}}$ for the case when variances are unknown:

$$\mathbb{E}[(W_j - W_i)^+ \mid \mathcal{E}] \approx \sqrt{\hat{\sigma}_j^2 + \hat{\sigma}_i^2}\Psi_{\nu_{ji}}\left[\frac{\mu_i - \mu_j}{\sqrt{\hat{\sigma}_j^2 + \hat{\sigma}_i^2}}\right], \tag{3.7}$$

where

$$\nu_{ji} = \frac{\left(\hat{\sigma}_j^2 + \hat{\sigma}_i^2\right)^2}{(\hat{\sigma}_j^2)^2/\nu_j + (\hat{\sigma}_i^2)^2/\nu_i},$$

and

$$\Psi_\nu[s] = \int_{u=s}^{\infty}(u-s)\phi_\nu(u)du = \frac{\nu + s^2}{\nu - 1}\phi_\nu(s) - s\Phi_\nu(-s),$$

and $\phi_\nu(\cdot)$ and $\Phi_\nu(\cdot)$ are the pdf and cdf of a Student t-distributed random variable with ν degrees of freedom. Thus $\Psi_\nu[s]$ offers a closed-form expression for $\mathbb{E}[X - s]^+$ where X is a Student t-distributed random variable with ν degrees of freedom.

By expressing the pEOC in terms of an integral of the pPGS, we present an upper bound on the pEOC derived from Slepian's inequality that is—to the best of our knowledge—the first of its kind. Our approach uses 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:

$$\begin{aligned}
\text{pEOC}_i &= \mathbb{E}[W_{[k]} - W_i \mid \mathcal{E}] \\
&= \int_0^\infty \mathbb{P}(W_{[k]} - W_i > \delta \mid \mathcal{E})\, d\delta \\
&= \int_0^\infty \mathbb{P}(W_i < W_j - \delta \text{ for some } j \neq i \mid \mathcal{E})\, d\delta \\
&= \int_0^\infty \left[1 - \mathbb{P}(W_i \geq W_j - \delta \text{ for all } j \neq i \mid \mathcal{E})\right] d\delta \\
&= \int_0^\infty \left[1 - \text{pPGS}_i\right] d\delta,
\end{aligned} \tag{3.8}$$

where pPGS_i is implicitly a function of δ. Proposition 5 implies that for the alternative with the highest posterior mean, $\text{pPGS}_{(k)} \geq \text{pPGS}^{\text{Slep}}_{(k)}$ for all $\delta \geq 0$. Substituting the definition of $\text{pPGS}^{\text{Slep}}_{(k)}$ into Equation (3.8) gives

$$\text{pEOC}_{(k)} \leq \int_0^\infty [1 - \text{pPGS}^{\text{Slep}}_{(k)}]\, d\delta = \int_0^\infty \left[1 - \prod_{j \neq (k)} \mathbb{P}(W_{(k)} \geq W_j - \delta \mid \mathcal{E})\right] d\delta =: \text{pEOC}^{\text{Slep}}_{(k)}.$$

The integrand in $\text{pEOC}^{\text{Slep}}_{(k)}$ contains a product of $k-1$ cdfs and should therefore take roughly as long to numerically integrate as Equation (3.6). One minor difference is that the cdfs in $\text{pEOC}^{\text{Slep}}_{(k)}$ deal with the difference of two normal or Student t-distributed random variables, necessitating some form of analytical (e.g., Welch 1938) or numerical approximation when the variances are unknown and the sample sizes are unequal. In contrast, the cdfs in Equation (3.6) are for normal or Student t-distributed random variables.

To illustrate the potential slack in the $\text{pEOC}^{\text{Bonf}}_{(k)}$ and $\text{pEOC}^{\text{Slep}}_{(k)}$ bounds, we again take the approach of evaluating them for slippage configurations of posterior means with a common posterior variance. 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 $\text{pEOC}^{\text{Bonf}}_{(k)}$ and $\text{pEOC}^{\text{Slep}}_{(k)}$ for the posterior distributions. Figure 3.6 shows that the tightness of both bounds deteriorates as the number of alternatives increases, with the $\text{pEOC}^{\text{Bonf}}_{(k)}$ 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 3.6.

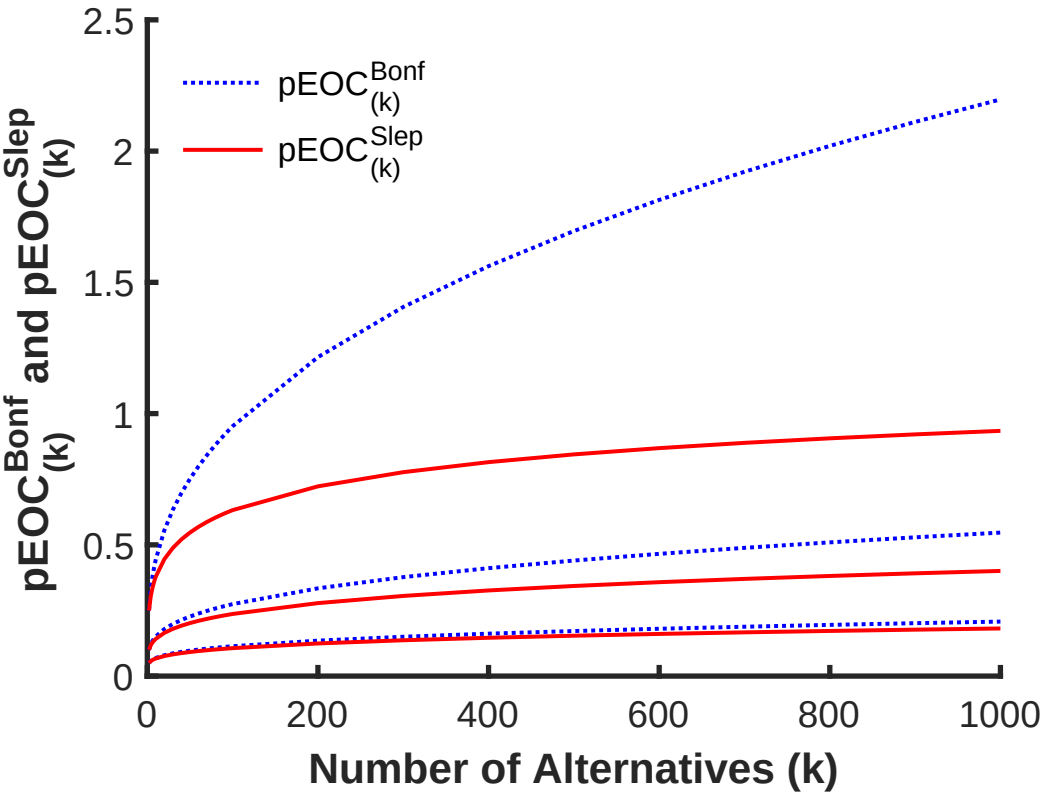

Figure 3.6: $\text{pEOC}_{(k)}^{\text{Bonf}}$ and $\text{pEOC}_{(k)}^{\text{Slep}}$ in a slippage configuration of posterior means in which $\text{pEOC}_{(k)} = \beta$ for $\beta = 0.05$, 0.10, and 0.25.

3.5.2 Numerical Integration

We now turn our attention to ways to compute the pEOC without resorting to evaluating a k-dimensional integral. The proof of Proposition 6 indicates that one way to compute pEOC_i 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 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

$$\begin{aligned}\mathbb{E}[W_{[k]}|\mathcal{E}] &= \int_{-\infty}^{\infty} w \sum_{i=1}^{k} f_{W_i|\mathcal{E}}(w) \prod_{j \neq i} F_{W_j|\mathcal{E}}(w)\, dw \\ &= \sum_{i=1}^{k} \int_{-\infty}^{\infty} w \left[\prod_{j \neq i} F_{W_j|\mathcal{E}}(w) \right] f_{W_i|\mathcal{E}}(w)\, dw. \end{aligned} \tag{3.9}$$

Each of the k integrals in Equation (3.9) resembles that of Equation (3.6) with $\delta = 0$ and an extra w multiplied in the integrand. We evaluate Equation (3.9) for the same experimental setup of random problem instances and report the timing results in Table 3.2. Unsurprisingly, the computational times are roughly equivalent to k times the computational times in Table 3.1. For more than a modest number of alternatives, numerically integrating Equation (3.9) becomes too expensive for the purposes of checking the pEOC stopping rule.

Table 3.2: Computational time to numerically integrate Equation (3.9) for the pEOC for 10000 random problem instances for different numbers of alternatives. All times are accurate to within ±0.05 seconds.

Number of Alternatives	10	100	1000
Average Time (seconds)	0.12	1.314	18.54

Another approach to computing pEOC_i is to substitute Equation (3.6) for pPGS_i into Equation (3.8) to yield a two-dimensional integral:

$$\begin{aligned}\text{pEOC}_i &= \int_0^\infty \left[1 - \int_{-\infty}^\infty \left[\prod_{j\neq i} F_{W_j|\mathcal{E}}(w+\delta)\right] f_{W_i|\mathcal{E}}(w)\, dw\right] d\delta \\ &= \int_0^\infty \int_{-\infty}^\infty \left[1 - \prod_{j\neq i} F_{W_j|\mathcal{E}}(w+\delta)\right] f_{W_i|\mathcal{E}}(w)\, dw\, d\delta. \end{aligned} \tag{3.10}$$

Table 3.3: Computational time to numerically integrate Equation (3.10) for the pEOC for 1000 random problem instances for different numbers of alternatives. All times are accurate to within ±0.05 seconds.

Number of Alternatives	10	100	1000
Average Time (seconds)	2.27	2.88	5.17

Table 3.3 shows that the computational time for evaluating Equation (3.10) scales better with k than does evaluating Equation (3.9), but for small numbers of alternatives it is more intensive.

3.6 Experimental Results

In this section, simulation experiments give a sense of the potential savings from using the methods proposed in Sections 3.4 and 3.5 to exactly check the pPGS and pEOC stopping rules. As mentioned earlier, there is a tradeoff between the the computational time spent checking the stopping rule and the total number of observations taken by a procedure. In the following experiments, we report the total number of observations taken by a procedure. Overall savings in terms of a procedure's run-time can be worked out by comparing the computational times reported in Sections 3.4 and 3.5—along with the frequency with which the stopping rule is checked—to the time required to generate a simulation replication and the number of observations taken.

Let N_{new} be the (random) number of observations taken by a procedure using the methods described in Sections 3.4 and 3.5. For comparison, we use the approach of Branke et al. (2007) for checking the stopping rules: terminating when $\text{pPGS}^{\text{Slep}}_{(k)} \geq 1 - \alpha$ or terminating when $\text{pEOC}^{\text{Bonf}}_{(k)} \leq \beta$; let $N_{standard}$ be the corresponding number of observations taken. Because of the slack in the $\text{pPGS}^{\text{Slep}}$ and $\text{pEOC}^{\text{Bonf}}$ bounds, $N_{new} \leq N_{standard}$ almost surely.

The run-lengths of a given procedure can vary greatly from run to run and from problem to problem. To facilitate meaningful comparisons across different procedures and settings, we consider the relative (as opposed to the absolute) difference in the number of observations taken by a procedure when using the aforementioned approaches. That is, we let $(N_{standard} - N_{new})/N_{standard}$ represent the fractional savings from exactly checking the stopping rule. The distribution of $(N_{standard} - N_{new})/N_{standard}$ describes the savings that can occur from run to run

with the average fractional savings being a useful summary statistic.

We test the proposed methods on three allocation rules: equal allocation (EA), Thompson sampling (TS), and optimal computing budget allocation (OCBA). While EA can be inefficient—it continues to sample alternatives that are clearly inferior—it provides a baseline for potential savings. The TS rule allocates the next simulation replication in proportion to the pPCS of each alternative. This can be achieved by generating a random problem instance from the posterior distribution and choosing to take the next simulation replication from the alternative with the highest performance. For the pPGS stopping rule, the OCBA rule chooses the alternative that would yield the highest value of $\text{pPGS}_{(k)}^{\text{Slep}}$ if an extra observation were taken from it and the posterior mean were unchanged. Likewise, for the pEOC stopping rule, the OCBA rule chooses the alternative that would yield the lowest value of $\text{pEOC}_{(k)}^{\text{Bonf}}$; see Branke et al. (2007) for full details. We run the Thompson sampling and OCBA rules fully sequentially, i.e., taking one observation at a time, and check the stopping rule after every allocation decision.

To speed up the numerical experiments, we use a conditional Monte Carlo method referred to as *splitting* (Asmussen and Glynn, 2007). On a given macroreplication of a procedure, the stopping rule that uses the exact pPGS or pEOC will be satisfied before the stopping rule that uses the $\text{pPGS}_{(k)}^{\text{Slep}}$ or $\text{pEOC}_{(k)}^{\text{Bonf}}$ bounds. When the exact stopping condition is met, we record the observations collected by that time. We then *split* the macroreplication by generating s independent realizations of the rest of the procedure that each run until the $\text{pPGS}_{(k)}^{\text{Slep}}$ or $\text{pEOC}_{(k)}^{\text{Bonf}}$ stopping condition is met. In this way, one macroreplication yields s observations of the fractional savings. While this method is not guaranteed to

reduce the variance of the Monte Carlo estimator of the average fractional savings, it allows us to quickly generate identically distributed (albeit dependent) observations of the fractional savings. In our experiments, we run $m = 100$ macroreplications of the procedures with $s = 50$ splits for a total of 5000 observations of the fractional savings. These observations are plotted in histograms to illustrate the distribution of the fractional savings.

In our experiments, we test the procedures on a fixed problem instance with $k = 50$ alternatives whose performances are in a slippage configuration with a spacing of $w_{[k]} - w_i = 1$ for all $i \neq [k]$. Admittedly, the slippage configuration is unlikely to arise in practice. It does however offer an indication of the greatest savings from exactly checking the stopping rule because of the slack of the $\text{pPGS}^{\text{Slep}}_{(k)}$ or $\text{pEOC}^{\text{Bonf}}_{(k)}$ bounds in this configuration. We chose the values of $\delta = 1$, $1 - \alpha = 0.90$ and $\beta = 0.25$.

Figures 3.7, 3.8, and 3.9 show the fractional savings from exactly checking the pPGS stopping rule when using the different allocation rules. In Figure 3.7, we see that the potential savings when using the EA rule are greater than those when using more sophisticated allocations. This can be explained by the fact that failing to stop as soon as possible incurs a cost of at least k additional observations. While the average fractional savings for the TS and OCBA rules seem modest ($\approx 5\%$), the histograms in Figures 3.8 and 3.9 show that the upper tails of the fractional savings exceeds 10% for a nontrivial proportion of runs.

Figures 3.10, 3.11, and 3.12 show the fractional savings from exactly checking the pEOC stopping rule. Compared to the experiments for the pPGS stopping rule, the average fractional savings for the pEOC stopping rule are roughly three times greater. This observation is likely explained by the looseness of the

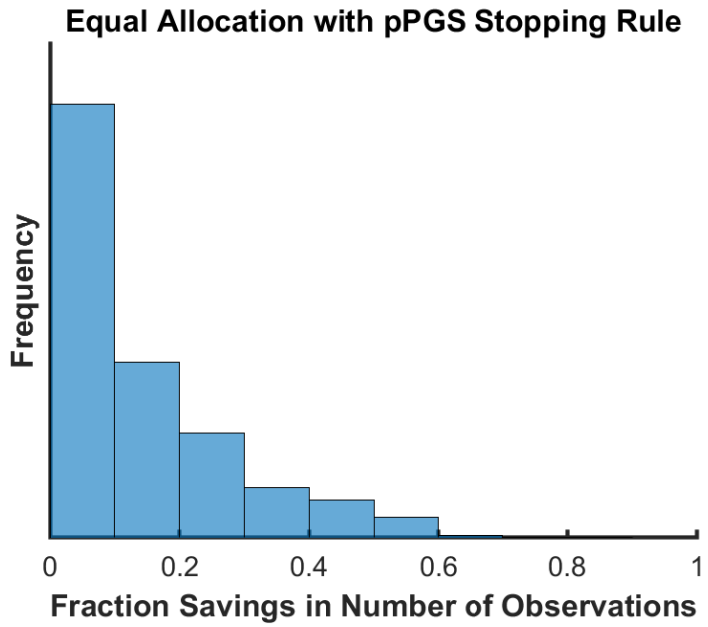

Figure 3.7: Histogram of fraction savings for equal allocation for the pPGS stopping rule with $1 - \alpha = 0.90$ in a slippage configuration. Average fraction savings are 13.55% with 95% CI [11.59%, 15.50%].

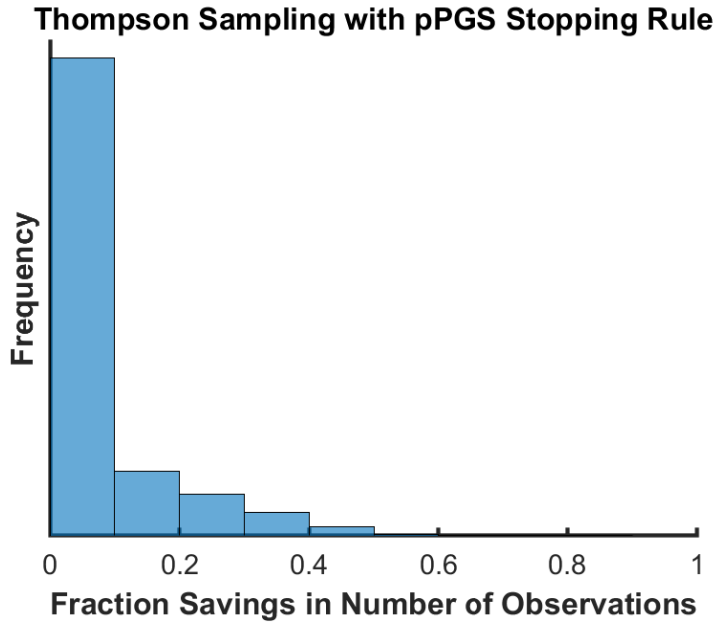

Figure 3.8: Histogram of fraction savings for Thompson sampling for the pPGS stopping rule with $1 - \alpha = 0.90$ in a slippage configuration. Average fraction savings are 6.44% with 95% CI [5.08%, 7.81%].

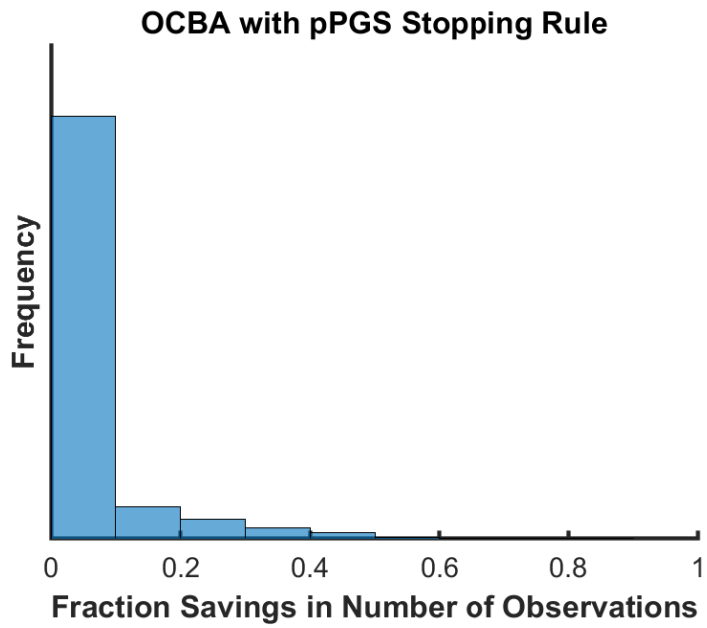

Figure 3.9: Histogram of fraction savings for OCBA for the pPGS stopping rule with $1-\alpha = 0.90$ in a slippage configuration. Average fraction savings are 4.45% with 95% CI [3.15%, 5.75%].

$\text{pEOC}^{\text{Bonf}}_{(k)}$ bound, as depicted in Figure 3.6. Figure 3.10 for the EA rule shows the potential for three- and four-fold savings in the number of observations just by exactly evaluating pEOC instead of the $\text{pEOC}^{\text{Bonf}}_{(k)}$ bound. The fractional savings for the TS or OCBA rules likewise have the potential for significant savings.

