# Fixed-Confidence, Fixed-Tolerance Guarantees for Selection-of-the-Best Procedures

David Eckman and Shane Henderson
School of Operations Research and Information Engineering
Cornell University
Ithaca, New York 14853

July 12, 2018

**Abstract**

Selection-of-the-best procedures designed under the indifference-zone (IZ) formulation provide a guarantee on the probability of correct selection (PCS) whenever the performance of the best system exceeds that of the second-best system by a specified amount. We discuss the shortcomings of this guarantee and argue that providing a guarantee on the probability of good selection (PGS)—selecting a system whose performance is within a specified tolerance of the best—is a more justifiable goal. Although this form of fixed-confidence, fixed-tolerance guarantee has been well studied in the multi-armed-bandit community, it has received far less attention in the simulation community. We examine numerous techniques for proving the PGS guarantee, including sufficient conditions under which selection and subset-selection procedures that deliver the IZ-inspired PCS guarantee also deliver the PGS guarantee. We also compare the frequentist PGS guarantee to its Bayesian counterpart and discuss the differences in how procedures are designed for these two goals.

# 1 Introduction

In many problems of decision-making under uncertainty, the objective is to select the best from among a finite number of systems, i.e., alternatives, where the performance of a system must be estimated via stochastic simulation. When each system can be simulated to some degree within the available computational budget, these problems go by the name of ranking and selection (R&S). 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 selection-of-the-best 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 when the performance of the best system is at least $\delta$ better than those of the other systems, the best system will be chosen with probability exceeding $1-\alpha$, where both $\delta$ and $1-\alpha$ 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 major 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, the decision-maker might be equally satisfied with selecting any system whose performance is, in some sense, "good."

For this reason, a more suitable goal is to guarantee that for *any* problem instance, a system with performance strictly within $\delta$ of the best will be chosen with probability exceeding $1-\alpha$. The value of $\delta$ thus 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 in this paper. 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 paper since, for this class of problems, the definition of good selection remains unsettled [Hunter et al., 2017, Branke et al., 2016].

Selection-of-the-best procedures have been implemented in commercial software and parallel environments, enabling R&S to be more widely used on large-scale problems. The application of R&S to problems with thousands, or even millions, of systems has necessitated the design of efficient procedures that scale well. Moreover, this large-scale setting justifies the use of procedures that return a subset of high-quality systems from which a final selection can be made. Motivated by these trends, we believe that the time is right for the PGS guarantee to displace the IZ-inspired PCS guarantee as the leading design goal for frequentist selection and subset-selection procedures.

With this objective in mind, the main contributions of this paper are threefold:

1. We explain the flaws of the IZ-inspired PCS guarantee and clarify sufficient conditions under which the IZ-inspired PCS guarantee implies the PGS guarantee.

2. We synthesize and extend past results to present a unified treatment of the PGS guarantee.

3. We elucidate the key ideas behind past proof techniques that could be used to prove PGS guarantees for existing procedures and to design future procedures.

This paper is a considerable outgrowth of an advanced tutorial that is to be presented at the 2018 Winter Simulation Conference [Eckman and Henderson, 2018a]. In addition to greatly expanding on the central ideas in the tutorial, this paper provides more precise definitions of key concepts, proofs of the main results, and extensions to subset-selection procedures.

For all of our focus on the PGS 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 the other $\alpha \times 100\%$ of the time. An even stronger goal that avoids 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. 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]. Selection-of-the-best procedures with guarantees on expected linear loss are typically designed under a Bayesian framework. Since we consider frequentist guarantees throughout much of this paper, we choose to focus on PGS as opposed to expected linear loss.

A related problem in which delivering a PGS guarantee is a featured goal is best-arm identification for multi-armed bandits [Audibert and Bubeck, 2010]. In this setting, the PGS guarantee is called the probably approximately correct (PAC) selection guarantee where "probably" refers to the fixed confidence, $1-\alpha$, and "approximately correct" refers to the fixed tolerance, $\delta$ [Even-Dar et al., 2002, 2006]. Research in this area has focused on the design and complexity analysis of efficient selection procedures [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]. With a few exceptions, e.g., Chandrasekaran and Karp [2014], the IZ formulation does not appear in the multi-armed-bandit literature.

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 $\delta$ of the best as a good system, as we do in this paper.

The remainder of this paper is outlined as follows. In Section 2, we argue that the PGS guarantee is superior in many ways to the IZ-inspired PCS guarantee. In Sections 3 and 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 5, we review other methods for proving the PGS guarantee and highlight several technical issues that arise. In Section 6, we compare the frequentist PGS guarantee with the Bayesian guarantee on posterior PGS and we discuss future research directions for the PGS guarantee in Section 7.

# 2 The IZ-Inspired PCS Guarantee versus the PGS Guarantee

Before mathematically defining the various fixed-confidence guarantees, we introduce some standard notation 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. When we later compare selection decisions under different configurations, it will be implicitly assumed that the ordered indices are with respect to a fixed configuration, $\mu$, unless otherwise stated.

We define a selection procedure as one that determines how many observations should be taken from each system and then ultimately selects the system with the best estimated performance. The index of the system chosen by a selection procedure, 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 $\mathrm{CS} := \{\mu_K = \mu_{[k]}\}$. 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

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

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 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, it will be selected with probability 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].

As a way around 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 $\mathrm{PZ}(\delta) := \{\mu : \mu_{[k]} - \mu_{[k-1]} \geq \delta\}$ and the indifference zone $\mathrm{IZ}(\delta) := \{\mu : \mu_{[k]} - \mu_{[k-1]} < \delta\}$. In the preference zone, the best system's performance is at least $\delta$ better than that of the second-best system, whereas in the indifference zone, there are systems with performances within $\delta$ 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:

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

Ever since the conception of the indifference zone, Goal PCS-PZ has been a popular 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 $\mu \in \mathrm{PZ}(\delta)$, 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 GS $:= \{\mu_K > \mu_{[k]} - \delta\}$. Under this definition, only systems with performances *strictly* within $\delta$ 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.

## 2.1 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 $\delta$, the configuration of systems' performances will be in the indifference zone. Likewise, in the case when a selection-of-the-best procedure is used to "clean-up" after a simulation-optimization search [Boesel et al., 2003], 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, 2018b].

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 either case, the implicit Bayesian assumption about the configuration lying in the preference zone and interest in the linear loss function $\mu_{[k]} - \mu_K$ suggests that a selection procedure with a Bayesian guarantee may be preferred.

