Spring 2011

Simulation ranking and selection procedures and applications in network reliability design

Andrew Paul Kiekhaefer
University of Iowa

SIMULATION RANKING AND SELECTION PROCEDURES AND APPLICATIONS IN NETWORK RELIABILITY DESIGN

by

Andrew Paul Kiekhaefer

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Industrial Engineering
in the Graduate College of
The University of Iowa

May 2011

Thesis Supervisor: Associate Professor Yong Chen

ABSTRACT

This thesis presents three novel contributions to the application as well as development of ranking and selection procedures. Ranking and selection is an important topic in the discrete event simulation literature concerned with the use of statistical approaches to select the best or set of best systems from a set of simulated alternatives. Ranking and selection is comprised of three different approaches: subset selection, indifference zone selection, and multiple comparisons. The methodology addressed in this thesis focuses primarily on the first two approaches: subset selection and indifference zone selection.

Our first contribution regards the application of existing ranking and selection procedures to an important body of literature known as system reliability design. If we are capable of modeling a system via a network of arcs and nodes, then the difficult problem of determining the most reliable network configuration, given a set of design constraints, is an optimization problem that we refer to as the network reliability design problem. In this thesis, we first present a novel solution approach for one type of network reliability design optimization problem where total enumeration of the solution space is feasible and desirable. This approach focuses on improving the efficiency of the evaluation of system reliabilities as well as quantifying the probability of correctly selecting the true best design based on the estimation of the expected system reliabilities through the use of ranking and selection procedures, both of which are novel ideas in the system reliability design literature. Altogether, this method eliminates the guess work that was previously associated with this design problem and maintains significant runtime improvements over the existing methodology.

Our second contribution regards the development of a new optimization framework for the network reliability design problem that is applicable to any topological and terminal configuration as well as solution sets of any sizes. This framework focuses

on improving the efficiency of the evaluation and comparison of system reliabilities, while providing a more robust performance and user-friendly procedure in terms of the input parameter level selection. This is accomplished through the introduction of two novel statistical sampling procedures based on the concepts of ranking and selection: Sequential Selection of the Best Subset and Duplicate Generation. Altogether, this framework achieves the same convergence and solution quality as the baseline cross-entropy approach, but achieves runtime and sample size improvements on the order of 450% to 1500% over the example networks tested.

Our final contribution regards the development and extension of the general ranking and selection literature with novel procedures for the problem concerned with the selection of the k-best systems, where system means and variances are unknown and potentially unequal. We present three new ranking and selection procedures: a subset selection procedure, an indifference zone selection procedure, and a combined two stage subset selection and indifference zone selection procedure. All procedures are backed by proofs of the theoretical guarantees as well as empirical results on the probability of correct selection. We also investigate the effect of various parameters on each procedure's overall performance.

Abstract Approved: _____
Thesis Supervisor

_____
Title and Department

_____
Date

SIMULATION RANKING AND SELECTION PROCEDURES AND APPLICATIONS IN NETWORK RELIABILITY DESIGN

by

Andrew Paul Kiekhaefer

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Industrial Engineering
in the Graduate College of
The University of Iowa

May 2011

Thesis Supervisor: Associate Professor Yong Chen

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

PH.D. THESIS

_____

This is to certify that the Ph.D. thesis of

Andrew Paul Kiekhaefer

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Industrial Engineering at the May 2011 graduation.

Thesis Committee: _____
Yong Chen, Thesis Supervisor

_____
Jeffrey Ohlmann

_____
Russell Lenth

_____
Andrew Kusiak

_____
Pavlo Krokhmal

To my family and friends.

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Professor Yong Chen, for all the time and effort he has invested in me. His encouragement and professional guidance were essential in helping me grow both personally and professionally. He has helped me to discover, understand and broaden my own professional interests as well as provided me opportunities to see and experience the world. I am grateful for the investment he has made in me and for believing in me. Thank you.

I would also like to thank my colleagues that I have shared an office with and the faculty that I have taken classes from during my studies. They have all made my time at Iowa an enjoyable and memorable experience. Thank you.

Lastly, I would like to thank my friends and family that have supported me throughout my studies. I could not have made it as far as I have without them. Thank you all for your love and support.

ABSTRACT

This thesis presents three novel contributions to the application as well as development of ranking and selection procedures. Ranking and selection is an important topic in the discrete event simulation literature concerned with the use of statistical approaches to select the best or set of best systems from a set of simulated alternatives. Ranking and selection is comprised of three different approaches: subset selection, indifference zone selection, and multiple comparisons. The methodology addressed in this thesis focuses primarily on the first two approaches: subset selection and indifference zone selection.

Our first contribution regards the application of existing ranking and selection procedures to an important body of literature known as system reliability design. If we are capable of modeling a system via a network of arcs and nodes, then the difficult problem of determining the most reliable network configuration, given a set of design constraints, is an optimization problem that we refer to as the network reliability design problem. In this thesis, we first present a novel solution approach for one type of network reliability design optimization problem where total enumeration of the solution space is feasible and desirable. This approach focuses on improving the efficiency of the evaluation of system reliabilities as well as quantifying the probability of correctly selecting the true best design based on the estimation of the expected system reliabilities through the use of ranking and selection procedures, both of which are novel ideas in the system reliability design literature. Altogether, this method eliminates the guess work that was previously associated with this design problem and maintains significant runtime improvements over the existing methodology.

Our second contribution regards the development of a new optimization framework for the network reliability design problem that is applicable to any topological and terminal configuration as well as solution sets of any sizes. This framework focuses

on improving the efficiency of the evaluation and comparison of system reliabilities, while providing a more robust performance and user-friendly procedure in terms of the input parameter level selection. This is accomplished through the introduction of two novel statistical sampling procedures based on the concepts of ranking and selection: Sequential Selection of the Best Subset and Duplicate Generation. Altogether, this framework achieves the same convergence and solution quality as the baseline cross-entropy approach, but achieves runtime and sample size improvements on the order of 450% to 1500% over the example networks tested.

Our final contribution regards the development and extension of the general ranking and selection literature with novel procedures for the problem concerned with the selection of the k-best systems, where system means and variances are unknown and potentially unequal. We present three new ranking and selection procedures: a subset selection procedure, an indifference zone selection procedure, and a combined two stage subset selection and indifference zone selection procedure. All procedures are backed by proofs of the theoretical guarantees as well as empirical results on the probability of correct selection. We also investigate the effect of various parameters on each procedure's overall performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Discrete-event simulation is the imitation of a hypothetical or real-world system of stochastic components through a computer generated model, where the state of the system changes through a series of discrete events [1]. This type of modeling often involves the use of mathematical, logical, statistical, and symbolic methods to generate artificial data, based on some real world data or distributions, to analyze various "what if" scenarios regarding system performance. These artificial results can then be used to draw conclusions or make inferences about the actual performance of the system being studied to recommend and justify structural changes to the actual system.

The importance and benefits of discrete-event simulation are numerous. In fact, over the last three decades, discrete-event simulation has been consistently reported as one of the most popular and important operations research tools to the operations research practitioner. Shannon, Long, and Buckles [2] conduct a survey regarding the importance of such tools among operations research practitioners in the Institute for Industrial Engineers, and simulation ranked first in utility and interest and second in familiarity. Forgionne [3]; Harpell, Lane, and Mansour [4]; and Lane, Mansour, and Harpell [5] conduct surveys and all reported that simulation ranked second in importance and usage behind statistical analysis for operations research practitioners in large corporations. Thomas and DaCosta [6] surveyed 137 large firms on the operations research tools they used most frequently, and simulation came in second behind statistical analysis. In fact, discrete-event simulation can be found and used in virtually any industry, from manufacturing environments to hospital resource management, from

business management strategies such as Six Sigma to investment analysis. The usage and benefits of this subject are limitless.

In practice, discrete-event simulation is used to address many different types of problems. For example, it is often used in: process improvements scenarios to identify and analyze bottlenecks, queuing systems to improve management of a given system, performance testing to evaluate ideas before implementation, optimization problems to evaluate competing alternatives and select the best solution, and stress testing to determine how a system will perform under extreme conditions or worst case scenarios. However, the central focus of most of these examples and in most discrete-event simulation studies is to select the best design scenario from among a set of competing alternatives.

To achieve the goal of selecting the best design and to do so in an efficient and justifiably correct way, a popular approach referred to as ranking and selection is often relied upon in discrete-event simulation studies. Ranking and selection (R&S) is a group of statistical techniques developed to address the optimization problem associated with the goal of selecting the "best" simulated system configuration or configurations from a given solution space, where "best" refers to the maximization or minimization of some user-specified performance measure. These methods are popular because of their ease of use and value in the simulation output analysis. In this thesis, we focus on the development of Ranking and Selection procedures as well as the application of R&S to the network reliability design problem.

Network reliability design (NRD) is the optimization problem associated with maximizing the system reliability of a given network subject to a set of design

constraints. In this thesis, we represent this problem via an undirected graph of reliable nodes and unreliable edges, where each edge maintains its own individual attributes, such as cost and reliability. System reliability of any candidate design is then defined as the probability that a given set of terminal nodes within the graph are connected via functioning edges. The objective of the problem is then to select the candidate design with maximum system reliability.

There are many applications for the NRD optimization problem in practice. For example, supply chains, sewage systems, flood mitigation systems and power grids can all be modeled as an unreliable network. However, the most prevalent use of NRD has been in the design of communications networks, as there is increasing importance on the reliability of the services these networks provide. Over the past ten years, communications networks have become one of the primary sources for information creation, storage, distribution and retrieval. Many aspects of daily life, be it: news and weather reports, journal paper submittals, databases information retrieval, job applications, social networking, stock purchases, meetings or communications in general (to name a few), now rely heavily on the services provided by communication networks. Reliability has become an increasingly important consideration in the design of communication networks.

Although the NRD problem is not exactly a discrete-event simulation problem, as it is more of a deterministic optimization problem, it still shares many of the same characteristics as a typical discrete-event simulation problem. For example, although there is an exact solution to every problem, the reliability function typically requires some sort of approximation or simulation technique to evaluate it, thus the problem

objective functions are evaluated much like a typical simulation optimization problem. However, in the literature, the NRD problem is typically classified as a combinatorial optimization problem, which can lead to a great deal of information being lost in the solution process as well as sub-optimal results. Therefore, in this thesis, we investigate the applications of simulation ranking and selection to NRD.

1.1 Problem Backgrounds and Literature Review

In this section we present literature reviews for both Ranking and Selection and the network reliability design problem.

1.1.1 Ranking and Selection

Ranking and selection is a group of statistical techniques used to justify the selection of the best or set of the best alternatives from a finite set of alternatives based on the estimation of their expected performance, where "best" refers to the alternative with the maximum or minimum expected value of a given performance criteria [7]. This group of techniques is unique in that they justify the selection of the best by ensuring that a user-specified probability of correct selection is obtained while minimizing the computational effort required to meet this probability. Therefore, these procedures act as a guideline for the sampling process, specifying how many samples are required to ensure the desired probability of selecting the best alternative as well as determining when alternatives can be designated as inferior and eliminated from further sampling.

R&S is a well developed and rather large body of literature with many places to find comprehensive reviews and overviews of R&S concepts and procedures. For example, Fu [8, 9] provides reviews concerning how R&S is used throughout simulation

optimization. Law and Kelton [10] provide a thorough yet simple introduction to R&S with references to more advanced concepts. James et al. [11] provides a comprehensive and up-to-date survey of the R&S literature. Goldsman and Nelson [7, 12, 13] provide new perspectives and state-of-the-art reviews of R&S, and Bechhofer, Santner, and Goldsman [14] and Kim and Nelson [15] provide texts regarding the overview of R&S.

The statistical techniques used in R&S literature to evaluate the simulation optimization problem can be classified into two groups: subset selection procedures (SSP) and indifference-zone selection procedures (IZP), where each group maintains its own distinguishing characteristics, advantages and limitations. We next provide a brief introduction and background regarding these two groups of procedures.

1.1.1.1 Subset Selection Procedures

Subset selection is a group of statistical procedures designed to select a random sized subset, that contains the best or a set of the best alternatives with a user-specified probability, from a finite population of alternatives, based on the estimation of each alternative's expected performance. As such, SSP are independent of the number of samples used, meaning they can be applied to a problem regardless of the sample size. However, because of this, these procedures generally maintain no mechanism to guarantee that the subset returned to the user will be smaller than the original population. Therefore, by themselves, SSP are limited in their use. However, despite their limitations, many useful SSP have been proposed in the literature.

Most SSP proposed in the literature focus on selection of the single best alternative. Gupta [16] introduces this idea proposing a single-stage SSP for populations with common unknown variance. Gupta and Santner [17, 18] extend Gupta's SSP to

allow for a user-specified subset size, and Sullivan and Wilson [19] propose an SSP that allows for unknown and unequal variances. Nelson et al. [20] later extends this work to the more general case of populations with unknown and unequal variance, and Wilson [21] proposes an new decomposition lemma to enhance the lower bound on SSP probability of correct selection.

For selection of a set of (possibly more than one) best alternatives, the literature is much smaller and less well developed. Carroll, Gupta and Huang [22] propose a conservative approach for selecting a subset of a population of systems with known and unknown yet equal variance that contains the k-best systems. Bofinger and Mengersen [23] introduce a less conservative choice of critical point for use with Carroll, Gupta and Huang's SSP. Koenig and Law [24] propose a subset selection procedure for selecting the k-best systems without regard to order for a population of systems with unknown and potentially unequal variance. To our knowledge, research beyond this remains open to investigation.

1.1.1.2 Indifference Zone Selection Procedures

Indifference zone selection is a group of statistical procedures designed to select the true best alternative or alternatives from a population of competing alternatives, based on the estimation of their expected performance, and with a user-specified probability. This selection is made through the use of a sample size calculation based on an indifference parameter, which indicates the user-specified practically significant difference that the experimenter is indifferent to. These procedures are thus synonymous to a power calculation, in that they are used to determine the number of samples required to detect this practically significant difference and select the true best design with a given

probability. However, although these procedures can be very useful, IZP by themselves can be overly conservative. The reason for this is that IZP must sample every alternative in a population a prescribed number of times before a selection can be made. Therefore, no elimination of inferior alternatives occurs until each alterative has been fully sampled. Despite this drawback, many useful IZP have been proposed in the literature.

For selection of the single best alternative, a large number of IZP have been proposed. Bechhofer [25] proposes the first single-stage IZ procedure for populations with known common variance. Paulson [26] and Bechhofer et al [27] later improve upon this initial work proposing two-stage procedures for populations with unknown common variance. Zinger and St. Pierre [28] propose an IZP for populations with unequal known variances. Dudewicz and Dalal [29] as well as Rinott [30] present a two-stage IZP for populations with unequal and unknown variances. Clark and Yang [31] extend Rinott's approach to account for the variance reduction technique known as common random numbers (CRN) for populations with a small number of candidate designs (i.e. around 20 or less). Nelson and Matejcik [32] extend the work of Clark and Yang in a two-stage procedure for systems with a relatively large number of candidate designs (i.e. around 1000 or less) using the assumption of sphericity.

For selection of the set of (more than one) best alternatives, the literature is very small. Bechhofer [33, 34] proposes a single stage IZP for common and known variance, to select the k-best systems with and without regard to order, as well as a two-stage IZP for selecting the k-best alternatives from a population of systems with common unknown variance. To our knowledge, research beyond this remains open to investigation.

1.1.1.3 Combined Procedures

In more recent years however, the literature has focused on the integration of these two R&S approaches as well as the inclusion of multiple comparison confidence intervals into a single combined procedure. These combined procedures are typically more efficient and powerful in selecting the best alternative than their individual counterparts, as the benefits of each individual procedure typically account for and outweigh the limitations of the other. However, these combined procedures are typically more conservative in their theoretical guarantee than their individual counterparts, thus are not always the best option to use in practice. However, if prior information concerning the population of alternatives under study is not available (which is typically the case for most simulation problems), combined procedures are the best option.

For selection of the single best alternative, the state-of-the-art ranking and selection procedures include Nelson's [20] two-stage combined SSP+IZP for populations with unknown and unequal variance, Nelson's [35] combined procedure for comparisons with a standard, Chick and Inoue's [36] Bayesian approach to the two-stage combined ranking and selection problem accounting for CRN, and Nelson and Kim's [37] fully sequential SSP+IZP for populations of systems with unknown and unequal mean and variance. To our knowledge, combined procedures regarding the selection of the set of (more than one) best alternatives remain open to investigation.

1.1.2 Network Reliability Design

NRD is the difficult combinatorial optimization problem associated with finding the topological configuration of an unreliable network, given certain design constraints, that maximizes the system reliability function. This optimization problem is considered difficult for two reasons. First, finding the optimal solution from within the designated solution space is NP-hard [38]. One reason for this, is that the solution space is discrete

and the number of candidate designs in the solution space increases exponentially with the number of edges in the graph. As a result, exact methods for searching the solution space such as total enumeration [39, 40], dynamic programming, and branch-and-bound [41] are only practical for small problems. For moderate to large sized problems, heuristic search methods are typically necessary.

Second, exact evaluation of the system reliability of any feasible design is also NP-hard [42]. One difficulty associated with this, is that the computational complexity of this calculation increases with the size and topology of the network, often rendering exact evaluation techniques such as total enumeration and network decomposition [43-47] impractical for even small topologies. [48-55] have also proposed using theoretical bounds on the system reliability as a substitute for the actual reliability measure, however deriving tight bounds is often difficult. As such, system reliability is typically approximated through simulation and estimation techniques such as Crude Monte Carlo simulation or more recently artificial neural networks [56]. For Crude Monte Carlo simulation, various approaches have been proposed to improve the efficiency of the simulation through different types of sampling techniques, such as dagger sampling [57], failure sampling [58], graph evolution [59], stratified sampling [48], importance sampling [60], and sequential construction and destruction methods [61, 62]. However, research regarding opportunities of additional savings of sampling efficiency in a particular optimization framework still remains an open area of study.

Despite these difficulties, various approaches have been proposed for solving the NRD optimization problem, including: construction heuristics [63], genetic algorithms [64-71], neural networks [56, 72], simulated annealing [73, 74], and the cross-entropy

method [75-80]. However, research focusing on efficient evaluation of system reliability within a given search mechanism is a relatively new concept. Hui et al. [76, 78] presents a cross entropy approach to solve the NRD problem using three alternative sampling methods for highly reliable networks, as well as introduces a hybrid cross-entropy approach to improve the efficiency of the convergence of the purchase probability vector. Altiparmak and Dengiz [80] propose a new generation algorithm and optimization framework for the cross entropy approach to eliminate the generation of infeasible solutions and only evaluate the most promising designs.

1.1.3 Limitations of Literature

This subsection summarizes the limitations of existing works in the NRD and R&S literature.

1.1.3.1 Network Reliability Design

Although a wide body of literature regarding the evaluation of system reliability and navigation of the solution space exist for the NRD optimization problem, we believe that there are still various opportunities for potential improvement. First, to our knowledge, no solution approach has addressed the evaluation of system reliability from the context of a comparison problem, but rather from the context of an evaluation of each individual design. Second, to our knowledge, no approach has provided a lower bound guarantee on the probability of making a correct selection given the number of samples collected when enumeration of the entire solution space is possible. And lastly, although various sampling methods to improve the efficiency of the Monte Carlo simulation to estimate system reliability have been proposed, to our knowledge, these methods are only

used within the context of the conventional approach to Monte Carlo simulation, which involves a user-specified number of Monte Carlo samples being collected from each design in consideration, regardless of solution quality.

1.1.3.2 Ranking and Selection

Although R&S is a well developed and rigorously tested set of statistical techniques, the primary focus in the literature has been on selection of the single best alternative. However, selection of the set of multiple best alternatives can be very useful in practice as no model is a perfect representation of the real world, and hidden constraints may exist that ultimately require alternate solutions. Although various individual procedures have been proposed for selecting the set of the best alternatives, to our knowledge, this specific body of literature is incomplete and has never considered the state-of-the-art combined procedure. Also, apart from their individual applications, new advances in this area could open up potential for entirely new meta-heuristics concerning the general simulation optimization problem, as selection of the set of the best is a neighborhood search criteria common to many leading meta-heuristics.