3.7 Conclusion

In this chapter, 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 has crossed a threshold. We discuss interpretations of Bayesian R&S guarantees and how they differ from frequentist guarantees, both in terms of their empirical performance and their efficacy in a variety of decision-making

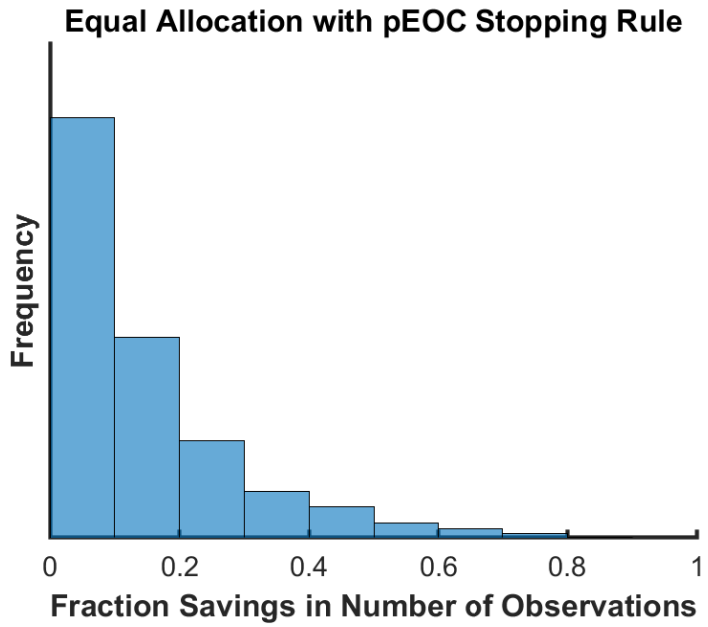

Figure 3.10: Histogram of fraction savings for equal allocation for the pEOC stopping rule with $\beta = 0.25$ in a slippage configuration. Average fraction savings are 36.10% with 95% CI [33.14%, 39.06%].

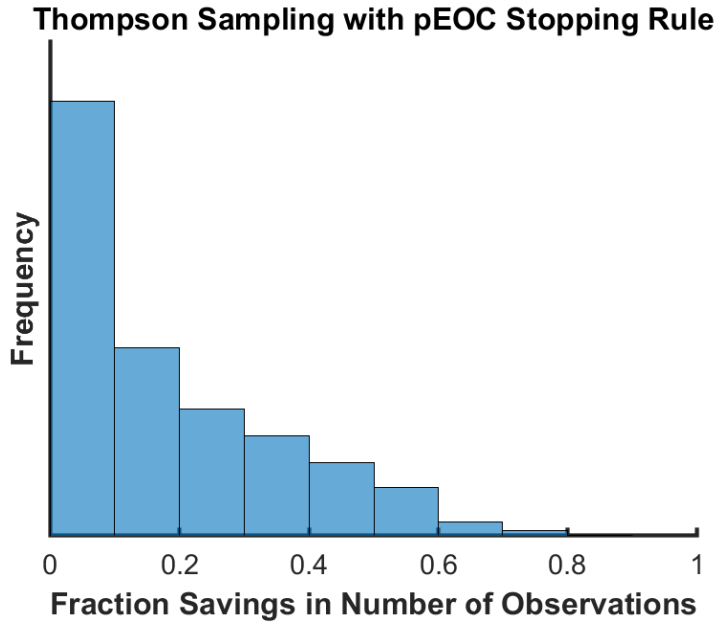

Figure 3.11: Histogram of fraction savings for Thompson sampling for the pEOC stopping rule with $\beta = 0.25$ in a slippage configuration. Average fraction savings are 18.59% with 95% CI [16.68%, 20.49%].

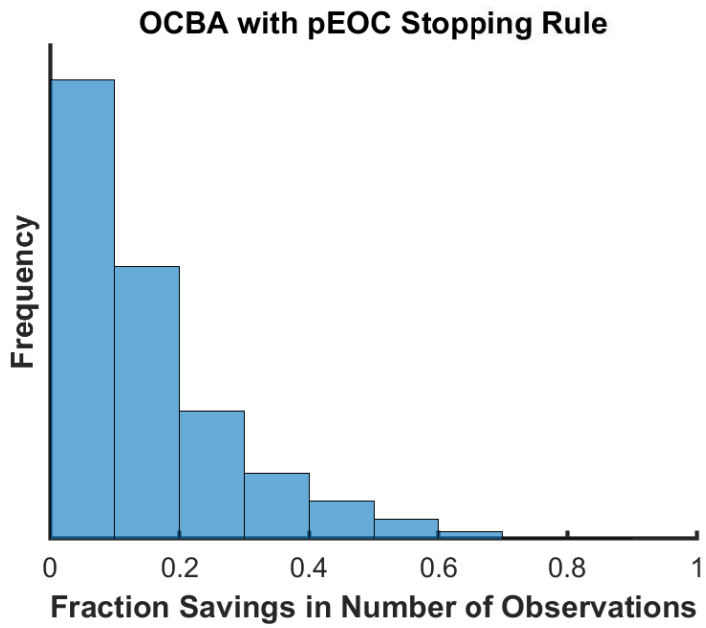

Figure 3.12: Histogram of fraction savings for OCBA for the pEOC stopping rule with $\beta = 0.25$ in a slippage configuration. Average fraction savings are 14.65% with 95% CI [13.18%, 16.13%].

settings. We also call attention to the negative consequences of the pPCS guarantee, namely its tendency to cause a procedure to incur long run-lengths for little gain in performance.

A crucial consideration in running a Bayesian R&S procedure is how to check the stopping rule. 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 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 methods can translate into savings in the number of simulation replications a procedure takes. While savings in random problem instances tend to be modest, they can be large on occasion, with the pEOC stopping rule showing greater potential savings. Depending on the time needed to

generate a simulation replication, even a small reduction in the number of simulation replications can translate to a meaningful reduction in the run time of a procedure.

In future work, we hope to explore the potential for using Monte Carlo methods to "pre-check" stopping rules. Specifically, only when a Monte Carlo estimate of the pPGS or pEOC crosses the desired threshold does one use numerical integration to exactly evaluate the posterior quantity of interest and check the stopping rule. A straightforward application of Monte Carlo methods would be to generate random problem instances from the posterior distribution of $\boldsymbol{W}$ and average the estimates $\mathbf{1}\{W_i \geq W_{[k]} - \delta\}$ for pPGS_i and $W_{[k]} - W_i$ for pEOC_i. The variances of these Monte Carlo estimators may be further reduced by using conditional Monte Carlo schemes that exploit properties of elliptical distributions to estimate integrals over intersections or unions of half-spaces (Jian and Henderson, 2017; Ahn and Kim, 2018). Alternatively, Monte Carlo integration could be used to estimate the one- and two-dimensional integrals in Equations (3.6), (3.9), and (3.10).

As R&S methods have come to be used to tackle problems with thousands of alternatives, an area of active research is the design of parallelizable procedures. Bayesian procedures naturally require some degree of synchronization to compute posterior quantities of interest given the most current data. While it makes sense to run simulation replications in parallel across processors, we see little benefit from parallelizing the operation of checking a stopping rule.

Subset selection is a related R&S approach wherein a procedure returns a subset of alternatives from which the decision-maker can make a final decision based on secondary performance measures or other factors. Natural analogs

to the pPCS and pPGS criteria are the posterior probabilities that a given subset contains an optimal or δ-optimal alternative. In this setting, a procedure may take observations until it can identify a fixed-size subset of alternatives for which the relevant posterior probability exceeds $1 - \alpha$. A procedure can otherwise terminate at any time and then identify a subset (of uncontrolled size) satisfying the guarantee. An interesting, open question is how to identify the smallest subset of alternatives for which the posterior probability that it contains a δ-optimal alternative is at least $1 - \alpha$.

CHAPTER 4

REUSING SEARCH DATA IN RANKING AND SELECTION

The majority of this chapter appeared as Eckman and Henderson (2018).

4.1 Introduction

Real-world simulation optimization problems often feature a large, possibly uncountable, number of systems and a limited computational budget with which to evaluate them. To have any hope of tackling such problems, practitioners must restrict their attention to evaluating only a subset of systems. In the absence of structural information about the objective function, candidate systems are usually identified via sampling or search.

Whereas sampling methods, e.g., random sampling, Latin hypercube sampling, and orthogonal designs, determine a set of candidate systems without regard to observed performance, search methods move from system to system, taking replications and using the observed performance of explored systems to identify the next system to evaluate. As a result, search methods are expected to return better systems than sampling. Search methods used in simulation optimization problems range from the naive, e.g., random search, to more advanced methods combined with optimization, e.g., stochastic approximation, sample average approximation; see, for example, Chapters 18, 20, and 21 of Henderson and Nelson (2006) and Chapters 10, 11, and 12 of Fu (2015) for more discussion.

As pointed out by Boesel et al. (2003b), heuristic search methods, e.g., simulated annealing, Nelder-Mead, and tabu search, do not typically guarantee convergence to the globally optimal system, and procedures with guarantees of

converging to a local optimum often require the simulation budget to approach infinity. In either case, no finite-time guarantee is made with respect to identifying the best system among those explored. To obtain a statistical guarantee of this kind, ranking-and-selection (R&S) procedures can be applied on the set of candidate systems (Boesel et al., 2003b).

R&S procedures have been embedded within various search methods to provide some degree of statistical control at each iteration of the search; examples include nested partitions (Olafsson, 1999), simulated annealing (Ahmed and Alkhamis, 2002), and pattern search (Sriver et al., 2009), all of which do not reuse past replications. In contrast, Boesel et al. (2003a,b) examine how R&S procedures can be applied on a database of explored systems to "clean up" *after* search. Boesel et al. (2003a) outline how simulation optimization software can be configured to permit R&S after search while Boesel et al. (2003b) further elucidate selection and subset-selection procedures that *reuse* replications taken during a heuristic search. The procedures give guarantees on the probability of correct selection for instances in which the configuration of the returned systems is in the preference zone, i.e., the best system is at least δ better than the others.

It is enticing to reuse the replications gathered during a search as input to a R&S procedure, especially when generating simulation replications is computationally expensive. Nevertheless, search data violates one of the key assumptions of most R&S procedures: that replications are independent within and across systems. More specifically, since the identities of new systems depend on the observed performance of previously explored systems, the replications taken during search are *conditionally dependent given the sequence of returned*

systems. This conditional dependence causes issues when attempting to make guarantees conditional on the output of a search.

Even though this dependence issue may be known to the R&S community, existing procedures that reuse search data are not designed to address it. Through simulation experiments, we demonstrate how this dependence can undermine the guarantees of R&S procedures that reuse search data. For practical simulation optimization problems, we find that the dependence in the search data likely does not lead to violated guarantees for such procedures. However, this conclusion is a result of the conservative inequalities used in the design of the procedures, and not of any focused effort to account for the dependence of the search data.

Similar issues will arise for other applications of R&S procedures in which search data is reused. In this setting, a notable R&S procedure that reuses previously collected replications is Sequential Selection with Memory (SSM) (Pichitlamken et al., 2006). SSM provides a guarantee on the probability of selecting the best candidate system in a given search iteration and has been applied within random search (Pichitlamken et al., 2006) and nested partitions for discrete optimization (Pichitlamken and Nelson, 2003). See Pichitlamken and Nelson (2001) for an earlier presentation of SSM.

Hong and Nelson (2007) address a related setting in which systems are revealed in batches by a generic "system-generating algorithm" that can be considered a search. In this framework, selection decisions are made at each iteration and early selection decisions are intended to direct the search in identifying new systems. In making each selection decision, the proposed procedures reuse the replications of systems returned in earlier batches. These procedures pro-

vide guarantees over the selection decisions at each search iteration as well as an overall guarantee on the system ultimately selected.

The remainder of this chapter is outlined as follows. In Section 4.2, we rigorously define the types of guarantees that arise from applying R&S after sampling or search and discuss conditions under which these guarantees hold. In the process, we explain why guarantees based on the indifference-zone formulation are unsatisfactory. In Section 4.3, we present a search-like method, called Adversarial Search, in which an adversary introduces new systems in a way that exploits the conditional dependence of search data. In Section 4.4, we test several selection and subset-selection procedures that reuse the replications taken by Adversarial Search. We also test a subset-selection procedure that reuses data from a more-realistic search. In Section 4.5, we summarize our findings and outline open research questions.

4.2 R&S after Sampling or Search

We consider the optimization problem $\max_{x\in\Theta} \mu(x)$ where Θ is a space of systems, presumably large, and $\mu : \Theta \mapsto \mathbb{R}$ is an objective function. Here x represents a vector of decision variables, e.g., simulation parameters, and is henceforth referred to as a *system*. In keeping with convention, we use x when the identity of a system is fixed and X when it is a random variable. At a given system x, the objective function can only be estimated from noisy replications $Y_j(x)$, $j = 1, 2, \ldots$, satisfying $\mu(x) = \mathbb{E}[Y_j(x)]$. We further assume that the replications are of the form $Y_j(x) = \mu(x) + \sigma(x)\xi_j$ where $\sigma(x)$ represents the variability at system x and the random error terms $\xi_j \sim \mathcal{N}(0, 1)$ are independent and identically distributed

(i.i.d.). In this chapter, we consider only R&S procedures for which the random error terms ξ_j are independent across systems. In Section 4.5, we briefly mention how our results can be extended to R&S procedures that make use of common random numbers.

4.2.1 Sampling vs. Search

We use the term *sampling* to describe methods that choose systems $X_1, \ldots, X_k \in \Theta$ without the need to take any replications, where k is specified in advance by the decision-maker. More precisely, sampling methods are of the form $X_i \in m\mathcal{F}_{i-1}$, i.e., X_i is measurable w.r.t. $\mathcal{F}_{i-1}$, for $i = 1, \ldots, k$, where $\mathcal{F}_{i-1}$ is the sigma-field generated by $\{X_1, \ldots, X_{i-1}, U_0, \ldots U_{i-1}\}$ (all information obtained just prior to the identification of system X_i) and U_{i-1} represents an (optional) random input associated with identifying system X_i. In other words, each system returned by a sampling method is a deterministic function of the sequence of previously returned systems and the random inputs used to identify them.

In contrast, we define *search* methods to be those that on the ith iteration take n_{0i} replications of a system X_i and identify the $(i+1)$st system based on these replications and those of previously explored systems. More precisely, let $\mathbf{Y}(X_i) := [Y_1(X_i), \ldots, Y_{n_{0i}}(X_i)]^T$ be the vector of n_{0i} replications of system X_i and let $\bar{Y}(X_i) := (n_{0i})^{-1} \sum_{j=1}^{n_{0i}} Y_j(X_i)$ denote the sample mean of these replications; thus $\bar{Y}(X_i)$ is a conditionally unbiased estimator of $g(X_i)$ conditioned on X_i. Then search methods are of the form $X_i \in m\mathcal{G}_{i-1}$, i.e., X_i is measureable w.r.t. $\mathcal{G}_{i-1}$, for $i = 1, \ldots, k$, where $\mathcal{G}_{i-1}$ is the sigma-field generated by $\{X_1, \ldots, X_{i-1}, \mathbf{Y}(X_1), \ldots, \mathbf{Y}(X_{i-1}), U_0, \ldots, U_{i-1}\}$. By containing all of the search data

$\mathbf{Y}(X_1), \ldots, \mathbf{Y}(X_{i-1})$, the sigma-field $\mathcal{G}_{i-1}$ is very general; all of the search methods we consider use only the sample means $\bar{Y}(X_1), \ldots, \bar{Y}(X_{i-1})$.

By using the replications $\mathbf{Y}(X_1), \ldots, \mathbf{Y}(X_{i-1})$ to choose system X_i, search methods may return better systems than sampling methods. However, this improvement in the quality of returned systems comes at a cost: the replications taken during search are dependent. In particular, because the identity of a new system is dependent on the observed performance—and not just the identities—of previously explored systems, *the search data are conditionally dependent given the sequence of returned systems*.

To help illustrate this dependence, consider a small example in which only two systems X_1 and X_2 are returned by a search. Their replications are

$$Y_j(X_1) = \mu(X_1) + \sigma(X_1)\xi_{1j} \quad \text{for } j = 1, \ldots, n_{01},$$

$$Y_\ell(X_2) = \mu(X_2) + \sigma(X_2)\xi_{2\ell} \quad \text{for } \ell = 1, \ldots, n_{02},$$

where ξ_{ij} denotes the random error term of the jth replication of the ith returned system. Although the random error terms ξ_{1j} and $\xi_{2\ell}$ are independent for all j and ℓ, the replications $Y_j(X_1)$ and $Y_\ell(X_2)$ are dependent because the identity of the second system X_2 depends—through the search—on the identity of the first system X_1 *and* its replications $\mathbf{Y}(X_1)$. Furthermore, when we condition on the identities of the returned systems, i.e., the event $\{X_1 = x_1, X_2 = x_2\}$, the fact that $X_2 = x_2$ provides information about the replications $Y_j(X_1)$ and the error terms ξ_{1j}, $j = 1, \ldots, n_{01}$. Because of this information about the observed performance of system X_1, the replications $Y_1(X_1), \ldots, Y_{n_{01}}(X_1)$ are conditionally dependent given the sequence of returned systems. We will discuss in Section 4.2.4 how this conditional dependence can affect the guarantees of R&S procedures that reuse search data.

In the preceding argument, by comparing the replications of systems X_1 and X_2, we implicitly assumed knowledge of the *order* in which the systems were returned by the search. While this information is often available in practice, we argue that knowledge of the order of returned systems is not necessary for search data to be conditionally dependent. That is, if the returned systems were randomly permuted after the search, so that their initial ordering was no longer known, the search data would still be dependent conditioned only on the identities of the returned systems. As justification for this claim, all of the R&S procedures tested in this chapter—and almost all R&S procedures in general—are invariant under permutation, i.e., they make the same selection decisions regardless of the order in which the systems are labeled. Our experimental results for these procedures when reusing search data (see Section 4.4) show lower probabilities of correct and good selection compared to the case of independent data.

4.2.2 Guarantees of R&S Procedures

We introduce notation necessary for describing aspects of R&S procedures. Let $\mathcal{X} := \{x_1, \ldots, x_k\} \subseteq \Theta$ denote a fixed set of systems and define the vectors $\mu := [\mu(x_1), \ldots, \mu(x_k)]^T$ and $\Sigma := [\sigma^2(x_1), \ldots, \sigma^2(x_k)]^T$. Thus μ and Σ are the means and variances, respectively, of the normal distributions from which replications of systems $x_1, \ldots, x_k$ are drawn. Let $x_{[i]}$ denote the ith system when the systems are ordered by the objective function μ, i.e., $\mu(x_{[1]}) \leq \mu(x_{[2]}) \leq \cdots \leq \mu(x_{[k]})$. We assume that larger objective function values are better, hence system $x_{[k]}$ is (one of) the best.

The vector μ is referred to as the *configuration* of the systems. Under the indifference-zone formulation of Bechhofer (1954), the decision-maker specifies a parameter $\delta > 0$ that partitions the space of configurations into the preference zone and the indifference zone. The preference zone is defined as $\text{PZ}(\delta) := \{\mu : \mu(x_{[k]}) - \mu(x_{[k-1]}) \geq \delta\}$ and the indifference zone $\text{IZ}(\delta)$ is defined as its complement.

We consider two kinds of R&S procedures: selection and subset-selection procedures. Selection procedures ultimately select a single system, denoted x_K, whereas subset-selection procedures return a subset of systems $I \subseteq \mathcal{X}$. Under both approaches, guarantees are defined with respect to the events of correct selection (CS), i.e., selecting or preserving the best system, and good selection (GS), i.e., selecting or preserving a system whose performance is strictly within δ of the best. For selection procedures, we define correct and good selections as $\text{CS} := \{\mu(x_K) = \mu(x_{[k]})\}$ and $\text{GS} := \{\mu(x_K) \geq \mu(x_{[k]}) - \delta\}$. For subset-selection procedures, we likewise define correct and good selections as $\text{CS} := \{x_{[k]} \in I\}$ and $\text{GS} := \{\exists x \in I \text{ s.t. } \mu(x) > \mu(x_{[k]}) - \delta\}$.

R&S procedures often give guarantees on the probability of correct selection (PCS) based on the indifference-zone formulation:

$$\mathbb{P}_{(\mu,\Sigma)}(\text{CS}) \geq 1 - \alpha \quad \text{for all } \mu \in \text{PZ}(\delta), \tag{4.1}$$

for $1/k < 1 - \alpha < 1$ where $\mathbb{P}_{(\mu,\Sigma)}$ is the probability measure with respect to the normal distributions specified by the elements of μ and Σ. That is, for any configuration in the preference zone, the R&S procedure guarantees a lower bound on PCS. Some procedures also give guarantees on the probability of good selection (PGS) (Bechhofer, 1954; Dudewicz and Dalal, 1975; Rinott, 1978), regardless of whether the configuration is in the preference zone:

$$\mathbb{P}_{(\mu,\Sigma)}(\text{GS}) \geq 1 - \alpha \quad \text{for all } \mu. \tag{4.2}$$

Guarantee (4.2) implies Guarantee (4.1) because in PZ(δ) the only good system is the best system.