Another concern with Goal PCS-PZ is the meaning of the IZ parameter. Under Goal PGS, $\delta$ 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 $\delta$ 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. 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 $\delta$ in an attempt to be more confident that the configuration lies in PZ($\delta$). Yet by choosing $\delta$ 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 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. 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 performance of the best and second-best systems is useful in proving PCS guarantees. As we will see, proofs of Goal PGS must 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. We do not address this matter any further in this paper, but instead leave it as an open research question.

## 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 $j$th 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 observations $X_{1j}, X_{2j}, \ldots, X_{kj}$ to be dependent across systems, as is the case when common random numbers are 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 $\mu_i$ are the corresponding means. This normality assumption 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 [2015]. 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 make a more relaxed assumption on the joint distribution of the observations, stated in Assumption 1.

**Assumption 1** *The joint distribution $F$ is fully specified by the configuration $\mu$.*

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 $\delta$ 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 the probability measure $\mathbb{P}_\mu$ is well-defined and unique. It will be important later when we manipulate the configuration that $F$ is unambiguously defined for each $\mu$.

Given the above setup, we further describe a selection procedure as one that (i) takes observations $X_{ij}$ from all systems, (ii) calculates estimators $Y_i$ of $\mu_i$, and (iii) selects a single system $K \in \arg\max_i Y_i$ as the best, while terminating in finite time. As an illustration, for the standard R&S setting with normally distributed observations, the performances $\mu_i$ are the means of the marginal distributions $F_i$, and the estimators $Y_i$ are naturally the sample means.

With regards to the estimators, we make an additional assumption, stated in Assumption 2.

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

Assumption 2 is made so that the event of ties among estimators occurs with probability zero. We can then 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. We make Assumption 2 as a matter of convenience; a careful accounting of ties should allow the results of Section 3 and 4 to extend to the case where the estimators are discrete random variables.

# 3 Goal PCS-PZ Implies Goal PGS

## 3.1 Counterexample to Goal PCS-PZ Implies Goal PGS

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 a contrived selection procedure (Procedure 1) that achieves Goal PCS-PZ but not Goal PGS. In the counterexample, it is assumed that observations from System $i$ are independent and identically distributed (i.i.d.) from normal distributions with means $\mu_i$ and known common variance $\sigma^2$.

**Proposition 1** *Procedure 1 achieves Goal PCS-PZ but not Goal PGS for $k > 2$.*

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. Since selection procedures typically choose the system with the highest estimator, one might yet expect that "reasonable" selection procedures that achieve Goal PCS-PZ also achieve Goal PGS.

**Procedure 1:** Counterexample to Goal PCS-PZ implies Goal PGS

**Setup:** 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.

**Sampling:** Take $n = \lceil 2(h_B + r)^2\sigma^2\delta^{-2} \rceil$ observations from each system, independent across systems, where $h_B$ is the constant of Bechhofer [1954].

**Estimation:** 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 $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(k)}$.

**Selection:** 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 $Y_{(k-1)}$.

## 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 preference-zone configuration. That is, for an arbitrary configuration $\mu \in \mathrm{IZ}(\delta)$, one finds a configuration $\mu^* \in \mathrm{PZ}(\delta)$ for which it can be shown that

$$\mathbb{P}_{\mu}(\mathrm{GS}) \geq \mathbb{P}_{\mu^*}(\mathrm{GS}), \tag{1}$$

where the notation $\mathbb{P}_{\mu}$ and $\mathbb{P}_{\mu^*}$ reflects the dependence of the probability measures on the configuration. Because $\mu^* \in \mathrm{PZ}(\delta)$, it follows from Goal PCS-PZ that $\mathbb{P}_{\mu^*}(\mathrm{GS}) = \mathbb{P}_{\mu^*}(\mathrm{CS}) \geq 1 - \alpha$ and so $\mathbb{P}_{\mu}(\mathrm{GS}) \geq 1 - \alpha$. Hence if it can be shown that Inequality (1) holds for any arbitrary configuration $\mu \in \mathrm{IZ}(\delta)$ and its corresponding configuration $\mu^* \in \mathrm{PZ}(\delta)$, then Goal PCS-PZ implies Goal PGS.

Some care is needed in constructing the related configuration $\mu^*$, as Inequality (1) should not be expected to hold for an arbitrary choice of $\mu^*$. Intuitively, the configuration $\mu^*$ should closely resemble $\mu$ so that the probabilities of good selection can be easily compared.

A simple choice for constructing $\mu^*$ is to only increase the performance of (one of) the best systems until it is exactly $\delta$ 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 $\mu$. 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 $\mu^* \in \mathrm{PZ}(\delta)$ 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 $\mu^*$ is instead to decrease the performances of the good systems of $\mu$ 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 $\mu$. The related configuration $\mu^*$ 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, $\mu^*$ is identical to $\mu$ except that the good systems of $\mu$ (other than the best) are now bad systems with knife-edge performance $\mu_{[k]} - \delta$. From this construction, $\mu^* \in \mathrm{PZ}(\delta)$.

In Section 3.3, we show that under various conditions on selection procedures, Inequality (1) is satisfied for any configuration $\mu \in \mathrm{IZ}(\delta)$ and this choice of $\mu^*$.

## 3.3 Sufficient Conditions

Dating back to Fabian [1962], numerous efforts have been made to identify sufficient conditions under which selection procedures that achieve Goal PCS-PZ simultaneously achieve Goal PGS. 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 imply Inequality (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 $\mu$ 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 $\mu$ 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 of Theorem 1.* Fix an arbitrary configuration $\mu$ and define $\mathcal{G}$, $\mathcal{B}$, and $\mu^*$ accordingly. Let $Y_{[k]}$ denote the estimator associated with System $[k]$ where the index $[k]$ is with respect to the configuration $\mu$. Then

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

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 $\mu^* \in \mathrm{PZ}(\delta)$. □

Condition (C1) may be difficult to verify. For this reason, we list four conditions in Corollary 1 that each imply Condition (C1), but may be easier to verify. Proofs of the conditions of Corollary 1 and their relationships can be found in Appendices A.2 and A.3.

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

*(C2) Let $B_1$ and $B_2$ be disjoint subsets of $\{1, 2, \dots, k\}$ 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 $(B_1, B_2, \mathrm{IP})$,*

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

*for all pairs of configurations $\mu$ 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 $A \subset \{1, \dots, k\}$, the joint distribution of the estimators $Y_i$ for $i \in A$ does not depend on $\mu_j$ for all $j \notin A$.*

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

*(C5) The joint distribution of the estimators $Y_1, \dots, Y_k$ is shift invariant, i.e., for any pair of configurations $\mu$ 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.4.