1.2 Thesis Objective and Outline

In this thesis, we investigate the prospect of using R&S procedures and concepts based on R&S theory to improve the sampling efficiency in solving the NRD optimization problem as well as providing some sort of statistical guarantee on making a correct selection. To do this, we classify the NRD problem according to the size of the solution space, and propose different solution methodology accordingly. For a small number of candidate designs, we introduce an approach to directly apply the state-of-the-

art approaches in R&S. For a large number of candidate designs, we use the theory and concepts of R&S to develop a new optimization framework to improve the sampling efficiency of a commonly used meta-heuristic. We then investigate the prospect of new generic combined R&S procedures backed by theoretical proofs to expand the R&S literature to encompass the selection of the k-best designs problem given unknown and unequal means and variances.

The remainder of this thesis is structured as follows. In Chapter 2 we propose a ranking and selection method for the small network reliability design problem. In Chapter 3 we propose a new optimization framework to improve the efficiency of the evaluation process of the cross entropy approach for the large network reliability design problem. In Chapter 4 we present three novel ranking and selection procedures for addressing the problem of selecting the k-best systems under the conditions of unknown and unequal means and variances. Finally, in Chapter 5 we discuss conclusions as well as areas for potential future work.

CHAPTER 2
A RANKING AND SELECTION APPROACH FOR SMALL NETWORK RELIABILITY DESIGN PROBLEMS

2.1 The sNRD Problem

In this chapter, we consider an important case of the network reliability design optimization problem in which the optimal design is selected by evaluating and comparing the estimated system reliability of every design within a finite solution space, i.e. any NRD where total enumeration of the solution space is practical and desirable. In this thesis, we refer to this problem as the small network reliability design problem (sNRD). Although this definition is subjective depending on the time each individual is willing to invest in the simulation, in this thesis we consider the sNRD as any NRD optimization problem with no more than 1000 candidate designs.

Although problems with more than 1000 candidate designs are typically the most common NRD problems in the literature and industry, there are still many important applications for this sNRD methodology. For example, network expansion [70, 77] is a common problem in practice as many existing networks are often redesigned or expanded rather than constructed from scratch. Also, most state-of-the-art methods for solving the large NRD problem rely on repeatedly solving some type of neighborhood search within the framework of the heuristic search method. These neighborhoods typically fall within the sNRD classification. Thus the sNRD is the basis for most of the NRD solution methodology proposed in the literature.

To illustrate the sNRD problem, consider the network displayed in Figure 2.1, from [77], such that each edge maintains its own individual reliability. In this example,

the existing network has been divided into two sets $\{A1, A2, A3\}$ and $\{B1, B2, B3, B4\}$ which require connecting via one additional edge. The goal of this problem is therefore to select the location for the additional edge that maximizes system reliability.

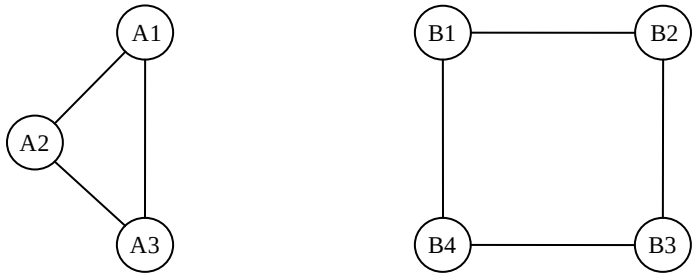

Figure 2.1: Example sNRD

2.1.1 Issues with Monte Carlo Approaches to NRD

The current approach to solving the sNRD problem involves selecting the design with the largest system reliability estimate based on the Monte Carlo evaluation of every design in the solution space, such that n Monte Carlo samples are observed from each design. Although this sampling approach is straightforward and simple to implement, the procedure has its drawbacks. First, the efficiency as well as accuracy of the procedure are highly dependent on the users prior knowledge of the example network and corresponding ability to select an efficient value of n. Second, even if an efficient value of n is selected, this method of sampling is inefficient as many designs do not require n samples in order to determine that they are indeed inferior designs. Third, this sampling method provides no a priori theoretical guarantee on the probability of making a correct selection. In other words, this approach does not account for type I error associated with making a correct selection based on approximations, where type I error for the NRD

problem is the error associated with incorrectly selecting the optimal solution due to the uncertainty of the reliability estimates. To our knowledge, no method within the NRD literature exists that tackles these three problems either individually or in combination.

2.1.2 Chapter Objective and Outline

The contribution of this chapter to the NRD literature is thus threefold. First, to present a new formulation of the sNRD problem that accounts for the corresponding type I error associated with making a correct selection. Second, to introduce the use of ranking and selection procedures to the reliability design literature by proposing novel solution methodology for the sNRD that quantifies and controls the type I error associated with correctly selecting the optimal solution based on approximations. Third, to provide a rule of thumb for selecting a batch size that insures a good approximation of the binomial approximation to the normal distribution in our proposed procedure, in order to satisfy the normality assumption associated with the ranking and selection procedures.

The remainder of this chapter is organized as follows. Section 2 describes the NRD optimization problem, how Monte Carlo simulation is applied to solve it, and our new objectives in solving the sNRD. Section 3 provides an overview of ranking and selection procedures. Section 4 introduces the proposed method. Section 5 justifies the extension and application of the chosen ranking and selection procedure to the NRD problem. Finally, Section 6 provides numerical examples of the proposed procedure.

2.2 Problem Formulation

For the NRD optimization problem, let $G(V, E, T)$ represent a network with reliable node set $V = \{1, \dots, \tilde{n}\}$, unreliable edge set $E = \{1, \dots, \tilde{m}\}$, and terminal node set

$T \subseteq V$ which represents the set of nodes that must be able to communicate with one another in order for the network to be considered operational. We then define the network reliability design optimization problem as follows,

$$\max_{\boldsymbol{x}_i \in \theta} R(\boldsymbol{x}_i) \tag{2.1}$$

where θ is the set of feasible solutions defined by the following constraints:

$$\begin{aligned} &\textstyle\sum_{e=1}^{\tilde{m}} x_{ie} c_e \leq C_{max} && \text{(budget constraint)} \\ &0 \leq x_{ie} \leq 1, e \in E, i && \text{(edge constraint)} \\ &x_{ie} \in \mathbb{Z}^+ \cup \{0\}, e \in E, i && \text{(integer constraint)} \end{aligned} \tag{2.2}$$

where c_e represents the individual cost of edge $e \in E$ and $\boldsymbol{x}_i = (x_{i1}, \dots, x_{i\tilde{m}})$ is the edge purchase vector of design i, where x_{ie} is the binary state of edge $e \in E$, such that

$$x_{ie} = \begin{cases} 1 & \text{if edge } e \text{ is purchased in design } i \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

$R(\boldsymbol{x}_i)$ is the evaluation of the system reliability of design i, and C_{max} is the maximum allowable budget for purchasing edges in each design.

In a sNRD problem, we assume θ is nonempty and finite, containing a total of K feasible designs, such that $K \leq 1000$. As such, we can index the set of all candidate designs via $\theta = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_K\}$, and represent the optimal solution(s) to (2.1) as $\theta^* = \{i : R_i \geq R_j \ \forall j = 1, \dots, K\}$, where $R_i \equiv R(\boldsymbol{x}_i)$.

System reliability for this problem is defined as the probability that a given set of terminal nodes $T \subseteq V$ are connected at any given time, where the system is connected if all nodes in T can communicate with one another via operational edges. Thus, if we let $\boldsymbol{y}_{ij} = (y_{ij1}, \dots, y_{ij\tilde{m}})$ represent the jth possible edge state of design i, where an edge state is a unique vector of individual edge realizations y_{ije}, such that

$$y_{ije} = \begin{cases} 1 \text{ if edge } e \text{ is operating in the } j\text{th edge state of design } i \\ 0 \text{ otherwise} \end{cases}$$

and S_i represent the set of all possible edge states for design i, then the system reliability of design i is mathematically defined as:

$$R_i = Pr\{\phi = 1\} = \sum_{\boldsymbol{y}_{ij} \in S_i} \phi(\boldsymbol{y}_{ij}) P(\boldsymbol{y}_{ij}) \tag{2.4}$$

where $\phi(\boldsymbol{y}_{ij})$ is the structure function for the evaluation of the connectivity of edge state $\boldsymbol{y}_{ij}$, such that

$$\phi(\boldsymbol{y}_{ij}) = \begin{cases} 1 & \text{if } T \text{ is connected in state } \boldsymbol{y}_{ij} \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

and $P(\boldsymbol{y}_{ij})$ is the probability of edge state $\boldsymbol{y}_{ij}$ occurring given independent edge failures, such that

$$P(\boldsymbol{y}_{ij}) = \prod_{e \in \{e' : y_{ije'} = 1\} \cap E_i'} r_e \prod_{e \in \{e' : y_{ije'} = 0\} \cap E_i'} 1 - r_e \tag{2.6}$$

where r_e represents the individual reliability of edge $e \in E$, and $E_i' = \{e' \in E : x_{ie'} = 1\}$. However, because the evaluation of (2.4) is difficult and cannot typically be solved in polynomial time for most networks, (2.1) is typically represented via:

$$\max_{\boldsymbol{x}_i \in \theta} \hat{R}(\boldsymbol{x}_i) \tag{2.7}$$

where $\hat{R}(\boldsymbol{x}_i)$ is the approximation of R_i, based on Monte Carlo simulation, such that

$$\hat{R}_i = \hat{R}(x_i) = \frac{\sum_{j=1}^{n_i} \phi(\boldsymbol{z}_{ij})}{n_i} \tag{2.8}$$

where $\boldsymbol{z}_{ij}$ is the jth randomly selected edge state of design i from state space S_i, such that the eth element of $\boldsymbol{z}_{ij}$ is generated as follows,

$$z_{ije} = \begin{cases} 1 & \text{if } \xi_{ije} \leq r_e x_{ie} \\ 0 & \text{if } \xi_{ije} > r_e x_{ie} \end{cases}$$

where ξ_{ije} is randomly generated according to $\xi_{ije} \sim \mathrm{U}(0,1)$, r_e is the individual reliability of edge e, n_i is the number of Monte Carlo samples observed in the evaluation of $\hat{R}_i$.

These n_i Monte Carlo samples are assumed to be independent and identically distributed for all j, and independent of $\boldsymbol{z_{i'j}}$ for $i \neq i'$. Accordingly, $\hat{R}_i$ is an unbiased estimator of R_i with corresponding variance,

$$\sigma_i^2 = Var(\hat{R}_i) = \frac{\hat{R}_i(1-\hat{R}_i)}{n_i}$$

2.2.1 Common Random Numbers

A common variant of Monte Carlo simulation is the use of common random numbers (CRN). CRN is a variance reduction technique used to compare simulated alternatives under homogenous experimental conditions in order to reduce the random error involved in their pair-wise comparison. For the network reliability design problem, CRN are implemented by using the same single stream of randomly generated numbers for the jth replication of each design, i.e. $\xi_{1je} = \xi_{2je} = \cdots = \xi_{Kje}\ \forall j$.

However, the variance reduction achieved by CRN is not free, as it induces a non-zero correlation between treatment means. Therefore, this correlation must specifically be accounted for within the corresponding parameters of the procedure it is used in, making the overhead cost of using CRN more expensive. Thus, whenever CRN is employed it is necessary to ensure that the tradeoff of increased overhead for reduced variance is beneficial to the overall procedure performance in some way.

2.2.2 Proposed Objective

Given a finite sampling horizon, no approximate can represent its true mean with absolute certainty. With this in mind, any solution obtained using (2.7) should have a confidence level associated with how likely it is to represent the true best design.

However, this issue has never been addressed in the network reliability design literature. Solutions selected via (2.7) are just assumed to be optimal.

Therefore, in this chapter we propose a more rigorous objective for the sNRD optimization problem in order to quantify this probability of correct selection (PCS), and provide a more meaningful result. This objective is as follows,

$$Pr\{select\ \hat{R}_k | R_k - R_i \geq \delta\ \forall i \neq k\} \geq 1 - \alpha \tag{2.9}$$

where α is the allowable type I family-wise error and δ is the smallest difference worth detecting between any two system reliabilities, such that any two system reliabilities that differ by less than δ are considered equivalent. This definition states that, when making a selection based on approximations of the true system reliabilities, we want to select the true best design with a user-specified probability of at least $1 - \alpha$.

We introduce the concept of δ in this definition for two reasons. First, it enables a more realistic representative of the real world, as one often only desires a certain level of accuracy within a design problem; and second, to make this new definition for the sNRD optimization problem more tractable, enabling the use of a sample size calculation to determine the number of Monte Carlo samples necessary to achieve a desired level of confidence.

In order to implement (2.9) in terms of the NRD problem, however, there are still two major issues that need to be addressed. First, we need to find a selection rule that can account for the type I error associated with the multiple comparisons involved in ranking the K system reliabilities with a Monte Carlo simulation. Second, we need a mechanism for determining the number of samples required to represent each individual design's system reliability, such that Monte Carlo samples are only taken when necessary. For the

sNRD problem, we believe that (2.9) is possible through the use of ranking and selection procedures from the discrete event simulation literature.

2.3 Methodology

In this section, we describe our approach to adapt R&S procedures to the sNRD optimization problem as well as the four R&S procedures we consider in this thesis.

2.3.1 Procedure Selection

In this chapter, we consider adapting four R&S procedures to solve the sNRD optimization problem: Procedure R [30], which is a baseline 2-stage IZP (see Appendix A); Procedure CY [31], which is the extension of Procedure R to account for the use of CRN (see Appendix B); Procedure 2SP [20], which is a 3-stage combined procedure (see Appendix C); and Procedure KN [37], which is a fully sequential combined procedure that also accounts for the effects of CRN (see Section 4).

Based on the published results in the ranking and selection literature, the approach we select for our proposed approach is Procedure KN, as it has been shown to maintain superior overall performance as compared to the other three procedures, for optimization problems similar to the sNRD problem [37]. For example, Procedure KN has been shown to make correct selections using significantly less total samples as well as maintain a more robust performance in terms of the selection of its input parameter levels. For a verification of this selection, we refer the reader to Section 6.

2.3.2 Satisfying Normality Assumption with Batch Means

In order to provide the theoretical guarantee associated with the user-specified probability of correct selection, all ranking and selection procedures discussed in this paper assume that the samples used in approximating the estimates of the expected performances for each individual alternative are normally distributed. As such, these techniques cannot be directly applied to the sNRD optimization problem, as Monte Carlo simulation is often necessary to evaluate the system reliabilities, and Monte Carlo samples for reliability evaluation are Bernoulli random variables. We account for this distribution mismatch, however, by approximating normally distributed samples, in accordance with the central limit theorem through the use of batch means of Monte Carlo samples.

We redefine our samples in terms of the corresponding batch mean, such that

$$d_{ij} = \frac{1}{B}\sum_{l=1}^{B} \phi(\boldsymbol{z_{il}}) \tag{2.10}$$

where d_{ij} represents the jth batch mean of the ith topological design of network G, and B is the number of i.i.d. Monte Carlo samples used in each batch. We then substitute b_i for all n_i in Procedure KN, where b_i is the number of batch means used in estimating the system reliability of design i, such that

$$b_i = \frac{n_i}{B} \;\; \forall i \in (1, \dots, K)$$

We then define the system reliability estimate for design i in terms of b_i and d_{ij}, such that

$$\hat{R}_i = \frac{1}{b_i}\sum_{j=1}^{b_i} d_{ij} \tag{2.11}$$

with the estimation of its corresponding variance as,

$$\hat{\sigma}_i^2 = Var(\hat{R}_i) = \frac{\hat{R}_i(1-\hat{R}_i)}{b_i}$$

Therefore, when B is large enough, d_{ij} $\forall j$ will approximately follow a normal distribution. For further discussion concerning this approximation and how to select a level for B that maintains a satisfactory normal approximation, we refer the reader to Section 2.5.1.

2.4 Proposed Procedure

In this section, we present our proposed approach, Procedure KN, as adjusted for network reliability, which we call: Procedure KN+NR; we then follow this with a discussion of the individual details involved in each step.

Proposed Approach: Procedure KN+NR

1) Using (2.10), take b_0 i.i.d. batch means d_{il} from each design i; set $b_{count} = b_0$, where b_{count} is the batch mean counter variable for Procedure KN+NR. And determine the system reliability estimates for all K designs based on their initial b_0 batch means, $\hat{R}_i(b_0)$ $\forall i \in \{1, \dots, K\}$.
2) Determine the sample variance of the difference of the system reliability estimates of design i and design j determined in Step 1, as follows

$$s_{ij}^2 = \frac{1}{b_0 - 1}\sum_{l=1}^{b}\left(d_{il} - d_{jl} - \left(\hat{R}_i(b_0) - \hat{R}_j(b_0)\right)\right)^2 \forall i, j \in \{1, \dots, K\}$$

3) Calculate $h = 2c\eta(b_0 - 1)$ where η is the solution to $g(\eta) = \beta$ (see Appendix D) represented by $\eta_c = \left(\frac{1}{2}\right)\left(\left(\frac{2\alpha}{(N-1)}\right)^{\frac{-2}{(b_0-1)}} - 1\right)$ if CRN are used and $\eta_I = \left(\frac{1}{2}\right)\left(2\left(1 - (1-\alpha)^{\frac{1}{(N-1)}}\right)^{\frac{-2}{(b_0-1)}} - 1\right)$ if independent replications are used, α is the user-specified allowable type I error, and c is a constant with recommended value $c = 1$. (see [37] for further details)

4) Compute $b_{ij} = \left\lfloor\left(\frac{hs_{ij}}{\delta}\right)^2\right\rfloor \forall i, j \in \{1, \dots, K\}$

5) Set $b_i = \max_{j \neq i}\{b_0, b_{ij}\} \forall i$, and $b = \max_i b_i$.

6) Initialize $I = \{1, \dots, K\}$, where I is defined as the set of candidate designs still in contention for the best.

7) Set $I^{old} = I$ and $I = \left(i : i \in I^{old} \text{ and } \hat{R}_i \geq \hat{R}_l - W_{il}, \forall l \in I^{old}, l \neq i\right)$ where $W_{il} = \max\left(0, \left(\frac{\delta}{2cb_{count}}\right)\left(\left(\frac{hS_{ij}}{\delta}\right)^2 - b_{count}\right)\right)$.

8) If $|I| = 1$, stop and select the candidate design in I as the best.

   Else if $b_{count} = b$, stop and select the design in I with the largest system reliability estimate as the best design.

   Else, take Δb additional batch means for all $i \in I$, set $b_{count} = b_{count} + \Delta b$, where Δb is the user-specified incremental increase for b_{count} between subsequent iterations, such that $1 \leq \Delta b$; set $\hat{R}_i = \hat{R}_i(b_{count})$ for all $i \in I$; and return to Step 7.

To break this procedure down, Steps 1-5 represent the indifference zone procedure (IZP) of Procedure KN+NR, Steps 6-7 represent the subset selection procedure (SSP) of Procedure KN+NR, and Step 8 represents the stopping criteria.

For the IZP, Step 1 and Step 2 are used to determine initial system reliability estimates and the variances associated with their pair-wise difference. Steps 3, 4, and 5 then initialize various parameters according to the system reliability estimates and variances in Steps 1 and 2, as well as various other user-specified inputs. The result of these steps is determining the level of b in Step 5, which represents the maximum number of batch means to observe from any candidate design in Procedure KN+NR.

For the SSP, Step 6 is used to initialize a set I to all N candidate designs, and Step 7 is used to screen out inferior designs based on a significant difference of their associated pair-wise comparisons. Thus, W_{il} represents the half-width of the confidence interval associated with determining if $\hat{R}_i$ is significantly inferior to $\hat{R}_l$.

Finally, Step 8 is the stopping and iteration criteria, such that if any of the specified objectives in the procedure are met, Procedure KN+NR terminates. However, if none of the specified objectives are met, Δb additional batch means are observed from each design in set I, and the procedure returns to Step 7. This process then continues to iterate in this manner until one of the stopping criteria are reached.

2.5 Parameter Selection

In this section, we propose a rule of thumb for selecting an efficient level of batch size that satisfies Procedure KN+NR's normality assumption. We then extend the proofs

from [37] for both the recommended level of c as well as when using CRN are beneficial for our desired range of K in Procedure KN+NR.