4.2.3 PGS and PCS Guarantees after Sampling or Search

In the R&S literature, the set of systems under consideration is usually fixed in advance, as in Section 4.2.2. We consider extensions of Guarantees (4.1) and (4.2) to instances in which the set of systems $\mathcal{X}$—and hence the configuration μ—is not fixed, but is instead random, namely, the output of a sampling or search method. We first ask the motivating question: what are meaningful guarantees on PGS and PCS when a R&S procedure $\mathcal{R}$ is run after an arbitrary sampling or search method $\mathcal{S}$ identifies a set of candidate systems?

For PGS, one might be interested in the guarantee that for an instance of the combined procedure $\mathcal{S} + \mathcal{R}$, a good selection is made with probability $\geq 1 - \alpha$, i.e.,

$$\mathbb{P}(\text{GS after } \mathcal{S}) \geq 1 - \alpha. \tag{4.3}$$

Guarantee (4.3) is particularly relevant in practical problems for which the combined procedure $\mathcal{S} + \mathcal{R}$ is run only once. In contrast to $\mathbb{P}_{(\mu,\Sigma)}$, the probability measure $\mathbb{P}$ in Guarantee (4.3) is defined with respect to the replications taken by $\mathcal{R}$ *and* the replications and random inputs of $\mathcal{S}$; hence we will refer to Guarantee (4.3) as an "overall" guarantee.

The analogous overall PCS guarantee is given by

$$\mathbb{P}(\text{CS after } \mathcal{S} \mid \mu(\mathcal{X}) \in \text{PZ}(\delta)) \geq 1 - \alpha, \tag{4.4}$$

where $\mu(\mathcal{X})$ denotes the configuration of the random set of returned systems $\mathcal{X}$.

Guarantee (4.4) is restrictive in the sense that it is only over instances in which the configuration of the returned systems is in the preference zone. When using search methods on practical problems, this is likely a rare event because search methods typically return systems with similar performance as they approach a local optimum. This is especially the case for problems in which the space of systems and the objective function are continuous. Furthermore, the decision-maker has little control over whether a search returns a configuration in PZ(δ) and no way of verifying this event with certainty. For these reasons, we strongly believe that PCS guarantees based on the indifference-zone assumption are inappropriate for the setting of R&S after search.

Guarantees (4.3) and (4.4) appear to be difficult to prove directly, the main obstacle being the probability measure $\mathbb{P}$. Whereas the probability measure $\mathbb{P}_{(\mu,\Sigma)}$ of Guarantees (4.1) and (4.2) is specific to the set of systems, $\mathbb{P}$ is defined on a greatly enlarged probability space. Proving statements involving $\mathbb{P}$ may require knowledge of the likelihood that $\mathcal{S}$ returns an arbitrary set of systems—knowledge that is unavailable to us.

A more promising approach to proving Guarantees (4.3) and (4.4) is to condition on the set of returned systems $\mathcal{X}$. This yields the conditional guarantees

$$\mathbb{P}(\text{GS after } \mathcal{S} \mid \mathcal{X}) \geq 1 - \alpha \quad \text{for all } \mathcal{X}, \tag{4.5}$$

$$\mathbb{P}(\text{CS after } \mathcal{S} \mid \mathcal{X}) \geq 1 - \alpha \quad \text{for all } \mathcal{X} \text{ s.t. } \mu(\mathcal{X}) \in \text{PZ}(\delta). \tag{4.6}$$

By the law of total expectation, Guarantees (4.5) and (4.6) imply Guarantees (4.3) and (4.4), respectively. Conditioning on the set of systems $\mathcal{X}$ has the advantage of fixing the distribution of the replications taken by $\mathcal{R}$. Therefore the probability measure in Guarantees (4.5) and (4.6) more closely resembles $\mathbb{P}_{(\mu,\Sigma)}$ of Guarantees (4.1) and (4.2).

Proposition 7 makes use of this observation to establish conditions under which Guarantee (4.5)—and hence Guarantee (4.3)—follows from Guarantee (4.2). An analogous result for the PCS guarantees follows from the same proof. Although Proposition 7 is trivial to prove, it is important because it shows how R&S procedures can be used safely after sampling or search.

Proposition 7. *Suppose that a R&S procedure $\mathcal{R}$ takes as input a random set of systems $\mathcal{X}$ returned by a sampling or search method $\mathcal{S}$, where $\mathcal{X}$ contains a fixed number of systems. If $\mathcal{R}$ does not reuse any replications taken by $\mathcal{S}$, but instead takes its own (new) replications and guarantees*

$$\mathbb{P}_{(\mu,\Sigma)}(\text{GS}) \geq 1 - \alpha \quad \textit{for all } \mu,$$

then

$$\mathbb{P}(\text{GS } \textit{after } \mathcal{S} \mid \mathcal{X}) \geq 1 - \alpha \quad \textit{for all } \mathcal{X}.$$

Proof. After running $\mathcal{S}$, the set of returned systems $\mathcal{X}$ is fixed. Therefore the replications taken by $\mathcal{R}$ are drawn from fixed distributions parameterized by the mean vector $\mu(\mathcal{X})$ and variance vector $\Sigma(\mathcal{X})$. It follows that for all $\mathcal{X}$,

$$\mathbb{P}(\text{GS after } \mathcal{S} \mid \mathcal{X}) = \mathbb{P}_{(\mu(\mathcal{X}),\Sigma(\mathcal{X}))}(\text{GS}) \geq 1 - \alpha.$$

□

An obvious deficiency of Proposition 7 is that $\mathcal{R}$ does not use any of the replications taken by $\mathcal{S}$. When $\mathcal{S}$ is a sampling method, any replications that may have been collected during $\mathcal{S}$ can in fact be reused in $\mathcal{R}$ without affecting Guarantees (4.5) and (4.6) since, conditional on the set of returned systems, these replications are statistically identical to those taken by $\mathcal{R}$, namely, i.i.d. and independent across systems. The more complicated case of reusing replications when $\mathcal{S}$ is a search method is discussed in the next section.

4.2.4 Reusing Search Replications

Reusing search data in a R&S procedure is desirable since computational effort has already been expended to generate them. For example, in parallel computing environments, great efficiency can be gained by having processors communicate the observed performance of explored systems to other processors to help with screening and search.

When a R&S procedure reuses search replications, Guarantees (4.5) and (4.6) are harder to prove directly because the selection decisions now depend on the data $\mathbf{Y}(X_1), \ldots, \mathbf{Y}(X_k)$, replications we have shown to be dependent. A potential approach to proving Guarantees (4.5) and (4.6) is to further condition on $\mathcal{G}_k$, i.e., the sigma-field generated by the sequence of returned systems, the search replications, and the random inputs used to identify the returned systems. Conditioning on $\mathcal{G}_k$ simplifies the probability measure over the selection decisions. However, this approach also runs into a major problem: guarantees conditional on $\mathcal{G}_k$ will not hold in the almost-sure sense for any fixed R&S procedure that takes a finite number of samples. For sets in $\mathcal{G}_k$ corresponding to realizations of the search in which the good systems perform poorly, the resulting PCS and PGS will be reduced. As normal random variables, the sample means of the good systems can be made arbitrarily negative. Thus the corresponding PCS and PGS can almost always be made arbitrarily small, so that eventually their guarantees are violated.

With no apparent approach to directly prove Guarantees (4.5) and (4.6), it is natural to wonder if they in fact hold for existing R&S procedures that reuse search data. To address this question, we attempt to find instances of optimization problems and searches for which Guarantees (4.3) and (4.4) are violated,

thereby implying that Guarantees (4.5) and (4.6) are violated.

4.3 Adversarial Search

We present a search-like method, Adversarial Search (AS), that is designed to exploit the dependence of the search data in a way that misleads R&S procedures. AS is designed to amplify the difficulties in ensuring Guarantees (4.3) and (4.4) in a contrived manner. Later we will do the same for a more-realistic search procedure.

A detailed description of AS is given in Algorithm 1. In short, AS introduces a δ-better system (relative to the current best) when the best system thus far looks best, i.e., has the highest sample mean, and a δ-worse system otherwise.

Algorithm 1: Adversarial Search (AS)

Data: An integer $k > 1$ and initial system X_1 (either fixed or random).
Result: A sequence of systems $X_2, \ldots, X_k$.
Take replications $Y_j(X_1)$, $j = 1, \ldots, n_{01}$, and calculate $\bar{Y}(X_1)$;
$i^*_{true} \leftarrow 1$ (index of system with highest true performance $\mu(X_i)$);
$i^*_{obs} \leftarrow 1$ (index of system with highest observed mean $\bar{Y}(X_i)$);
for $i \leftarrow 2$ **to** k **do**
  if $i^*_{obs} = i^*_{true}$ **then**
    Introduce a system X_i s.t. $\mu(X_i) = \mu(X_{i^*_{true}}) + \delta$;
    $i^*_{true} \leftarrow i$;
  else
    Introduce a system X_i s.t. $\mu(X_i) = \mu(X_{i^*_{true}}) - \delta$;
  end
  Take replications $Y_j(X_i)$, $j = 1, \ldots, n_{0i}$, and calculate $\bar{Y}(X_i)$;
  if $\bar{Y}(X_i) > \bar{Y}(X_{i^*_{obs}})$ **then**
    $i^*_{obs} \leftarrow i$;
  end
end

The intuition behind AS is as follows. When the best system thus far has

the highest observed performance, AS introduces a new best system, thereby leaving the previous best system as a tough competitor. And when the best system thus far does not have the highest observed performance, systems that are δ worse than the best system are introduced, thereby leaving the best system as a weak competitor. In this way, AS attempts to return configurations with search data for which the best system does not look best.

Every set of systems returned by AS has a unique best system that is at least δ better than all of the others, i.e., $\mu(\mathcal{X}) \in \text{PZ}(\delta)$ for every $\mathcal{X}$. Therefore the events of correct selection and good selection are equivalent—as are Guarantees (4.3) and (4.4)—when applying a R&S procedure after AS.

We call AS a "search-like" method because it relies on two unworkable assumptions that are beyond our formal definition of search: (i) a sufficiently large space of systems from which to draw systems of a given performance and (ii) knowledge of the objective function. Although AS does not satisfy our definition of search, it can serve as a near-worst-case benchmark for testing if the guarantees of R&S procedures reusing search data are robust to all optimization problems and all search methods. This claim is justified by the observation that there exist instances of optimization problems and search methods that behave exactly the same as AS.

To demonstrate, consider an example in which there are four systems in Θ. Without loss of generality, let them be labeled 1, 2, 3, and 4 with performances $\mu(1) = 0$, $\mu(2) = 0$, $\mu(3) = \delta$ and $\mu(4) = 2\delta$. Suppose that the simulation budget allows only three of the four systems to be evaluated, i.e., $k = 3$. We consider a particular search method over these four systems. First, Systems 2 and 3 are evaluated by taking n_{02} and n_{03} samples, respectively. If System 2 has a higher

sample mean, System 1 is next evaluated by taking n_{01} samples, and if System 3 has a higher sample mean, System 4 is next evaluated by taking n_{04} replications. One possible reasoning for identifying the third system in this fashion is that the decision-maker believes System 1 will have performance similar to System 2 while System 4 will have performance similar to System 3. One can verify that this search method behaves exactly the same as AS when letting $X_1 = 2$ w.p. 1.

In this fashion, one can construct similar optimization problems and search methods that mimic AS for larger numbers of returned systems ($k > 3$). Hence finding violated guarantees in our experiments with AS will show that the guarantees of R&S procedures that reuse search data do not hold for all optimization problems and all search methods.

4.4 Experiments with Reusing Search Replications

In Sections 4.4.1 and 4.4.2, we present experimental results for two selection procedures (Bechhofer and Rinott) and two subset-selection procedures (Modified Gupta and Screen-to-the-Best). Although these procedures are not specifically designed to be used on search data, we believe they effectively illustrate the negative impact that dependent search data can have on PCS guarantees. Some of the R&S procedures that have been proposed for use after search (e.g., Screen-and-Continue of Nelson et al. (2001) and Screen-Restart-and-Select and Sort-and-Iterative-Screen of Boesel et al. (2003b)) are built around combinations of the Screen-to-the-Best and Rinott procedures. Therefore our findings provide some insight into the performance of such procedures.

We test the four procedures under two settings: (i) when applied after AS,

reusing search data, and (ii) on a fixed set of systems in the slippage configuration (SC): $\mu(x_{[i]}) = \mu(x_{[k]}) - \delta$ for $1 \leq i < k$. For the R&S procedures we consider, PCS is minimized in the slippage configuration; therefore testing the slippage configuration gives a lower bound on PCS for the procedures when applied after sampling. For both settings, we perform 10 000 macroreplications at each value of k and calculate the empirical PCS—an unbiased estimate of the left-hand sides of Guarantees (4.3) and (4.4). In Section 4.4.3, we test a subset-selection procedure that reuses replications from a realistic search.

In all of our experiments, we choose a desired PCS of $1 - \alpha = 0.95$ and fix an indifference-zone parameter $\delta = 1$, an initial sample size $n_0 = 10$, and a common variance $\sigma^2 = 1$. Although a common variance is unrealistic for simulation optimization problems, we should expect the guarantees of R&S procedures that reuse search data to hold even in this stylized setting.

4.4.1 Selection Procedures

In our experiments, we test the procedure of Bechhofer (1954) for common, known variance and the procedure of Rinott (1978) for uncommon, unknown variances. We choose to test the latter because it is easier to implement than the procedure of Dudewicz and Dalal (1975) that also handles uncommon, unknown variances. The Bechhofer procedure is designed to ensure a tight PCS of $1 - \alpha$ in the slippage configuration while the Rinott procedure is more conservative because it must allow for unequal, unknown variances.

We focus on single- and multi-stage selection procedures—of which the Bechhofer and Rinott procedures are two examples—because all of the required

replications can be taken during the search. On the other hand, fully sequential procedures take replications from systems one at a time. Some fully sequential procedures, e.g., Kim and Nelson (2001) and Frazier (2014) for unknown variances, can be easily incorporated into a selection-after-search framework by taking only n_0 replications of each system during the search. For other fully sequential procedures that do not require an initial sample size of n_0 replications from each system, e.g., Frazier (2014) for the case of common, known variance, it remains an open question how they might be applied in conjunction with search.

The details of the two selection procedures we test are as follows:

Bechhofer Take $N = \lceil (2h_B^2\sigma^2)/\delta^2 \rceil$ replications of each system where h_B is the α-upper quantile of the maximum of a $k-1$ dimensional multivariate normal vector with means 0, variance 1, and pairwise correlations 1/2. The constant h_B can be derived from the values of N tabulated in Table 2.1 on page 19 of Bechhofer et al. (1995). Select the system with the highest sample mean.

Rinott Take n_0 replications of each system and calculate S_i^2, the sample variance of System i based on the initial replications. Take additional replications so that System i has a total of N_i replications where

$$N_i = \max\left\{n_0, \left\lceil \frac{h_R^2 S_i^2}{\delta^2} \right\rceil\right\}$$

and $h_R = h(k, 1-\alpha, n_0)$ is the solution to

$$\int_0^\infty \left[\int_0^\infty \Phi\left(\frac{h}{\sqrt{\nu\,(1/x + 1/y)}}\right) f_\nu(x)dx\right]^{k-1} f_\nu(y)\,dy = 1-\alpha,$$

where $\nu = n_0 - 1$ and f_ν is the probability density function (p.d.f.) of a chi-squared random variable with ν degrees of freedom. The constant h_R is tabulated in Table 2.13 on pages 62–63 of Bechhofer et al. (1995). Select the system with the highest overall sample mean.

When testing the Bechhofer and Rinott procedures applied after AS, we consider two cases: (i) "AS All": all of the required replications are taken during AS, i.e., $n_{0i} = N_i$ for $i = 1, \ldots, k$, and (ii) "AS n_0": only n_0 replications of each system are taken during AS and the remaining replications are taken afterwards, i.e., $n_{0i} = 10$ for $i = 1, \ldots, k$.

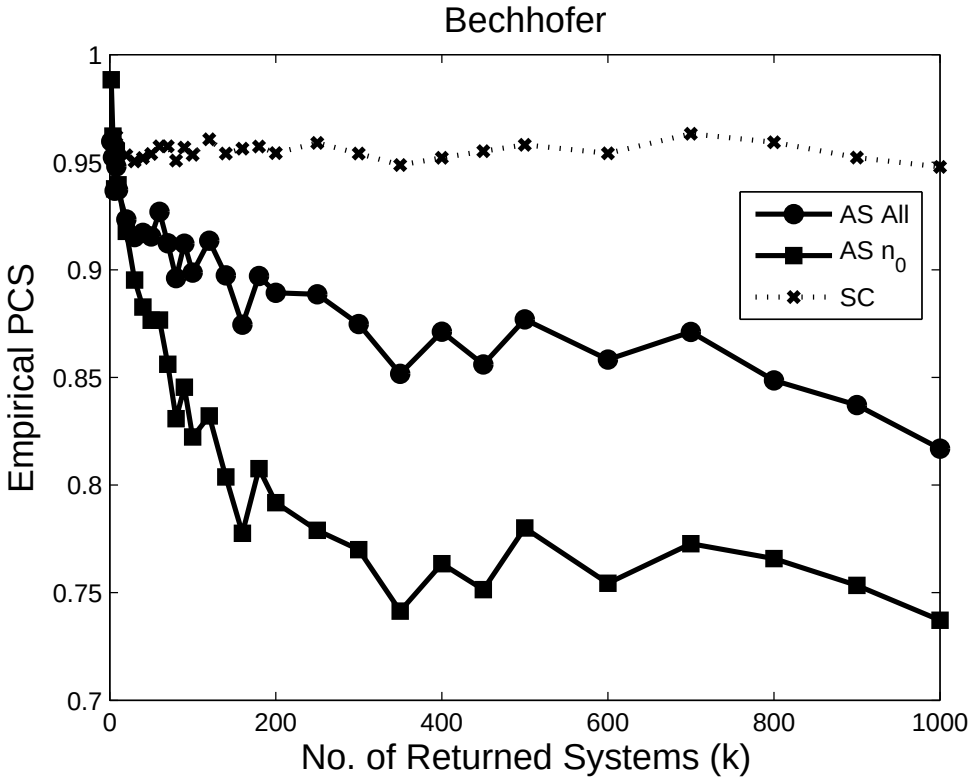

Figure 4.1: Empirical PCS of the Bechhofer procedure in the slippage configuration (SC) and when applied after AS, reusing replications taken during search. Based on the usual normal confidence intervals, the empirical PCS of the Bechhofer procedure is accurate to within ±0.01.

In Figure 4.1, the empirical PCS for the Bechhofer procedure reusing repli-

cations from AS quickly falls below the guaranteed PCS of 0.95. The empirical PCS for the case of taking n_0 replications during AS deteriorates more rapidly than when all of the replications are taken during AS. One possible explanation is that after taking only n_0 replications, the initial ranking of systems is more variable and so AS likely returns less-favorable configurations of systems, such as the slippage configuration.

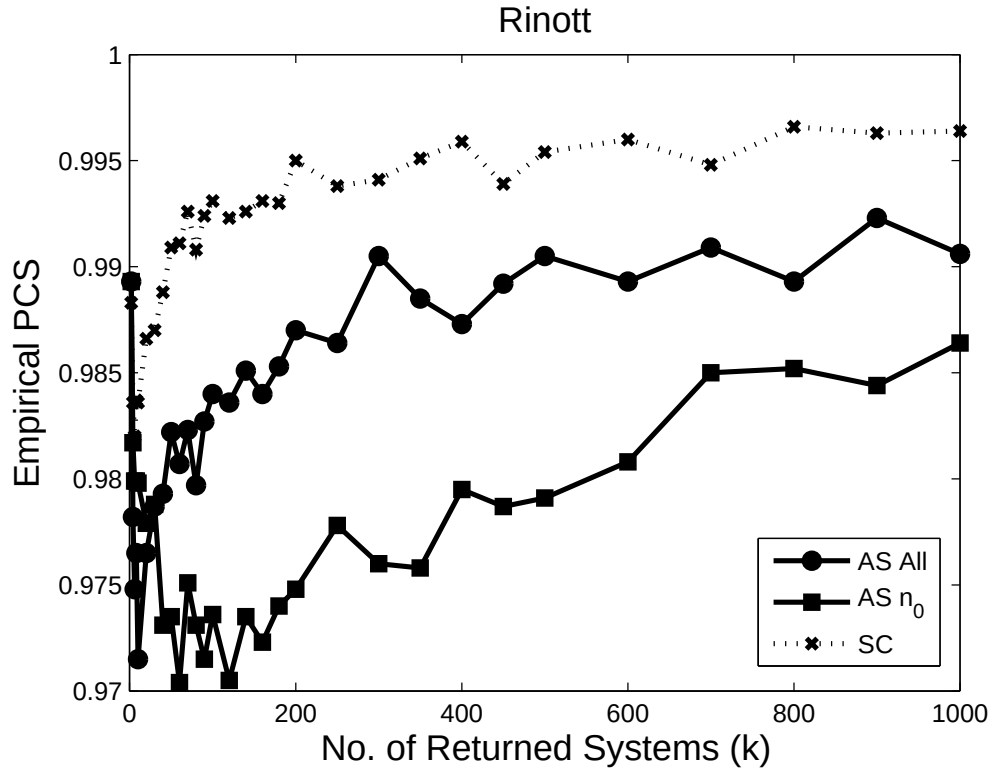

Figure 4.2: Empirical PCS of the Rinott procedure in the slippage configuration (SC) and when applied after AS, reusing replications taken during search. Based on the usual normal confidence intervals, the empirical PCS of the Rinott procedure is accurate to within ±0.005.