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 $\mu^{sc}$ where $\mu^{sc}_{[k]} = \mu_{[k]}$ and $\mu^{sc}_i = \mu^{sc}_{[k]} - \delta$ for all $i \neq [k]$:

$$\begin{pmatrix} Y_{[k]} \\ Y_{[k-1]} + (\mu_{[k]} - \mu_{[k-1]} - \delta) \\ \vdots \\ Y_{[1]} + (\mu_{[k]} - \mu_{[1]} - \delta) \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} Y^{sc}_{[k]} \\ Y^{sc}_{[k-1]} \\ \vdots \\ Y^{sc}_{[1]} \end{pmatrix}. \tag{2}$$

The left-hand side of Equation (2) features estimators under an arbitrary configuration $\mu$ and the right-hand side features estimators under the corresponding slippage configuration $\mu^{sc}$. In contrast to Equation (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 "permutability condition," procedures achieving Goal PCS-PZ also guarantee the stronger probability statement

$$\mathbb{P}_\mu(\mu_K \geq \mu_{[k]} - D) \geq 1 - \alpha \quad \text{for all } \mu, \tag{3}$$

where $D = \max\{0, \delta - (Y_{(k)} - Y_{(k-1)})\}$ and $Y_{(i)}$ denotes the $i$th-lowest estimator. Equation (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. Nevertheless, 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. But these conditions 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]. Procedures such as these tend to be more efficient than multi-stage procedures because they can quickly eliminate inferior systems without taking unnecessary observations. Yet by iteratively eliminating systems from contention, procedures of this kind 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 $-\infty$, making the probability statements in Conditions (C1)–(C5) harder to verify. 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.

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 a proof.

**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 $i = 1, \ldots, k$, $\mathbb{P}_\mu(\text{Select } i)$ is nonincreasing in $\mu_j$ 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 $\mu_i$—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].

*Proof of Theorem 2.* Fix an arbitrary configuration $\mu$ and define $\mathcal{G}$ and $\mathcal{B}$ accordingly. If $|\mathcal{B}| = k - 1$, then there is only one good system in $\mu$ 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

$$\mathbb{P}_\mu(\text{Select } [|\mathcal{B}|+1]) + \sum_{i \neq [|\mathcal{B}|+1]} \mathbb{P}_\mu(\text{Select } i) = 1 = \mathbb{P}_{\mu^{(1)}}(\text{Select } [|\mathcal{B}|+1]) + \sum_{i \neq [|\mathcal{B}|+1]} \mathbb{P}_{\mu^{(1)}}(\text{Select } i). \tag{4}$$

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 (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 $\mu_i^{(\ell)} = \mu_i^{(\ell-1)}$ for all $i \neq [|\mathcal{B}|+\ell]$; i.e., the worst $\ell$ good systems of $\mu$ have been made bad. From this definition, $\mu^{(|\mathcal{G}|-1)} = \mu^*$. Therefore the inequalities all together yield

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

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 2:** Two-Stage Sampling

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

To show that Procedure 2 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 $n_1$ samples. Since systems are simulated independently, the probabilities of selecting Systems 2 and 3 will decrease as $\mu_1$ increases.

On the other hand, when $n_1 = \infty$, the procedure effectively 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 $\mu_1$ increases. This implies that the probability that System 3 is selected decreases as $\mu_1$ increases. Thus the probability that System 2 is selected must increase as $\mu_1$ 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 2.

An example of a selection procedure that may violate Condition (C6) is that of Fairweather [1968], which closely resembles Procedure 2. The NSGS procedure of Nelson et al. [2001] might also fail to satisfy Condition (C6); selecting a high value of $\alpha_0$ for the screening stage and a small value of $\alpha_1$ for the selection stage would lead to large second-stage sample sizes, possibly mimicking the behavior of Procedure 2 for the case $n_1 = \infty$.

The proof of Condition (C6) suggests 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 arms 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 $i = 1, \dots, k$, $\mathbb{P}_\mu$(Select some $j$ for which $\mu_j < \mu_i$) is nonincreasing in $\mu_i$.*

*Proof of Theorem 3.* The proof follows that of Theorem 2 with a few small changes.

Fix an arbitrary configuration $\mu$ 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. In terms of PGS, 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 2 where increasing the performance of a system increased the probability that a superior system was selected; this relationship is permitted under Condition (C7). Ultimately, Condition (C7) suffers from the same difficulties as Condition (C6), namely, deriving monotonicity relationships for the probabilities of selecting (subsets of) systems.

## 3.4 Sequential Procedures

As pointed out in Section 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 defined later in Section 6. 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] and the SSM procedure of Pichitlamken et al. [2006] are incorrect. 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 [2017] 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] and Jennison et al. [1982] 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 excessively conservative procedures.

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. Many sequential selection procedures are designed to eliminate systems—or terminate—when the estimators of pairs or groups of systems become well-separated, suggesting 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 probability of correct selection 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. Thus 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 problem instances in the indifference zone; however, these two procedures exert some control on the maximum number of observations that are taken before terminating.

It remains an open question whether taking more observations for problem instances in the indifference zone is a consequence of the IZ formulation, or if it is instead an unavoidable byproduct of Goal PGS.

# 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

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

where the event of correct subset selection is defined as 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 designed 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]. Gao and Chen [2015] and Kaufmann et al. [2016] study the related problem of identifying the top $m$ systems given a fixed sampling budget.

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 $j$th 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 $\ell$, then System $i$ also eliminates System $\ell$. 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 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.5 and A.6, 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 $A \subset \{1, \dots, k\}$, the joint distribution of the estimators $Y_i$ for $i \in A$ does not depend on $\mu_j$ 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 $A \subset \{1, \dots, k\}$, 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 $\mu_j$ 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 $A \subset \{1, \dots, k\}$, $\mathbb{P}_\mu(A \cap I \neq \emptyset)$ is nondecreasing in $\mu_i$ for all $i \in A$ and nonincreasing in $\mu_j$ for all $j \notin A$ when all other components of $\mu$ are held fixed.*

Condition (C10) is a generalization of Condition (C6) of Theorem 2 since for singleton subsets $A$ and the choice of $I = \{i : Y_i > Y_j$ 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 $\mu_i$ and nonincreasing in $\mu_j$ for all $j \neq i$ when all other components of $\mu$ 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., 2003]. 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 3 and 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 3 and 4 can be combined in this setting to prove an overall PGS guarantee is left as a direction for future research.