2.5.1 Batch Size that Maintains an Acceptable Binomial Approximation to the Normal for sNRD

In section 2.4, we presented our proposed approach for solving the sNRD optimization problem, in accordance with the new objective defined in section 2.2. In this procedure, we satisfy the normality assumption associated with the theoretical guarantee of Procedure KN by introducing the use of batch means of Monte Carlo samples to approximate a normal distribution. However, in doing so, we are left with the task of selecting an appropriate level of batch size that still maintains an efficient simulation, which is not easy. Therefore, in this section we develop a model to help select an efficient level of batch size for Procedure KN+NR based on the user-specified input parameters levels.

Before we present this model, however, we must first introduce the concept of critical batch size. Critical batch size, denoted by B^*, is the minimum level of batch size that maintains an appropriate normal distribution for Procedure KN+NR; and is the level of batch size for Procedure KN+NR that we want to predict. This is because in [37], it is shown that when using batch means with Procedure KN, the smaller the batch size used the more efficient the simulation will be; however, in terms of the accuracy of the binomial approximation to the normal, the larger the batch size used the more accurate the resulting normal approximation will be, meaning the lower the risk of violating the normality assumption of Procedure KN. Thus, B^* is the equilibrium point for these two

competing batch size objectives, and what we consider to be the optimal level of batch size for Procedure KN+NR. We therefore construct our model to predict this value.

Unfortunately, though, determining the critical batch size for any set of input parameter levels is not easy for the sNRD optimization problem, as B^* cannot be analytically derived, only approximated through simulation. Therefore, we are not able to derive or specify an exact formula for this value, only a rule of thumb for its approximation. The rule of thumb we developed is as follows,

$$\frac{\kappa \ln K}{\sqrt{b_0 \delta \left(1 - \tilde{R}\right)}} \equiv RoT \tag{2.12}$$

where RoT gives an appropriate upper bound for B^*, $\tilde{R}$ is an estimate of the reliability of the best true candidate design, K is the number of candidate designs in the solution space, b_0 is the initial number of batches, and κ is a constant with range $0.5 \leq \kappa$ such that the larger κ is the more conservative the approximation for RoT will be. When $\kappa = 0.5$, (2.12) is valid for the parameter range: $\delta \in [0.001, 0.05]$, $K \in [30, 1000]$, $\tilde{R} \in [0.5, 0.995]$, and $b_0 \in [10, 20]$.

We developed (2.12) based on results from the experiment described in Appendix E, using a binomial random number generator, $d_{ij} \sim Bin(b; R_i)\ \forall j$, as a substitute for each batch mean of Monte Carlo samples. This not only greatly reduced the runtime of this experiment, but also provided a means of constructing a solution space in accordance with the worst case scenario for Procedure KN+NR, which is a scenario where the system reliability of the true best design is exactly δ greater than the rest, i.e. $R_2 = R_3 = \cdots = R_K = R_1 - \delta$.

Table 2.1: Observed Critical Batch Size vs. Rule of Thumb Batch Size

$K = 30$										$K = 60$									
b_0	δ	$\tilde{R}$	B^*	RoT	b_0	δ	$\tilde{R}$	B^*	RoT	b_0	δ	$\tilde{R}$	B^*	RoT	b_0	δ	$\tilde{R}$	B^*	RoT
10	0.05	0.65	2	5	20	0.05	0.65	1	3	10	0.05	0.65	3	5	20	0.05	0.65	2	4
10	0.05	0.8	3	6	20	0.05	0.8	2	4	10	0.05	0.8	4	7	20	0.05	0.8	3	5
10	0.05	0.95	7	11	20	0.05	0.95	5	8	10	0.05	0.95	8	13	20	0.05	0.95	6	10
10	0.05	0.99	11	25	20	0.05	0.99	8	18	10	0.05	0.99	14	29	20	0.05	0.99	10	21
10	0.05	0.995	10	35	20	0.05	0.995	7	25	10	0.05	0.995	12	41	20	0.05	0.995	7	29
10	0.03	0.65	3	6	20	0.03	0.65	2	4	10	0.03	0.65	4	7	20	0.03	0.65	3	5
10	0.03	0.8	4	7	20	0.03	0.8	3	5	10	0.03	0.8	5	9	20	0.03	0.8	4	6
10	0.03	0.95	10	14	20	0.03	0.95	8	10	10	0.03	0.95	12	17	20	0.03	0.95	10	12
10	0.03	0.99	18	32	20	0.03	0.99	14	22	10	0.03	0.99	22	38	20	0.03	0.99	18	27
10	0.03	0.995	20	44	20	0.03	0.995	16	32	10	0.03	0.995	24	53	20	0.03	0.995	20	38
10	0.01	0.65	4	10	20	0.01	0.65	3	7	10	0.01	0.65	5	11	20	0.01	0.65	4	8
10	0.01	0.8	7	13	20	0.01	0.8	6	9	10	0.01	0.8	8	15	20	0.01	0.8	7	11
10	0.01	0.95	20	25	20	0.01	0.95	16	18	10	0.01	0.95	22	29	20	0.01	0.95	18	21
10	0.01	0.99	38	54	20	0.01	0.99	33	39	10	0.01	0.99	46	65	20	0.01	0.99	36	46
10	0.01	0.995	70	77	20	0.01	0.995	50	54	10	0.01	0.995	70	92	20	0.01	0.995	50	65
10	0.005	0.65	5	13	20	0.005	0.65	4	7	10	0.005	0.65	6	16	20	0.005	0.65	5	11
10	0.005	0.8	9	18	20	0.005	0.8	7	13	10	0.005	0.8	12	21	20	0.005	0.8	9	15
10	0.005	0.95	26	35	20	0.005	0.95	19	25	10	0.005	0.95	30	41	20	0.005	0.95	24	29
10	0.005	0.99	52	77	20	0.005	0.99	38	54	10	0.005	0.99	60	92	20	0.005	0.99	50	65
10	0.005	0.995	100	108	20	0.005	0.995	75	77	10	0.005	0.995	100	130	20	0.005	0.995	90	92
10	0.001	0.99	85	171	20	0.001	0.99	75	121	10	0.001	0.99	100	205	20	0.001	0.99	70	145
10	0.001	0.995	170	241	20	0.001	0.995	100	171	10	0.001	0.995	140	290	20	0.001	0.995	100	205

Table 2.1 summarizes the results corresponding to each of the 88 treatments used in the experiment in Appendix E, as well as corresponding approximation from (2.12) using $\kappa = 0.5$, indicated by column 'RoT'. We refer to RoT in (2.12) as an upper bound to B^* as it never underestimates B^* for the given range of interest of the parameters,

according to our simulation results in Appendix E, yet follows the same trends and stays relatively close to the simulated values of B^*.

To verify that (2.12) is valid for the range of K we specify, not just the range used to construct it in Appendix E, we randomly select ten of the sixty possible treatments of parameters: $\tilde{R}$, δ, and b_0 from the range specified in Appendix E, while setting $K = 500$, $\alpha = 0.05$, and B according to its RoT upper bound in (2.12). Each treatment is then replicated 500 times in order to empirically determine the experimental probability of correct selection, which we denote as ExPCS in Table 2.2. A summary of the results of these treatments are shown in Table 2.2.

Table 2.2: Validation of Batch Size Rule of Thumb

Treatment	$\tilde{R}$	δ	b_0	K	Replications	Avg Samples	ExPCS
1	0.95	0.05	0	500	500	2.0056E+05	0.962
2	0.8	0.05	20	500	500	4.2437E+05	0.976
3	0.999	0.001	20	500	500	1.0708E+07	0.970
4	0.995	0.001	10	500	500	4.3662E+07	0.964
5	0.65	0.05	10	500	500	7.2629E+05	0.976
6	0.99	0.01	10	500	500	1.4212E+06	0.974
7	0.995	0.03	10	500	500	4.3062E+05	1.00
8	0.999	0.005	10	500	500	2.4436E+06	1.00
9	0.8	0.01	10	500	500	1.2488E+07	0.976
10	0.995	0.005	20	500	500	2.1319E+06	0.96

Since the acceptable type I error rate was set to $\alpha = 0.05$, the nominal PCS for this problem was 0.95. Comparing this to the results in Table 2.2, the ExPCS is greater

than the nominal PCS in every treatment. Since these treatments are all randomly selected from our range of interest, this is a good indication that (2.12) is indeed a satisfactory and efficient upper bound approximation for B^* for our given range of interest of the parameters.

2.5.2 Selection of c

Kim and Nelson [37] recommend the selection of $c = 1$ for Step 3 of Procedure KN+NR when the experimenter has no prior knowledge of the candidate designs in the solution space. They support this claim by stating that under these conditions, the best choice of c should be the level that minimizes:

$$c\eta^2 \times \frac{2(b_0 - 1)S_{iK}^4}{\delta^3} \tag{2.13}$$

and determine the minimum value for (2.13) for the range $K \in [2,100]$, as indicated by the gray values in Table 2.3.

However, since in the sNRD we apply Procedure KN+NR to a much larger range of K than tested in [37], we extend Kim and Nelson's justification for the recommended level of c to much larger range of K, to see if the same results apply. Table 2.3 is a reproduction of the results from [37] extended to the range of K relevant to the sNRD, i.e. $K \in [2,1000]$.

As one can see from the results in Table 2.3, $c = 1$ still maintains the smallest area for all treatments, thus using the same logic as in [37] we conclude that $c = 1$ remains the preferred choice for Procedure KN+NR in the sNRD problem, given no prior

knowledge of the candidate designs in the solution space. For further details on this explanation, we refer the reader to [37].

Table 2.3: Area $c\eta^2$ in Units of $2(b_0-1)S_{iK}^4/\delta^3$, when $g(\eta)=0.05/(K-1)$ for $K \in [2,1000]$

n_0	c	K									
		2	5	10	20	100	200	400	600	800	1000
5	1	1.169	7.088	18.007	40.858	232.02	475.45	966.17	1459.1	1953.1	2447.8
	2	1.593	9.325	23.446	52.88	298.09	609.84	1237.9	1868.4	2500.3	3133
	3	2.096	12.14	30.433	68.517	385.43	788.15	1599.3	2413.6	3229.6	4046.6
	4	2.618	15.087	37.769	84.972	477.5	976.24	1980.7	2989	3999.3	5010.9
	5	3.147	18.084	45.237	101.7	571.36	1168	2369.5	3575.7	4784.2	5994.3
	10	5.822	33.284	83.142	186.8	1048.2	2142.4	4345.6	6557.2	8773.1	10992
10	1	0.112	0.403	9.738	1.221	3.296	4.858	7.056	8.729	10.131	11.359
	2	0.148	0.504	0.903	1.469	3.886	5.69	8.22	10.143	11.751	13.16
	3	0.193	0.645	1.148	1.863	4.893	7.153	10.32	12.725	14.736	16.498
	4	0.24	0.796	1.411	2.286	5.993	8.75	12.616	15.553	18.008	20.158
	5	0.288	0.946	1.682	2.723	7.116	10.395	14.985	18.469	21.384	23.935
	10	0.529	1.731	3.058	4.942	12.905	18.823	27.12	33.418	38.686	43.296
15	1	0.038	0.12	0.203	0.311	0.705	0.96	1.287	1.518	1.703	1.859
	2	0.05	0.148	0.243	0.366	0.804	1.085	1.443	1.696	1.898	2.068
	3	0.065	0.188	0.307	0.459	1.002	1.35	1.793	2.105	2.354	2.564
	4	0.08	0.231	0.376	0.563	1.222	1.645	2.183	2.562	2.864	3.119
	5	0.096	0.275	0.447	0.666	1.449	1.949	2.587	3.035	3.393	3.695
	10	0.176	0.499	0.81	1.204	2.611	3.516	4.663	5.471	6.114	6.657
20	1	0.019	0.056	0.092	0.136	0.285	0.374	0.485	0.561	0.62	0.669
	2	0.025	0.068	0.108	0.158	0.319	0.416	0.534	0.614	0.677	0.729
	3	0.032	0.087	0.136	0.197	0.395	0.515	0.66	0.758	0.835	0.899
	4	0.039	0.106	0.166	0.24	0.482	0.626	0.801	0.921	1.014	1.092
	5	0.047	0.126	0.198	0.286	0.571	0.741	0.949	1.09	1.2	1.292
	10	0.086	0.231	0.357	0.515	1.03	1.335	1.708	1.962	2.16	2.324

2.5.3 Extension of CRN Justification

Kim and Nelson [37] show that using CRN in Procedure KN is beneficial as long as ρ, the correlation induced between designs by CRN, is greater than $1-\frac{\eta_I}{\eta_c}$ (where η_c

and η_I are parameters of Procedure KN given in section 2.4). They further prove that η_c is at least 1.02 times η_I for $K \in [2,100]$ so that CRN is beneficial as long as $\rho > 0.02$. In this section, we extend their proof to a larger range of K using the same justification. To do this, let Ψ be the ratio of η_c to η_I, i.e.

$$\Psi = \frac{\eta_c}{\eta_I} = \frac{\left(\frac{1}{2}\right)\left(\left(\frac{2\alpha}{K-1}\right)^{\frac{-2}{b_0-1}} - 1\right)}{\left(\frac{1}{2}\right)\left(\left(2\left((1-\alpha)^{\frac{1}{K-1}}\right)\right)^{\frac{-2}{b_0-1}} - 1\right)} \tag{2.14}$$

Therefore, Ψ is a function of the initial number of batches b_0, the acceptable type I error α, and the number of designs in the solution space K, for the range of interest: $10 \leq b_0 \leq 20$, $0 < \alpha \leq 0.1$, and $2 \leq K \leq 1000$. We then find an upper bound for Ψ by evaluating its partial derivatives over the given ranges of interest.

To begin, we evaluate $\frac{\delta\Psi}{\delta b_0}$ for $K = \{2,3,..,1000\}$, $b = \{2,3,...,20\}$, and $\alpha = \{0.01, 0.05, 0.1\}$. For this range, $\frac{\delta\Psi}{\delta b_0}$ is always negative indicating that Ψ is a decreasing function for $2 \leq b_0 \leq 20$, which implies that we only need to consider the smallest value of b_0 in our range of interest for our upper bound, i.e. $b_0 = 10$.

Next, setting $b_0 = 10$, we evaluate $\frac{\delta\Psi}{\delta\alpha}$ for $K = \{2,3,..,1000\}$ and $\alpha = \{0.01, 0.05, 0.1\}$. For this range, $\frac{\delta\Psi}{\delta\alpha}$ is always positive indicating that Ψ is an increasing function for $0.01 \leq \alpha \leq 0.1$, which implies that we only need to consider the largest value of α in our range of interest for our upper bound, i.e. $\alpha = 0.1$.

Finally setting $b_0 = 10$ and $\alpha = 0.1$, we can then express an upper bound for Ψ as a function of K, as follows

$$\Psi \leq \frac{-1 + 1.42997\left(\frac{1}{K-1}\right)^{\frac{-2}{9}}}{-1 + \left(2\left(1 - 0.9^{\frac{1}{K-1}}\right)\right)^{\frac{-2}{9}}} \tag{2.15}$$

We then evaluate (2.15) for our desired range of $K \in [2,1000]$, where the continuous curve of this evaluation is displayed in Figure 2.2.

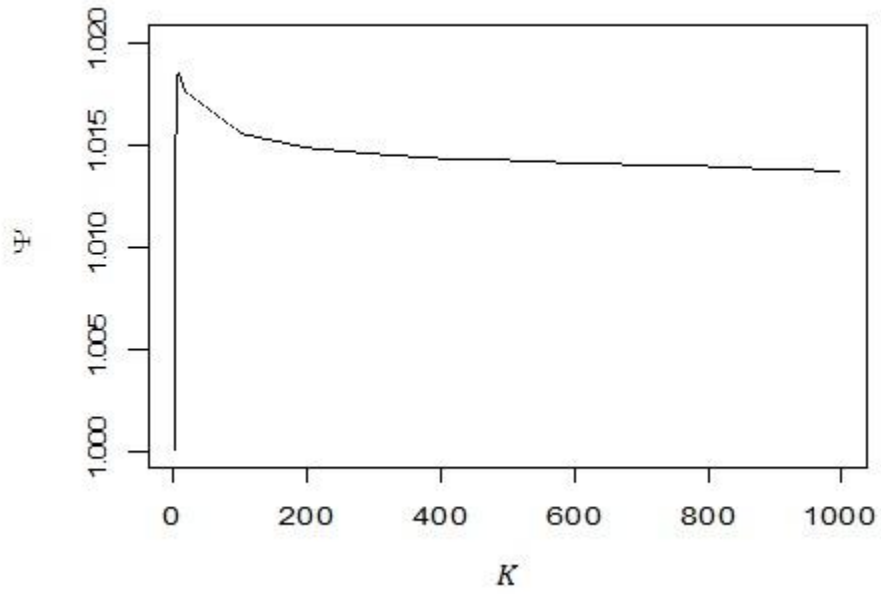

Figure 2.2: Upper bound of Ψ

Clearly, Ψ still peaks at $K = 7$ as it did when evaluated for the smaller range of K in [37]. Therefore, for the parameter range of interest in sNRD problems, we still conclude that $\Psi < 1.02$, implying that using CRN in Procedure KN+NR will be beneficial so long as $\rho > 0.02$. And for larger K, say $K > 200$, $\rho > 0.015$ will be sufficient for CRN to be beneficial

2.6 Numerical Examples

In this section we present three numerical examples to verify the selection of Procedure KN as our method of choice, validate the batch size rule of thumb model, and test the efficiency of Procedure KN+NR according to our RoT approximation, which we denote as Procedure KN+NR+RoT.

2.6.1 Example Networks

Throughout this section, we use three example networks to test our proposed procedures, which we refer to as: the 30-node chain, the 35-node chain, and the complex network. For the two chain network examples, each network consists of a base network defined as a path spanning all nodes in V and having $|V| - 1$ total edges, where all individual edge reliabilities are 0.99. The goal of the problem is to add one additional edge to the base network in such a way that system reliability is maximized. Figure 2.3 shows a generic example of this problem, where the base network is represented by the solid edges connecting nodes 1 through q and the dashed edges represent instances of how an additional edge could be added to the base network. For the 30-node and 35-node base networks, this creates a solution space of 406 and 561 candidate designs respectively.

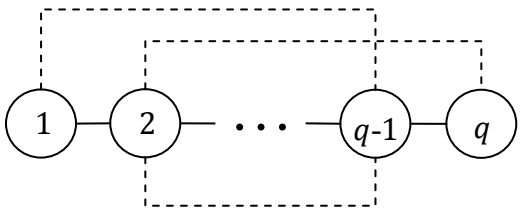

Figure 2.3: Example Chain Network

For the complex network example, we consider the network expansion problem depicted in Figure 2.4, which is a network that cannot analytically be solved and whose system reliability is difficult to evaluate. This example consists of a base network of 13 nodes and 18 edges, whose individual edge reliabilities are all 0.95. The goal of the problem is to expand the network to include node 0 as well as two additional edges in such a way that system reliability is maximized. Therefore, excluding edge redundancy, there are a total of 858 candidate designs in this solution space.

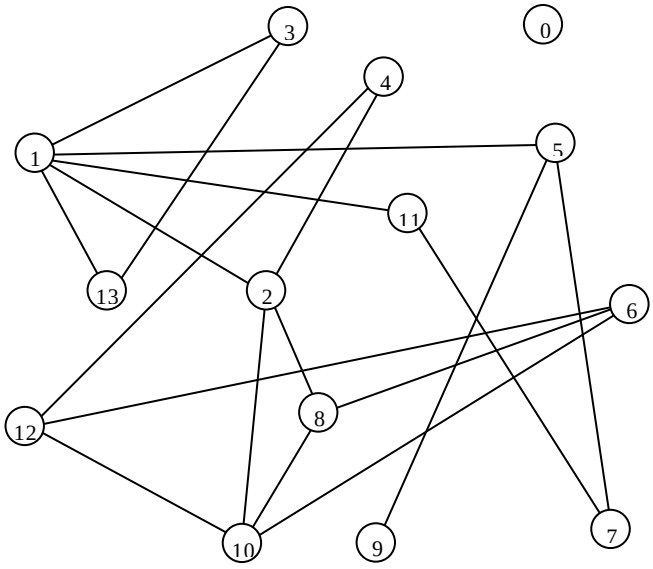

Figure 2.4: Complex Network Expansion Example

2.6.2 Verification of Efficiency of Procedure KN

In this section, we verify that the published results from discrete event simulation concerning the performance of: Procedure R, Procedure CY, Procedure 2SP, and Procedure KN still apply to the sNRD problem, i.e. Procedure KN is the overall most efficient method. To do this, we apply and compare these four procedures using the 30-

node and 35-node chain networks. We do not use the complex network example for this experiment as its run time was too great for both Procedure R and Procedure CY.