The conservativeness of the Rinott procedure is evident in Figure 4.2 where the empirical PCS for the slippage configuration increases to above 0.995 as the number of returned systems increases to 1000. The empirical PCS for the Rinott procedure applied after AS initially decreases until around 50 to 100 returned systems before increasing again. This trend suggests that the conservativeness

of the Rinott procedure offsets most of the impact of AS on decreasing PCS. Although the empirical PCS for both cases of the Rinott procedure applied after AS never drops below the desired PCS of 0.95, it stays below the empirical PCS of the slippage configuration.

4.4.2 Subset-Selection Procedures

We also test subset-selection procedures, which were first developed by Gupta (1965) as an alternative to the indifference-zone formulation of Bechhofer (1954). Since we wish to handle the cases of correct selection and good selection simultaneously, we will rely on variations of the Gupta procedure that provide PCS guarantees when the indifference-zone assumption is satisfied. That is, they deliver the guarantee

$$\mathbb{P}_{(\mu,\Sigma)}\{x_{[k]} \in I\} \geq 1 - \alpha \quad \text{for all } \mu \in \text{PZ}(\delta).$$

Two such procedures are as follows:

Modified Gupta A modified version of the original Gupta procedure for common, known variance σ^2 that uses a different "yardstick" based on the indifference-zone assumption. Take n_0 replications from each system and return the set of systems

$$I = \{X_i : \bar{Y}(X_i) \geq \bar{Y}(X_j) - (W - \delta)^+ \text{ for all } j \neq i\},$$

where $W = h_B \sigma \sqrt{2/n_0}$, and h_B is Bechhofer's constant, as mentioned in Section 4.4.1. This procedure was developed by van der Laan (1992), but the original version did not have the necessary positive-part operator in the

term $(W-\delta)^+$. A complete proof of the procedure's PCS guarantee is given in the appendix.

Screen-to-the-Best An extension of Gupta's procedure to handle unknown and uncommon variances developed by Nelson et al. (2001). Take n_0 replications from each system and return the set of systems

$$I = \{X_i : \bar{Y}(X_i) \geq \bar{Y}(X_j) - (W_{ij} - \delta)^+ \text{ for all } j \neq i\},$$

where

$$W_{ij} = t\left(\frac{S_i^2}{n_0} + \frac{S_j^2}{n_0}\right)^{1/2}$$

and $t = t_{(1-\alpha)^{1/(k-1)},n_0-1}$ is the upper $(1-\alpha)^{1/(k-1)}$ quantile of a t distribution with $n_0 - 1$ degrees of freedom. The Screen-to-the-Best procedure was further extended to unequal sample sizes (Boesel et al., 2003b), but we do not consider the extension here.

The Modified Gupta procedure is designed to be tight in the slippage configuration while the Screen-to-the-Best procedure is more conservative.

As seen in Figure 4.3, the empirical PCS for the Modified Gupta procedure after AS quickly drops below the desired PCS of 0.95. The empirical PCS further deteriorates as the number of systems increases, even to the point where PCS is only about half of its guaranteed value!

In Figure 4.4 the empirical PCS of the Screen-to-the-Best procedure after AS stays well below the empirical PCS of the slippage configuration, dropping to just above its guarantee of 0.95 at around 100 systems, before increasing again. Nevertheless, we suspect that there exist problem instances and parameter set-

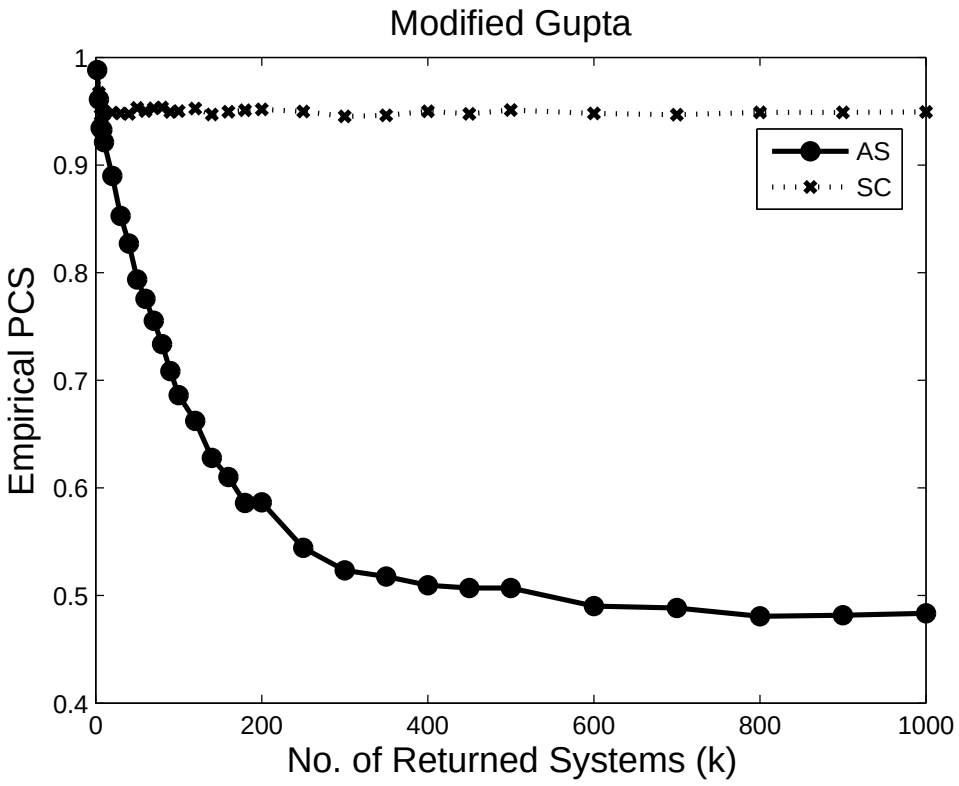

Figure 4.3: Empirical PCS of the Modified Gupta procedure in the slippage configuration (SC) and when applied after AS, reusing replications taken during search. Based on the usual normal confidence intervals, the empirical PCS of the Modified Gupta is accurate to within ±0.01.

tings for which the PCS for Screen-to-the-Best will fall below its guaranteed level.

4.4.3 A Realistic Search

The AS framework in Sections 4.4.1 and 4.4.2 is clearly unrealistic, yet effective at showing that R&S guarantees can suffer from reusing search data. We now consider a more-realistic optimization problem and search method and study the performance of a subset-selection procedure that reuses the replications from search. The optimization problem we consider is to maximize

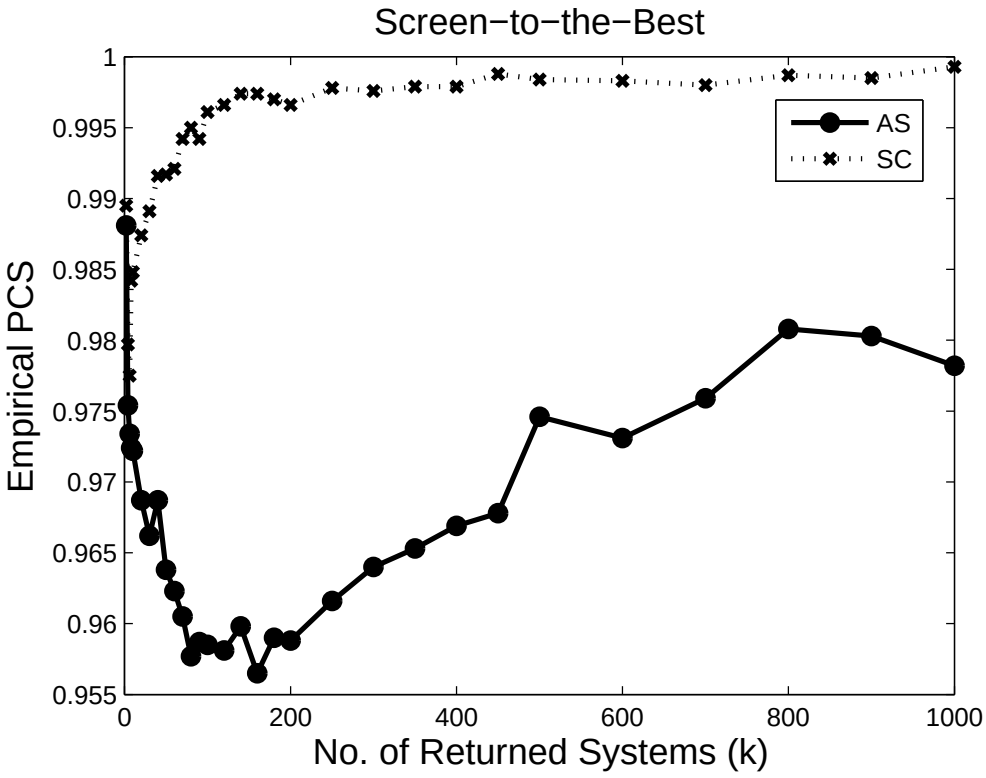

Figure 4.4: Empirical PCS of the Screen-to-the-Best procedure in the slippage configuration (SC) and when applied after AS, reusing replications taken during search. Based on the usual normal confidence intervals, the empirical PCS of the Screen-to-the-Best procedure is accurate to within ±0.01.

$\mu(x) = \lceil \log_2 x \rceil$ on the interval $[1/16, 16]$, a step-like function with ever-widening steps whose values range on the integers from -4 to 4, as plotted in Figure 4.5. We assume that replications taken at a point x are normally distributed with mean $g(x)$ and variance 1.

The search method we consider first evaluates the system $x_1 = 0.75$ by taking $n_0 = 10$ samples. Each new system is chosen uniformly at random from an interval of width 2 centered around the system with the highest sample mean among the explored systems, but otherwise independently of past observations. When the search attempts to return a system from outside the interval $[1/16, 16]$, the nearest endpoint is returned as the next system. In this way, the search finds

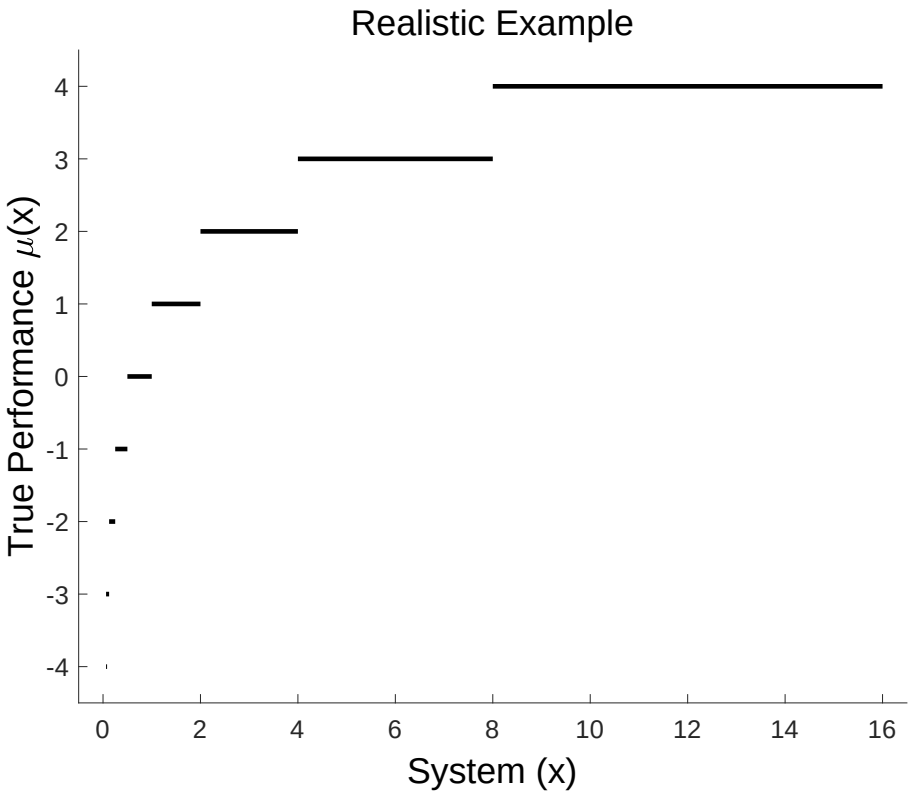

Figure 4.5: The true objective function $\mu(x) = \lceil \log_2 x \rceil$.

new systems from within a neighborhood of the system with the best observed performance.

We tested the Modified Gupta subset-selection procedure for 100 000 macroreplications and calculated the overall PGS and PCS of Guarantees (4.3) and (4.4), respectively, at a range of values of k. The empirical PCS of Guarantee (4.4) was averaged over only the macroreplications for which the configuration of the returned systems was in the preference zone. For $k < 10$, this amounted to roughly 1 in 4 macroreplications, while for $k = 100$, it was roughly 1 in 100. As seen in Figure 4.6 the empirical PGS remains well above its guaranteed level of 0.95, while the empirical PCS in the preference zone falls below this threshold.

The experimental results suggest that the event of the returned configuration

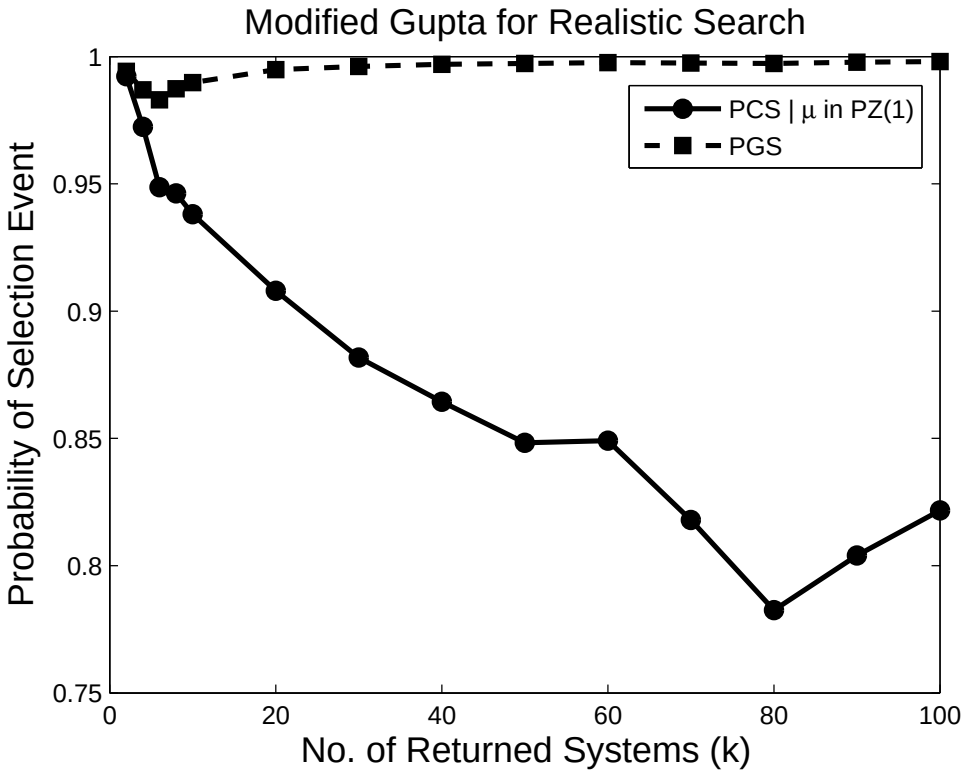

Figure 4.6: The empirical PCS and PGS of the Modified Gupta procedure when applied after AS, reusing replications taken during search. Based on the usual normal confidence intervals, the empirical PCS of the Modified Gupta procedure is accurate to within ±0.003 for $k \leq 10$ and accurate to within ±0.03 for $k \geq 80$.

being in the preference zone makes correct decisions less likely. And indeed, the structure of the objective function is responsible. When the search returns a configuration in the preference zone, it means that only one system was simulated from the highest explored "step" of the objective function. Because the search looks for new systems in an interval centered around the best-looking system, returning a configuration in the preference zone likely means the true best system did not have the highest sample mean immediately after it was evaluated. For this reason, the probability of correct selection suffers in these instances.

4.5 Conclusion

In this chapter, we consider the framework of applying a R&S procedure on a random set of systems returned by a sampling or search method. We formulate purposeful guarantees on PCS and PGS and showed how they can immediately follow from the traditional guarantees of R&S procedures. We then argue that, in this setting, PGS guarantees are superior to PCS guarantees based on the indifference-zone assumption.

We study issues that arise from reusing replications taken during search as input to a R&S procedure. We explain how search methods inherently produce replications that are conditionally dependent on the sequence of returned systems, an aspect that has been overlooked in the proofs of existing procedures that reuse search data. Furthermore, we design a search-like method that exploits this dependence to demonstrate how the PCS and PGS of R&S procedures can suffer as a consequence.

In our experiments, we observe empirical PCS well below the guaranteed level for R&S procedures with tight guarantees, e.g., Bechhofer and Modified Gupta. For more conservative procedures that handle unequal and unknown variances, e.g., Rinott and Screen-to-the-Best, we also observe lower-than-expected PCS, but not enough to violate the guarantees. These results suggest that for practical simulation optimization problems, the procedures proposed in the literature that reuse search replications can likely be performed with little fear of encountering violated guarantees. Since realistic simulation optimization searches return systems with ever-similar performance, we believe that PGS guarantees will in practice be robust to reusing search data.

How R&S procedures might be designed to safely reuse search data while offering provable guarantees remains an open question. One way forward may be procedures with asymptotic PCS and PGS guarantees, e.g., $\mathcal{KN}+$ and $\mathcal{KN}++$ of Kim and Nelson (2006a).

Because search methods can yield replications that are conditionally dependent *within* a given system—as discussed in Section 4.2.1—our findings should extend to R&S procedures that make use of common random numbers, e.g., Nelson and Matejcik (1995). For similar reasons, one might expect green simulation estimators to be biased when the systems have been identified via search (Feng and Staum, 2017).

CHAPTER 5

CONCLUSION

Much of the R&S literature focuses on the design of procedures that efficiently deliver a statistical guarantee specified by a decision-maker. We instead direct our attention to the guarantees themselves. While R&S guarantees are usually taken for granted, we hope to convey the importance of carefully choosing the guarantee. The choice of guarantee determines how a procedure takes observations, how it uses those observations to make a selection, and what kinds of assurances are provided to decision-makers regarding the quality of their selection.

In examining R&S guarantees, we observe that prevalent guarantees do not always align well with a decision-maker's goals. For instance, the IZ-inspired PCS guarantee is conditional on an untenable assumption about the unknown problem instance. In addition, the posterior PCS guarantee can cause a procedure to take many observations to detect differences in performances that are practically insignificant.

We advocate for the frequentist PGS and Bayesian posterior PGS guarantees which accommodate the decision-maker's tolerance with respect to a system's performance. For the PGS guarantee, there is an opportunity to design procedures that sequentially eliminate systems while performing efficiently in the indifference zone. Bootstrapping and other approaches that asymptotically deliver the PGS guarantee also merit further study. For the posterior PGS guarantee, evaluating the posterior PGS for problems with an extremely large number of systems remains computationally challenging.

In both the frequentist and Bayesian treatments, subset selection—returning a subset of system, rather than a single system—deserves greater attention. Subset selection is used for a variety of purposes, such as screening out inferior systems before running a selection procedure, or for instances in which the number of alternatives is large, but the available simulation budget is small. An intriguing problem in this space is identifying the smallest subset that contains a near-optimal system with high posterior probability. Evaluating this posterior probability gives rise to its own interesting computational considerations.

Another fascinating research question that we reopen is how to deliver frequentist R&S guarantees when a procedure reuses search observations. The answer could have profound implications for procedures that sequentially select systems based on the performances of others. Integrating R&S and search in this way—perhaps while exploiting structural information about the objective function—could dramatically strengthen the delivered guarantees.

APPENDIX A

PROOFS OF CHAPTER 2

A.1 Proof of Proposition 1

Proof. We first prove that Procedure 1 achieves Goal PCS-PZ.