# 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.

## 5.1 Multiple Comparisons

In Section 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 Assumption (2), the following joint probability statement holds:

$$\mathbb{P}_\mu(Y_{[k]} - Y_i - (\mu_{[k]} - \mu_i) > -\delta, \text{ for all } i \neq [k]) \geq 1 - \alpha \quad \text{for all } \mu. \tag{5}$$

From Assumption (2) and Goal PCS-PZ, for an arbitrary configuration $\mu$,

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

Goal PGS then follows directly from Equation (5):

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

where the first inequality comes from considering only the statement for $i = K$, and the second inequality comes from the definition of $K$ as the index of the system with the highest estimator.

The use of Equation (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 (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 (5) holds with a prespecified confidence is hard to achieve for procedures that take observations sequentially. A recent development in this area is the idea of using bootstrapping to estimate the probabilities of multiple comparison events and then stop sampling when the estimated probability exceeds $1 - \alpha$ [Lee and Nelson, 2017]. This bootstrapping approach sacrifices any kind of finite-sample-size guarantee, but can deliver different asymptotic versions of Goal PGS.

## 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 $\mu_i$ 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

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

where the last inequality is the result of applying Hoeffding's inequality twice. From the choice of $n$, $\mathbb{P}_\mu(Y_i > Y_{[k]}) \leq \alpha/k$. 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 alternative'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 $\delta$, the procedure terminates. Selecting the alternative with the best estimated performance thereby guarantees that a good selection is made with probability exceeding $1 - \alpha$.

# 6 Bayesian PGS

In Sections 2–5, we studied the R&S problem from the frequentist perspective in which the problem instance is assumed to be fixed and the probability is with respect to repeated runs of a procedure. Alternatively, the R&S problem has been studied from a Bayesian perspective [Berger and Deely, 1988, Gupta and Miescke, 1996]. Under the Bayesian framework, the decision-maker's a priori uncertainty about the unknown problem instance is described by a prior distribution on the parameters of the joint distribution $F$. By taking observations, the decision-maker is able to update their beliefs in the form of a posterior distribution for these parameters. The resulting posterior distribution on the problem instance is then used to make statements about the (Bayesian) probability of making a good selection by selecting a given system.

To make these ideas more concrete, we will consider the standard Bayesian R&S setting in which observations are assumed to be normally distributed, i.e., $X_{ij} \sim \mathcal{N}(W_i, \sigma_i^2)$ for $i = 1, \ldots, k$ where the mean, $W_i$, and variance, $\sigma_i^2$, are treated as random variables. Furthermore, the performance of a given system is commonly taken to be its mean performance, i.e., $\mu_i = \mathbb{E}[X_{ij}] = W_i$. Hence the (random) vector $W = (W_1, \ldots, W_k)$ represents the problem instance. When the systems are simulated independently, a normal-gamma conjugate pair on $(W_i, 1/\sigma_i^2)$ simplifies the computation of the joint posterior distribution [Chick and Inoue, 2001b]. If the systems are instead simulated using common random numbers and the joint distribution of the observations is a multivariate normal, a normal-Wishart conjugate pair for the mean vector and precision matrix similarly simplifies the analysis [Chick and Inoue, 2001a].

Under this setup, the marginal posterior distribution of $W$ is a multivariate normal or multivariate $t$-distribution, depending on whether the variances are known or unknown. From this posterior distribution, a probability can be assigned to the event that a given system, $i$, is good:

$$\text{PGS}_i^{\text{Bayes}} = \mathbb{P}(W_i > W_j - \delta, \text{ for all } j \neq i \,|\, \mathcal{E}), \tag{6}$$

where $\mathcal{E}$ denotes the evidence, i.e., the collected observations. The posterior PGS of System $i$, defined in Equation (6), can be regarded as the probability that—based on the evidence—the *random* problem instance $W$ is one for which the *fixed* System $i$ is good. Conversely, the frequentist notion of PGS referred to in our definition of Goal PGS is the probability that the *random* system chosen by a selection procedure will be good for a *fixed* problem instance.

Computing the posterior PGS in Equation (6) involves integrating the posterior distribution of $W$ over the region in which the performance of System $i$ is no worse than the maximum of the performances of the other systems minus $\delta$—a polyhedron described by the $k-1$ inequalities $W_i - W_j > -\delta$ for all $j \neq i$. For small numbers of systems, this integral can be computed numerically while for larger numbers of systems, it can be estimated via Monte

Carlo integration. The Monte Carlo approach, however, introduces a frequentist error in the estimation that complicates statistical conclusions. Another approach is to use Slepian's inequality [Slepian, 1962] to cheaply compute a lower bound on the posterior PGS for a given system [Branke et al., 2007].

An important advantage of the Bayesian formulation is that the posterior PGS can be computed for any system at any time, meaning Bayesian selection procedures can use it in a stopping condition. Under the frequentist formulation, repeatedly looking at the data to estimate the frequentist PGS could invalidate the procedure's statistical guarantees. For Bayesian selection procedures, the posterior PGS can be used in a stopping condition in several ways. For example, a procedure could terminate when the posterior PGS of the system with the highest estimator exceeds $1 - \alpha$, or when the maximum posterior PGS of any system exceeds $1 - \alpha$. The former of these two approaches requires less computation but could lead to more observations being taken before termination. Branke et al. [2005] propose using the lower bound from Slepian's inequality for the posterior PGS of the system with the highest estimator as a stopping rule for the expected value of information (VIP) [Chick and Inoue, 2001b] and optimal computing budget allocation (OCBA) [Chen et al., 2000] procedures. The various top-two sampling algorithms of Russo [2016] could also be modified to use the posterior PGS in a stopping condition.

By allowing the posterior PGS to be used in a stopping condition, the Bayesian formulation offers much flexibility in how selection procedures allocate observations across systems, thereby enabling greater efficiency. In contrast, frequentist selection procedures must be designed to take observations in a way that will guarantee Goal PGS for all configurations. Consequently, frequentist procedures will be inherently conservative for some problem instances. Empirical experiments have shown that Bayesian selection procedures tend to be more efficient than their frequentist counterparts [Branke et al., 2007].

# 7 Conclusions and Future Work

In this paper, we give a comprehensive overview of fixed-confidence, fixed tolerance guarantees, with the objective of reorienting the operations research community towards designing selection-of-the-best 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 modified 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 PGS guarantees.