For each example network, the acceptable type I error rate is set to $\alpha = 0.05$, the indifference parameter is set to $\delta = 0.005$, the initial number of batches is set to $b_0 = 15$, and the batch size is set to the conservative level of $B = 100$. Each procedure is then subject to ten independent replications for each example network.

Table 2.4 summarizes the results of the ten independent replications for each treatment. The comparison metric used here is the average total number of un-batched Monte Carlo samples to make a selection, as this provided an unbiased measure of each procedure's efficiency. Average total runtime, however, is in general directly proportional to average total number of Monte Carlo samples, with Procedure KN requiring 491 and 981 seconds to make a correct selection for the 30-node and 35-node chain networks respectively. All replications of all treatments select the true best design.

Table 2.4: Comparison of Ranking and Selection Methods on sNRD

Procedure	30-node	35-node
R	140,350,600	239,859,940
CY	114,224,040	232,983,300
2SP	3,899,734	5,858,560
KN	568,820	799,740

Therefore, using Procedure KN as a baseline, Procedure R, Procedure CY, and Procedure 2SP required 247, 201, and 7 times more total Monte Carlo samples to make a correct selection for the 30-node chain network; and 300, 292, and 8 times more total

Monte Carlo samples to make a correct selection for the 35-node chain network. These results reflect those published in the ranking and selection literature for discrete event simulation [20, 37], indicating that the performance of these procedures as applied to the sNRD problem are indeed consistent with the general results. Based on this, we conclude that Procedure KN was indeed the best ranking and selection approach to use in the sNRD optimization problem.

2.6.3 Validation of Batch Size Rule of Thumb and the Probability of Correct Selection

In Section 2.5.1, we present a batch size rule of thumb for Procedure KN+NR which is constructed using a binomial random number generator in place of Monte Carlo simulation. In this section, we verify that this was indeed an acceptable substitution for developing our rule of thumb, by empirically determining the PCS of Procedure KN+NR+RoT, based on 100 independent replications using actual Monte Carlo samples, for each of the three example networks. We then compare the efficiency of Procedure KN+NR+RoT with the conventional approach described in Section 1, in terms of their corresponding PCS based on an equal number of total Monte Carlo samples. This is done by using $n_j = \frac{total_j}{K}$ as the number of un-batched Monte Carlo samples to evaluate each candidate design in the conventional approach, where $total_j$ represents the total number of Monte Carlo samples required for the proposed procedure to make a correct selection in replication j. We then empirically determine the probability of correct selection for the conventional approach, for each of the three example networks, based on the solution quality of the corresponding 100 independent replications.

The parameters used in this experiment are: $\alpha = 0.05$, $\delta = 0.05$, $b_0 = 15$, and $\tilde{R} = 0.965$ for the 30-node chain network; $\alpha = 0.05$, $\delta = 0.05$, $b_0 = 15$, and $\tilde{R} = 0.955$ for the 35-node chain network; and $\alpha = 0.05$, $\delta = 0.01$, $b_0 = 15$, and $\tilde{R} = 0.999$ for the complex network. A summary of these results are displayed in Table 2.5.

Table 2.5: PCS of Conventional vs Proposed Procedure

Network	Number of Designs	# of Monte Carlo samples per design (n_j)	Conventional Approach Empirical PCS	KN Empirical PCS
30-node chain	406	1402	78%	100%
35-node chain	561	1426	64%	100%
Complex	858	8714	50%	100%

Since we are unable to analytically verify the optimal design of the complex network, we accept the majority selection from both Procedure KN+NR+RoT and the conventional approach as the true best design, as shown in Table 2.6. For each of the tables in Table 2.6 through Table 2.8, "Design" specifies the node pairs, in parentheses, of the base network to be connected using the additional edge/s; "Reliability" is the estimated system reliability of the specific design based on 25 million Monte Carlo samples; and "Percent Selection from ..." is the percentage of replications that the designated procedure selected that particular design as the best.

For each example network in this experiment, Procedure KN+NR+RoT correctly select the true best design 100% of the time, whereas the conventional approach correctly select the true best design only 76%, 64%, and 50% of the time for the 30-node, 35-node,

and complex networks respectively. This pattern seems to indicate that the larger the solution space and the more complex the topology of the example network, the more efficient Procedure KN+NR+RoT is as compared to the conventional approach. These results both validate our batch size rule of thumb, and provide some insight into the efficiency and sample size savings associated with using Procedure KN+NR+RoT.

Table 2.6: Complex Network PCS Breakdown

Design	Reliability	Percent Selection from KN	Percent Selection from Conventional Approach
(0,2) & (0,9)	0.9674	0	11
(0,4) & (0,9)	0.9696	100	50
(0,6) & (0,9)	0.9677	0	7
(0,8) & (0,9)	0.9677	0	15
(0,9) & (0,10)	0.9676	0	8
(0,9) & (0,12)	0.9678	0	9

Table 2.7: 30-node Chain Network PCS Breakdown

Design	Reliability	Percent Selection from KN	Percent Selection from Conventional Approach
(1,29)	0.9564	0	11
(1,30)	0.9639	100	78
(2,30)	0.9564	0	9
(3,30)	0.9489	0	2

Table 2.8: 35-node Chain Network PCS Breakdown

Design	Reliability	Percent Selection from KN	Percent Selection from Conventional Approach
(1,33)	0.9379	0	2
(1,34)	0.9450	0	12
(2,34)	0.9379	0	3
(1,35)	0.9521	100	64
(2,35)	0.9450	0	14
(3,35)	0.9308	0	5

CHAPTER 3
A NEW OPTIMIZATION FRAMEWORK FOR THE LARGE NETWORK RELIABILITY DESIGN PROBLEM

3.1 Introduction

In this chapter, we consider the most common case of the NRD problem in the literature which can also be considered the complement to the sNRD. This problem is the NRD optimization problem where total enumeration of the solution space is impractical or infeasible, i.e. the optimal design is selected by evaluating and comparing the estimated system reliability of a subset of the solution space. In this thesis, we refer to this problem as the large network reliability design problem (large NRD). Although this definition is subjective depending on the time each individual is willing to invest in the simulation, in this thesis we consider the large NRD as any NRD optimization problem with more than 1000 candidate designs.

The purpose of this chapter and subsequent contribution to the network reliability design literature is to provide a new optimization framework for the large NRD optimization problem that focuses on efficient and dynamic sampling within a given search heuristic. This framework consists of a population-based global guidance system, a selection-of-the-best subset procedure, and a duplicate evaluation process for eliminating redundant design evaluations. This procedure is geared toward ease of use, in that its performance is robust in terms of the input parameter selection, and the methodology is straightforward and easy to understand.

The specific global guidance system we use is the cross-entropy method. The cross-entropy method is a population-based meta-heuristic developed in 1997 by

Rubinstein [81] and often described as a Monte Carlo approach to combinatorial optimization. This method is a global search heuristic with backtracking, meaning it is capable of escaping local optimums due to its random search mechanism. The cross-entropy method is defined as the following two step iterative process:

1) Generate a sample of N designs from the solution space according to some random mechanism.
2) Update the parameters of the random mechanism, on the basis of the data, in order to produce a "better" sample in the next iteration.

We select this method as our global guidance system because it has already been proven both successful and easy to implement for the network reliability design problem [76-79, 82].

We couple the cross-entropy method with two novel statistical procedures designed to improve the sampling efficiency of the Monte Carlo simulation: Sequential Selection of the Best Subset (SSBS) and Duplicate Generation (DG). SSBS is a sequential sampling procedure and is based on the ideas of ranking and selection from the simulation optimization literature. Its purpose is to select a subset of each population of designs, using as few samples as statistically possible, that contains only the candidate designs better than the upper ρ-quantile of the population. Duplicate Generation on the other hand, is a statistical procedure designed to eliminate the evaluation of redundant designs yet still provide an independent system reliability estimate for each duplicate design index.

The remainder of this paper is organized as follows: Section 2 describes the large NRD optimization problem. Section 3 provides an overview of the cross entropy approach for network reliability design. Section 4 introduces our proposed optimization

framework and the details of both SSBS and DG. Section 5 provides numerical examples to quantify the performance and significance of our proposed framework and Section 6 provides a sensitivity analysis on the newly introduced input parameters as well as recommended settings.

3.2 Problem Formulation

For the large NRD optimization problem, we use the identical formulation as that of the sNRD in Chapter 2 but for $K > 1000$. Therefore, using the same variable, constant and function definitions as in Chapter 2, we summarize the formulation of the large NRD:

Let $G(V, E, T)$ represent a network with reliable node set $V = \{1, \dots, \tilde{n}\}$, unreliable edge set $E = \{1, \dots, \tilde{m}\}$, and terminal node set $T \subseteq V$. We then define the large NRD optimization problem as follows,

$$\max_{\boldsymbol{x}_i \in \theta} \hat{R}(\boldsymbol{x}_i) \tag{3.1}$$

where θ is the set of feasible solutions defined by the following constraints:

$$\begin{aligned} &\textstyle\sum_{e=1}^{\tilde{m}} x_{ie} c_e \le C_{max} && \text{(budget constraint)} \\ &0 \le x_{ie} \le 1, e \in E, i && \text{(edge constraint)} \\ &x_{ie} \in \mathbb{Z}^+ \cup \{0\}, e \in E, i && \text{(integer constraint)} \end{aligned} \tag{3.2}$$

where x_{ie}, c_e and C_{max} are the same as in chapter 2, $\hat{R}(\boldsymbol{x}_i)$ is the approximation of R_i based on Monte Carlo simulation using (2.8), θ is nonempty and finite containing a total of $K > 1000$ feasible designs, and the true optimal solution/s can be denoted by $\theta^* = \{i : R_i \ge R_j \ \forall j = 1, \dots, K\}$ where R_i is evaluated via (2.4).

3.3 The Cross Entropy Approach

In this section we describe the basic cross entropy (CE) approach for NRD [78]. This approach consists of five steps: initialization, generation, evaluation, updating, and stopping. We first present the entire procedure as a whole, shown in Algorithm 3.1, then explain the details involved in each step.

3.3.1 The Cross Entropy Approach for Network Reliability Design

Let $\boldsymbol{a}_t = (a_{t1}, a_{t2}, \dots, a_{t\tilde{m}})$ represent the vector of the edge purchase probabilities during iteration t for network $G(V, E, T)$, such that a_{te} is the probability of purchasing/selecting edge e during iteration t when generating a candidate design. The cross entropy approach can then be described by the five-step procedure in Algorithm 3.1 to iteratively update $\boldsymbol{a}_t$ based on the ρN best designs in iteration $t-1$, where $0 < \rho < 1$ and N is the number of randomly generated designs in each iteration, until $\boldsymbol{a}_t$ converges to the degenerate probability vector $\boldsymbol{a}^*$. This degenerate probability vector then corresponds to what is assumed to be θ^*.

Algorithm 3.1: The Cross Entropy Approach

Step 1) Initialization: Initialize $\hat{\boldsymbol{a}}_0$ such that $\hat{a}_{0e} = 0.5\ \forall e \in E$. Set $t = 1$.

Step 2) Generation: Generate a population of N candidate designs $[\boldsymbol{x}_1, \dots, \boldsymbol{x}_N]$ from θ using Algorithm 3.2 with $\boldsymbol{a}_t = \hat{\boldsymbol{a}}_{t-1}$.

Step 3) Evaluation: Using Monte Carlo simulation, evaluate the system reliability of each design in Step 2 using the same number of samples for each candidate design evaluation. Then determine $\hat{\gamma}_t$, the value of the candidate design with the $\lceil \rho N \rceil^{\text{th}}$ largest estimated performance for iteration t, using (3.3).

Step 4) Updating: Update $\hat{\boldsymbol{a}}_t$ using (3.4), where $\tilde{a}_{te}$ is described by (3.5).

Step 5) Stopping: If $\max_{e \in E}(\min(\hat{a}_{te}, 1 - \hat{a}_{te})) \leq \beta$, then stop and select $\hat{\boldsymbol{a}}_t$ as the optimal solution; else, set $t = t + 1$ and return to Step 2.

We next describe the details of each step in Algorithm 3.1.

3.3.2 Step 1 of Algorithm 3.1: Generation

In the generation step, a random mechanism is used to generate N candidate designs from θ, based on the edge purchase probability vector $\boldsymbol{a}_t = (a_{t1}, a_{t2}, \dots, a_{t\tilde{m}})$. The mechanism begins by generating a random permutation $\pi = (e_1, \dots, e_{\tilde{m}})$ of all $\tilde{m}$ edges in E, where e_k represents the k^{th} edge index of permutation π. Then starting with $k = 1$, purchase edge e_k with probability $a_{te_k} \in \boldsymbol{a}_t$ so long as the cost of edge e_k (i.e. c_{e_k}) does not exceed the remaining allowable budget. If edge e_k is purchased, set $x_{ie_k} = 1$ where x_{ie_k} is the binary state of edge e_k; otherwise set $x_{ie_k} = 0$. Repeat this procedure sequentially for each edge in π or until the allowable budget is exhausted. The outline for this procedure is given in Algorithm 3.2.

Algorithm 3.2: Design Generation

Step 1) Set $k = 1$, $Budget = C_{max}$, and $x_{ie_k} = 0$ for $k = 1, \dots, \tilde{m}$.

Step 2) Generate a uniform random permutation $\pi = (e_1, \dots, e_{\tilde{m}})$ of edges $(1, \dots, \tilde{m})$.

Step 3) If $c_{e_k} \leq Budget$ with probability a_{te_k} set $x_{ie_k} = 1$ and $Budget = Budget - c_{e_j}$.

Step 4) If $k = \tilde{m}$ or $Budget = 0$, stop; else set $k = k + 1$ and return to Step 3.

3.3.3 Steps 3 and 4 of Algorithm 3.1: Evaluation & Updating

Steps 3 and 4 of the cross entropy approach concern the evaluation and updating of $\boldsymbol{a}_t$ to direct the search to improving regions of θ. Step 3 evaluates the system reliability for each of the N candidate designs via a Monte Carlo simulation and sets $\hat{\gamma}_t$ to

the $\lceil \rho N \rceil^{\text{th}}$ largest system reliability estimate of the N candidate designs in iteration t, such that

$$\hat{\gamma}_t = \hat{R}_{(\lceil \rho N \rceil)} \tag{3.3}$$

$\hat{a}_{te}$ is then updated in Step 4 as follows,

$$\hat{\boldsymbol{a}}_t = \lambda \tilde{\boldsymbol{a}}_t + (1-\lambda)\hat{\boldsymbol{a}}_{t-1} \tag{3.4}$$

where λ is a smoothing parameter typically assuming values between $0 < \lambda < 1$ and $\tilde{\boldsymbol{a}}_t = (\tilde{a}_{t1}, \tilde{a}_{t2}, \ldots, \tilde{a}_{t\tilde{m}})$, such that

$$\tilde{a}_{te} = \frac{\sum_{i=1}^{N} I_{\{\hat{R}(\boldsymbol{x}_i) \geq \hat{\gamma}_t\}} x_{ie}}{\sum_{i=1}^{N} I_{\{\hat{R}(\boldsymbol{x}_i) \geq \hat{\gamma}_t\}}} \; \forall e = 1, \ldots, \tilde{m} \tag{3.5}$$

where $\tilde{a}_{te}$ is the approximation of a_{te}, such that

$$a_{te} = E_{a_{t-1}}[X_e | R(\boldsymbol{x}_i) \geq \gamma_t] \;\; \forall e \in E \tag{3.6}$$

and $I_{\{\hat{R}(\boldsymbol{x}_i) \geq \hat{\gamma}_t\}}$ is an indicator function, such that

$$I_{\{\hat{R}(\boldsymbol{x}_i) \geq \hat{\gamma}_t\}} = \begin{cases} 1 & \text{if } \hat{R}(\boldsymbol{x}_i) \geq \hat{\gamma}_t \\ 0 & \text{otherwise} \end{cases}$$

Although (3.5) is capable of directing the search to find θ^*, depending on the size of N, direct use of (3.5) can still lead to suboptimal convergence by enabling $\hat{\boldsymbol{a}}_t$ to converge too quickly. To avoid this, $\hat{\boldsymbol{a}}_t$ is updated via the method of exponential smoothing in (3.4) as this enables all past iteration data to be accounted for in the updating of $\hat{\boldsymbol{a}}_t$ and subsequently reduce the potential for suboptimal convergence.

3.3.4 Step 5 of Algorithm 3.1: Stopping

The cross entropy approach terminates when its stopping criterion, given in step 5, has been fulfilled. This criterion specifies that if the algorithm is to terminate in iteration t, then every individual purchase probability $\hat{a}_{te}$ must be within β of 0 or 1,

where β is a user-specified parameter with range $0 < \beta < 1$ and typically assuming values within $0.01 < \beta < 0.1$.

3.4 Proposed Framework

In this section, we present our proposed optimization framework: Procedure CE+SSBS+DG. This framework consists of seven steps: initialization, generation, classification, SSBS, DG, updating, and stopping. We first present the procedure as a whole, shown in Algorithm 3.3, then explain the details involved in each newly added step and corresponding motivation.

Algorithm 3.3: Procedure CE+SSBS+DG

Step 1) Initialization: Set the iteration counter to $t = 1$ and the purchase probability vector to $\widehat{\boldsymbol{a}}_0 = \{0.5, 0.5, \ldots, 0.5\}$ such that $\hat{a}_{te} = 0.5 \; \forall e \in E$.

Step 2) Generation: Generate a population of N candidate designs $[\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N]$ from θ using Algorithm 3.2 with $\boldsymbol{a}_t = \widehat{\boldsymbol{a}}_{t-1}$.

Step 3) Classification: Partition the set of design indices generated in Step 2 into the pair-wise disjoint sets: I, J, $\boldsymbol{G}$ and $G_i \; \forall i \in J$, according to

Step 4) Algorithm 3.4 in Section 3.4.1. The classification for these sets is based on the number of duplicates of each design generated in Step 2.

Step 5) SSBS: Evaluate the designs in $I \cup J$ using Algorithm 3.5 in Section 3.4.2.

Step 6) DG: Generate a pseudo-random system reliability estimate for each design index in $\boldsymbol{G}$ using Algorithm 3.6 in section 3.4.3.

Step 7) Updating: Determine $\widehat{\boldsymbol{a}}_t$ using (3.4), where $\tilde{a}_{te}$ is described by (3.5) and $\hat{\gamma}_t$ is described by (3.3) for the set of designs in $H = I \cup J \cup \boldsymbol{G}$.

Step 8) Stopping: If $\max_{e \in E}(\min(\hat{a}_{te}, 1 - \hat{a}_{te})) \leq \beta$, then stop and select $\widehat{\boldsymbol{a}}_t$ as the optimal solution; else, set $t = t + 1$ and return to Step 2.

Thus, Algorithm 3.3 is basically an extension of Algorithm 3.1 that replaces Step 3 of Algorithm 3.1 with the Classification phase in Step 3, SSBS in Step 4, and DG in Step 5 of Algorithm 3.3. Step 6 of Algorithm 3.3 then selects the ρN designs with the largest system reliability estimates as the ρN best, just as done in Algorithm 3.1 when using the conventional approach. Therefore, in the worst case scenario where no designs are eliminated in Steps 3-5 of Algorithm 3.3, CE+SSBS+DG will perform no worse than Algorithm 3.1.

We next describe the details of each new step in Algorithm 3.3 as well as discuss the motivation for their inclusion in the optimization framework.