We modify a proof for the procedure of Bechhofer (1954), shown in Kim and Nelson (2006b). For any arbitrary $\mu \in \text{PZ}(\delta)$,

$$
\begin{aligned}
\mathbb{P}_\mu(\text{CS}) &= \mathbb{P}_\mu(\text{Select } [k]) \\
&\geq \mathbb{P}_\mu\left(\frac{Y_{[k]} - Y_i}{\sigma\sqrt{2/n}} > r \text{ for all } i \neq [k]\right) \\
&= \mathbb{P}_\mu\left(\frac{Y_i - Y_{[k]} - (\mu_i - \mu_{[k]})}{\sigma\sqrt{2/n}} < -r - \frac{\mu_i - \mu_{[k]}}{\sigma\sqrt{2/n}} \text{ for all } i \neq [k]\right) \\
&\geq \mathbb{P}_\mu\left(\frac{Y_i - Y_{[k]} - (\mu_i - \mu_{[k]})}{\sigma\sqrt{2/n}} < -r + \frac{\delta}{\sigma\sqrt{2/n}} \text{ for all } i \neq [k]\right) \\
&= \mathbb{P}\left(Z_i < -r + \frac{\delta}{\sigma\sqrt{2/n}} \text{ for all } i \neq [k]\right) \\
&\geq \mathbb{P}(Z_i < h_B \quad \text{for all } i \neq [k]) \\
&= 1 - \alpha,
\end{aligned}
$$

where $(Z_i : i \neq [k])$ have a joint multivariate normal distribution with means 0, variances 1, and common pairwise correlations 1/2. The first inequality comes from the fact that System $[k]$ will be selected if $Y_{[k]} > \max_{i \neq [k]} Y_i + r\sigma\sqrt{2/n}$. The second inequality comes from the fact that $\mu_{[k]} - \mu_i \geq \delta$ since $\mu \in \text{PZ}(\delta)$. The last inequality follows from the relationship between r, n and h_B and the last equality follows from the definition of h_B by Bechhofer (1954).

We next show that Procedure 1 does not achieve Goal PGS for $k > 2$.

Consider the configuration $\tilde{\mu}$ defined by $\tilde{\mu}_k = \tilde{\mu}_{k-1}$ and $\tilde{\mu}_i = \tilde{\mu}_k - \Delta$ for $i = 1, \ldots, k-2$ where $\Delta > \delta$. That is, Systems k and $k-1$ are tied as the best and Systems 1 through $k-2$ are all bad. From this construction, $\tilde{\mu} \in \text{IZ}(\delta)$.

$$
\begin{aligned}
1 - \mathbb{P}_{\tilde{\mu}}(\text{GS}) &= \mathbb{P}_{\tilde{\mu}}(\text{Select neither } k-1 \text{ nor } k) \\
&\geq \mathbb{P}_{\tilde{\mu}}\left(\frac{|Y_k - Y_{k-1}|}{\sigma\sqrt{2/n}} < r \text{ and } \max(Y_1, \ldots, Y_{k-2}) < \min(Y_{k-1}, Y_k) \right) \\
&\geq 1 - \mathbb{P}_{\tilde{\mu}}\left(\frac{|Y_k - Y_{k-1}|}{\sigma\sqrt{2/n}} \geq r \right) - 2(k-2)\mathbb{P}_{\tilde{\mu}}(Y_1 > Y_k) \\
&= 1 - 2\Phi(-r) - 2(k-2)\mathbb{P}_{\tilde{\mu}}\left(\frac{Y_k - Y_1 - (\mu_k - \mu_1)}{\sigma\sqrt{2/n}} < -\frac{\mu_k - \mu_1}{\sigma\sqrt{2/n}} \right) \\
&= 1 - 2\Phi(-r) - 2(k-2)\Phi\left(\frac{-\Delta}{\sigma\sqrt{2/n}} \right).
\end{aligned}
$$

The first inequality follows from the fact that one of the systems 1 through $k-2$ will be selected if systems $k-1$ and k have the two highest estimators and they are within $r\sigma\sqrt{2/n}$ of each other. The second inequality follows from applying Boole's inequality over the intersection of $\{|Y_k - Y_{k-1}| < r\sigma\sqrt{2/n}\}$ and the $2(k-2)$ events contained in $\{\max(Y_1, \ldots Y_{k-2}) \leq \min(Y_{k-1}, Y_k)\}$.

We can now take r so large that $2\Phi(-r)$ is as small as desired. Fixing r also fixes the sample size n. We can then take Δ so large that $2(k-2)\Phi(-\Delta/(\sigma\sqrt{2/n}))$ is as small as desired. Thus we can make $1 - \mathbb{P}_{\tilde{\mu}}(\text{GS})$ arbitrarily close to 1, while retaining Goal PCS-PZ. □

A.2 Proof of Proposition 2

Proof. We first prove that Procedure 2 achieves Goal PCS-PZ.

Fix an arbitrary configuration $\mu \in \text{PZ}(\delta)$ and consider the event

$$A = \{|Y_i - \mu_i| \leq \delta/4 \ \ \forall i = 1, \ldots, k\},$$

where the first-stage estimators Y_i are distributed as $Y_i \sim \mathcal{N}(\mu_i, \sigma^2/n_0)$. On the event A, $(k) = [k]$, $(k-1) = j$ for some $j \neq [k]$, and for these two systems,

$$Y_{(k)} - Y_{(k-1)} = Y_{[k]} - Y_j = (Y_{[k]} - \mu_{[k]}) - (Y_j - \mu_j) + (\mu_{[k]} - \mu_j) \geq -\delta/4 - \delta/4 + \delta = \delta/2.$$

Therefore $A \subseteq \{Y_{(k)} - Y_{(k-1)} \geq \delta/2\}$.

By the independence of the first-stage estimators,

$$\begin{aligned}
\mathbb{P}_\mu(A) &= \mathbb{P}_\mu\left(|Y_i - \mu_i| \leq \delta/4 \ \ \forall i = 1, \ldots, k\right) \\
&= \prod_{i=1}^{k} \mathbb{P}_\mu\left(|Y_i - \mu_i| \leq \delta/4\right) \\
&= \mathbb{P}_\mu\left(\left|\mathcal{N}(0, \sigma^2/n_0)\right| \leq \delta/4\right)^k \\
&= \mathbb{P}_\mu\left(|\mathcal{N}(0,1)| \leq \frac{\delta}{4\sigma/\sqrt{n_0}}\right)^k \\
&= \left(1 - 2\Phi\left(\frac{-\delta}{4\sigma/\sqrt{n_0}}\right)\right)^k.
\end{aligned}$$

From the choice of n_0,

$$\begin{aligned}
\mathbb{P}_\mu(A) &= \left(1 - 2\Phi\left(\frac{-\delta}{4\sigma/\sqrt{n_0}}\right)\right)^k \\
&\geq \left(1 - 2\Phi\left(\frac{-\delta}{4\sigma}\left(\frac{4\sigma}{\delta}\Phi^{-1}\left(\frac{1-(1-\alpha)^{1/(2k)}}{2}\right)\right)\right)\right)^k \\
&= \left(1 - 2\left(1 - \Phi\left(\Phi^{-1}\left(\frac{1-(1-\alpha)^{1/(2k)}}{2}\right)\right)\right)\right)^k \\
&= \left(1 - 2\left(\frac{1+(1-\alpha)^{1/(2k)}}{2}\right)\right)^k \\
&= \sqrt{1-\alpha}.
\end{aligned}$$

Therefore $\mathbb{P}_\mu(Y_{(k)} - Y_{(k-1)} \geq \delta/2) \geq \mathbb{P}_\mu(A) \geq \sqrt{1-\alpha}$.

When $Y_{(k)} - Y_{(k-1)} \geq \delta/2$, the first-stage observations are ignored and Bechhofer's procedure is run with the confidence parameter $\sqrt{1-\alpha}$. Since Bechhofer's procedure delivers Goal PCS-PZ, $\mathbb{P}_\mu(\text{CS} \mid Y_{(k)} - Y_{(k-1)} \geq \delta/2) \geq \sqrt{1-\alpha}$.

All together,

$$\begin{aligned}\mathbb{P}_\mu(\text{CS}) &= \mathbb{P}_\mu(\text{CS} \mid Y_{(k)} - Y_{(k-1)} \geq \delta/2)\,\mathbb{P}_\mu(Y_{(k)} - Y_{(k-1)} \geq \delta/2) \\ &\quad + \mathbb{P}_\mu(\text{CS} \mid Y_{(k)} - Y_{(k-1)} < \delta/2)\,\mathbb{P}_\mu(Y_{(k)} - Y_{(k-1)} < \delta/2) \\ &\geq \left(\sqrt{1-\alpha}\right)\left(\sqrt{1-\alpha}\right) = 1 - \alpha.\end{aligned}$$

We next show that Procedure 2 does not achieve Goal PGS for some values of α and k. Set $\alpha = 1/3$ and consider the configuration $\tilde{\mu}$ defined by $\tilde{\mu}_k = \tilde{\mu}_{k-1}$ and $\tilde{\mu}_i = \tilde{\mu}_k - \delta$ for $i = 1, \ldots, k-2$. On the event A, $\{(k), (k-1)\} = \{k, k-1\}$ and for these two systems,

$$Y_{(k)} - Y_{(k-1)} = |Y_k - Y_{k-1}| \leq |Y_k - \tilde{\mu}_k| + |Y_{k-1} - \tilde{\mu}_{k-1}| + |\tilde{\mu}_k - \tilde{\mu}_{k-1}| \leq \delta/4 + \delta/4 + 0 = \delta/2.$$

Since the event $\{Y_{(k)} - Y_{(k-1)} = \delta/2\}$ occurs with probability zero,

$$\mathbb{P}_{\tilde{\mu}}(Y_{(k)} - Y_{(k-1)} \geq \delta/2) \leq \mathbb{P}_{\tilde{\mu}}(A^c) \leq 1 - \sqrt{1-\alpha} \leq \frac{1}{3}.$$

Using simulation, one can find a large enough value of k such that when taking only a single observation from each system,

$$\mathbb{P}_{\tilde{\mu}}(\max\{Y_{k-1}, Y_k\} > \max\{Y_1, \ldots, Y_{k-2}\}) < \frac{1}{3},$$

which implies that for this choice of k, $\mathbb{P}_{\tilde{\mu}}(\text{GS} \mid Y_{(k)} - Y_{(k-1)} < \delta/2) < 1/3$.

All together,

$$\begin{aligned}\mathbb{P}_{\tilde{\mu}}(\text{GS}) &= \mathbb{P}_{\tilde{\mu}}(\text{GS} \mid Y_{(k)} - Y_{(k-1)} \geq \delta/2)\,\mathbb{P}_{\tilde{\mu}}(Y_{(k)} - Y_{(k-1)} \geq \delta/2) \\ &\quad + \mathbb{P}_{\tilde{\mu}}(\text{GS} \mid Y_{(k)} - Y_{(k-1)} < \delta/2)\,\mathbb{P}_{\tilde{\mu}}(Y_{(k)} - Y_{(k-1)} < \delta/2) \\ &< (1)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)(1) = \frac{2}{3} = 1 - \alpha.\end{aligned}$$

□

A.3 Proof of Corollary 1

We prove the result of Corollary 1 for each of the four conditions.

Direct proofs that selection procedures achieving Goal PCS-PZ and satisfying either Condition (C2) or (C3) also achieve Goal PGS can be found in those of Lemmas 2 and 1, respectively, of Guiard (1996). Instead, we show that Conditions (C2) and (C3) each imply Condition (C1).

A.3.1 Proof that Condition (C2) implies Condition (C1)

Proof. Fix an arbitrary subset A and configurations μ and $\tilde{\mu}$ as specified in the statement of Condition (C1). Fix an arbitrary $i \in A$ and set $B_1 = \{i\}$, $B_2 = A\backslash\{i\}$, and IP $= \{(i, j) : j \in B_2\}$, i.e. IP $= B_1 \times B_2$. Then for all index pairs (i, j) in IP, $\mu_i - \mu_j = \tilde{\mu}_i - \tilde{\mu}_j$ since $i, j \in A$. Thus by Condition (C2),

$$\mathbb{P}_\mu(Y_i > Y_j \text{ for all } j \in A\backslash\{i\}) \geq \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j \text{ for all } j \in A\backslash\{i\}).$$

Since the choice of $i \in A$ was arbitrary, we simultaneously have that

$$\mathbb{P}_\mu(Y_i = \max_{j\in A} Y_j) \geq \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i = \max_{j\in A} \tilde{Y}_j), \quad \text{for all } i \in A. \tag{A.1}$$

where we have used the fact that ties in the estimators Y_j occur with probability zero. Summing both sides of (A.1) over all $i \in A$ gives $1 = 1$. Thus it must be the case that

$$\mathbb{P}_\mu(Y_i > Y_j \text{ for all } j \in A\backslash\{i\}) = \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j \text{ for all } j \in A\backslash\{i\}) \quad \text{for all } i \in A.$$

Since the choices of A, μ, and $\tilde{\mu}$ were arbitrary, we have proven the result. □

A.3.2 Proof that Condition (C3) implies Condition (C1)

Proof. Fix an arbitrary subset A and configurations μ and $\tilde{\mu}$ as specified in the statement of Condition (C1). Take $S = A$. Since μ and $\tilde{\mu}$ only differ for indices $i \notin A$, Condition (C3) implies that $Y_A \stackrel{d}{=} \tilde{Y}_A$ where Y_A (respectively $\tilde{Y}_A$) denotes the vector of estimators Y_i ($\tilde{Y}_i$) for $i \in A$. Therefore

$$\mathbb{P}_\mu(Y_i > Y_j \text{ for all } j \in A\backslash\{i\}) = \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j \text{ for all } j \in A\backslash\{i\}) \quad \text{for all } i \in A.$$

Since the choices of A, μ, and $\tilde{\mu}$ were arbitrary, we have proven the result. □

A.3.3 Proof that Condition (C4) implies Condition (C1)

Proof. From Condition (C4), the estimators $Y_1, \ldots, Y_k$ are mutually independent and so the joint distribution of the estimators is the product of the marginal distributions. Therefore for an arbitrary subset $A \subseteq \{1, \ldots, k\}$, the joint distribution of Y_i for $i \in A$ is the product of the marginal distributions for Y_i for $i \in A$. The remainder of the proof follows from that of Condition (C3). □

A.3.4 Proof that Condition (C5) implies Condition (C1)

Proof. Fix an arbitrary subset A and configurations μ and $\tilde{\mu}$ as specified in the statement of Condition (C1). By Condition (C5), $Y_A - \mu_A \stackrel{d}{=} \tilde{Y}_A - \tilde{\mu}_A$ where μ_A is the vector of components μ_i for $i \in A$. Since $\mu_i = \tilde{\mu}_i$ for all $i \in A$, $\mu_A = \tilde{\mu}_A$ and so $Y_A \stackrel{d}{=} \tilde{Y}_A$. Therefore

$$\mathbb{P}_\mu(Y_i > Y_j \text{ for all } j \in A\backslash\{i\}) = \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j \text{ for all } j \in A\backslash\{i\}) \quad \text{for all } i \in A.$$

Since the choices of A, μ, and $\tilde{\mu}$ were arbitrary, we have proven the result. □

A.4 Relationships of Conditions in Corollary 1

We prove the four relationships in Figure 2.1 that were not proven in Corollary 1, namely (i) (C4) implies (C3), (ii) (C5) implies (C6), (iii) (C5) implies (C2), and (iv) (C5) implies (C3).

A.4.1 Proof that Condition (C4) implies Condition (C3)

Proof. From Condition (C4), the estimators $Y_1, \ldots, Y_k$ are mutually independent and so the joint distribution of the estimators is the product of the marginal distributions. Therefore for an arbitrary subset $A \subset \{1, \ldots, k\}$, the joint distribution of Y_i for $i \in A$ is the product of the marginal distributions for Y_i for $i \in A$. This joint distribution thus does not depend on μ_j for $j \notin A$, hence Condition (C3) is satisfied. □

A.4.2 Proof that Condition (C5) implies Condition (C6)

Proof. Fix an arbitrary configuration μ and arbitrary system $i \in \{1, \ldots, k\}$. For arbitrary $\ell \neq i$, define $\tilde{\mu} = \mu + \epsilon e_\ell$ for $\epsilon > 0$ where e_ℓ is k-vector of zeros with a one as the ℓth element. By Condition (C5), $Y + \epsilon e_\ell \stackrel{d}{=} \tilde{Y}$ where Y (respectively $\tilde{Y}$) is the complete vector of estimators Y_i ($\tilde{Y}_i$) for $i = 1, \ldots, k$. Thus

$$\mathbb{P}_{\tilde{\mu}}(\text{Select } i) = \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j \text{ for all } j \neq i) = \mathbb{P}_{\mu}(Y_i > Y_j \text{ for all } j \neq i, \ell \text{ and } Y_i > Y_\ell + \epsilon)$$

which is nonincreasing in ϵ. Since increasing ϵ is equivalent to increasing the true mean of the ℓth system, we have shown that $\mathbb{P}_{\mu}(\text{Select } i)$ is nonincreasing

in μ_ℓ for all $\ell \neq i$. Then since the choice of i was arbitrary, we have proven the result. □

A.4.3 Proof that Condition (C5) implies Condition (C2)

Proof. Define arbitrary B_1, B_2, and IP as in the statement of Condition (C2). Using the substitutions $Z = Y - \mu$ and $\tilde{Z} = \tilde{Y} - \tilde{\mu}$,

$$\begin{aligned}\mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j, \text{ for all } (i,j) \in \text{IP}) &= \mathbb{P}(\tilde{Z}_i + \tilde{\mu}_i > \tilde{Z}_j + \tilde{\mu}_j, \text{ for all } (i,j) \in \text{IP}) \\ &= \mathbb{P}(\tilde{Z}_i + (\tilde{\mu}_i - \tilde{\mu}_j) > \tilde{Z}_j, \text{ for all } (i,j) \in \text{IP}) \\ &\geq \mathbb{P}(\tilde{Z}_i + (\mu_i - \mu_j) > \tilde{Z}_j, \text{ for all } (i,j) \in \text{IP}),\end{aligned}$$

where the last inequality follows from the fact the $\tilde{\mu}_i - \tilde{\mu}_j \geq \mu_i - \mu_j$ for all $(i,j) \in \text{IP}$. From Condition (C5), $\tilde{Z} \stackrel{d}{=} Z$. Thus

$$\begin{aligned}\mathbb{P}(\tilde{Z}_i + (\mu_i - \mu_j) > \tilde{Z}_j, \text{ for all } (i,j) \in \text{IP}) &= \mathbb{P}(Z_i + (\mu_i - \mu_j) > Z_j, \text{ for all } (i,j) \in \text{IP}) \\ &= \mathbb{P}_\mu(Y_i > Y_j, \text{ for all } (i,j) \in \text{IP}).\end{aligned}$$

Putting everything together, we have

$$\mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i > \tilde{Y}_j, \text{ for all } (i,j) \in \text{IP}) \geq \mathbb{P}_\mu(Y_i > Y_j, \text{ for all } (i,j) \in \text{IP}).$$

Since the choice of B_1, B_2 and IP was arbitrary, Condition (C2) holds. □

A.4.4 Proof that Condition (C5) implies Condition (C3)

Proof. From Condition (C5), the random vector $Z = Y - \mu$ has a distribution $H(z)$ that does not depend on μ. Thus for an arbitrary set A, the distribution of $Z_A := (Z_i : i \in A)$ does not depend on any μ_j for $j \notin A$. Therefore Condition (C3) is satisfied. □

A.5 Proof of PGS Guarantee of Sobel and Huyett (1957) Procedure

Sobel and Huyett (1957) presented tables for the common number of observations needed from each of k systems with Bernoulli-distributed rewards in order to select the system with the highest success probability μ_i with high probability. The procedure takes n observations from each system, calculates the average number of successes, i.e., $Y_i = n^{-1} \sum_{j=1}^{n} X_{ij}$, and selects the system with the highest estimator. In the event of ties, it selects at random from among the tied systems. The procedure is designed to deliver Goal PCS-PZ.

Systems are sampled independently and therefore the estimators are independent, implying that Condition (C4) is satisfied. Yet the result of Corollary 1 cannot be immediately applied since the sample means of Bernoulli observations are discrete random variables, violating Assumption 2. Most of the proof that Condition (C4) and Goal PCS-PZ imply Goal PGS can still be used in this setting, but the issue of tied estimators must be handled.

Proof. Let $Z = (Z_1, \ldots, Z_k)$ be a random permutation of $(1, \ldots, k)$ that is generated before the experiment. Thus Z is independent of the observations X_{ij} and the estimators Y_i. When there are multiple systems that are tied for having the largest estimator, the procedure will select the one having the highest Z_i among them. Because the permutation Z is chosen uniformly at random, this tie-breaking rule is equivalent to choosing uniformly from among the tied systems.