A strength of the multi-armed-bandit literature is the provision of complexity bounds exploring the computational requirement of a procedure analytically. Section 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 in Ma and Henderson [2018], building on results in 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 have PGS guarantees. Moreover, there is an opportunity for designing procedures that deliver the PGS guarantee more efficiently than existing IZ-inspired procedures. Other approaches that merit further investigation include using bootstrapping to obtain asymptotic guarantees and using posterior PGS as a stopping condition for delivering the Bayesian PGS guarantee.

Although not discussed in this paper, fixed-confidence, fixed tolerance guarantees can be a worthy aim for stochastic optimization algorithms for problems with continuous domains. A future research direction in this setting is the design of algorithms that search over the space of solutions and deliver such guarantees upon termination. Proving guarantees under this framework would likely require some assumptions about the distribution of the observation error and the structure of the unknown objective function.


# Acknowledgments

We thank Steve Chick, Jeff Hong, and Barry Nelson for valuable conversations. This work is supported by the National Science Foundation under grants DGE–1650441 and CMMI–1537394 and by the Army Research Office under grant W911NF-17-1-0094.


# Appendix

## A.1 Proof of Proposition 1

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 [2006]. 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 $i = 1, \dots, k-2$ 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, $\tilde{\mu} \in \mathrm{IZ}(\delta)$.

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

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 $\{\max(Y_1, \dots Y_{k-2}) \leq \min(Y_{k-1}, Y_k)\}$.

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 $\Delta$ so large that $2(k-2)\Phi(-\Delta/(\sigma\sqrt{2/n}))$ is as small as desired. Thus we can make $1 - \mathbb{P}_{\tilde{\mu}}(\mathrm{GS})$ arbitrarily close to 1, while retaining Goal PCS-PZ. $\square$

## A.2 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).

*Proof of Condition (C2) implies Condition (C1).* Fix an arbitrary subset $A$ and configurations $\mu$ 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 $\mathrm{IP} = \{(i, j) : j \in B_2\}$, i.e. $\mathrm{IP} = B_1 \times B_2$. 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

$$\mathbb{P}_{\mu}(Y_i = \max_{j \in A} Y_j) \geq \mathbb{P}_{\tilde{\mu}}(\tilde{Y}_i = \max_{j \in A} \tilde{Y}_j), \quad \text{for all } i \in A. \tag{7}$$

where we have used the fact that ties in the estimators $Y_j$ occur with probability zero. Summing both sides of (7) 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 $A$, $\mu$, and $\tilde{\mu}$ were arbitrary, we have proven the result. $\square$

*Proof of Condition (C3) implies Condition (C1).* Fix an arbitrary subset $A$ and configurations $\mu$ and $\tilde{\mu}$ as specified in the statement of Condition (C1). Take $S = A$. Since $\mu$ 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 $A$, $\mu$, and $\tilde{\mu}$ were arbitrary, we have proven the result. $\square$

*Proof of Condition (C4) implies Condition (C1).* 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). $\square$

*Proof of Condition (C5) implies Condition (C1).* Fix an arbitrary subset $A$ and configurations $\mu$ 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 $\mu_A$ is the vector of components $\mu_i$ 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 $A$, $\mu$, and $\tilde{\mu}$ were arbitrary, we have proven the result. $\square$

## A.3 Relationships of Conditions in Corollary 1

$$\begin{tikzcd}
 & \text{(C1)} & \\
\text{(C3)} \arrow[ur] & \text{(C2)} \arrow[u] & \text{(C6)} \\
\text{(C4)} \arrow[u] & \text{(C5)} \arrow[u] \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

Figure 1: Relations of Conditions of Corollary 1 and Theorems 1 and 2.

Figure 1 shows the relationships among the conditions in Corollary 1 and Theorems 1 and 2. Conditions (C1) and (C6) are the most general. We prove the four relationships in Figure 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).

*Proof of Condition (C4) implies Condition (C3).* 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 $\mu_j$ for $j \notin A$, hence Condition (C3) is satisfied. □

*Proof of Condition (C5) implies Condition (C6).* Fix an arbitrary configuration $\mu$ and arbitrary system $i \in \{1, \ldots, k\}$. For arbitrary $\ell \neq i$, define $\tilde{\mu} = \mu + \epsilon e_\ell$ for $\epsilon > 0$ where $e_\ell$ is $k$-vector of zeros with a one as the $\ell$th 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 $\epsilon$. Since increasing $\epsilon$ is equivalent to increasing the true mean of the $\ell$th system, we have shown that $\mathbb{P}_\mu(\text{Select } i)$ is nonincreasing in $\mu_\ell$ for all $\ell \neq i$. Then since the choice of $i$ was arbitrary, we have proven the result. □

*Proof of Condition (C5) implies Condition (C2).* 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. □

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

## A.4 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 $\mu_i$ 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. 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.

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, $\mu$, define $\mathcal{G}$, $\mathcal{B}$, and $\mu^*$ accordingly. From the definition of the event of good selection for the procedure,

$$\begin{aligned}\mathbb{P}_\mu(\mathrm{GS}) \geq \mathbb{P}_\mu \big(&\{Y_{[k]} > Y_i \text{ for all } i \in \mathcal{B}\} \\ &\cup \{Y_{[k]} \geq Y_i \text{ for all } i \in \mathcal{B} \text{ and } Y_{[k]} = Y_j \text{ for some } j \in \mathcal{B} \\ &\quad \text{and } Z_{[k]} > Z_i \text{ for all } i \in T([k]) \cap \mathcal{B}\}\big),\end{aligned} \tag{8}$$

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 (8) 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 (8) 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 $\mu_j$ for $j \in \mathcal{G}\backslash\{[k]\}$. Consequently, the distribution of $T([k]) \cap \mathcal{B}$ also does not depend on the performances $\mu_j$ for $j \in \mathcal{G}\backslash\{[k]\}$. Therefore we can relate the probability of the event in Inequality (8) under configuration $\mu$ to that of a similar event under configuration $\mu^*$:

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

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

In addition,

$$
\begin{aligned}
&\mathbb{P}_{\mu^*}\Big(\{Y^*_{[k]} > Y^*_i \text{ for all } i \in \mathcal{B}\} \\
&\quad \cup \{Y^*_{[k]} \geq Y^*_i \text{ for all } i \in \mathcal{B} \text{ and } Y^*_{[k]} = Y^*_j \text{ for some } j \in \mathcal{B} \text{ and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k]) \cap \mathcal{B}\}\Big) \\
&\geq \mathbb{P}_{\mu^*}\Big(\{Y^*_{[k]} > Y^*_i \text{ for all } i \neq [k]\} \\
&\quad \cup \{Y^*_{[k]} \geq Y^*_i \text{ for all } i \neq [k] \text{ and } Y^*_{[k]} = Y^*_j \text{ for some } j \neq [k] \text{ and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k])\}\Big).
\end{aligned} \tag{9}
$$

We justify Inequality (9) 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 (9),

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

For the second term on the right-hand side of Inequality (9), we separate outcomes into two cases. If $Y^*_{[k]} = Y^*_j$ for some $j \in \mathcal{B}$, then the outcome belongs to the event

$$\{Y^*_{[k]} \geq Y^*_i \text{ for all } i \in \mathcal{B} \text{ and } Y^*_{[k]} = Y^*_j \text{ for some } j \in \mathcal{B} \text{ and } Z_{[k]} > Z_i \text{ for all } i \in T^*([k]) \cap \mathcal{B}\}.$$

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,

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

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.5 Proof of Theorem 4

We prove the two conditions in Theorem 4 separately.

*Proof of Condition (C8) in Theorem 4 for restricted subset-selection.* Fix an arbitrary configuration $\mu$ and define the subsets $\mathcal{G}$ and $\mathcal{B}$ and the configuration $\mu^*$ 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 $j$th 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 $i \in \mathcal{G}\backslash\{[k]\}$.

$$\begin{aligned}\{Y_{[k]} \geq & \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|+1\rangle} - d\}\} \\ &= \{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|+1\rangle} - d\} \text{ and } Y_{[k]} \geq Y_j \;\; \forall j \in \mathcal{G}\backslash\{[k]\}\} \\ &\quad \cup \{Y_{[k]} \geq \max\{Y_{\langle|\mathcal{B}|-m+2\rangle}, Y_{\langle|\mathcal{B}|+1\rangle} - d\} \text{ and } \exists j \in \mathcal{G}\backslash\{[k]\} \text{ s.t. } Y_j > Y_{[k]}\}.\end{aligned} \tag{10}$$

For the first event on the right-hand side of Equation (10),

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

For the second event on the right-hand side of Equation (10), let $j' = \arg\max_{j \in \mathcal{G}\backslash\{[k]\}} Y_j$, the system in $\mathcal{G}\backslash\{[k]\}$ with the highest estimator. Then

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

Since $j' \in \mathcal{G}\backslash\{[k]\}$, both events on the right-hand side of Equation (10) 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 $\mu$. $\square$

*Proof of Condition (C8) in Theorem 4 for pairwise comparisons.* Fix an arbitrary configuration $\mu$ and define the subsets $\mathcal{G}$ and $\mathcal{B}$ and the configuration $\mu^*$ 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 $i \in \mathcal{G}\backslash\{[k]\}$:

$$\begin{aligned}\{Y_{[k]} &\geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B}\} \\ &= \{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B} \text{ and } Y_{[k]} \geq Y_i - W_{[k]i} \text{ for all } i \in \mathcal{G}\backslash\{[k]\}\} \\ &\quad \cup \{Y_{[k]} \geq Y_j - W_{[k]j} \text{ for all } j \in \mathcal{B} \text{ and } \exists i \in \mathcal{G}\backslash\{[k]\} \text{ s.t. } Y_{[k]} < Y_i - W_{[k]i}\}.\end{aligned} \tag{11}$$

For the first event on the right-hand side of Equation (11),

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

For the second event on the right-hand side of Equation (11), let $i' := \arg\max_{i \in \mathcal{G}\setminus\{[k]\}}\{Y_i : Y_{[k]} < Y_i - W_{[k]i}\}$, the index of the system with the highest estimator among those in $\mathcal{G}\setminus\{[k]\}$ 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 $\mathcal{G}\setminus\{[k]\}$ 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 $\mathcal{G}\setminus\{[k]\}$ that eliminate System $[k]$. Therefore System $i'$ is not eliminated by any other $i \in \mathcal{G}\setminus\{[k]\}$. 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 $I$,

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

Both events on the right-hand side of Equation (11) 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 $\mu$. $\square$

## A.6 Proof of Theorem 5

Fix an arbitrary configuration $\mu$ and define the subsets $\mathcal{G}$ and $\mathcal{B}$ and the configuration $\mu^*$ 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 $\mathbb{P}_\mu(\text{GS}) \geq 1 - \alpha$ for every $\mu$. $\square$

# References

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

Jean-Yves Audibert and Sebastien Bubeck. Best arm identification in multi-armed bandits. In *Proceedings of the 23rd Conference on Learning Theory (COLT)*, 2010.

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

Robert E. Bechhofer, Jack Kiefer, and Milton Sobel. *Sequential Identification and Ranking Procedures: With Special Reference to Koopman-Darmois Populations*. University of Chicago Press, Chicago, 1968.

James O. Berger and John Deely. 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, 1988.

Jose Blanchet, Jingchen Liu, and Bert Zwart. Large deviations perspective on ordinal optimization of heavy-tailed systems. In Scott J. Mason, Ray R. Hill, Lars Moünch, Oliver Rose, Thomas Jefferson, and John W. Fowler, editors, *Proceedings of the 2008 Winter Simulation Conference*, pages 489–494, Piscataway, New Jersey, 2008. IEEE.

Justin Boesel, Barry L. Nelson, and Seong-Hee Kim. Using ranking and selection to "clean up" after simulation optimization. *Operations Research*, 51(5):814–825, 2003.

Jürgen Branke, Stephen E. Chick, and Christian Schmidt. New developments in ranking and selection: An empirical comparison of the three main approaches. In Michael E. Kuhl, Natalie M. Steiger, Frank B. Armstrong, and Jeffrey A. Joines, editors, *Proceedings of the 2005 Winter Simulation Conference*, pages 708–717, Piscataway, New Jersey, 2005. IEEE.

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

Jurgen Branke, Wen Zhang, and Yang Tao. Multiobjective ranking and selection based on hypervolume. In Theresa M. K. Roeder, Peter I. Frazier, Roberto Szechtman, Enlu Zhou, Todd R. Huschka, and Stephen E. Chick, editors, *Proceedings of the 2016 Winter Simulation Conference*, pages 859–870, Piscataway, New Jersey, 2016. IEEE.