3.4.1 Step 3 of Procedure CE+SSBS+DG: Classification

Step 3 of Algorithm 3.3 serves as a means of partitioning the population of N design indices generated in Step 2 into the pair-wise disjoint sets: I, J, $\boldsymbol{G}$ and $G_i\ \forall i \in J$; where I is the set of design indices with less than δ total duplicates, J is a set of design indices representing all unique designs in the population that have δ or more total duplicates, and G_i is the set of design indices not in J that are duplicates of design $i \in J$. We use ω to denote the set of unassigned design indices and $\boldsymbol{G}$ to denote the set of all duplicate designs not in J, i.e. $\boldsymbol{G} = \left(\bigcup_{i \in J} G_i\right)$. The outline for this procedure is given in Algorithm 3.4.

Algorithm 3.4: Design Classification Procedure

Step 3.1) Set $I = \emptyset, J = \emptyset, \omega = (1, \ldots, N)$

Step 3.2) for $i \in \omega$

set $G_i = \{j : j \in \omega \ and \ \boldsymbol{x}_j = \boldsymbol{x}_i\}$ and $\omega = \omega \backslash G_i$.

if $|G_i| < \delta$

$I = I \cup G_i$

$G_i = \emptyset$

if $|G_i| \geq \delta$

$J = J \cup \{i\}$

$G_i = \mathrm{G_i} \backslash \{\mathrm{i}\}$

end

end

Step 3.3) Set $\boldsymbol{G} = \bigcup_{i \in J} G_i$

We illustrate the use of

Algorithm 3.4 through the following example: suppose a population of seven randomly generated topologies $[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_7]$ represent a total of four designs $[A, B, C, D]$ such that $[\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3]$ each represent design A, $[\boldsymbol{x}_4, \boldsymbol{x}_5]$ each represent design B, $[\boldsymbol{x}_6]$ represents design C, and $[\boldsymbol{x}_7]$ represents design D. If $\delta = 2$, then following

Algorithm 3.4, the indices of these seven designs can be partitioned as follows:

1) All indices are unassigned.

   $I = \emptyset, J = \emptyset, \omega = \{1,2,3,4,5,6,7\}$

2) For design A: since $|\{1,2,3\}| \geq \delta$, assign the first index to set J and the remaining two indices to the corresponding set G_i.

   $I = \emptyset, J = \{1\}, G_1 = \{2,3\}, \omega = \{4,5,6,7\}$

3) For design B: since $|\{4,5\}| \geq \delta$, assign the first index to set J and the remaining index to the corresponding set G_i.

   $I = \emptyset, J = \{1,4\}, G_1 = \{2,3\}, G_4 = \{5\}, \omega = \{6,7\}$

4) For design C: since $|\{6\}| < \delta$, assign this index to set I.

   $I = \{6\}, J = \{1,4\}, G_1 = \{2,3\}, G_4 = \{5\}, \omega = \{7\}$

5) For design D: since $|\{7\}| < \delta$, assign this index to set I, and compute $\boldsymbol{G}$.

   $I = \{6,7\}, J = \{1,4\}, G_1 = \{2,3\}, G_4 = \{5\}, \boldsymbol{G} = \{2,3,5\}\ \omega = \emptyset$

3.4.2 Step 4 of Procedure CE+SSBS+DG: Sequential Selection of the Best Subset

In Algorithm 3.1, the conventional cross entropy approach based on Monte Carlo simulation uses an equal number of Monte Carlo samples in the evaluation of each design, regardless of solution quality. However, this type of sampling approach is often inefficient as many inferior designs will not require this level of accuracy to determine if they are indeed inferior. To improve upon this, we introduce a procedure called SSBS – Sequential Selection of the Best Subset, to provide efficient selection of the ρN best designs, based on statistical significance, using as few samples as possible.

SSBS is a statistical sampling procedure designed for simple yet efficient selection of the ρN best designs from a population of N competing designs. SSBS is fully sequential with elimination, meaning it simultaneously samples every design in consideration, one observation at a time, and eliminates designs from further sampling and contention for the best as soon as they are deemed inferior by a significant difference. This procedure thus improves upon this original approach by minimizing the number of Monte Carlo samples required to find the ρN best designs by eliminating inferior designs from further sampling and contention for the best at the earliest possible time.

Step 4 of Algorithm 3.3 is where we employ SSBS. In this step, the system reliability of the indices in sets I and J (generated in Step 2 and classified in Step 3) are evaluated using SSBS with screening and the Monte Carlo simulation without screening respectively, while the indices in set $\boldsymbol{G}$ are not evaluated as they will be dealt with later in Step 5. The reason that the indices in J are evaluated differently than the indices in I is because the evaluation of the duplicates in $G_i \; \forall i \in J$ in Step 5 are dependent on the evaluation of $i \in J$, thus, we do not want any designs in J to potentially be eliminated by the screening mechanism in SSBS.

Algorithm 3.5: Sequential Selection of the Best Subset

Step 4.1) Set $H = I \cup J \cup \boldsymbol{G}$.

Step 4.2) Take n_0 samples for each design in $I \cup J$. This is necessary to ensure that the initial reliability estimates are approximated appropriately. Set $n_{count} = n_0$.

Step 4.3) Set $\hat{R}_i = \begin{cases} \hat{R}_i(n_{count}) & \forall i \in I \cup J \\ \hat{R}_j(n_{count}) & \forall j \in J, i \in G_j \end{cases}$

Step 4.4) Determine the sets of the upper and lower confidence limits $\boldsymbol{UCL}$ and $\boldsymbol{LCL}$ respectively, where $\boldsymbol{LCL} = \{LCL_i : i \in H\}$, such that

$$LCL_i = \hat{R}_i - z_{1-\alpha/2}\sqrt{\frac{\hat{R}_i(1-\hat{R}_i)}{n_{count}}}$$

and $\boldsymbol{UCL} = \{UCL_i : i \in H\}$, such that

$$UCL_i = \hat{R}_i + z_{1-\alpha/2}\sqrt{\frac{\hat{R}_i(1-\hat{R}_i)}{n_{count}}}$$

Step 4.5) Set $\hat{\gamma}_t^{LCL}$ to the $\lceil \rho N \rceil^{\text{th}}$ largest lower confidence limit of the population of N designs, such that $\hat{\gamma}_t^{LCL} = LCL_{(\lceil \rho N \rceil)}$

Step 4.6) Eliminate the designs in I with UCL_i smaller than $\hat{\gamma}_t^{LCL}$, that is: set $I_0 = I$ and $I = \{i : i \in I_0 \ and \ UCL_i \geq \hat{\gamma}_t^{LCL}\}$.

Step 4.7) Update H in accordance with the new set of indices defined by I.

Step 4.8) If $n_{count} < n \ and \ |H| > \rho N$, where n is the allocated maximum sample size for each iteration, take Δn additional samples for each design in $I \cup J$, set $n_{count} = n_{count} + \Delta n$, and return to Step 4.3. Else, stop and return the indices in H as the set of designs still in contention for the ρN best, and their corresponding reliability estimates ($\hat{R}_i$) as the evaluation of their system reliabilities.

Procedure SSBS begins with an initial sampling phase where n_0 Monte Carlo samples are collected for every design in $I \cup J$ in order to generate an initial system

reliability estimate for every unique design in consideration. This initial sampling phase is used as a precautionary measure to ensure that all system reliability estimates are approximated appropriately before any elimination occurs. It then creates copies of the reliability estimates for the design indices in J as temporary reliability estimates for the corresponding design indices in $\boldsymbol{G}$ in order to determine $\hat{\gamma}_t^{LCL}$.

The procedure then begins what we call the screening phase, designated by Steps 4.4-4.6. This phase begins by calculating the upper and lower confidence limits associated with the system reliability estimate of each design index in $H = I \cup J \cup \boldsymbol{G}$. It then determines the $\lceil \rho N \rceil^{\text{th}}$ largest lower confidence limit, designated by $\hat{\gamma}_t^{LCL}$, and eliminates all design indices in I whose upper confidence limit of its system reliability estimate are less than $\hat{\gamma}_t^{LCL}$. The procedure then collects an additional Δn samples and returns to Step 4.3 to repeat the screening phase. This sampling and screening process then iterate sequentially until either the user-allocated simulation effort has been exhausted (i.e. the maximum number of samples has been collected, $n_{count} = n$), or the desired number of candidate designs has been obtained within the best subset. This maximum sample criterion is used as a failsafe to keep SSBS from potentially iterating indefinitely when trying to differentiate between two very similar or duplicate designs. It then returns the indices in H, and their corresponding system reliability estimates, as the set of designs still in contention for ρN best.

3.4.2.1 Confidence Limits in SSBS

Since the method of evaluation used in SSBS is Monte Carlo simulation, whose individual samples follow a Bernoulli distribution, the screening device in Step 5 of Procedure SSBS requires the computation and comparison of binomial confidence limits. However, exact evaluation of binomial confidence limits, otherwise known as the Clopper-Pearson Confidence Interval, requires the following calculation [83]:

$$\{\phi | P[Bin(n;\phi) \leq X] \geq \alpha/2\} \cap \{\phi | P[Bin(n;\phi) \geq X] \geq \alpha/2\} \tag{3.7}$$

where X is the number of successes observed in the sample, α is the uncertainty level associated with the confidence interval, and $Bin(n;\phi)$ is a binomial random variable with n trials and ϕ probability of success. However, calculation of the Clopper-Pearson confidence interval is not easy, thus given that SSBS requires evaluation of the binomial confidence interval thousands of times for even small problems, (3.7) is not a feasible option for this procedure. Therefore, approximation methods are necessary.

The two most common approaches for approximating the confidence limits are the normal approximation or Wald confidence interval:

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{3.8}$$

and the Wilson score confidence interval [84]:

$$\frac{\hat{p} + \frac{1}{2n}z_{1-\alpha/2}^2 \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{1-\alpha/2}^2}{4n^2}}}{1 + \frac{1}{n}z_{1-\alpha/2}^2} \tag{3.9}$$

where $\hat{p}$ is the outcome of a Bernoulli trials process, n is the number of Monte Carlo samples observed, and $z_{1-\alpha/2}$ is the $1-\alpha/2$ percentile of the standard normal distribution. Although both approximations can be used interchangeably in SSBS, we present SSBS using (3.8) because it is simpler. For further discussion on this selection, we refer the reader to section 3.5.

3.4.3 Step 5 of Procedure CE+SSBS+DG: Duplicate Generation

In many iterations of the cross-entropy method, especially ones close to convergence, many designs generated by the random mechanism are duplicates (i.e.

different design indices that represent the same topology), yet still require independent evaluation in order to minimize the effects of incurring any type I errors associated with approximating the system reliabilities in the optimization problem. The purpose of this step therefore, is to provide fast pseudo-sampling of duplicate designs without the use of the expensive Monte Carlo simulation. This is accomplished through a procedure we call Duplicate Generation (DG).

DG is a very simple and straight forward procedure that uses the distribution of each system reliability estimate in J to quickly generate i.i.d. system reliability estimates for each design index in G_i $\forall i \in J$ by approximating a Monte Carlo simulation via a random-number generator.

The idea behind this is that since all indices in $G_i \cup \{i\}$ represent the same design with unknown true reliability R_i, each system reliability estimate generated from a Monte Carlo simulation will be an i.i.d. random number according to $\frac{Bin(n;R_i)}{n}$. Therefore, instead of using a Monte Carlo simulation to generate estimates from this unknown distribution, DG approximates R_i of design $i \in J$ via $\hat{R}_i(n)$, then generates pseudo system reliability estimates for each design index in G_i via $\frac{Bin(n;\hat{R}_i(n))}{n}$. The outline for DG is provided in Algorithm 3.6.

Algorithm 3.6 : Duplicate Elimination – Regeneration

Step 5.1) Set $\hat{R}_i = \hat{R}_i(n) \;\; \forall i \in I \cup J$

Step 5.2) Set $\hat{R}_i = \frac{Bin\left(n;\hat{R}_j(n)\right)}{n} \;\; \forall j \in J, \; i \in G_j$

3.5 Numerical Experiments

In this section, we present three numerical examples to quantify the efficiency and selection ability of our proposed optimization framework. We first compare the

performance of CE+SSBS+DG with CE, and assess the individual as well as combined effect of DG and SSBS on CE. We then compare the performance of CE+SSBS+DG when using the two different confidence intervals outlined in (3.8) and (3.9) and how this choice affects the procedures performance.

3.5.1 Example Networks

We represent each example network by a graph of nodes and edges, where the nodes are labeled alphabetically, the edges are labeled numerically, and the terminal nodes (i.e. the nodes whose connectivity define the networks system reliability state) are represented via darkened nodes. Individual edge reliabilities and costs are then provided in the corresponding tables according to each edge index.

Example 1: Example 1 is the 6-node complete network shown in Figure 3.1. This is a common problem in the literature referred to as a two-terminal network, as only two nodes, a source (a) and a sink (f) need to be able to communicate for the network to be considered operational. For this problem, the individual link costs and reliabilities are provided in Table 3.1, the maximum allowable budget was set to 1500, and the remaining framework and cross-entropy parameters were set to: $n = 20{,}000$, $N = 750$, $\rho = 0.1$, $\Delta n = 50$, and $n_0 = 200$. In all, there are a total of 2^{15} candidate designs in this network, with optimal design displayed in Figure 3.2 having system reliability of 0.999928033.

Example 2: Example 2 is the 7-node complete network shown in Figure 3.3. This is another common problem in the literature referred to as an all-terminal network, as every node in the network must be able to communicate with every other node in the network for the system to be considered operational. For this example, the individual link costs and reliabilities are randomly generated according to $c_i \sim \mathrm{U}(10,30)$ and $r_i \sim \mathrm{U}(0.91,0.99)$ respectively and are provided in Table 3.2, the maximum allowable budget was set to 125, and the remaining framework and cross-entropy parameters were set to: $n = 4000$, $N = 1000$, $\Delta n = 20$, $n_0 = 100$, and $\rho = 0.05$. Altogether, there are a

total of 2^{21} candidate designs in this example and the optimal solution is unknown. However, the best found design (based on our simulation) is depicted in Figure 3.4 with estimated system reliability of 0.9934, based on a 10-million sample Monte Carlo simulation.

Example 3: Example 3 is the 9-node complete network shown in Figure 3.5. This example is another all-terminal design problem but with a much larger solution space. For this example, the individual link costs and reliabilities were randomly generated according to $c_i \sim \mathrm{U}(10,99)$ and $r_i \sim \mathrm{U}(0.80,0.99)$ respectively and are provided in Table 3.3, the maximum allowable budget was set to 300, and the other framework and cross-entropy parameters were set to: $n = 2000$, $\rho = 0.025$, $N = 2000$, $\Delta n = 20$, and $n_0 = 100$. Altogether, there are a total of 2^{36} candidate designs for this network and the optimal design is unknown. However, the best found design (based on our simulation) is depicted in Figure 3.6 having estimated system reliability of 0.9490, based on a 10-million sample Monte Carlo simulation.

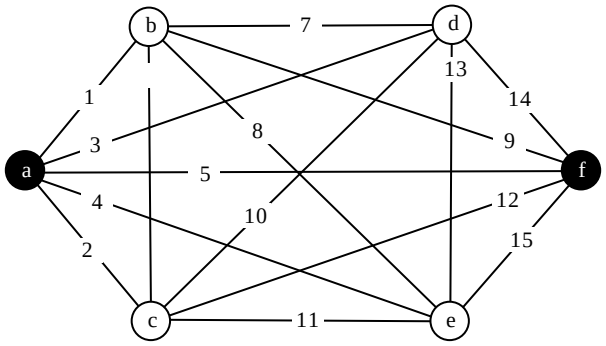

Figure 3.1: 6-node complete two-terminal graph.

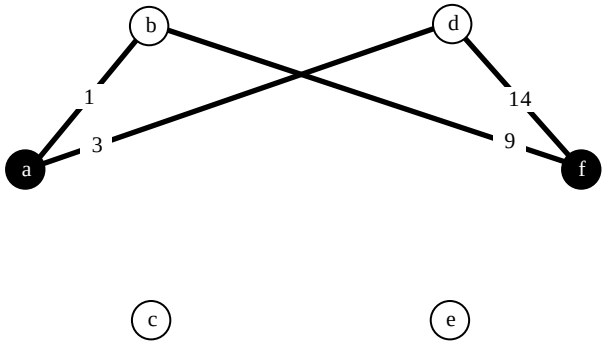

Figure 3.2: Optimal design for Example 1.

Table 3.1: Example 1 Individual Link Cost and Reliabilities.

i	c_i	r_i	i	c_i	r_i	i	c_i	r_i
1	331	0.9951	6	335	0.9958	11	330	0.9947
2	347	0.9968	7	332	0.9952	12	325	0.9937
3	327	0.9942	8	302	0.9902	13	324	0.9935
4	340	0.9959	9	344	0.9964	14	350	0.9973
5	2000	0.9908	10	315	0.9917	15	312	0.9912

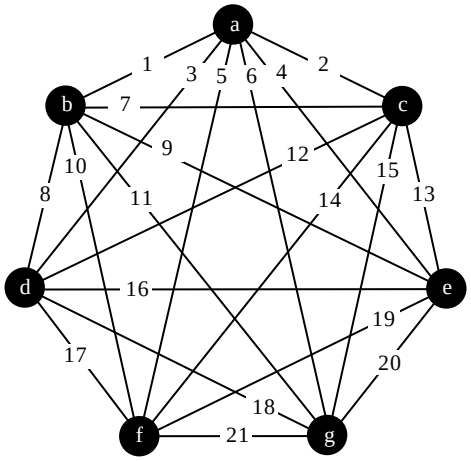

Figure 3.3: 7-node complete all-terminal graph.

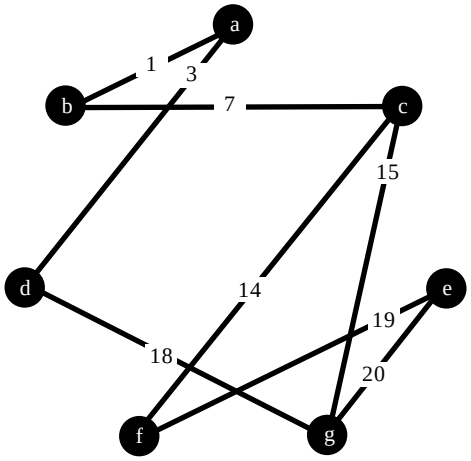

Figure 3.4: Best found design for Example 2.

Table 3.2: Example 2 Individual Link Cost and Reliabilities.

i	c_i	r_i	i	c_i	r_i	i	c_i	r_i
1	14	0.99	8	25	0.94	15	11	0.96
2	30	0.98	9	26	0.93	16	18	0.92
3	25	0.95	10	15	0.94	17	16	0.99
4	23	0.99	11	15	0.91	18	13	0.92
5	15	0.91	12	26	0.96	19	13	0.98
6	24	0.94	13	20	0.94	20	15	0.98
7	20	0.95	14	24	0.92	21	27	0.91

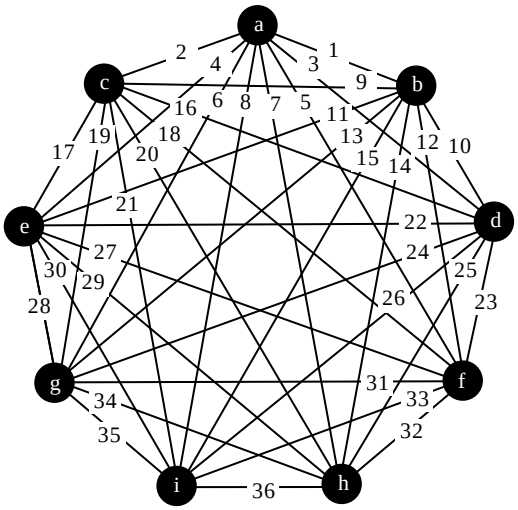

Figure 3.5: 9-node complete all-terminal graph.