Fix an arbitrary configuration $\mu \in \mathrm{IZ}(\delta)$. Because of the restrictions that $\mu_i \in [0, 1]$ for all $i = 1, \ldots, k$, it must be the case that $\mu_{[k]} \geq \delta$, otherwise all systems

would be good. For this configuration, μ, define $\mathcal{G}$, $\mathcal{B}$, and μ^* accordingly. From the definition of the event of good selection for the procedure,

$$\begin{aligned}\mathbb{P}_\mu(\text{GS}) \geq \mathbb{P}_\mu\big(&\{Y_{[k]} > Y_i \text{ for all } i \in \mathcal{B}\} \\ &\cup \{Y_{[k]} \geq Y_i \text{ for all } i \in \mathcal{B} \text{ and } Y_{[k]} = Y_j \text{ for some } j \in \mathcal{B} \\ &\text{ and } Z_{[k]} > Z_i \text{ for all } i \in T([k]) \cap \mathcal{B}\}\big), \end{aligned} \tag{A.2}$$

where $T([k])$ denotes the set of systems other than system $[k]$ whose estimators are tied with $Y_{[k]}$, i.e., $T([k]) := \{i \neq [k] : Y_i = Y_{[k]}\}$.

The first term on the right-hand side of Inequality (A.2) is the event that System $[k]$ clearly outperforms all of the bad systems. Thus no matter the performance of the other good systems, a bad system will not be selected. The second term on the right-hand side of Inequality (A.2) is the event that the best system performs no worse than all the bad systems, ties at least one of them, and is preferred to all tying bad systems based on the tie-break ranking.

By Condition (C4), the joint distribution of the estimators Y_i for $i \in \mathcal{B} \cup \{[k]\}$ does not depend of the performances μ_j for $j \in \mathcal{G}\backslash\{[k]\}$. Consequently, the distribution of $T([k]) \cap \mathcal{B}$ also does not depend on the performances μ_j for $j \in \mathcal{G}\backslash\{[k]\}$. Therefore we can relate the probability of the event in Inequality (A.2) under configuration μ to that of a similar event under configuration μ^*:

$$\begin{aligned}&\mathbb{P}_\mu\big(\{Y_{[k]} > Y_i \text{ for all } i \in \mathcal{B}\} \\ &\quad \cup \{Y_{[k]} \geq Y_i \text{ for all } i \in \mathcal{B} \text{ and } Y_{[k]} = Y_j \text{ for some } j \in \mathcal{B} \\ &\quad \text{ and } Z_{[k]} > Z_i \text{ for all } i \in T([k]) \cap \mathcal{B}\}\big) \\ &= \mathbb{P}_{\mu^*}\Big(\{Y^*_{[k]} > Y^*_i \text{ for all } i \in \mathcal{B}\} \\ &\quad \cup \{Y^*_{[k]} \geq Y^*_i \text{ for all } i \in \mathcal{B} \text{ and } Y^*_{[k]} = Y^*_j \text{ for some } j \in \mathcal{B} \\ &\quad \text{ and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k]) \cap \mathcal{B}\}\Big), \end{aligned}$$

where Y_i^* is the estimator of System i under configuration μ^* and $T^*([k])$ is the random set of systems tied with System $[k]$ under configuration μ^*. Here the index $[k]$ is still with respect to the systems' performances in configuration μ and not μ^*.

In addition,

$$\begin{aligned}
&\mathbb{P}_{\mu^*}\Big(\{Y_{[k]}^* > Y_i^* \text{ for all } i \in \mathcal{B}\} \\
&\quad \cup \{Y_{[k]}^* \geq Y_i^* \text{ for all } i \in \mathcal{B} \text{ and } Y_{[k]}^* = Y_j^* \text{ for some } j \in \mathcal{B} \\
&\quad \text{and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k]) \cap \mathcal{B}\}) \\
&\geq \mathbb{P}_{\mu^*}\Big(\{Y_{[k]}^* > Y_i^* \text{ for all } i \neq [k]\} \\
&\quad \cup \{Y_{[k]}^* \geq Y_i^* \text{ for all } i \neq [k] \text{ and } Y_{[k]}^* = Y_j^* \text{ for some } j \neq [k] \\
&\quad \text{and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k])\}\Big). \qquad \text{(A.3)}
\end{aligned}$$

We justify Inequality (A.3) by showing that every outcome on the right-hand side is contained in the event on the left-hand side. For the first term on the right-hand side of Inequality (A.3),

$$\{Y_{[k]}^* > Y_i^* \text{ for all } i \neq [k]\} \subseteq \{Y_{[k]}^* > Y_i^* \text{ for all } i \in \mathcal{B}\}.$$

For the second term on the right-hand side of Inequality (A.3), we separate outcomes into two cases. If $Y_{[k]}^* = Y_j^*$ for some $j \in \mathcal{B}$, then the outcome belongs to the event

$$\begin{aligned}
&\{Y_{[k]}^* \geq Y_i^* \text{ for all } i \in \mathcal{B} \text{ and } Y_{[k]}^* = Y_j^* \text{ for some } j \in \mathcal{B} \\
&\qquad \text{and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k]) \cap \mathcal{B}\}.
\end{aligned}$$

If instead $Y_{[k]}^* \neq Y_j^*$ for any $j \in \mathcal{B}$, then the outcome belongs to the event

$$\{Y_{[k]}^* > Y_i^* \text{ for all } i \in \mathcal{B}\}.$$

Finally, from the definition of correct selection and Goal PCS-PZ,

$$\begin{aligned}
&\mathbb{P}_{\mu^*}\Big(\{Y^*_{[k]} > Y^*_i \text{ for all } i \neq [k]\} \\
&\quad \cup \{Y^*_{[k]} \geq Y^*_i \text{ for all } i \neq k \text{ and } Y^*_{[k]} = Y^*_j \text{ for some } j \neq [k] \\
&\qquad \text{and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k])\}\Big) \\
&= \mathbb{P}_{\mu^*}(\mathrm{CS}) \geq 1 - \alpha.
\end{aligned}$$

Altogether, we have shown that $\mathbb{P}_{\mu}(\mathrm{GS}) \geq 1 - \alpha$, i.e., the procedure of Sobel and Huyett (1957) achieves Goal PGS. □

A.6 Proof of Theorem 4

We prove the two conditions in Theorem 4 separately.

Proof. Proof of Condition (C8) in Theorem 4 for restricted subset-selection. Fix an arbitrary configuration μ and define the subsets $\mathcal{G}$ and $\mathcal{B}$ and the configuration μ^* accordingly. Then

$$\begin{aligned}
1 - \alpha \leq \mathbb{P}_{\mu^*}(\mathrm{CS}) &= \mathbb{P}_{\mu^*}(\mathrm{GS}) \\
&= \mathbb{P}_{\mu^*}\{Y^*_{[k]} \geq \max\{Y^*_{(k-m+1)}, Y^*_{(k)} - d\}\} \\
&\leq \mathbb{P}_{\mu^*}\{Y^*_{[k]} \geq \max\{Y^*_{\langle |\mathcal{B}|-m+2\rangle}, Y^*_{\langle |\mathcal{B}|+1\rangle} - d\}\},
\end{aligned}$$

where $Y^*_{\langle j\rangle}$ is the jth lowest among those for systems belonging to the subset $\mathcal{B} \cup \{[k]\}$. Taking $A = \mathcal{B} \cup \{[k]\}$ in Condition (C8),

$$\mathbb{P}_{\mu^*}\{Y^*_{[k]} \geq \max\{Y^*_{\langle |\mathcal{B}|-m+2\rangle}, Y^*_{\langle |\mathcal{B}|+1\rangle} - d\}\} = \mathbb{P}_{\mu}\{Y_{[k]} \geq \max\{Y_{\langle |\mathcal{B}|-m+2\rangle}, Y_{\langle |\mathcal{B}|+1\rangle} - d\}\}.$$

We now partition the event on the right-hand side to factor in the estimators Y_i for $i \in \mathcal{G}\backslash\{[k]\}$.

$$\begin{aligned}
&\{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|+1\rangle} - d\}\} \\
&\quad = \{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|+1\rangle} - d\} \text{ and } Y_{[k]} \geq Y_j \ \ \forall j \in \mathcal{G}\backslash\{[k]\}\} \\
&\qquad \cup \{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|+1\rangle} - d\} \text{ and } \exists j \in \mathcal{G}\backslash\{[k]\} \text{ s.t. } Y_j > Y_{[k]}\}. \qquad \text{(A.4)}
\end{aligned}$$

For the first event on the right-hand side of Equation (A.4),

$$\begin{aligned}
&\{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}+1|\rangle} - d\} \text{ and } Y_{[k]} \geq Y_j \ \ \forall j \in \mathcal{G}\backslash\{[k]\}\} \\
&\quad \subseteq \{Y_{[k]} \geq \max\{Y_{(k-m+1)}, Y_{(k)} - d\}\} = \{[k] \in I\}.
\end{aligned}$$

For the second event on the right-hand side of Equation (A.4), let $j' = \arg\max_{j \in \mathcal{G}\backslash\{[k]\}} Y_j$, the system in $\mathcal{G}\backslash\{[k]\}$ with the highest estimator. Then

$$\begin{aligned}
&\{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|+1\rangle} - d\} \text{ and } \exists j \in \mathcal{G}\backslash\{[k]\} \text{ s.t. } Y_j > Y_k\} \\
&\quad \subseteq \{Y_{j'} \geq \max\{Y_{(k-m+1)}, Y_{(k)} - d\}\} = \{j' \in I\}.
\end{aligned}$$

Since $j' \in \mathcal{G}\backslash\{[k]\}$, both events on the right-hand side of Equation (A.4) are contained in the event of good selection. Thus

$$\mathbb{P}_\mu\{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|\rangle} - d\}\} \leq \mathbb{P}_\mu(\text{GS}),$$

from which it follows that $\mathbb{P}_\mu(\text{GS}) \geq 1 - \alpha$ for every μ. □

Proof. Proof of Condition (C8) in Theorem 4 for pairwise comparisons. Fix an arbitrary configuration μ and define the subsets $\mathcal{G}$ and $\mathcal{B}$ and the configuration μ^* accordingly. Then

$$\begin{aligned}
1 - \alpha \leq \mathbb{P}_{\mu^*}(\text{CS}) &= \mathbb{P}_{\mu^*}(\text{GS}) \\
&= \mathbb{P}_{\mu^*}\{Y^*_{[k]} \geq Y^*_j - W^*_{[k]j} \text{ for all } j \neq [k]\} \\
&\leq \mathbb{P}_{\mu^*}\{Y^*_{[k]} \geq Y^*_j - W^*_{[k]j} \text{ for all } j \in \mathcal{B}\} \\
&= \mathbb{P}_{\mu}\{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B}\},
\end{aligned}$$

where the last equality follows from Condition (C8) with $A = \mathcal{B} \cup \{[k]\}$.

We now partition this last event based on the estimators Y_i and the terms $W_{[k]i}$ for $i \in \mathcal{G}\backslash\{[k]\}$:

$$\begin{aligned}
&\{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B}\} \\
&\quad = \{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B} \text{ and } Y_{[k]} \geq Y_i - W_{[k]i} \text{ for all } i \in \mathcal{G}\backslash\{[k]\}\} \\
&\qquad \cup \{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B} \text{ and } \exists i \in \mathcal{G}\backslash\{[k]\} \text{ s.t. } Y_{[k]} < Y_i - W_{[k]i}\}. \quad \text{(A.5)}
\end{aligned}$$

For the first event on the right-hand side of Equation (A.5),

$$\begin{aligned}
&\{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B} \text{ and } Y_{[k]} \geq Y_i - W_{[k]i} \text{ for all } i \in \mathcal{G}\backslash\{[k]\}\} \\
&\qquad = \{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \neq [k]\} \\
&\qquad = \{[k] \in I\}.
\end{aligned}$$

For the second event on the right-hand side of Equation (A.5), let $i' := \arg\max_{i \in \mathcal{G}\backslash\{[k]\}} \{Y_i : Y_{[k]} < Y_i - W_{[k]i}\}$, the index of the system with the highest estimator among those in $\mathcal{G}\backslash\{[k]\}$ that eliminate System $[k]$. Following an argument similar to that given by Nelson et al. (2001), we claim that there exists no other system in $\mathcal{G}\backslash\{[k]\}$ that eliminates System i'. Towards a contradiction, suppose there was such a System i''. To eliminate System i', its estimator $Y_{i''}$ would have to be greater than $Y_{i'}$. By the transitive property of eliminations, System i'' would also eliminate System $[k]$. But from how i' is defined, it has the highest estimator among the systems in $\mathcal{G}\backslash\{[k]\}$ that eliminate System $[k]$. Therefore System i' is not eliminated by any other $i \in \mathcal{G}\backslash\{[k]\}$. By a similar transitive argument, one can show that there exists no system in $\mathcal{B}$ that eliminates System i' since that system would also have eliminated System $[k]$. Therefore we can

conclude that System i' will be retained in the set I,

$$\{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B} \text{ and } \exists i \in \mathcal{G}\backslash\{[k]\} \text{ s.t. } Y_{[k]} < Y_i - W_{[k]i}\}$$
$$\subseteq \{\exists i' \in \mathcal{G}\backslash\{[k]\} \text{ s.t. } i' \in I\}$$

Both events on the right-hand side of Equation (A.5) are contained in the event of good selection. Altogether, we have

$$\mathbb{P}_\mu\{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B}\} \leq \mathbb{P}_\mu(\text{GS}),$$

from which it follows that $\mathbb{P}_\mu(\text{GS}) \geq 1 - \alpha$ for every μ. □

A.7 Proof of Theorem 5

Proof. Fix an arbitrary configuration μ and define the subsets $\mathcal{G}$ and $\mathcal{B}$ and the configuration μ^* accordingly. Then

$$1 - \alpha \leq \mathbb{P}_{\mu^*}(\text{CS}) = \mathbb{P}_{\mu^*}(\text{GS}) = \mathbb{P}_{\mu^*}\{[k] \in I\} \leq \mathbb{P}_{\mu^*}\{\exists i \in \mathcal{G} \text{ s.t. } i \in I\}.$$

Let $\mu^{(0)} := \mu$ and for $\ell = 1, \ldots, |\mathcal{G}| - 1$ recursively define the related configuration $\mu^{(\ell)}$ by $\mu^{(\ell)}_{[|\mathcal{B}|+\ell]} = \mu_{[k]} - \delta$ and $\mu^{(\ell)}_i = \mu^{(\ell-1)}_i$ for all $i \neq [|\mathcal{B}| + \ell]$. Repeatedly applying Condition (C10) with $A = \mathcal{G}$ yields a chain of inequalities,

$$\begin{aligned}
\mathbb{P}_{\mu^*}\{\exists i \in \mathcal{G} \text{ s.t. } i \in I\} &= \mathbb{P}_{\mu^{(|\mathcal{G}|-1)}}\{\exists i \in \mathcal{G} \text{ s.t. } i \in I\} \\
&\leq \mathbb{P}_{\mu^{(|\mathcal{G}|-2)}}\{\exists i \in \mathcal{G} \text{ s.t. } i \in I\} \\
&\leq \cdots \\
&\leq \mathbb{P}_{\mu^{(1)}}\{\exists i \in \mathcal{G} \text{ s.t. } i \in I\} \\
&\leq \mathbb{P}_{\mu}\{\exists i \in \mathcal{G} \text{ s.t. } i \in I\} \\
&= \mathbb{P}_{\mu}(\text{GS}).
\end{aligned}$$

from which it follows that $\mathbb{P}_\mu(\mathrm{GS}) \geq 1 - \alpha$ for every μ. □

APPENDIX B

PROOFS OF CHAPTER 3

B.1 Proof of Proposition 4

Let $Y := \max_{\ell \neq i,j} W_\ell$, the highest performance of the alternatives, excluding Alternatives i and j. From the definition of the pPGS:

$$\begin{aligned} \text{pPGS}_i &= \mathbb{P}(W_i \geq Y - \delta, W_i \geq W_j - \delta \mid \mathcal{E}) \\ &= \mathbb{P}(W_i \geq Y - \delta \mid W_i \geq W_j - \delta, \mathcal{E})\, \mathbb{P}(W_i \geq W_j - \delta \mid \mathcal{E}), \end{aligned} \tag{B.1}$$

$$\begin{aligned} \text{pPGS}_j &= \mathbb{P}(W_j \geq Y - \delta, W_j \geq W_i - \delta \mid \mathcal{E}) \\ &= \mathbb{P}(W_j \geq Y - \delta \mid W_j \geq W_i - \delta, \mathcal{E})\, \mathbb{P}(W_j \geq W_i - \delta \mid \mathcal{E}). \end{aligned} \tag{B.2}$$

We show that $\text{pPGS}_i < \text{pPGS}_j$ by proving that the two terms in Equation (B.1) are each less than their corresponding terms in Equation (B.2).

For the second terms,

$$\begin{aligned} \mathbb{P}(W_i \geq W_j - \delta \mid \mathcal{E}) = \mathbb{P}(W_j - W_i \leq \delta \mid \mathcal{E}) = \mathbb{P}(\mathcal{N}(\mu_j - \mu_i, \hat{\sigma}_i^2 + \hat{\sigma}_j^2) \leq \delta) = \Phi\left(\frac{\delta - \mu_j + \mu_i}{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}}\right) \\ \leq \Phi\left(\frac{\delta - \mu_i + \mu_j}{\sqrt{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}}\right) = \mathbb{P}(\mathcal{N}(\mu_i - \mu_j, \hat{\sigma}_i^2 + \hat{\sigma}_j^2) \leq \delta) = \mathbb{P}(W_i - W_j \leq \delta \mid \mathcal{E}) \\ = \mathbb{P}(W_j \geq W_i - \delta \mid \mathcal{E}). \end{aligned}$$

For the first terms, we condition on the value of the difference between W_i and W_j:

$$\mathbb{P}(W_i \geq Y - \delta \mid W_i \geq W_j - \delta, \mathcal{E}) = \int_{-\delta}^{\infty} \mathbb{P}(W_i \geq Y - \delta \mid W_i - W_j = z, \mathcal{E}) f_{W_i - W_j \mid \mathcal{E}}(z)\, dz, \tag{B.3}$$

where $f_{W_i - W_j | \mathcal{E}}(\cdot)$ is the pdf of $W_i - W_j \mid \mathcal{E}$. Relating Equation (B.3) to the integral that arises from the first term in Equation (B.2) involves several steps. First, we relate the pdf of $W_i - W_j \mid \mathcal{E}$ to that of $W_j - W_i \mid \mathcal{E}$ via a shift. Next, we show that $W_i \mid \{W_i - W_j = z, \mathcal{E}\}$ is stochastically monotone increasing in z. Finally, we show that for $z \geq 0$, $W_i \mid \{W_i - W_j = z, \mathcal{E}\}$ is first-order stochastically dominated by $W_j \mid \{W_j - W_i = z, \mathcal{E}\}$.

For the first step, $W_i - W_j \mid \mathcal{E} \sim \mathcal{N}(\mu_i - \mu_j, \hat{\sigma}_i^2 + \hat{\sigma}_j^2)$ and $W_j - W_i \mid \mathcal{E} \sim \mathcal{N}(\mu_j - \mu_i, \hat{\sigma}_i^2 + \hat{\sigma}_j^2)$. Thus the pdf of $W_j - W_i \mid \mathcal{E}$ is identical to that of $W_i - W_j \mid \mathcal{E}$, but shifted $2(\mu_j - \mu_i)$ to the right, i.e., $f_{W_i - W_j | \mathcal{E}}(z) = f_{W_j - W_i | \mathcal{E}}(z + 2(\mu_j - \mu_i))$.