Mark Broadie, Minsup Han, and Assaf Zeevi. Implications of heavy tails on simulation-based ordinal optimization. In Shane G. Henderson, Bahar Biller, Ming-Hua Hsieh, John Shortle, Jeffrey D. Tew, and Russell R. Barton, editors, *Proceedings of the 2007 Winter Simulation Conference*, pages 439–447, Piscataway, New Jersey, 2007. IEEE.

Karthekeyan Chandrasekaran and Richard Karp. Finding a most biased coin with fewest flips. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, pages 394–407, 2014.

Chun-Hung Chen, Jianwu Lin, Enver Yücesan, and Stephen E. Chick. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems*, 10(3):251–270, 2000.

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

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

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

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

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

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

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

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

Gordon M. Clark and Wei-Ning Yang. A Bonferroni selection procedure when using common random numbers with unknown variances. In James R. Wilson, James O. Henriksen, and Stephen D. Roberts, editors, *Proceedings of the 1986 Winter Simulation Conference*, pages 313–315, New York, 1986. IEEE.

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

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

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

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

David J. Eckman and Shane G. Henderson. Guarantees on the probability of good selection. In Markus Rabe, Angel A. Juan, Navonil Mustafee, Anders Skoogh, Sanjay Jain, and Björn Johansson, editors, *Proceedings of the 2018 Winter Simulation Conference*, Piscataway, New Jersey, 2018a. IEEE. To appear.

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

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. PAC bounds for multi-armed bandit and markov decision processes. In Jyrki Kivinen and Robert H. Sloan, editors, *2002 International Conference on Computational Learning Theory (COLT)*, pages 255–270, Berlin, 2002. Springer.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of Machine Learning Research*, 7(Jun):1079–1105, 2006.

Václav Fabian. On multiple decision methods for ranking population means. *Annals of Mathematical Statistics*, 33(1):248–254, 1962.

William R. Fairweather. Some extensions of Somerville's procedure for ranking means of normal populations. *Biometrika*, 55(2):411–418, 1968.

Weiwei Fan, L. Jeff Hong, and Barry L. Nelson. Indifference-zone-free selection of the best. *Operations Research*, 64(6):1499–1514, 2016.

Paul D. Feigin and Ishay Weissman. On the indifference zone approach to selection—A consistency result. *Annals of the Institute of Statistical Mathematics*, 33(1):471–474, 1981.

Peter Frazier. 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, 2014.

Siyang Gao and Weiwei Chen. A note on the subset selection for simulation optimization. In Levent Yilmaz, Wai Kin Victor Chan, Il-Chul Moon, Theresa M. K. Roeder, Charles Macal, and Manuel D. Rossetti, editors, *Proceedings of the 2015 Winter Simulation Conference*, pages 3768–3776, Piscataway, New Jersey, 2015. IEEE.

Guido Giani. On a strengthening of the indifference zone approach to a generalized selection goal. *Communications in Statistics – Theory and Methods*, 15(10):3163–3171, 1986.

Peter Glynn and Sandeep Juneja. A large deviations perspective on ordinal optimization. In Rick G. Ingalls, Manuel D. Rossetti, Jeffrey S. Smith, and Brett A. Peters, editors, *Proceedings of the 2004 Winter Simulation Conference*, pages 577–585, Piscataway, New Jersey, 2004. IEEE.

Peter Glynn and Sandeep Juneja. Selecting the best system, large deviations, and multi-armed bandits. *ArXiv e-prints*, 2015. Preprint arXiv:1507.04564v2, `http://arxiv.org/abs/1507.04564v2`.

V. Guiard. Different definitions of $\Delta$-correct selection for the indifference zone formulation. *Journal of Statistical Planning and Inference*, 54(2):175–199, 1996.

Shanti S. Gupta. On some multiple decision (selection and ranking) rules. *Technometrics*, 7 (2):225–245, 1965.

Shanti S. Gupta and Klaus J. Miescke. Bayesian look ahead one-stage sampling allocations for selection of the best population. *Journal of Statistical Planning and Inference*, 54(2): 229–244, 1996.

J. Hayter. On the selection probabilities of two-stage decision procedures. *Journal of Statistical Planning and Inference*, 38(2):223–236, 1994.

Y.-C. Ho, C. G. Cassandras, C.-H. Chen, and L. Dai. Ordinal optimisation and simulation. *Journal of the Operational Research Society*, 51(4):490–500, 2000.

Yosef Hochberg and Ajit C. Tamhane. *Multiple Comparison Procedures*. Wiley, New York, 2009.

Jason C. Hsu. Constrained simultaneous confidence intervals for multiple comparisons with the best. *Annals of Statistics*, 12(3):1136–1144, 1984.

Jason C. Hsu. *Multiple Comparisons: Theory and Methods*. CRC Press, New York, 1996.

Susan R. Hunter and Raghu Pasupathy. Large-deviation sampling laws for constrained simulation optimization on finite sets. In Björn Johansson, Sanjay Jain, Jairo Montoya-Torres, Joe Hugan, and Enver Yücesan, editors, *Proceedings of the 2010 Winter Simulation Conference*, pages 995–1002, Piscataway, New Jersey, 2010. IEEE.

Susan R. Hunter, Eric A. Applegate, Viplove Arora, Byran Chong, Kyle Cooper, Oscar Rincón-Guevara, and Carolina Vivas-Valencia. An introduction to multi-objective simulation optimization. Working paper, 2017.

Christopher Jennison, Iain M. Johnstone, and Bruce W. Turnbull. Asymptotically optimal procedures for sequential adaptive selection of the best of several normal means. In *Statistical Decision Theory and Related Topics III*, pages 55–86. Elsevier, New York, 1982.

Shivaram Kalyanakrishnan and Peter Stone. Efficient selection of multiple bandit arms: Theory and practice. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, pages 511–518, 2010.

Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. PAC subset selection in stochastic multi-armed bandits. In *Proceedings of the 29th International Conference on Machine Learning (ICML)*, pages 655–662, 2012.

Samuel C. Kao and Tze Leung Lai. Sequential selection procedures based on confidence sequences for normal populations. *Communications in Statistics – Theory and Methods*, A9(16):1657–1676, 1980.

Zohar Karnin, Tomer Koren, and Oren Somekh. Almost optimal exploration in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, pages 1238–1246, 2013.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of A/B testing. In *Proceedings of the 27th Conference on Learning Theory (COLT)*, 2014.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, 17 (1):1–42, 2016.