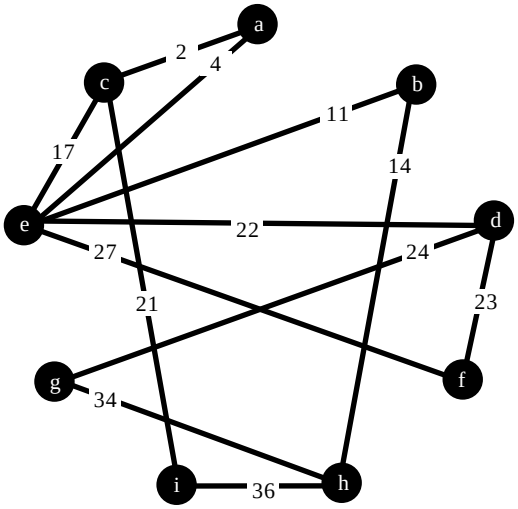

Figure 3.6: Best found design for Example 3.

Table 3.3: Example 3 Individual Link Cost and Reliabilities.

i	c_i	r	i	c_i	r_i	i	c_i	r_i
1	78	0.94	13	29	0.87	25	73	0.86
2	51	0.93	14	13	0.98	26	49	0.89
3	87	0.81	15	52	0.88	27	26	0.93
4	20	0.82	16	62	0.82	28	32	0.93
5	90	0.88	17	17	0.98	29	49	0.97
6	71	0.81	18	44	0.88	30	62	0.87
7	80	0.83	19	25	0.80	31	42	0.80
8	55	0.82	20	92	0.91	32	58	0.87
9	81	0.86	21	23	0.93	33	68	0.91
10	55	0.82	22	23	0.82	34	24	0.91
11	23	0.87	23	12	0.90	35	40	0.90
12	97	0.89	24	30	0.90	36	37	0.86

For each example, the allowable type I error used in SSBS is set to $\alpha = 0.05$, the smoothing parameter is set to $\lambda = 0.7$, the stopping parameter is set to $\beta = 0.1$, and the maximum allowable duplicates is set to $\delta = 10$ (when applicable). All simulations are run on an HP xw4600 workstation with Intel® Core™ Duo CPU E8400@3.0Ghz processor, and each method is replicated a total of ten times for each example network. All corresponding data for the independent replications can be found in Appendix F through Appendix H.

3.5.2 Breakdown of Performance

To quantify the performance of CE+SSBS+DG and assess the individual as well as combined effects of DG and SSBS on CE+SSBS+DG, we compare the performance of CE with that of: CE+DG, CE+SSBS, and CE+SSBS+DG on each of the three example networks.

Table 3.4 through Table 3.6 summarize the results of the ten independent replications for each method in Examples 1-3 respectively, where the raw data for each iteration is provided in Appendix F through Appendix H respectively. In these tables, "Method" refers to the optimization method used in the simulation, "Frequency of Best" refers to the percentage of time the optimal design was found in the ten independent replications, "Average Iterations" refers to the average number of cross-entropy iterations required for $\boldsymbol{a}$ to converge to $\boldsymbol{a}^*$, and "Average Time" and "Average Samples" refer to the average total time and average total number of Monte Carlo samples required for the simulation to converge to $\boldsymbol{a}^*$ respectively.

In all three examples, each method converges to the optimal design in most replications, as indicated by 'Frequency of Best', where any suboptimal convergence could be attributed to the standard error inherent to the simulation. This suggests that all methods perform similarly in terms of their ability to correctly select the best design, thus can be compared on even terms.

Average cross-entropy iterations to convergence are also fairly consistent from method to method in all three examples, requiring an average of 6.8, 10.6, and 10.5 iterations for Examples 1-3 respectively. This consistency suggests that the cross-entropy search heuristic portion of each method is fairly stable and not much affected by the type of evaluation method used. Methods using DG do tend to have slightly higher average iterations, requiring: 1.053, 1.261, and 1.138 more iterations on average for Examples 1-3 respectively. However, any cost associated with this increase in Average Iterations due to DG is always less than its benefit, as indicated by the overall reduction in Average Time and Average Samples between CE and CE+DG as well as CE+SSBS and CE+SSBS+DG, shown in Tables 3.4-3.6.

Of the four methods we tested, CE+SSBS+DG has the overall best performance in every replication of every example, and CE consistently has the worst. To put these results into better perspective, when using CE as the basis for comparison: CE+DG reduces Average Samples by a factor of 1.27, 1.18, and 1.18 for Examples 1-3 respectively; CE+SSBS reduces Average Samples by a factor of 2.96, 4.24, and 7.29 for Examples 1-3 respectively; and CE+SSBS+DG reduces Average Samples by a factor of 4.64, 9.29, and 15.39 for Examples 1-3 respectively; where Average Time follows a pattern similar and comparable to Average Samples.

Overall, the effect of DG on average samples seems to be independent of the size of the solution space, reducing average samples by approximately 20% in all three examples. The effect of SSBS, however, seems to be dependent on the size of the solution space, such that the larger the solution space is the more effective SSBS will be in reducing average samples.

To test the significance of the effects of DG and SSBS in CE+SSBS+DG, we look at an ANOVA for the full factorial model for each example network, using the data provided in Appendix F through Appendix H. Using the log transformation for Average Samples as our dependent variable, based on the Box-Cox transformation and its

Table 3.4: Example 1 Summary Statistics.

Method	Frequency of Best	Average Iterations	Average Time (seconds)	Average Samples
CE	100%	6.5	37,024	82,500,000
CE+DG	100%	7	29,141	65,132,000
CE+SSBS	90%	6.8	10,891	27,894,494
CE+SSBS+DG	100%	7	8,328	17,787,130

Table 3.5: Example 2 Summary Statistics.

Method	Frequency of Best	Average Iterations	Average Time (seconds)	Average Samples
CE	100%	9.2	15,024	32,800,000
CE+DG	90%	11.6	13,568	27,785,200
CE+SSBS	100%	9.6	2,391	6,225,738
CE+SSBS+DG	100%	12.1	1,962	3,530,522

Table 3.6: Example 3 Summary Statistics.

Method	Frequency of Best	Average Iterations	Average Time (seconds)	Average Samples
CE	100%	9.8	26,621	70,400,000
CE+DG	90%	11.3	25,347	59,652,200
CE+SSBS	100%	9.8	2,825	9,660,030
CE+SSBS+DG	100%	11	1,962	4,573,660

associated diagnostic plots, the individual main effects of DG and SSBS are both significant at the $p \ll 0.01$ level in all three examples. In addition to this, the interaction effect of DG with SSBS is also significant at the $p < 0.01$ level for all three examples, indicating that there is an added benefit of using these two procedures together.

3.5.3 Comparison of Confidence Intervals

To determine how the choice of confidence interval will affect the overall performance of CE+SSBS+DG, we compare the performance of CE+SSBS+DG using the binomial approximations described in (3.8) and (3.9) for all three example networks. Each example is conducted using the same experimental conditions described in section 3.5.1.

Table 3.7 summarizes the results of the ten independent replications for each method for each example network (for the raw results see Appendix I). The optimal design is again selected in most replications. Average iterations are very consistent from method to method with little to no difference between methods. Average samples for CE+SSBS+DG are slightly lower when using (3.8), requiring on average 7.5%, 6.5%, 5.1% less samples than (3.9) for Examples 1-3 respectively; however based on ten replications, this difference in Average Samples is not significant, having p-values of $p = 0.1031$, $p = 0.3359$, and $p = 0.4636$ for Examples 1-3 respectively. This suggests that there is little to no benefit to using one approximation over the other. Despite this, we still recommend (3.8), as although the two methods are statistically equivalent in terms of the performance of CE+SSBS+DG, (3.8) is simpler, easier to implement, and more common.

Table 3.7: Summary Statistics for Confidence Interval Analysis

Example	Method	CI	Frequency of Best	Average Iterations	Average Samples
1	CE+SSBS+DG	(3.8)	100%	7	17,787,130
		(3.9)	100%	6.6	19,116,660
2	CE+SSBS+DG	(3.8)	100%	12.1	3,530,522
		(3.9)	100%	12.1	3,760,074
3	CE+SSBS+DG	(3.8)	100%	11	4,573,660
		(3.9)	90%	11.3	4,808,214

3.6 Parameter Sensitivity

In this section, we investigate how the level of α, the allowable type I error associated with each individual confidence interval used in SSBS, as well as how the level of n, the maximum number of Monte Carlo samples allowed for each design estimate, effect the overall performance of CE+SSBS+DG.

3.6.1 α Sensitivity

Recall that for CE+SSBS+DG, α only represents the allowable type I error inherent in pair-wise comparisons, not the entire procedure. Therefore, for the proposed framework, α is considered an input parameter, not a lower bound on the probability of making a correct selection. Thus, the purpose of this experiment is to see how the level of α effects the overall performance of CE+SSBS+DG, and if we can recommend a level of α for general use within the proposed framework.

For this experiment, we return to the network in Figure 3.3 using the same costs, reliabilities, and parameter settings as in Example 2, but this time varying the level of α over the range $\alpha = [0.01,0.60]$. Ten independent replications are observed for each level and used to determine the frequency of selecting the best design, average number of

iterations required to converge, and average total number of Monte Carlo samples required to converge.

Table 3.8: α Sensitivity Summary Statistics.

α	Frequency of Best	Average Number of Iterations	Average Total Samples
0.01	100%	11.6	3,898,040
0.05	100%	12.1	3,530,522
0.10	100%	11.9	2,663,818
0.20	100%	14.8	2,426,542
0.30	90%	11.7	2,043,940
0.40	90%	12.5	1,744,512
0.50	80%	12.9	1,573,616
0.60	70%	11.8	1,136,604

Table 3.8 summarizes the results in this experiment (for the raw results see Appendix J). For this example, it appears that the larger α gets the smaller Total Samples becomes. This is only true to a point however, as for $\alpha \geq 0.3$, the Frequency of Best starts declining as α increases, as the confidence intervals are failing to represent the true means thus leading to frequent suboptimal convergence. This relationship is as expected though, because we know that for a two-sided confidence interval, as α approaches 1.0 the half-width for the corresponding confidence interval approaches 0, meaning each confidence interval will be more and more likely to misrepresent it's corresponding true mean, resulting in a higher likelihood of incurring errors in the selection of the ρN best designs and subsequently increasing the risk of suboptimal convergence.

This presents us with the typical trade-off: efficiency vs. accuracy, which leads to the question: what range of α is acceptable for general use in this type of problem?

Although we know, based on this example, that large values of α will likely work for many networks, we still recommend a conservative setting for this parameter such that $\alpha \sim (0.05, 0.10)$. This is because we cannot generalize the results from one example to every network, so we believe that it is better to err on the side of caution in this case in order to reduce the risk of suboptimal convergence, despite this parameter setting being possibly less efficient in some cases.

3.6.2 n Sensitivity

When using Monte Carlo simulation as the evaluation mechanism within a heuristic search such as the Cross Entropy Approach, selection of n is important in terms of the procedure efficiency as well as the solution quality. This is because if too few samples are used, approximates will not accurately represent the true system reliabilities and subsequently lead to suboptimal convergence; whereas if too many samples are used, the evaluation process can be unnecessarily expensive. However, a-priori selection of efficient n is typically not possible in this type of problem without prior information of the given network. Therefore, in this experiment, we investigate how the level of n effects the overall performance of CE+SSBS+DG as compared to CE.

To do this, we again return to the network in Figure 3.3 using the same costs, reliabilities, and parameter settings as in Example 2, but varying the level of n for each method over the values $n = \{3000, 4000, 5000, 6000, 8000, 10000\}$. Ten independent replications were observed for each treatment of parameters for each method and used to determine the frequency of selecting the best design, average number of CE iterations required to converge, and average total number of Monte Carlo samples required to converge.

Table 3.9 summarizes the results of ten independent replications for each level of n for each method (for the raw results see Appendix K). For this experiment, we make our comparison over the range $3000 \leq n \leq 10000$ as for $n < 3000$ both methods begin

having consistent suboptimal convergence, indicating that the sample size was too small to appropriately approximate the system reliabilities. Figure 3.7 shows a plot of n vs Total Samples for this range, where the dashed and solid lines represent the CE and CE+SSBS+DG methods respectively.

Table 3.9: n Sensitivity Summary Statistics.

Method	n	Frequency of Best	Average Iterations	Average Total Samples
CE+SSBS+DG	3,000	100%	12.2	2,906,994
	4,000	100%	12.1	3,530,522
	5,000	100%	11.8	3,949,232
	6,000	100%	10.4	4,243,072
	8,000	100%	10.9	5,453,556
	10,000	100%	10	5,974,300
CE	3,000	90%	9.9	26,700,000
	4,000	100%	9.2	32,800,000
	5,000	100%	9.3	41,500,000
	6,000	100%	9.4	50,400,000
	8,000	100%	9.4	67,200,000
	10,000	100%	9.4	84,000,000

Clearly, CE+SSBS+DG requires much fewer samples to converge to the optimal solution than that of CE. At their minimum values, total samples for CE is still approximately three times that of total samples for CE+SSBS+DG. Also, the rate of increase in total samples with respect to n is much smaller for CE+SSBS+DG than CE, where total samples for CE is approximately proportional to n, and total samples for CE+SSBS+DG is approximately logarithmically proportional to n. Thus, CE+SSBS+DG is much more robust in its overall performance in terms of the level of n selected, as n

does not have nearly as strong of an influence on total samples in CE+SSBS+DG as it does in CE.

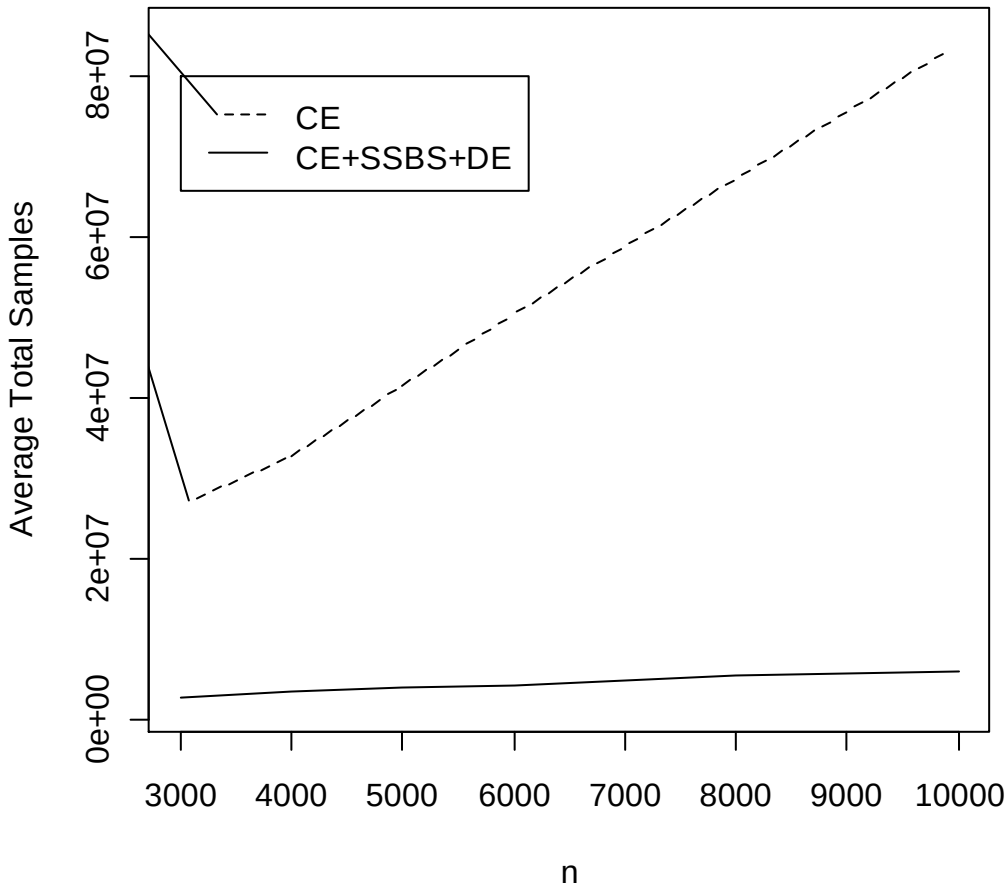

Figure 3.7: Average Total Samples vs n

CHAPTER 4

NEW RANKING AND SELECTION PROCEDURES FOR SELECTING THE k-BEST

4.1 Introduction

In this thesis, we have shown that despite being designed for discrete event simulation, ranking and selection (R&S) procedures are applicable and effective in numerous other fields, such as the NRD problem. Therefore, in this chapter, we turn our attention from the application of R&S procedures to that of their development, namely, we focus on the problem of selecting k-best alternative designs for $k > 1$, which is henceforth referred to as the selection of the k-best problem.

Over the past fifty years, a large body of literature has been devoted to the research and development of ranking and selection procedures, concerning selection of the single best alternative. However, to our knowledge, very little research has been devoted to the selection of the k-best problem. For a detailed review of the related literature, we refer the reader to Chapter 1. Selection of the k-best alternatives in simulation can be a very useful tool for the practical applications of R&S procedures, however, as hidden or un-modeled parameters can potentially invalid certain simulation solutions and many state-of-the-art global search heuristics rely on a neighborhood selection of the k-best search, such as the cross-entropy method introduced in chapter 3. Therefore, in this chapter, we propose a variety of new and more general R&S procedures for the selection of the k-best problem.

The remainder of this chapter proceeds as follows: section 4.2 introduces our problem framework, section 4.3 introduces the proposed indifference zone procedure (IZP) and subset selection procedure (SSP) for selecting the k-best, section 4.4 introduces an important decomposition lemma for developing combined R&S procedures, section 4.5 introduces our combined approach for the selection of the k-best, section 4.6 provides proofs for the theoretical guarantees associated with our procedures, and section 4.7

provides empirical results to investigate the actual efficiency, probability of correct selection (PCS), and overall performance of these new procedures.

4.2 Problem Framework

In this section, we define the R&S problem of interest as well as the goals we seek to achieve through our new procedures. To do this, we use the notations consistent with that of [85].

Let π_i represent the normally distributed random output associated with the i^{th} alternative design, for $i = 1, \dots, K$. Let X_{ij} represent the j^{th} observed output of alternative i, for $j = 1, \dots, N_i$, where N_i is the total number of outputs observed from alternative i. These outputs are assumed to be independent and identically distributed for fixed i, as they represent either outputs of independent replications of simulations or approximately independent batch means from a single replication.

Let $\mu_i = \mathrm{E}[X_{ij}]$ represent the expected value of an output from the i^{th} alternative, and $\sigma_i^2 = Var[X_{ij}]$ represent its variance. Then the order of the expected values of the K alternatives can be denoted as follows,

$$\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[K]}$$

where a larger expected value is indicative of a better or more desirable solution and $[i]$ represents the index of the i^{th} smallest alternative. However, since outputs are random, we estimate the true mean of each alternative through their sample mean, which we denote by $\bar{X}_i$, such that

$$\bar{X}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} X_{ij}$$

Therefore, $\bar{X}_{[i]}$ represents the sample mean for the i^{th} smallest alternative.

Let $\mathcal{K}$ represent the set of indices of the k-best alternatives, such that $\mathcal{K} = \{[K-k+1],[K-k+2],\dots,[K]\}$. The probability of correct selection requirement can be written as

$$\Pr\{CS\} \geq 1-\alpha \tag{4.1}$$

which states that the probability of making a *correct selection* is greater than or equal to a user-specified confidence level $1-\alpha$, where a *correct selection* indicates that the goal of the specific procedure was met. For our purposes we define the individual goals of the SSP and IZP using (4.2) and (4.3) respectively, where δ is the minimum practically significant difference, i.e. the largest difference that the experimenter is indifferent to.

Goal SSP: Use a given sample size n_0 for each alternative to select a random-sized subset I of the alternatives that contains the k-best alternatives associated with $\mu_{[K-k+1]},\mu_{[K-k+2]},\dots,\mu_{[K]}$, where $2 \leq k \leq K-1$, assuming $\mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta$. (4.2)

Goal IZP: Determine the sample size for each alternative to select exactly the subset of k-best alternatives $\mathcal{K}$, assuming $\mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta$. (4.3)

4.3 Proposed Individual R&S Procedures

In this section we introduce two new ranking and selection procedures, a SSP and an IZP, for selection of the k-best problems. These procedures extend the literature of selection of the k-best problems to the more general, and more practical, case of unknown and unequal variance, that is, the variance of the output from each alternative is unknown and may be different from each other. As such, these procedures assume that all samples are i.i.d. $N(\mu_i,\sigma_i^2)$ random variables where μ_i and σ_i^2 are unknown and potentially unequal for all i.