By substitution into Equation (B.3),

$$\mathbb{P}(W_i \geq Y - \delta \mid W_i \geq W_j - \delta, \mathcal{E}) = \int_{-\delta}^{\infty} \mathbb{P}(W_i \geq Y - \delta \mid W_i - W_j = z, \mathcal{E}) f_{W_j - W_i | \mathcal{E}}(z + 2(\mu_j - \mu_i))\, dz, \tag{B.4}$$

For the second and third steps, we consider the joint distribution of W_i and $W_i - W_j$ given $\mathcal{E}$:

$$\begin{pmatrix} W_i \\ W_i - W_j \end{pmatrix} \sim \mathcal{MVN}\left(\begin{pmatrix} \mu_i \\ \mu_i - \mu_j \end{pmatrix}, \begin{pmatrix} \lambda_i^2 & \lambda_i^2 \\ \lambda_i^2 & \lambda_i^2 + \lambda_j^2 \end{pmatrix}\right) =: \mathcal{MVN}\left(\begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right).$$

The conditional means and variances of W_i given $W_i - W_j = z$ and $\mathcal{E}$ can be obtained from the conditional formula for the multivariate normal distribution:

$$\begin{aligned} \mathbb{E}[W_i \mid W_i - W_j = z, \mathcal{E}] &= \nu_1 + \Sigma_{12}\Sigma_{22}^{-1}(z - \nu_2) \\ &= \mu_i - \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}(z - (\mu_i - \mu_j)) \\ &= \frac{\hat{\sigma}_j^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}\mu_i + \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}\mu_j + \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}z, \end{aligned}$$

$$\begin{aligned}
\mathrm{Var}[W_i \mid W_i - W_j = z, \mathcal{E}] &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \\
&= \hat{\sigma}_i^2 - \hat{\sigma}_i^2 \left(\frac{1}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2} \right) \hat{\sigma}_i^2 \\
&= \hat{\sigma}_i^2 \left(1 - \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2} \right) \\
&= \frac{\hat{\sigma}_i^2 \hat{\sigma}_j^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}.
\end{aligned}$$

Hence

$$W_i \mid \{W_i - W_j = z, \mathcal{E}\} \sim \mathcal{N}\left(\frac{\hat{\sigma}_j^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}\mu_i + \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}\mu_j + \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2} z, \frac{\hat{\sigma}_i^2 \hat{\sigma}_j^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2} \right), \tag{B.5}$$

and similarly

$$W_j \mid \{W_j - W_i = z, \mathcal{E}\} \sim \mathcal{N}\left(\frac{\hat{\sigma}_j^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}\mu_i + \frac{\hat{\sigma}_i^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2}\mu_j + \frac{\hat{\sigma}_j^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2} z, \frac{\hat{\sigma}_i^2 \hat{\sigma}_j^2}{\hat{\sigma}_i^2 + \hat{\sigma}_j^2} \right). \tag{B.6}$$

Because the conditional distributions in Equations (B.5) and (B.6) have the same variance—which is not a function of z—and different means, two forms of stochastic dominance can be shown. First, we see from Equation (B.5) that $W_i \mid \{W_i - W_j = z, \mathcal{E}\}$ is stochastically monotone in z. In other words, for values z_1 and z_2 satisfying $z_1 < z_2$,

$$W_i \mid \{W_i - W_j = z_1, \mathcal{E}\} \leq_{st} W_i \mid \{W_i - W_j = z_2, \mathcal{E}\},$$

where $\leq_{st}$ denotes first-order stochastic dominance. This means that for any value y,

$$\mathbb{P}(W_i \geq y \mid W_i - W_j = z_1, \mathcal{E}) \leq \mathbb{P}(W_i \geq y \mid W_i - W_j = z_2, \mathcal{E}).$$

Thus for the random variable Y, which is independent of W_i and W_j,

$$\mathbb{P}(W_i \geq Y \mid W_i - W_j = z_1, \mathcal{E}) \leq \mathbb{P}(W_i \geq Y \mid W_i - W_j = z_2, \mathcal{E}).$$

Second, it follows from Equations (B.5) and (B.6) that for $z > 0$,

$$W_i \mid \{W_i - W_j = z, \mathcal{E}\} \leq_{st} W_j \mid \{W_j - W_i = z, \mathcal{E}\}.$$

Hence for $z > 0$,

$$\mathbb{P}(W_i \geq Y \mid W_i - W_j = z, \mathcal{E}) \leq \mathbb{P}(W_j \geq Y \mid W_j - W_i = z, \mathcal{E}).$$

Applying these two properties to Equation (B.4) yields,

$$\begin{aligned}
&\mathbb{P}(W_i \geq Y - \delta \mid W_i \geq W_j - \delta, \mathcal{E}) \\
&\quad = \int_{-\delta}^{\infty} \mathbb{P}(W_i \geq Y - \delta \mid W_i - W_j = z, \mathcal{E}) f_{W_j - W_i}(z + 2(\mu_j - \mu_i))\, dz \\
&\quad \leq \int_{-\delta}^{\infty} \mathbb{P}(W_i \geq Y - \delta \mid W_i - W_j = z + 2(\mu_j - \mu_i), \mathcal{E}) f_{W_j - W_i}(z + 2(\mu_j - \mu_i))\, dz \\
&\quad \leq \int_{-\delta}^{\infty} \mathbb{P}(W_j \geq Y - \delta \mid W_j - W_i = z + 2(\mu_j - \mu_i), \mathcal{E}) f_{W_j - W_i}(z + 2(\mu_j - \mu_i))\, dz.
\end{aligned} \tag{B.7}$$

By a change of variables and adding a positive term to the integral in Equation (B.7),

$$\begin{aligned}
\mathbb{P}(W_i \geq Y - \delta \mid W_i \geq W_j - \delta, \mathcal{E}) &\leq \int_{2(\mu_j - \mu_i) - \delta}^{\infty} \mathbb{P}(W_j \geq Y - \delta \mid W_j - W_i = z', \mathcal{E}) f_{W_j - W_i}(z')\, dz' \\
&< \int_{-\delta}^{\infty} \mathbb{P}(W_j \geq Y - \delta \mid W_j - W_i = z', \mathcal{E}) f_{W_j - W_i}(z')\, dz' \\
&= \mathbb{P}(W_j \geq Y - \delta \mid W_j \geq W_i - \delta, \mathcal{E}).
\end{aligned}$$

Having shown that $\mathbb{P}(W_i \geq W_j - \delta \mid \mathcal{E}) < \mathbb{P}(W_j \geq W_i - \delta \mid \mathcal{E})$ and

$$\mathbb{P}(W_i \geq Y - \delta \mid W_i \geq W_j - \delta, \mathcal{E}) < \mathbb{P}(W_j \geq Y - \delta \mid W_j \geq W_i - \delta, \mathcal{E}),$$

it follows from Equations (B.1) and (B.2) that $\text{pPGS}_i < \text{pPGS}_j$.

B.2 Proof of Proposition 5

To prove Proposition 5, we will make use of the following two results:

Lemma 1. *(Slepian, 1962) Let $Z_1, \ldots, Z_d$ be normally distributed random variables with* $\mathrm{Cov}(Z_i, Z_j) \geq 0$ *for all i, j. For any constants c_j, $j = 1, \ldots, d$,*

$$\mathbb{P}\left(\bigcap_{j=1}^{d}\{Z_j \geq c_j\}\right) \geq \prod_{i=1}^{d} \mathbb{P}(Z_j \geq c_j).$$

Lemma 2. *(Tamhane, 1977) Let $V_1, \ldots, V_k$ be independent random variables. For any nonnegative, real-valued functions $g_j(v_1, \ldots, v_k)$ for $j = 1, \ldots, d$ that are each nondecreasing in each of their arguments,*

$$\mathbb{E}\left[\prod_{j=1}^{d} g_j(V_1, \ldots, V_k)\right] \geq \prod_{j=1}^{d} \mathbb{E}[g_j(V_1, \ldots, V_k)].$$

We now prove Proposition 5.

Proof. Under Assumptions 4–7,

$$W_i \mid \{\lambda_i, \mathcal{E}\} \sim \mathcal{N}\left(\frac{\nu_i^0 \mu_i^0 + n_i \bar{x}_i}{\nu_i^0 + n_i}, \frac{1}{(\nu_i^0 + n_i)\lambda_i}\right),$$

and therefore

$$W_i - W_j \mid \{\lambda_i, \lambda_j, \mathcal{E}\} \sim \mathcal{N}\left(\frac{\nu_i^0 \mu_i^0 + n_i \bar{x}_i}{\nu_i^0 + n_i} - \frac{\nu_j^0 \mu_j^0 + n_j \bar{x}_j}{\nu_j^0 + n_j}, \frac{1}{(\nu_i^0 + n_i)\lambda_i} + \frac{1}{(\nu_j^0 + n_j)\lambda_j}\right).$$

From the definition of pPGS_i,

$$\mathrm{pPGS}_i = \mathbb{P}\left(\bigcap_{j \neq i}\{W_i - W_j \geq -\delta\} \mid \mathcal{E}\right).$$

Conditioning on the posterior precisions $\Lambda_1, \ldots, \Lambda_k$ allows Lemma 1 to be applied with $Z_j = W_i - W_j \mid \{\lambda_1, \ldots, \lambda_k, \mathcal{E}\}$:

$$\begin{aligned}\mathbb{P}\left(\bigcap_{j \neq i}\{W_i - W_j \geq -\delta\} \mid \mathcal{E}\right) &= \mathbb{E}\left[\mathbb{P}\left(\bigcap_{j \neq i}\{W_i - W_j \geq -\delta\} \mid \Lambda_1, \ldots, \Lambda_k, \mathcal{E}\right) \mid \mathcal{E}\right] \\ &\geq \mathbb{E}\left[\prod_{j \neq i} \mathbb{P}\left(W_i - W_j \geq -\delta \mid \Lambda_1, \ldots, \Lambda_k, \mathcal{E}\right) \mid \mathcal{E}\right]\end{aligned}$$

Since $\mu_i - \mu_j \geq -\delta$ for all $j \neq i$, the functions

$$g_j(\lambda_1, \ldots, \lambda_k) = \mathbb{P}\left(W_i - W_j \geq -\delta \mid \lambda_1, \ldots, \lambda_k, \mathcal{E}\right),$$

are nonnegative and nondecreasing in $\lambda_1, \ldots, \lambda_k$ for all $j \neq i$. By Lemma 2,

$$\begin{aligned}\mathbb{E}\left[\prod_{j\neq i} \mathbb{P}\left(W_i - W_j \geq -\delta \mid \Lambda_1, \ldots, \Lambda_k, \mathcal{E}\right) \mid \mathcal{E}\right] &\geq \prod_{j\neq i} \mathbb{E}\left[\mathbb{P}\left(W_i - W_j \geq -\delta \mid \Lambda_1, \ldots, \Lambda_k, \mathcal{E}\right) \mid \mathcal{E}\right] \\ &= \prod_{j\neq i} \mathbb{P}\left(W_i \geq W_j - \delta \mid \mathcal{E}\right).\end{aligned}$$

□

APPENDIX C

PROOFS OF CHAPTER 4

In this appendix we prove the validity of the Modified Gupta procedure, namely, that it satisfies

$$\mathbb{P}_{(\mu,\Sigma)}\{x_{[k]} \in I\} \geq 1 - \alpha \quad \text{for all } \mu \in \text{PZ}(\delta),$$

where Σ is a vector whose elements are all σ^2.

Fix an arbitrary set of systems $\{x_1, \ldots, x_k\}$ with configuration $\mu \in \text{PZ}(\delta)$ and a common, known variance $\sigma^2 < \infty$. The procedure takes n_0 i.i.d. replications $Y_1(x_i), \ldots, Y_{n_0}(x_i)$ for $i = 1, \ldots, k$ and computes the corresponding sample means $\bar{Y}(x_i)$ for $i = 1, \ldots, k$. The Modified Gupta procedure preserves a set of systems I such that for each system $i \in I$,

$$\bar{Y}(x_i) \geq \max_{j \neq i} \bar{Y}(x_j) - (W - \delta)^+,$$

where $W = h_B \sigma \sqrt{2/n_0}$ and h_B is Bechhofer's constant. Note that the positive part operator on the term $(W - \delta)^+$ is needed to ensure that at least one system—the one with the largest sample mean—is included in the set I.

We follow the proof of Gupta's procedure presented by Kim and Nelson

(2006b) and modify where necessary:

$$
\begin{aligned}
\mathbb{P}_{(\mu,\Sigma)}\{x_{[k]} \in I\} &= \mathbb{P}_{(\mu,\Sigma)}\left\{\bar{Y}(x_{[k]}) \geq \max_{j\neq k} \bar{Y}(x_{[j]}) - (W-\delta)^+\right\} \\
&= \mathbb{P}_{(\mu,\Sigma)}\left\{\bar{Y}(x_{[k]}) \geq \max_{j\neq k} \bar{Y}(x_{[j]}) - \left(h_B\sigma\sqrt{2/n_0} - \delta\right)^+\right\} \\
&= \mathbb{P}_{(\mu,\Sigma)}\left\{\bar{Y}(x_{[k]}) \geq \bar{Y}(x_{[j]}) - \left(h_B\sigma\sqrt{2/n_0} - \delta\right)^+ \text{ for all } j\neq k\right\} \\
&= \mathbb{P}_{(\mu,\Sigma)}\left\{\frac{\bar{Y}(x_{[j]}) - \bar{Y}(x_{[k]}) - \left(\mu(x_{[j]}) - \mu(x_{[k]})\right)}{\sigma\sqrt{2/n_0}}\right. \\
&\qquad \left.\leq \left(h_B - \frac{\delta}{\sigma\sqrt{2/n_0}}\right)^+ - \frac{\left(\mu(x_{[j]}) - \mu(x_{[k]})\right)}{\sigma\sqrt{2/n_0}} \text{ for all } j\neq k\right\} \\
&\geq \mathbb{P}_{(\mu,\Sigma)}\left\{\frac{\bar{Y}(x_{[j]}) - \bar{Y}(x_{[k]}) - \left(\mu(x_{[j]}) - \mu(x_{[k]})\right)}{\sigma\sqrt{2/n_0}}\right. \\
&\qquad \left.\leq h_B - \frac{\delta + \left(\mu(x_{[j]}) - \mu(x_{[k]})\right)}{\sigma\sqrt{2/n_0}} \text{ for all } j\neq k\right\} \\
&\geq \mathbb{P}_{(\mu,\Sigma)}\left\{Z_j \leq h_B \text{ for all } j\neq k\right\} \\
&= 1-\alpha\,,
\end{aligned}
$$

where the variables $Z_1,\ldots,Z_{k-1}$ are distributed according to a multivariate normal with means 0, variances 1, and common pairwise correlations $1/2$. The last equality follows from the definition of Bechhofer's constant h_B.

BIBLIOGRAPHY

Ahmed, M. A. and Alkhamis, T. M. (2002). Simulation-based optimization using simulated annealing with ranking and selection. *Computers and Operations Research*, 29(4):387–402.

Ahn, D. and Kim, K.-K. (2018). Efficient simulation for expectations over the union of half-spaces. *ACM Transactions on Modeling and Computer Simulation*, 28(3):Article 23. 1–20.

Alam, K. and Rizvi, M. H. (1966). Selection from multivariate normal populations. *Annals of the Institute of Statistical Mathematics*, 18(1):307–318.

Andradottir, S. and Bier, V. (2000). Applying Bayesian ideas in simulation. *Simulation Practice and Theory*, 8(3-4):253–280.

Andradóttir, S. and Kim, S.-H. (2010). Fully sequential procedures for comparing constrained systems via simulation. *Naval Research Logistics*, 57(5):403–421.

Asmussen, S. and Glynn, P. W. (2007). *Stochastic simulation: algorithms and analysis*, volume 57. Springer Science & Business Media.

Audibert, J.-Y., Bubeck, S., and Munos, R. (2010). Best arm identification in multi-armed bandits. In Kalai, A. and Mohri, M., editors, *Proceedings of the 23rd Conference on Learning Theory (COLT)*, pages 41–53.

Bechhofer, R. E. (1954). A single-sample multiple decision procedure for ranking means of normal populations with known variances. *Annals of Mathematical Statistics*, 25(1):16–39.

Bechhofer, R. E., Kiefer, J., and Sobel, M. (1968). *Sequential Identification and Ranking Procedures: With Special Reference to Koopman-Darmois Populations*. University of Chicago Press, Chicago.

Bechhofer, R. E., Santner, T. J., and Goldsman, D. M. (1995). *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*. Wiley, New York.

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

Berger, J. O. and Deely, J. (1988). A Bayesian approach to ranking and selection of related means with alternatives to analysis-of-variance methodology. *Journal of the American Statistical Association*, 83(402):364–373.

Birge, J. R. and Louveaux, F. (2011). *Introduction to Stochastic Programming*. Springer-Verlag, New York.

Blanchet, J., Liu, J., and Zwart, B. (2008). Large deviations perspective on ordinal optimization of heavy-tailed systems. In Mason, S. J., Hill, R. R., Moünch, L., Rose, O., Jefferson, T., and Fowler, J. W., editors, *Proceedings of the 2008 Winter Simulation Conference*, pages 489–494, Piscataway, New Jersey. IEEE.

Boesel, J., Nelson, B. L., and Ishii, N. (2003a). A framework for simulation-optimization software. *IIE Transactions*, 35(3):221–229.

Boesel, J., Nelson, B. L., and Kim, S.-H. (2003b). Using ranking and selection to "clean up" after simulation optimization. *Operations Research*, 51(5):814–825.

Bofinger, E. (1990). Posterior probability of correct selection. *Communications in Statistics – Theory and Methods*, 19(2):599–616.

Bofinger, E. (1994). Should we estimate the probability of correct selection? *Australian Journal of Statistics*, 36(1):59–66.

Branke, J., Chick, S. E., and Schmidt, C. (2005). New developments in ranking and selection: An empirical comparison of the three main approaches. In Kuhl, M. E., Steiger, N. M., Armstrong, F. B., and Joines, J. A., editors, *Proceedings of the 2005 Winter Simulation Conference*, pages 708–717, Piscataway, New Jersey. IEEE.

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

Branke, J., Zhang, W., and Tao, Y. (2016). Multiobjective ranking and selection based on hypervolume. In Roeder, T. M. K., Frazier, P. I., Szechtman, R., Zhou, E., Huschka, T. R., and Chick, S. E., editors, *Proceedings of the 2016 Winter Simulation Conference*, pages 859–870, Piscataway, New Jersey. IEEE.

Broadie, M., Cicek, D., and Zeevi, A. (2011). General bounds and finite-time improvement for the Kiefer-Wolfowitz stochastic approximation algorithm. *Operations Research*, 59(5):1211–1224.

Broadie, M., Han, M., and Zeevi, A. (2007). Implications of heavy tails on simulation-based ordinal optimization. In Henderson, S. G., Biller, B., Hsieh, M.-H., Shortle, J., Tew, J. D., and Barton, R. R., editors, *Proceedings of the 2007 Winter Simulation Conference*, pages 439–447, Piscataway, New Jersey. IEEE.

Bubeck, S. and Cesa-Bianchi, N. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends in Machine Learning*, 5(1):1–122.

Chandrasekaran, K. and Karp, R. (2014). Finding a most biased coin with fewest flips. In Florina Balcan, M., Feldman, V., and Szepesvári, C., editors, *Proceedings of the 27th Conference on Learning Theory (COLT)*, pages 394–407.

Chang, K. H., Hong, L. J., and Wan, H. (2013). Stochastic trust-region response-surface method (STRONG)—A new response-surface framework for simulation optimization. 25(2):230–243.

Chen, C.-H., Chick, S. E., Lee, L. H., and Pujowidianto, N. A. (2015). Ranking and selection: Efficient simulation budget allocation. In Fu, M., editor, *Handbook of Simulation Optimization*, pages 45–80.

Chen, C.-H., He, D., Fu, M., and Lee, L. H. (2008). Efficient simulation budget allocation for selecting an optimal subset. *INFORMS Journal on Computing*, 20(4):579–595.

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

Chen, E. J. and Lee, L. H. (2009). A multi-objective selection procedure of determining a Pareto set. *Computers and Operations Research*, 36(6):1872–1879.

Chen, P. (1982). *An Alternative Definition of Correct Selection in Ranking and Selection Problems*. PhD thesis, UC Santa Barbara, Santa Barbara, California.

Chick, S. E. (2006a). Bayesian ideas and discrete event simulation: why, what, and how. In *Proceedings of the 2006 Winter Simulation Conference*, pages 96–106.

Chick, S. E. (2006b). Subjective probability and Bayesian methodology. In *Handbooks in Operations Research and Management Science*, volume 13, pages 225–257.

Chick, S. E., Branke, J., and Schmidt, C. (2010). Sequential sampling to myopically maximize the expected value of information. *INFORMS Journal on Computing*, 22(1):71–80.

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

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

Chick, S. E. and Wu, Y. (2005). Selection procedures with frequentist expected opportunity cost bounds. *Operations Research*, 53(5):867–878.

Chiu, W. K. (1974a). A generalized selection goal for a ranking problem. *South African Statistical Journal*, 8(1):45–48.

Chiu, W. K. (1974b). The ranking of means of normal populations for a generalized selection goal. *Biometrika*, 61(3):579–584.

Chiu, W. K. (1977). On correct selection for a ranking problem. *Annals of the Institute of Statistical Mathematics*, 29(Part A):59–66.

Clark, G. M. and Yang, W.-N. (1986). A Bonferroni selection procedure when using common random numbers with unknown variances. In Wilson, J. R., Henriksen, J. O., and Roberts, S. D., editors, *Proceedings of the 1986 Winter Simulation Conference*, pages 313–315, New York. IEEE.

Degroot, M. H. (2004). *Optimal Statistical Decisions*, pages 169–170.

Deng, A. (2015). Objective Bayesian two sample hypothesis testing for online

controlled experiments. In *Proceedings of the 24th International Conference on World Wide Web*, pages 923–928.

Deng, A., Lu, J., and Chen, S. (2016). Continuous monitoring of A/B tests without pain: Optional stopping in Bayesian testing. In *Data Science and Advanced Analytics*.