Seong-Hee Kim and A. B. Dieker. Selecting the best by comparing simulated systems in a group of three. In Sanjay Jain, Roy R. Creasey, Jan Himmelspach, K. Preston White, and Michael Fu, editors, *Proceedings of the 2011 Winter Simulation Conference*, pages 3987–3997, Piscataway, New Jersey, 2011. IEEE.

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

Seong-Hee Kim and Barry L. Nelson. Selecting the best system. In Shane G. Henderson and Barry L. Nelson, editors, *Simulation*, volume 13 of *Handbooks in Operations Research and Management Science*. Elsevier, 2006.

K. Lam. A new procedure for selecting good populations. *Biometrika*, 73(1):201–206, 1986.

T. W. Edward Lau and Y.-C. Ho. Universal alignment probabilities and subset selection for ordinal optimization. *Journal of Optimization Theory and Applications*, 93(3):455–489, 1997.

Soonhui Lee and Barry L. Nelson. An omnibus ranking-and-selection procedure. Working paper, 2017.

Sijia Ma and Shane G. Henderson. An efficient fully sequential selection procedure guaranteeing probably approximately correct selection. In W. K. Victor Chan, Andrea D'Ambrogio, Gregory Zacharewicz, Navonil Mustafee, Gabriel Wainer, and Ernest H. Page, editors, *Proceedings of the 2017 Winter Simulation Conference*, pages 2225–2236, Piscataway, New Jersey, 2017. IEEE.

Sijia Ma and Shane G. Henderson. A sequential selection procedure delivering a probably-approximately-correct selection using confidence bands. Working paper, 2018.

D. M. Mahamunulu. Some fixed-sample ranking and selection problems. *The Annals of Mathematical Statistics*, 38(4):1079–1091, 1967.

Shie Mannor and John N. Tsitsiklis. The sample complexity of exploration in the multi-armed bandit problem. *Journal of Machine Learning Research*, 5(Jun):623–648, 2004.

Frank J. Matejcik and Barry L. Nelson. Two-stage multiple comparisons with the best for computer simulation. *Operations Research*, 43(4):633–640, 1995.

Barry L. Nelson and Souvik Banerjee. Selecting a good system: procedures and inference. *IIE Transactions*, 33(3):149–166, 2001.

Barry L. Nelson and David Goldsman. Comparisons with a standard in simulation experiments. *Management Science*, 47(3):449, 2001.

Barry L. Nelson and Frank Matejcik. Using common random numbers for indifference-zone selection and multiple comparisons in simulation. *Management Science*, 41(12):1935–1945, 1995.

Barry L. Nelson and Jeremy Staum. Control variates for screening, selection, and estimation of the best. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 16 (1):52–75, 2006.

Barry L. Nelson, Julie Swann, David Goldsman, and Wheyming Song. Simple procedures for selecting the best simulated system when the number of alternatives is large. *Operations Research*, 49(6):950–963, 2001.

Eric Ni, Dragos Ciocan, Shane Henderson, and Susan Hunter. Efficient ranking and selection in parallel computing environments. *Operations Research*, 65(3):821–836, 2017.

Takayuki Osogami. Finding probably best systems quickly via simulations. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 19(3):12:1–12:19, 2009.

Milton Parnes and R. Srinivasan. Some inconsistencies and errors in the indifference zone formulation of selection. *Sankhyā: The Indian Journal of Statistics*, 48(1):86–97, 1986.

Edward Paulson. A sequential procedure for selecting the population with the largest mean from $k$ normal populations. *Annals of Mathematical Statistics*, 35(1):174–180, 1964.

Juta Pichitlamken, Barry L. Nelson, and L. Jeff Hong. A sequential procedure for neighborhood selection-of-the-best in optimization via simulation. *European Journal of Operational Research*, 173(1):283–298, 2006.

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

Daniel Russo. Simple Bayesian algorithms for best-arm identification. *ArXiv e-prints*, 2016. Preprint arXiv:1602.08448v4, `http://arxiv.org/abs/1602.08448v4`.

Thomas J. Santner. A restricted subset selection approach to ranking and selection problems. *Annals of Statistics*, 3(2):334–349, 1975.

Thomas J. Santner. A two-stage procedure for selecting $\delta^*$-optimal means in the normal model. *Communications in Statistics – Theory and Methods*, 5(3):283–292, 1976.

Dale Schuurmans and Russell Greiner. Sequential PAC learning. In Wolfgang Maass, editor, *Proceedings of the Eighth Annual Conference on Computational Learning Theory (COLT)*, New York, 1995. ACM.

Sally Sievers. *Topics in Ranking Procedures*. PhD thesis, Cornell University, Ithaca, New York, 1972.

David Slepian. The one-sided barrier problem for Gaussian noise. *Bell Labs Technical Journal*, 41(2):463–501, 1962.

Milton Sobel and Marilyn J. Huyett. Selecting the best one of several binomial populations. *Bell Labs Technical Journal*, 36(2):537–576, 1957.

David W. Sullivan and James R. Wilson. Restricted subset selection procedures for simulation. *Operations Research*, 37(1):52–71, 1989.

Suyan Teng, Loo Hay Lee, and Ek Peng Chew. Integration of the indifference-zone with multi-objective computing budget allocation. *European Journal of Operational Research*, 2(203):419–429, 2010.

Shing Chih Tsai. Selecting the best simulated system with weighted control-variate estimators. *Mathematics and Computers in Simulation*, 82(4):705–717, 2011.

Paul van der Laan. Subset selection of an almost best treatment. *Biometrical Journal*, 34 (6):647–656, 1992.

Shu-Fei Wu and Yuh-Ru Yu. A procedure for selecting all good populations under heteroscedasticity. *Communications in Statistics – Simulation and Computation*, 38(1):46–57, 2008.

Wei-Ning Yang and Barry L. Nelson. Using common random numbers and control variates in multiple-comparison procedures. *Operations Research*, 39(4):583–591, 1991.

Ying Zhong and L. Jeff Hong. A new framework of designing sequential ranking-and-selection procedures. In W. K. Victor Chan, Andrea D'Ambrogio, Gregory Zacharewicz, Navonil Mustafee, Gabriel Wainer, and Ernest H. Page, editors, *Proceedings of the 2017 Winter Simulation Conference*, pages 2237–2244, Piscataway, New Jersey, 2017. IEEE.

Yuan Zhou, Xi Chen, and Jian Li. Optimal PAC multiple arm identification with applications to crowdsourcing. In *Proceedings of the 31st International Conference on Machine Learning (ICML)*, pages 217–225, 2014.