4.3.1 Subset Selection Procedure

Based on section 4.2, the goal of a SSP is to select $I \subseteq \{1,2,\dots,K\}$ of random size under conditions that satisfy (4.2), using an equal number (n_0) of samples from each system in consideration, such that

$$\Pr\{\mathcal{K} \subset I | \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta\} \geq 1-\alpha$$

Our proposed SSP is given as follows.

Algorithm 4.1: Subset Selection of the k-Best

Step 1) Select the overall confidence level $1-\alpha$. Then select the practically significant difference δ, the sample size $n_0 \geq 2$, and the number of the best systems k to return to the user, such that $2 \leq k < K$.

Step 2) Sample X_{ij} for $i = 1,\dots,K$, $j = 1,\dots,n_0$.

Step 3) Compute the sample means $\bar{X}_i$, sample variances S_i^2 and $w_{ij}\ \forall i, i \neq j$ for each system, where $w_{ij} = t\left(\frac{S_j^2+S_i^2}{n_0}\right)^{1/2}$ and $t = t_{(1-\alpha)^{\frac{1}{k(K-k)}}, n_0-1}$ is the $(1-\alpha)^{\frac{1}{k(K-k)}}$ quantile of the t-distribution with $n_0 - 1$ degrees of freedom.

Step 4) Let $B_{ij} = \bar{X}_i - (w_{ij} - \delta)^+$, where $y^+ = \max(0, y)$, and $B_j^{(K-k)}$ be the $(K–k)^{\text{th}}$ smallest $B_{ij}, i = 1,2,\dots,K, i \neq j$. Set $I = \left\{j: 1 \leq j \leq K, \bar{X}_j \geq B_j^{(K-k)}\right\}$

Step 5) Return I.

Notice that when $k = 1$, this procedure reduces to Nelson's Screen-to-the-Best procedure [20], thus this procedure can be considered an extension or a more general version of Nelson's SSP.

4.3.2 Indifference Zone Selection

The following indifference zone procedure determines the total number of samples necessary to obtain from each of the K alternatives in order to select the k-best systems in accordance with (4.3), such that

$$\Pr\{\text{select } \mathcal{K} | \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta\} \geq 1 - \alpha$$

In order to accommodate for unequal and unknown σ_i^2, this is a two-stage procedure, meaning that at the first stage an equal number of initial samples must be obtained from each system to estimate the variance of its output. Then the second stage sample size for each system is calculated and the best k systems will be selected based on all the samples from both stages. Our proposed IZP is given as follows.

Algorithm 4.2: Indifference Zone Procedure for Selection of the k-Best

Step 1) Select the overall confidence level $1 - \alpha$, the practically significant difference δ, the sample size n_0 where $n_0 \geq 2$, and the number of the best systems k to return to the user, such that $1 \leq k < K$.

Step 2) Sample X_{ij} for $i = 1, \dots, K, j = 1, \dots, n_0$.

Step 3) Determine the sample means $(\bar{X}_i)$ and sample variances (S_i^2) for each system i.

Step 4) Compute $N_i = \max\left\{n_0, \left\lceil \left(\frac{hS_i}{\delta}\right)^2 \right\rceil\right\}$ where $h = h_{(1-\alpha)^{1/k}, n_0, K-k+1}$ is Rinott's constant (see [14] for tables).

Step 5) Take $N_i - n_0$ samples for system i, and compute the second-stage sample means $\bar{X}_i$ based on the N_i total samples.

Step 6) Select the k indices with the k largest second-stage sample means as the k-best systems.

Notice that when $k = 1$, this procedure reduces to Rinott's two-stage IZP for selection of the single best, thus this procedure can be considered an extension or a more general version of Rinott's IZP.

4.4 Decomposition Lemma

In this section we introduce an important lemma from [20] that simplifies the development of combined subset selection and indifferent zone selection procedures. This lemma establishes that we can guarantee a probability of correct selection by combining a subset selection procedure with an indifference zone procedure that were developed in isolation from one another, even if the indifference zone procedure uses the same data as the subset selection procedure. This lemma is as follows, (from [20] yet augmented for selection of the k-best).

Let $\mathcal{S}$ be a subset selection procedure that collects data from each system and determines a random sized subset $I \subseteq \{1,2,\dots,K\}$ based on this data, such that $\Pr\{\mathcal{A}\} \geq 1 - \alpha_1$, where $\mathcal{A} = \{\mathcal{K} \subset I\}$ is the event that I contains the k-best alternatives.

Let J_ℓ, $\ell = 1,2,\dots,L$ be the $L = 2^{K-k}$ distinct subsets of $\{1,2,\dots,K\}$ that contain $\mathcal{K}$. Then let $\mathcal{R}$ be an indifference zone procedure that uses I, as well as the data used to determine I, as its initial stage of sampling to determine the number of additional samples required to select a set of indices $g \subset J_\ell$ such that $\Pr\{\mathcal{B}(J_\ell)\} \geq 1 - \alpha_2$ for any such subset J_ℓ, where $\mathcal{B}(J_\ell) = \{g = \mathcal{K}\}$ is the event that g is the set of the k best alternatives.

Then since $\mathcal{B}(\{1,2,\dots,K\}) \subseteq \mathcal{B}(J_\ell)$ for all J_ℓ, if a correct selection is made when $\mathcal{R}$ is applied to the entire set $\{1,2,\dots,K\}$, then a correct selection would also be made if $\mathcal{R}$

is applied to any of the 2^{K-k} subsets that contains $\mathcal{K}$. This property will hold for any procedure whose sampling for system i depends only on the data generated for system i, as the procedure given in section 4.3.2.

Let $\mathcal{B} = \bigcap_{\ell=1}^{L} \mathcal{B}(J_\ell)$, the event that $\mathcal{K}$ is selected for all subsets J_ℓ to which $\mathcal{R}$ can be applied.

Lemma 1: For the combined procedure $\mathcal{S} + \mathcal{R}$,

$$\Pr\{CS\} \geq \Pr\{\mathcal{A} \cap \mathcal{B}\} \geq 1 - (\alpha_1 + \alpha_2)$$

Proof: Any outcome belonging to the event $\mathcal{B}$ results in a correct selection, provided that the subset of systems considered by $\mathcal{R}$ contains $\mathcal{K}$. The event $\mathcal{A}$ only provides outcomes for which this is the case. Any outcome that satisfies both conditions will certainly result in a correct selection. Next notice that

$$\Pr\{\mathcal{A} \cap \mathcal{B}\} = \Pr\{\mathcal{A}\} + \Pr\{\mathcal{B}\} - \Pr\{\mathcal{A} \cup \mathcal{B}\}$$

$$\geq \Pr\{\mathcal{A}\} + \Pr\{\mathcal{B}(\{1, \dots, K\})\} - \Pr\{\mathcal{A} \cup \mathcal{B}\}$$

$$\geq (1 - \alpha_1) + (1 - \alpha_2) - 1$$

where the first inequality follows because $\Pr\{\bigcap_{\ell=1}^{L} \mathcal{B}(J_\ell)\} \geq \Pr\{\mathcal{B}(\{1,2, \dots, K\})\}$.

4.5 Combined Procedure

In this section we introduce a combined procedure using the SSP and IZP from section 4.3 as well as Nelson's decomposition lemma from section 4.4. The combined procedure has the same goal as an IZP. But the potential benefit of the combined procedure over an individual IZP is that systems with obviously inferior performances can be screened out quickly based on SSP such that additional second stage samples for IZP are not needed for these systems. This combined procedure is given below.

Algorithm 4.3: Combined Procedure

Step 1) Select the overall, first-stage and second-stage confidence levels: $1-\alpha$, $1-\alpha_1$ and $1-\alpha_2$ such that $\alpha = \alpha_1 + \alpha_2$. A convenient choice is $\alpha_1 = \alpha_2 = \alpha/2$. Then select the practically significant difference δ, the initial number of samples $n_0 \geq 2$, and the number of the best systems k to return to the user.

Step 2) Sample X_{ij} for $i = 1, \ldots, K$, $j = 1, \ldots, n_0$.

Step 3) Compute first stage sample means $(\bar{X}_i)$, sample variances (S_i^2) and $w_{ij}\ \forall i, i \neq j$ based on the n_0 first-stage samples for each system, where $w_{ij} = t\left(\frac{S_j^2 + S_i^2}{n_0}\right)^{1/2}$ and $t = t_{(1-\alpha_1)^{\frac{1}{k(K-k)}}, n_0 - 1}$ is the $(1-\alpha_1)^{\frac{1}{k(K-k)}}$ quantile of the *t*-distribution with $n_0 - 1$ degrees of freedom.

Step 4) Let $B_{ij} = \bar{X}_i - \left(w_{ij} - \delta\right)^+$ and $B_j^{(K-k)}$ be the $(K–k)^{\text{th}}$ smallest $B_{ij}, i = 1, 2, \ldots, K, i \neq j$. Set $I = \left\{j: 1 \leq j \leq K, \bar{X}_j \geq B_j^{(K-k)}\right\}$.

Step 5) If $|I| = k$ then stop. Otherwise compute the second-stage sample size $N_i,\ \forall i \in I,$ where $N_i = \max\left\{n_0, \left\lceil\left(\frac{hS_i}{\delta}\right)^2\right\rceil\right\}$ and $h = h_{(1-\alpha_2)^{1/k}, n_0, K-k+1}$, where h is Rinott's constant (see [14] for tables).

Step 6) At the second-stage, we take $N_i - n_0$ samples for system $i \in I$, Compute the second-stage sample means $\bar{X}_i$ based on the N_i total samples for each $i \in I$.

Step 7) Select the k indices with the k largest second-stage sample means as the k-best systems.

The combined procedure satisfies $\Pr\{\text{CS}|\mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta\} \geq 1-\alpha$ for selection of the k-best following Lemma 1, in that the SSP makes a CS with probability $\geq 1-\alpha_1$ and the IZP makes a CS with probability $\geq 1-\alpha_2$. Thus the overall probability of making a correct selection for the combined procedure is at least $1-(\alpha_1+\alpha_2) = 1-\alpha$.

4.6 Proofs

In this section we provide proofs for the PCS given in our individual subset selection and indifference zone procedures for selecting the k-best, given in Section 4.3. These proofs are given in Sections 4.6.1 and 4.6.2, respectively.

The following lemmas are used in the derivation of these proofs:

Lemma 2 (Banerjee [86]): Let Z be an $N(0,1)$ random variable that is independent of $Y_1, Y_2, \ldots, Y_K$, which are independent chi-quared random variables, with Y_i having degrees of freedom v_i. Let $\gamma_1, \gamma_2, \ldots, \gamma_K$ be arbitrary weights such that $\sum_{i=1}^{K} \gamma_i = 1$ and all $\gamma_i \geq 0$. Then,

$\Pr\left\{Z^2 \leq \sum_{i=1}^{K} t_i^2 \gamma_i \frac{Y_i}{v_i}\right\} \geq 1-\alpha$, where $t_i = t_{1-\alpha/2, v_i}$.

Lemma 3 (Slepian [87]): If $(X_1, \ldots, X_n)$ has the multivariate normal distribution with nonsingular covariance matrix $\Sigma = \left\|\sigma_{ij}^2\right\|_{i,j=1}^{n}$ then for any constants $c_1, \ldots, c_n$ the probability $\Pr\{X_1 \leq c_1, \ldots, X_n \leq c_n\}$ is strictly increasing as a function of σ_{ij} for $i \neq j$. In particular if $\sigma_{ij} > 0 \ \forall i, j = 1, \ldots, n$ then

$$\Pr\{X_1 \leq c_1, \ldots, X_n \leq c_n\} > \prod_{i=1}^{n} \Pr\{X_i \leq c_i\}$$

Lemma 4 (Tamhane [88]): Let $V_1, V_2, \ldots, V_k$ be independent random variables, and let $g_j(V_1, V_2, \ldots, V_k)$, $j = 1,2, \ldots, p$, be nonnegative, real-valued functions, each one nondecreasing in each of its arguments. Then

$$\mathrm{E}\left[\prod_{j=1}^{p} g_j(V_1, V_2, \dots, V_k)\right] \geq \prod_{j=1}^{p} \mathrm{E}\left[g_j(V_1, V_2, \dots, V_k)\right]$$

4.6.1 Subset Selection Procedure Proof

We first briefly summarize the SSP given in section 4.3.1 as follows:

Suppose there are n_0 independent outputs for each of the K alternatives under consideration. In this procedure, we guarantee to select a subset $I \subseteq \{1,2,\dots,K\}$ of the K alternatives such that:

$$\Pr\{\mathcal{K} \subseteq I\} \geq 1 - \alpha \tag{4.4}$$

where $0 < \alpha < 1$. Let $B_{ij} = \bar{X}_i - \left(w_{ij} - \delta\right)^+$ and $B_j^{(K-k)}$ represent the (K-k)$^{\text{th}}$ smallest $B_{ij}, i = 1,\dots,K, i \neq j$. Then (4.4) is achieved by including all alternatives in I that satisfy:

$$\bar{X}_j \geq B_j^{(K-k)} \tag{4.5}$$

The proof that (4.5) guarantees (4.4) is as follows,

Proof: First, it can be shown that

$$B_{[j]}^{(K-k)} \leq \max_{i \leq K-k} \bar{X}_{[i]} - \left(w_{[i][j]} - \delta\right)^+, \text{for all } j = 1,\dots,K \tag{4.6}$$

Then, under the condition that $\mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta$, (4.4) can be described as:

$$\Pr\{CS\} = \Pr\left\{\mathcal{K} \subseteq I \middle| \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta\right\}$$

$$= \Pr\left\{\bar{X}_{[j]} \geq B_{[j]}^{(K-k)}, \forall j \geq K-k+1 \middle| \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta\right\}$$

$$\geq \Pr\left\{\bar{X}_{[j]} \geq \max_{i \leq K-k} \bar{X}_{[i]} - \left(w_{[i][j]} - \delta\right)^+ \ \forall j \geq K-k+1 \middle| \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta\right\}$$

$$= \Pr\left\{\bar{X}_{[j]} \geq \bar{X}_{[i]} - \left(w_{[i][j]} - \delta\right)^+ \ \forall i \leq K-k, \forall j \geq K-k+1 \middle| \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta\right\}$$

$$= \Pr\left\{ \frac{\bar{X}_{[j]} - \bar{X}_{[i]} - \delta_{[j][i]}}{\left(\frac{\sigma_{[i]}^2 + \sigma_{[j]}^2}{n_0}\right)^{1/2}} \geq \frac{-\left(w_{[i][j]} - \delta\right)^+ - \delta_{[j][i]}}{\left(\frac{\sigma_{[i]}^2 + \sigma_{[j]}^2}{n_0}\right)^{1/2}} \; \forall i \leq K-k, \forall j \geq K-k+1 \middle| \ldots \right.$$

$$\left. \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta \right\}$$

where $\delta_{ij} = \mu_i - \mu_j$, and the inequality is true because of (4.6).

Next, to simplify terms, let

$$Z_{ij} = \frac{\bar{X}_{[j]} - \bar{X}_{[i]} - \delta_{[j][i]}}{\left(\frac{\sigma_{[i]}^2 + \sigma_{[j]}^2}{n_0}\right)^{1/2}}$$

and

$$v_{ij} = \left(\frac{\sigma_{[i]}^2 + \sigma_{[j]}^2}{n_0}\right)^{1/2}$$

We can then rewrite the previous probability statement as

$$= \Pr\left\{ Z_{ij} \geq \frac{-\left(\left(w_{[i][j]} - \delta\right)^+ + \delta_{[j][i]}\right)}{v_{ij}} \; \forall i \leq K-k, \forall j \geq K-k+1 \middle| \ldots \right.$$

$$\left. \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta \right\} \tag{4.7}$$

Due to the symmetry of the normal distribution (4.7) can be rewritten as

$$\Pr\left\{ Z_{ij} \leq \frac{\left(w_{[i][j]} - \delta\right)^+ + \delta_{[j][i]}}{v_{ij}} \; \forall i \leq K-k, \forall j \geq K-k+1 \middle| \mu_{[K-k+1]} - \mu_{[K-k]} \geq \delta \right\}$$

$$\geq \Pr\left\{ Z_{ij} \leq \frac{w_{[i][j]}}{v_{ij}} \; \forall i \leq K-k, \forall j \geq K-k+1 \right\} \tag{4.8}$$

where (4.8) is true because $\delta \leq \delta_{[j][i]}$. To further simplify, let

$$Q_{ij} = \frac{w_{[i][j]}}{v_{ij}} = t_{(1-\alpha)^{\frac{1}{k(K-k)}}, n_0 - 1} \left(\frac{S_{[j]}^2 + S_{[i]}^2}{\sigma_{[j]}^2 + \sigma_{[i]}^2}\right)^{1/2}$$

where t is the $(1-\alpha)^{\frac{1}{k(K-k)}}$ quantile of the t-distribution with $n_0 - 1$ degrees of freedom. Then, by conditioning on $S_1^2, \ldots, S_K^2$, we can rewrite (4.8) as

$$\begin{aligned}
&\mathrm{E}\left[\Pr\left\{Z_{ij} \le Q_{ij}\ \forall i \le K-k, \forall j \ge K-k+1 \middle| S_1^2, \ldots, S_K^2\right\}\right] \\
&\ge \mathrm{E}\left[\prod_{i=1}^{K-k} \prod_{j=K-k+1}^{K} \Pr\left\{Z_{ij} \le Q_{ij} \middle| S_1^2, \ldots, S_K^2\right\}\right] \\
&\ge \prod_{i=1}^{K-k} \prod_{j=K-k+1}^{K} \mathrm{E}\left[\Pr\left\{Z_{ij} \le Q_{ij} \middle| S_1^2, \ldots, S_K^2\right\}\right] \\
&= \prod_{i=1}^{K-k} \prod_{j=K-k+1}^{K} \Pr\{Z_{ij} \le Q_{ij}\} \qquad (4.9)
\end{aligned}$$

where the first inequality is true because:

$$\mathrm{cov}\left(Z_{ij_1}, Z_{ij_2}\right) = \frac{\sigma_{[i]}^2}{\left(\sigma_{[i]}^2 + \sigma_{[j_1]}^2\right)^{1/2}\left(\sigma_{[i]}^2 + \sigma_{[j_2]}^2\right)^{1/2}} \ge 0$$

$$\mathrm{cov}\left(Z_{i_1 j}, Z_{i_2 j}\right) = \frac{\sigma_{[j]}^2}{\left(\sigma_{[i_1]}^2 + \sigma_{[j]}^2\right)^{1/2}\left(\sigma_{[j_2]}^2 + \sigma_{[j]}^2\right)^{1/2}} \ge 0$$

$$\mathrm{cov}\left(Z_{i_1 j_1}, Z_{i_2 j_2}\right) = 0, \qquad \forall i_1 \ne i_2, j_1 \ne j_2$$

and Lemma 3 (Slepian's inequality). And the second inequality is true because of Lemma 4, where

$$\left(g_i(S_1^2, \ldots, S_K^2)\right) \equiv \Pr\left\{Z_{ij} < Q_{ij} \middle| S_1^2, \ldots, S_K^2\right\}$$

and $g_i(\cdot)$ is nonnegative and non-decreasing in $S_1^2, \ldots, S_K^2$ since Q_{ij} is non-decreasing in $S_1^2, \ldots, S_K^2$.