Desu, M. M. (1970). A selection problem. *Annals of Mathematical Statistics*, 41(5):1596–1603.

Desu, M. M. and Sobel, M. (1968). A fixed subset-size approach to the selection problem. *Biometrika*, 55(2):401–410.

Dmitriev, P., Gupta, S., Dong Woo, K., and Vaz, G. (2017). A dirty dozen: Twelve common metric interpretation pitfalls in online controlled experiments. *Proceedings of the 23rd ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '17*, pages 1427–1436.

Domingo, C., Gavalda, R., and Watanabe, O. (2002). Adaptive sampling methods for scaling up knowledge discovery algorithms. *Data Mining and Knowledge Discovery*, 6(2):131–152.

Dudewicz, E. J. and Dalal, S. R. (1975). Allocation of observations in ranking and selection with unequal variances. *Sankhyā: The Indian Journal of Statistics*, 37(1):28–78.

Eckman, D. J. and Henderson, S. G. (2018). Reusing search data in ranking and selection: What could possibly go wrong? *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 28(3):Article 18, 1–15.

Eckman, D. J. and Henderson, S. G. (2019a). Fixed-confidence, fixed-tolerance guarantees for ranking-and-selection procedures.

Eckman, D. J. and Henderson, S. G. (2019b). Posterior-based stopping rules for Bayesian sequential selection procedures.

Even-Dar, E., Mannor, S., and Mansour, Y. (2002). PAC bounds for multi-armed bandit and markov decision processes. In Kivinen, J. and Sloan, R. H., editors, *2002 International Conference on Computational Learning Theory (COLT)*, pages 255–270, Berlin. Springer.

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

Fabian, V. (1962). On multiple decision methods for ranking population means. *Annals of Mathematical Statistics*, 33(1):248–254.

Fairweather, W. R. (1968). Some extensions of Somerville's procedure for ranking means of normal populations. *Biometrika*, 55(2):411–418.

Faltin, F. W. and McCulloch, C. E. (1983). On the small-sample properties of the Olkin-Sobel-Tong estimator of the probability of correct selection. *Journal of the American Statistical Association*, 78(382):464–467.

Fan, W., Hong, L. J., and Nelson, B. L. (2016). Indifference-zone-free selection of the best. *Operations Research*, 64(6):1499–1514.

Feigin, P. D. and Weissman, I. (1981). On the indifference zone approach to selection—A consistency result. *Annals of the Institute of Statistical Mathematics*, 33(1):471–474.

Feng, M. and Staum, J. (2017). Green simulation: Reusing the output of repeated experiments. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 27(4):23:1–23:28.

Frazier, P. (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.

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

Frazier, P. I. (2010). Decision-theoretic foundations of simulation optimization.

Fu, M., Jian-Qiang Hu, Chun-Hung Chen, and Xiaoping Xiong (2004). Optimal computing budget allocation under correlated sampling. In *Proceedings of the 2004 Winter Simulation Conference*, pages 595–603.

Fu, M. C. (2015). *Handbook of Simulation Optimization*, volume 216. Springer, New York.

Fu, M. C. and Henderson, S. G. (2017). History of seeking better solutions, aka simulation optimization. In Chan, W. K. V., D'Ambrogio, A., Zacharewicz, G., Mustafee, N., Wainer, G., and Page, E. H., editors, *Proceedings of the 2017 Winter Simulation Conference*, pages 131–157, Piscataway, New Jersey. IEEE.

Gao, S. and Chen, W. (2015). A note on the subset selection for simulation optimization. In Yilmaz, L., Chan, W. K. V., Moon, I.-C., Roeder, T. M. K., Macal, C., and Rossetti, M. D., editors, *Proceedings of the 2015 Winter Simulation Conference*, pages 3768–3776, Piscataway, New Jersey. IEEE.

Garivier, A. and Kaufmann, E. (2016). Optimal best arm identification with fixed confidence. *Proceedings of the 29th Conference on Learning Theory (COLT)*, 49:998–1027.

Giani, G. (1986). On a strengthening of the indifference zone approach to a

generalized selection goal. *Communications in Statistics – Theory and Methods*, 15(10):3163–3171.

Glynn, P. and Juneja, S. (2004). A large deviations perspective on ordinal optimization. In Ingalls, R. G., Rossetti, M. D., Smith, J. S., and Peters, B. A., editors, *Proceedings of the 2004 Winter Simulation Conference*, pages 577–585, Piscataway, New Jersey. IEEE.

Glynn, P. and Juneja, S. (2018). Selecting the best system, large deviations, and multi-armed bandits. *ArXiv e-prints*. Preprint arXiv:1507.04564v3, http://arxiv.org/abs/1507.04564v3.

Gorder, B. and Kolonko, M. (2019). Ranking and selection: A new sequential Bayesian procedure for use with common random numbers. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 29(1):Article 2, 1–24.

Guiard, V. (1996). Different definitions of Δ-correct selection for the indifference zone formulation. *Journal of Statistical Planning and Inference*, 54(2):175–199.

Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. *Technometrics*, 7(2):225–245.

Gupta, S. S. and Kim, W.-C. (1980). On the problem of selecting good populations. Technical report.

Gupta, S. S. and Miescke, K. J. (1996). Bayesian look ahead one-stage sampling allocations for selection of the best population. *Journal of Statistical Planning and Inference*, 54(2):229–244.

Hayter, J. (1994). On the selection probabilities of two-stage decision procedures. *Journal of Statistical Planning and Inference*, 38(2):223–236.

Henderson, S. G. and Nelson, B. L., editors (2006). *Handbooks in Operations Research and Management Science: Simulation*, volume 13. Elsevier, Amsterdam.

Ho, Y.-C., Cassandras, C. G., Chen, C.-H., and Dai, L. (2000). Ordinal optimisation and simulation. *Journal of the Operational Research Society*, 51(4):490–500.

Hochberg, Y. and Tamhane, A. C. (2009). *Multiple Comparison Procedures*. Wiley, New York.

Hong, L. J., Luo, J., and Nelson, B. L. (2015). Chance constrained selection of the best. *INFORMS Journal on Computing*, 27(2):317–334.

Hong, L. J. and Nelson, B. L. (2007). Selecting the best system when systems are revealed sequentially. *IIE Transactions*, 39(7):723–734.

Hsu, J. C. (1984). Constrained simultaneous confidence intervals for multiple comparisons with the best. *Annals of Statistics*, 12(3):1136–1144.

Hsu, J. C. (1996). *Multiple Comparisons: Theory and Methods*. CRC Press, New York.

Hunter, S. R., Applegate, E. A., Arora, V., Chong, B., Cooper, K., Rincón-Guevara, O., and Vivas-Valencia, C. (2019). An introduction to multiobjective optimization. *ACM Transactions on Modeling and Computer Simulation*, 29(1):Article 7, 1–36.

Hunter, S. R. and Nelson, B. L. (2017). Parallel ranking and selection. In A. Tolk et al., editor, *Advances in Modeling and Simulation: Seminal Research from 50 Years of Winter Simulation Conferences*, chapter 12, pages 249–275. Springer International, Switzerland.

Hunter, S. R. and Pasupathy, R. (2010). Large-deviation sampling laws for constrained simulation optimization on finite sets. In Johansson, B., Jain, S., Montoya-Torres, J., Hugan, J., and Yücesan, E., editors, *Proceedings of the 2010 Winter Simulation Conference*, pages 995–1002, Piscataway, New Jersey. IEEE.

Inoue, K. and Chick, S. E. (1998). Comparison of Bayesian and frequentist assessments of uncertainty for selecting the best system. In *Proceedings of the 1998 Winter Simulation Conference*, pages 727–734.

Jennison, C., Johnstone, I. M., and Turnbull, B. W. (1982). Asymptotically optimal procedures for sequential adaptive selection of the best of several normal means. In Gupta, S. S. and Berger, J. O., editors, *Statistical Decision Theory and Related Topics III*, pages 55–86. Academic Press, New York.

Jia, Q.-S., Zhou, E., and Chen, C.-H. (2013). Efficient computing budget allocation for finding simplest good designs. *IIE Transactions*, 45(7):736–750.

Jian, N. and Henderson, S. G. (2017). Estimating the probability that a function observed with noise is convex.

Johari, R., Pekelis, L., Koomen, P., and Walsh, D. (2017). Peeking at A/B tests: Why it matters, and what to do about it. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1517–1525.

Johari, R., Pekelis, L., and Walsh, D. J. (2015). Always valid inference: Bringing sequential analysis to A/B testing.

Kadane, J. B., Schervish, M. J., and Seidenfeld, T. (1996). Reasoning to a foregone conclusion. *Journal of the American Statistical Association*, 91(435):1228–1235.

Kalyanakrishnan, S. and Stone, P. (2010). Efficient selection of multiple bandit arms: Theory and practice. In Fürnkranz, J. and Joachims, T., editors, *Proceedings of the 27th International Conference on Machine Learning (ICML)*, pages 511–518.

Kalyanakrishnan, S., Tewari, A., Auer, P., and Stone, P. (2012). PAC subset selection in stochastic multi-armed bandits. In Langford, J. and Pineau, J., editors, *Proceedings of the 29th International Conference on Machine Learning (ICML)*, pages 655–662.

Kao, S. C. and Lai, T. L. (1980). Sequential selection procedures based on confidence sequences for normal populations. *Communications in Statistics – Theory and Methods*, 9(16):1657–1676.

Karnin, Z., Koren, T., and Somekh, O. (2013). Almost optimal exploration in multi-armed bandits. In Dasgupta, S. and McAllester, D., editors, *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pages 1238–1246.

Kaufmann, E., Cappé, O., and Garivier, A. (2014). On the complexity of A/B testing. In Florina Balcan, M., Feldman, V., and Szepesvári, C., editors, *Proceedings of the 27th Conference on Learning Theory (COLT)*, pages 461–481.

Kaufmann, E., Cappé, O., and Garivier, A. (2016). On the complexity of best-arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, 17(1):1–42.

Kim, S.-H. and Dieker, A. B. (2011). Selecting the best by comparing simulated systems in a group of three. In Jain, S., Creasey, R. R., Himmelspach, J., White,

K. P., and Fu, M., editors, *Proceedings of the 2011 Winter Simulation Conference*, pages 3987–3997, Piscataway, New Jersey. IEEE.

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

Kim, S.-H. and Nelson, B. L. (2006a). On the asymptotic validity of fully sequential selection procedures for steady-state simulation. *Operations Research*, 54(3):475–488.

Kim, S.-H. and Nelson, B. L. (2006b). Selecting the best system. In Henderson, S. G. and Nelson, B. L., editors, *Simulation*, volume 13 of *Handbooks in Operations Research and Management Science*. Elsevier.

Lam, K. (1986). A new procedure for selecting good populations. *Biometrika*, 73(1):201–206.

Lau, T. W. E. and Ho, Y.-C. (1997). Universal alignment probabilities and subset selection for ordinal optimization. *Journal of Optimization Theory and Applications*, 93(3):455–489.

Lee, S. and Nelson, B. L. (2019). An omnibus ranking-and-selection procedure. Working paper.

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

Ma, S. and Henderson, S. G. (2017). An efficient fully sequential selection procedure guaranteeing probably approximately correct selection. In Chan, W.

K. V., D'Ambrogio, A., Zacharewicz, G., Mustafee, N., Wainer, G., and Page, E. H., editors, *Proceedings of the 2017 Winter Simulation Conference*, pages 2225–2236, Piscataway, New Jersey. IEEE.

Ma, S. and Henderson, S. G. (2018). Predicting the simulation budget in ranking and selection procedures.

Ma, S. and Henderson, S. G. (2019). A sequential selection procedure delivering a probably-approximately-correct selection using confidence bands. Working paper.

Mahamunulu, D. M. (1967). Some fixed-sample ranking and selection problems. *The Annals of Mathematical Statistics*, 38(4):1079–1091.

Mannor, S. and Tsitsiklis, J. N. (2004). The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, 5(Jun):623–648.

Matejcik, F. J. and Nelson, B. L. (1995). Two-stage multiple comparisons with the best for computer simulation. *Operations Research*, 43(4):633–640.

Merrick, J. R. W. (2009). Bayesian simulation and decision analysis: An expository survey. *Decision Analysis*, 6(4):222–238.

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

Nelson, B. L. and Goldsman, D. (2001). Comparisons with a standard in simulation experiments. *Management Science*, 47(3):449–463.

Nelson, B. L. and Matejcik, F. (1995). Using common random numbers for

indifference-zone selection and multiple comparisons in simulation. *Management Science*, 41(12):1935–1945.

Nelson, B. L. and Staum, J. (2006). Control variates for screening, selection, and estimation of the best. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 16(1):52–75.

Nelson, B. L., Swann, J., Goldsman, D., and Song, W. (2001). Simple procedures for selecting the best simulated system when the number of alternatives is large. *Operations Research*, 49(6):950–963.

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

Ni, E. C., Henderson, S. G., and Hunter, S. R. (2014). A comparison of two parallel ranking and selection procedures. In Tolk, A., Diallo, S. Y., Ryzhov, I. O., Yilmaz, L., Buckley, S., and Miller, J. A., editors, *Proceedings of the 2014 Winter Simulation Conference*, pages 3761–3772, Piscataway, New Jersey. IEEE.

Olafsson, S. (1999). Iterative ranking-and-selection for large-scale optimization. In *Proceedings of the 1999 Winter Simulation Conference*, pages 479–485.

Olkin, I., Sobel, M., and Tong, Y. L. (1976). Estimating the true probability of correct selection for location and scale parameter families. Technical report.

Olkin, I., Sobel, M., and Tong, Y. L. (1982). Bounds for a *k*-fold integral for location and scale parameter models with applications to statistical ranking and selection problems. *Statistical Decision Theory and Related Topics III*, 2:193–212.

Osogami, T. (2009). Finding probably best systems quickly via simulations. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 19(3):Article 12, 1–19.

Parnes, M. and Srinivasan, R. (1986). Some inconsistencies and errors in the indifference zone formulation of selection. *Sankhyā: The Indian Journal of Statistics*, 48(1):86–97.

Paulson, E. (1964). A sequential procedure for selecting the population with the largest mean from k normal populations. *Annals of Mathematical Statistics*, 35(1):174–180.

Peng, Y., Chen, C. H., Fu, M. C., and Hu, J. Q. (2016). Dynamic sampling allocation and design selection. *INFORMS Journal on Computing*, 28(2):195–208.

Peng, Y., Chen, C. H., Fu, M. C., and Hu, J. Q. (2018). Gradient-based myopic allocation policy: An efficient sampling procedure in a low-confidence scenario. *IEEE Transactions on Automatic Control*, 63(9):3091–3097.

Pichitlamken, J. and Nelson, B. L. (2001). Selection-of-the-best procedures for optimization via simulation. In *Proceedings of the 2001 Winter Simulation Conference*, pages 401–407.

Pichitlamken, J. and Nelson, B. L. (2003). A combined procedure for optimization via simulation. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 13(2):155–179.

Pichitlamken, J., Nelson, B. L., and Hong, L. J. (2006). A sequential procedure for neighborhood selection-of-the-best in optimization via simulation. *European Journal of Operational Research*, 173(1):283–298.

Powell, W. B. (2007). *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, volume 703. John Wiley & Sons, Hoboken, New Jersey.

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

Rouder, J. N. (2014). Optional stopping: no problem for Bayesians. *Psychonomic Bulletin & Review*, 21(2):301–308.

Russo, D. (2016). Simple Bayesian algorithms for best-arm identification. *ArXiv e-prints*. Preprint arXiv:1602.08448v4, http://arxiv.org/abs/1602.08448v4.

Sanborn, A. N. and Hills, T. T. (2014). The frequentist implications of optional stopping on Bayesian hypothesis tests. *Psychonomic Bulletin and Review*, 21(2):283–300.

Sanborn, A. N., Hills, T. T., Dougherty, M. R., Thomas, R. P., Yu, E. C., and Sprenger, A. M. (2014). Reply to Rouder (2014): Good frequentist properties raise confidence. *Psychonomic bulletin & review*, 21(2):309–311.

Santner, T. J. (1975). A restricted subset selection approach to ranking and selection problems. *Annals of Statistics*, 3(2):334–349.

Santner, T. J. (1976). A two-stage procedure for selecting δ^*-optimal means in the normal model. *Communications in Statistics – Theory and Methods*, 5(3):283–292.

Schuurmans, D. and Greiner, R. (1995). Sequential PAC learning. In Maass, W., editor, *Proceedings of the Eighth Annual Conference on Computational Learning Theory (COLT)*, pages 377–384, New York. ACM.

Shin, D., Broadie, M., and Zeevi, A. (2016). Tractable sampling strategies for quantile-based ordinal optimization. In Roeder, T. M. K., Frazier, P. I., Szecht-

man, R., Zhou, E., Huschka, T. R., and Chick, S. E., editors, *Proceedings of the 2016 Winter Simulation Conference*, pages 847–858, Piscataway, New Jersey. IEEE.

Sievers, S. (1972). *Topics in Ranking Procedures*. PhD thesis, Cornell University, Ithaca, New York.

Simio (2019). Simio: Simulation, Production Planning and Scheduling Software. Simio, http://www.simio.com/index.php.

Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell Labs Technical Journal*, 41(2):463–501.

Sobel, M. and Huyett, M. J. (1957). Selecting the best one of several binomial populations. *Bell Labs Technical Journal*, 36(2):537–576.

Sohn, J. K. and Kang, S. G. (1992). On the estimation of the true probability of a correct selection. *Communications in Statistics – Simulation and Computation*, 21(2):445–462.

Sriver, T. A., Chrissis, J. W., and Abramson, M. A. (2009). Pattern search ranking and selection algorithms for mixed variable simulation-based optimization. *European Journal of Operations Research*, 198(3):878–890.

Sullivan, D. W. and Wilson, J. R. (1989). Restricted subset selection procedures for simulation. *Operations Research*, 37(1):52–71.

Tamhane, A. C. (1977). Multiple comparisons in Model I one-way ANOVA with unequal variances. *Communications in Statistics – Theory and Methods*, A6(1):15–32.

Teng, S., Lee, L. H., and Chew, E. P. (2010). Integration of the indifference-zone with multi-objective computing budget allocation. *European Journal of Operational Research*, 203(2):419–429.

Tsai, S. C. (2011). Selecting the best simulated system with weighted control-variate estimators. *Mathematics and Computers in Simulation*, 82(4):705–717.

van der Laan, P. (1992). Subset selection of an almost best treatment. *Biometrical Journal*, 34(6):647–656.

Venter, J. H. (1989). Estimating the probability of correct selection. *Journal of Statistical Computation and Simulation*, 32(3):165–176.

Wald, A. (1973). *Sequential analysis*. Courier Corporation.

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

Wu, S.-F. and Yu, Y.-R. (2008). A procedure for selecting all good populations under heteroscedasticity. *Communications in Statistics – Simulation and Computation*, 38(1):46–57.

Xie, J., Frazier, P. I., and Chick, S. E. (2016). Bayesian optimization via simulation with pairwise sampling and correlated prior beliefs. *Operations Research*, 64(2):542–559.

Yang, W.-N. and Nelson, B. L. (1991). Using common random numbers and control variates in multiple-comparison procedures. *Operations Research*, 39(4):583–591.

Yu, E. C., Sprenger, A. M., Thomas, R. P., and Dougherty, M. R. (2014). When

decision heuristics and science collide. *Psychonomic Bulletin and Review*, 21(2):268–282.

Zabinsky, Z. B. (2015). Stochastic adaptive search methods: Theory and implementation. In Fu, M., editor, *Handbook of Simulation Optimization*, pages 293–318. Springer-Verlag, New York.

Zhang, S., Lee, L. H., Chew, E. P., Xu, J., and Chen, C.-H. (2016). A simulation budget allocation procedure for enhancing the efficiency of optimal subset selection. *IEEE Transactions on Automatic Control*, 61(1):62–75.

Zhong, Y. and Hong, L. J. (2017). A new framework of designing sequential ranking-and-selection procedures. In Chan, W. K. V., D'Ambrogio, A., Zacharewicz, G., Mustafee, N., Wainer, G., and Page, E. H., editors, *Proceedings of the 2017 Winter Simulation Conference*, pages 2237–2244, Piscataway, New Jersey. IEEE.

Zhong, Y. and Hong, L. J. (2018). Fully sequential ranking and selection procedures with PAC guarantee. In Rabe, M., Juan, A. A., Mustafee, N., Skoogh, A., Jain, S., and Johansson, B., editors, *Proceedings of the 2018 Winter Simulation Conference*, pages 1898–1908, Piscataway, New Jersey. IEEE.

Zhou, Y., Chen, X., and Li, J. (2014). Optimal PAC multiple arm identification with applications to crowdsourcing. In Xing, E. P. and Jebara, T., editors, *Proceedings of the 31st International Conference on Machine Learning (ICML)*, pages 217–225.