Finally, to complete the proof, we look at the individual probability $\Pr\{Z_{ij} \le Q_{ij}\}$. Since $Z_{ij} \sim N(0,1)$ and $Q_{ij} \ge 0$, we have

$$\Pr\left\{Z_{ij} \le Q_{ij}\right\} = \frac{1}{2} + \Pr\left\{0 \le Z_{ij} \le Q_{ij}\right\}$$

$$= \frac{1}{2} + \frac{1}{2}\Pr\{Z_{ij}^2 \leq Q_{ij}^2\}$$

$$= \frac{1}{2} + \frac{1}{2}\Pr\left\{Z_{ij}^2 \leq t^2\gamma_1 \frac{S_{[i]}^2}{\sigma_{[i]}^2} + t^2\gamma_2 \frac{S_{[j]}^2}{\sigma_{[j]}^2}\right\}$$

$$\geq \frac{1}{2} + \frac{1}{2}\left(1 - 2\left(1 - (1-\alpha)^{\frac{1}{k(K-k)}}\right)\right)$$

$$= (1-\alpha)^{\frac{1}{k(K-k)}} \tag{4.10}$$

where $\gamma_1 = \frac{\sigma_{[i]}^2}{\sigma_{[i]}^2 + \sigma_{[j]}^2}$ and $\gamma_2 = \frac{\sigma_{[j]}^2}{\sigma_{[i]}^2 + \sigma_{[j]}^2}$ and the inequality in (4.10) follows from Lemma 2. Substituting (4.10) into (4.9) we get

$$\Pr\{CS\} \geq \prod_{i=1}^{K-k} \prod_{j=K-k+1}^{K} \left((1-\alpha)^{\frac{1}{k(K-k)}}\right) = 1 - \alpha$$

4.6.2 Indifference Zone Procedure Proof

For the IZP in section 4.3.2, we guarantee to select k-best systems of the K alternatives such that:

$$Pr\{\text{select } \mathcal{K}\} \geq 1 - \alpha \tag{4.11}$$

This is achieved by taking N_i independent outputs for each of the K alternatives such that:

$$N_i = \max\left\{n_0, \left\lceil\left(\frac{hS_i}{\delta}\right)^2\right\rceil\right\} \tag{4.12}$$

where $h = h_{(1-\alpha)^{1/k}, n_0, K-k+1}$. The proof for the theoretical guarantee of the IZP is as follows:

Proof: For the selection of k-best problem, it can be shown that the mean configuration known as the slippage configuration, i.e. $\mu_{[K]} = \mu_{[K-1]} = \cdots = \mu_{[K-k+1]} = \mu_{[K-k]} +$

δ, and $\mu_{[K-k]} = \mu_{[K-k-1]} = \cdots = \mu_{[1]}$, minimizes the PCS. Therefore, in the following proof we base our probability calculations on this worst case scenario:

$$\Pr\{CS\} = \Pr\{\text{select systems } K, K-1, \ldots, K-k+1\}$$

$$= \Pr\left\{\bar{X}_{[j]} < \bar{X}_{[i]}, \forall i \in \mathcal{K}, j \notin \mathcal{K}\right\}$$

$$= \Pr\left\{\frac{\bar{X}_{[j]} - \left(\bar{X}_{[i]} - \delta\right)}{\left(\sigma_{[i]}^2/N_{[i]} + \sigma_{[j]}^2/N_{[j]}\right)^{1/2}} < \frac{\delta}{\left(\sigma_{[i]}^2/N_{[i]} + \sigma_{[j]}^2/N_{[j]}\right)^{1/2}}, \forall i \in \mathcal{K}, j \notin \mathcal{K}\right\}$$

To simplify, let

$$Z_{ij} = \frac{\bar{X}_{[j]} - \left(\bar{X}_{[i]} - \delta\right)}{\left(\sigma_{[i]}^2/N_{[i]} + \sigma_{[j]}^2/N_{[j]}\right)^{1/2}} \sim N(0,1)$$

and

$$Q_{ij} = \frac{h}{\left(\sigma_{[i]}^2/S_{[i]}^2 + \sigma_{[j]}^2/S_{[j]}^2\right)^{1/2}}$$

where $h = h_{1-(1-\alpha)^{1/k}, n_0, K-k+1}$. Since $N_i = \max\left\{n_0, \left\lceil \left(\frac{h}{\delta}\right) S_i^2 \right\rceil\right\}$, it follows that

$$Q_{ij} \leq \frac{\delta}{\left(\sigma_{[i]}^2/N_{[i]} + \sigma_{[j]}^2/N_{[j]}\right)^{1/2}}$$

Therefore we have

$$\Pr\{CS\} \geq \Pr\left\{Z_{ij} < Q_{ij}, \forall i \in \mathcal{K}, j \notin \mathcal{K}\right\}$$

$$= \mathrm{E}\left[\Pr\left\{Z_{ij} < Q_{ij}, \forall i \in \mathcal{K}, j \notin \mathcal{K} \middle| S_1^2, \ldots, S_K^2\right\}\right]$$

$$\geq \mathrm{E}\left[\prod_{i=K-k+1}^{K} \prod_{j=1}^{K-k} \Pr\left\{Z_{ij} < Q_{ij} \middle| S_1^2, \ldots, S_K^2\right\}\right]$$

$$\geq \prod_{i=K-k+1}^{K} \mathrm{E}\left[\prod_{j=1}^{K-k} \Pr\left\{Z_{ij} < Q_{ij} \middle| S_1^2, \ldots, S_K^2\right\}\right]$$

where the first inequality is true because:

$$\operatorname{cov}\left(Z_{ij_1}, Z_{ij_2}\right) = \frac{\sigma_{[i]}^2/n_0}{\left(\sigma_{[i]}^2/N_{[i]} + \sigma_{[j_1]}^2/N_{[j_1]}\right)^{1/2}\left(\sigma_{[i]}^2/N_{[i]} + \sigma_{[j_2]}^2/N_{[j_2]}\right)^{1/2}} \geq 0$$

$$\operatorname{cov}\left(Z_{i_1 j}, Z_{i_2 j}\right) = \frac{\sigma_{[j]}^2/n_0}{\left(\sigma_{[i_1]}^2/N_{[i_1]} + \sigma_{[j]}^2/N_{[j]}\right)^{1/2}\left(\sigma_{[i_2]}^2/N_{[i_2]} + \sigma_{[j]}^2/N_{[j]}\right)^{1/2}} \geq 0$$

$$\operatorname{cov}\left(Z_{i_1 j_1}, Z_{i_2 j_2}\right) = 0 \ \forall i_1 \neq i_2, j_1 \neq j_2$$

and Lemma 3 (Slepian's inequality). The second inequality is true because of Lemma 4, where

$$\left(g_i(S_1^2, \ldots, S_K^2)\right) \equiv \prod_{j=1}^{K-k} \Pr\left\{Z_{ij} < Q_{ij} \middle| S_1^2, \ldots, S_K^2\right\}$$

and $g_i(\cdot)$ is nonnegative and non-decreasing in $S_1^2, \ldots, S_K^2$, since Q_{ij} is non-decreasing in $S_1^2, \ldots, S_K^2$.

Then we have

$$\Pr\{CS\} \geq \prod_{i=K-k+1}^{K} \mathrm{E}\left[\prod_{j=1}^{K-k} \Phi\left(Q_{ij}\right)\right]$$

$$= \prod_{i=K-k+1}^{K} \mathrm{E}\left[\prod_{j=1}^{K-k} \Phi\left(\frac{h}{\left(\sigma_{[j]}^2/S_{[j]}^2 + \sigma_{[i]}^2/S_{[i]}^2\right)^{1/2}}\right)\right]$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

Let $Y_j \equiv (n_0 - 1)\frac{S_j^2}{\sigma_j^2}$ are independent χ^2 variables with $n_0 - 1$ degrees of freedom. We have

$$\Pr\{CS\} \geq \prod_{i=K-k+1}^{K} \mathrm{E}\left[\prod_{j=1}^{K-k} \Phi\left(\frac{h}{\left[(n_0 - 1)\left(1/Y_j + 1/Y_i\right)\right]^{1/2}}\right)\right]$$

Let $f(\cdot)$ be the density function of a χ^2 distribution with $(n_0 - 1)$ degrees of freedom. The right hand side (RHS) of the above equation can be simplified as

$$RHS = \prod_{i=K-k+1}^{K} \mathrm{E}\left[\prod_{j=1}^{K-k} \Phi\left(\frac{h}{\left[(n_0-1)\left(1/Y_j + 1/Y_i\right)\right]^{1/2}}\right)\right]$$

$$\prod_{i=K-k+1}^{K} \mathrm{E}\left[\mathrm{E}\left[\prod_{j=1}^{K-k} \Phi\left(\frac{h}{\left[(n_0-1)\left(1/Y_j + 1/Y_i\right)\right]^{1/2}}\right)\middle| Y_i\right]\right]$$

$$= \prod_{i=K-k+1}^{K} \int_0^\infty \left(\int_0^\infty \Phi\left(\frac{h}{[(n_0-1)(1/x+1/y)]^{1/2}}\right) f(x)dx\right)^{K-k} f(y)dy$$

$$= \left(\int_0^\infty \left(\int_0^\infty \Phi\left(\frac{h}{[(n_0-1)(1/x+1/y)]^{1/2}}\right) f(x)dx\right)^{K-k} f(y)dy\right)^{k} \qquad (4.13)$$

Since $h = h_{(1-\alpha)^{1/k},n_0,K-k+1}$ is the Rinott's constant, we have

$$\int_0^\infty \left(\int_0^\infty \Phi\left(\frac{h}{[(n_0-1)(1/x+1/y)]^{1/2}}\right) f(x)dx\right)^{K-k} f(y)dy = (1-\alpha)^{1/k}$$

Substituting the above equation into (4.13), we have

$$\Pr\{CS\} \geq 1-\alpha$$

4.7 Empirical Results

In this section, we provide empirical results regarding the theoretical guarantees we proved in section 4.6 as well as the overall performance of our ranking and selection procedures in sections 4.3 and 4.5. Since these procedures were designed to reduce to Nelson and Rinott when $k = 1$, and since to our knowledge these are novel procedures in the ranking and selection literature, the following experiments were designed to closely resemble those of [20] with the additional parameter k. This was done to enable us to use many of the parameter level recommendations made by [20], as well as to provide us with a basis for comparison for our new procedures when $k > 1$. The purpose of the

following experiments was to investigate the actual PCS, the benefit of including the screening step, and the effect of k on the overall performance.

For these experiments, we used various configurations of normal distributions instead of using systems simulation examples. This was done to provide us more control over the design of the experiment and enable us to draw conclusions on the performance of our procedures under different mean and variance configurations. This is also the common method used in the ranking and selection literature.

In all experiments, we designate the systems in positions $\{1, \dots, k\}$ as the k-best systems, while the systems in positions $\{k+1, \dots, K\}$ represent the inferior systems. The experimental factors of interest that we test are: the number of systems K, the number of the best systems to-be-selected k, the configuration of the population variances, and the configuration of the population means. The practical significant difference δ, the initial number of samples n_0, and nominal PCS were set to the recommended levels in [20].

We next describe the variance and mean configurations as well as the parameter levels and the design of our experiments.

4.7.1 Mean and Variance Configurations

In the following experiments we consider two mean configurations, the slippage configuration (SC) and the monotone-decreasing means configuration (MDM). These configurations are defined as follows,

1. SC is a mean configuration where the means of the k-best systems are all equal and exactly δ greater than the inferior systems, i.e. $\mu_{[K]} = \mu_{[K-1]} = \cdots = \mu_{[K-k+1]} = \mu_{[K-k]} + \delta$ and $\mu_{[K-k]} = \mu_{[K-k-1]} = \cdots = \mu_{[1]}$. This configuration is considered the "worst case scenario" for the selection of the k-best problem.
2. MDM is a mean configuration where the mean of the i^{th} best system is exactly $(i-1) \times \Delta$ less than the best system, where Δ is some constant. This is, in a

sense, considered the "favorable scenario" for a ranking and selection procedure and is used to showcase the benefit of the screening procedure.

We also investigate three different configurations of the variance: equal variance across all systems, decreasing variance across all systems, and increasing variance across all systems. The details regarding how we implement these various combinations of mean and variance configurations are described in the next section.

4.7.2 Experiment Design

For this experiment, we investigat the 2×3×2 factorial experiment for the twelve different combinations of the mean, variance and screening configurations. For each mean, variance, and screening configuration, the number of systems was varied over $K = 5, 10, 100, 500$ and the number of the best systems-to-be-selected was varied over $k = 1,2,5,10$ for $k < K$. The initial number of samples is set to $n_0 = 10$ as recommended in [20], and the practically significant difference is set to $\delta = \frac{\sigma_1}{\sqrt{n_0}}$.

When $k > 1$, we use Algorithm 4.3 when screening is used, and Algorithm 4.2 when screening is not used. When $k = 1$, we use Nelson's 2-stage combined procedure for selection of the single best system when screening is used, and Rinott's 2-stage indifference zone procedure for selection of the single best system when screening is not used.

For the MDM configuration, all systems are spaced evenly such that the mean associated with index i is exactly one practically significant difference less than the mean associated with index $i-1$, where the mean of the first index is set to δ, i.e. $\mu_i = \delta(2-i)$ for $i = 1,2,3,\dots,K$.

The variance configurations differ based on the type of mean configuration used. In SC, the variance is set such that: $\sigma_{k+1}^2 = \sigma_{k+2}^2 = \cdots = \sigma_K^2 = \rho\sigma^2$ with $\rho = 0.5, 2$ depending on whether the variance is decreasing or increasing, where σ^2 is the common

variance of the k-best systems. In the MDM configuration, the variance is set to $\sigma_i^2 = |\mu_i - \delta| + 1$ for increasing variance and $\sigma_i^2 = 1/|\mu_i - \delta| + 1$ for decreasing variance. For both mean configurations, $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_K^2 = 1$ is used for common variance.

For each treatment, we perform 500 macroreplications (complete replications) using $\alpha = 0.05$ and $\alpha_1 = \alpha_2 = 0.025$ as the nominal PCS. Therefore, the standard error of the actual PCS is $\sqrt{0.95(0.05)/500} \cong 0.00975$ in all treatments.

4.7.3 Results

The results from these experiments are summarized in Table 4.1, where PCS refers to the observed probability of correct selection from the 500 macroreplications, ANS refers to the average number of samples per system, and PSS refers to the percentage of systems receiving second stage sampling. We use PCS to provide us with an estimate of the conservativeness of the procedure, ANS to provide us with a measure of the computation load of a procedure, and PSS to provide a measure of the effectiveness of the screening procedure. Figure 4.1 to Figure 4.3 also summarize the results for $K = 100$ and $K = 500$ from Table 4.1, showing the graphical depiction of the effect of k on ANS.

For the combined and non-screening procedures, the PCS varies from close to the nominal value of 0.95 to 1. In most configurations, $k > 1$ has a higher PCS than $k = 1$ for the combined procedures, this is as expected as the selection of the k-best is generally more conservative than the single best equivalent. On average, the PCS of the SC configuration is lower than that of the MDM configuration because the indifference zone procedures were designed for the SC configuration, which is the worst-case scenario. Consequently, the results for MDM configurations become more conservative.

In certain scenarios, the combined procedure performs very well as compared to the non-screening equivalent, where as in other scenarios it does not. For MDM, the

combined procedure is very successful in eliminating inferior systems from receiving second stage sampling, as indicated by the PSS and ANS values in Table 4.1. For SC, however, the combined procedure performs consistently worse than the non-screening equivalent in all configurations. This is because, as indicated by the PSS values in Table 4.1, the screening stage of the combined procedure is not very effective for this type of mean configuration, and is essentially running the non-screening equivalent for a nominal PCS of $1 - \alpha_2$ instead of $1 - \alpha$.

For the MDM configurations, PSS increases steadily with k given a fixed K. This is as expected as more systems must be retained during the screening stage in order to make a correct selection. For SC when $k > 1$, PSS is fairly consistent around 1. This indicates that the screening procedure is too conservative to be beneficial in the worst case scenario mean configuration.

In general, ANS increases with k. The caveat to this is seen in Table 4.1 when k becomes a significant proportion of K, i.e. $k = 5$ and $K = 10$, for both screening configurations of SC with increasing variance. In this situation, since screening is ineffective due to the mean configuration, ANS depends primarily on the system variances. Since the k-best alternatives have lower variance, when k becomes large enough as compared to K, ANS begins to decrease as the overall variance of the K systems is lower.

For all configurations, ANS increases in a slow sublinear trend with increasing k except for MDM with increasing variance where ANS increases in a faster sublinear trend with increasing k for the combined procedures. This larger increase in ANS is partly because S_i^2 increases linearly with the system index for MDM with increasing variance, thus ANS grows much more quickly. Also, since N_i is directly proportional to S_i, ANS grows more rapidly as variance increase with the system index.

For all SC and MDM without screening configurations, ANS is increased by no more than 2 times when k is increased from 1 to 10. For MDM with screening, ANS

increases by no more than 4 times when k is increased from 1 to 10. Therefore, in general, our procedure requires only a small number of extra samples, as compared to selecting the single best design, to select multiple best systems simultaneously.

Table 4.1: Comparisons across all mean and variance configuration combinations, with $n_0 = 10$, $\delta = 1/\sqrt{n_0}$, and $PCS = 0.95$.

		K	5		10			100				500			
		k	1	2	1	2	5	1	2	5	10	1	2	5	10
MDM	increasing variance	PCS	0.996	0.998	0.998	0.998	1	1	0.998	1	1	1	1	1	1
		ANS	242	322	324	557	755	206	391	1012	2257	121	211	494	1035
		PSS	81%	96%	61%	83%	99%	12%	16%	25%	37%	3%	4%	7%	10%
	decreasing variance	PCS	0.990	0.994	0.998	1	1	0.996	1	1	1	1	1	1	1
		ANS	90	124	70	97	134	24	29	38	50	15	16	18	22
		PSS	60%	82%	35%	48%	73%	5%	6%	9%	13%	1%	2%	2%	3%
	constant variance	PCS	0.986	0.994	0.998	0.992	1	1	1	1	1	1	1	1	1
		ANS	133	193	113	175	298	40	55	91	146	20	25	36	52
		PSS	68%	91%	43%	60%	92%	6%	8.5%	13%	19%	2%	2%	3%	4%
SC	increasing variance	PCS	0.972	0.976	0.97	0.98	0.98	0.98	0.984	0.992	0.994	0.982	0.972	0.992	0.992
		ANS	310	332	433	497	479	857	1009	1214	1365	1217	1422	1727	1988
		PSS	94%	99%	96%	99%	100%	99%	100%	100%	100%	100%	100%	100%	100%
	decreasing variance	PCS	0.96	0.98	0.974	0.984	0.976	0.98	0.996	0.988	0.982	0.988	0.984	0.986	0.996
		ANS	99	143	120	165	235	216	260	328	396	305	358	439	512
		PSS	86%	98%	90%	99%	100%	97%	100%	100%	100%	99%	100%	100%	100%
	constant variance	PCS	0.972	0.964	0.98	0.984	0.982	0.978	0.99	0.99	0.994	0.986	0.992	0.994	0.994
		ANS	170	206	223	276	310	430	509	623	719	609	713	868	1002
		PSS	92%	99%	94%	99%	100%	98%	100%	100%	100%	99%	100%	100%	100%
MDM no screening	increasing variance	PCS	0.992	0.986	0.998	0.998	1	0.998	1	1	1	1	1	1	1
		ANS	222	262	446	543	626	6003	7127	8838	10254	41362	48597	59529	68864
		PSS	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	decreasing variance	PCS	0.992	0.994	0.994	0.998	1	1	1	1	1	1	1	1	1
		ANS	91	106	91	110	127	43	50	62	72	21	23	27	31
		PSS	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	constant variance	PCS	0.988	0.99	0.988	0.998	1	1	1	1	1	1	1	1	1
		ANS	138	166	183	224	259	362	430	532	614	519	608	745	864
		PSS	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
SC no screening	increasing variance	PCS	0.958	0.962	0.97	0.95	0.992	0.95	0.96	0.99	0.998	0.982	0.962	0.988	0.992
		ANS	245	257	351	400	385	719	853	1033	1170	1035	1216	1481	1708
		PSS	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

	decreasing variance	PCS	0.97	0.948	0.962	0.954	0.982	0.968	0.972	0.986	1	0.966	0.98	0.998	0.99
		ANS	82	111	101	134	194	183	220	279	338	260	305	377	441
		PSS	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
	constant variance	PCS	0.966	0.952	0.964	0.968	0.97	0.962	0.974	0.992	0.994	0.95	0.98	0.992	0.99
		ANS	136	159	183	225	259	362	430	530	615	519	608	745	864
		PSS	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

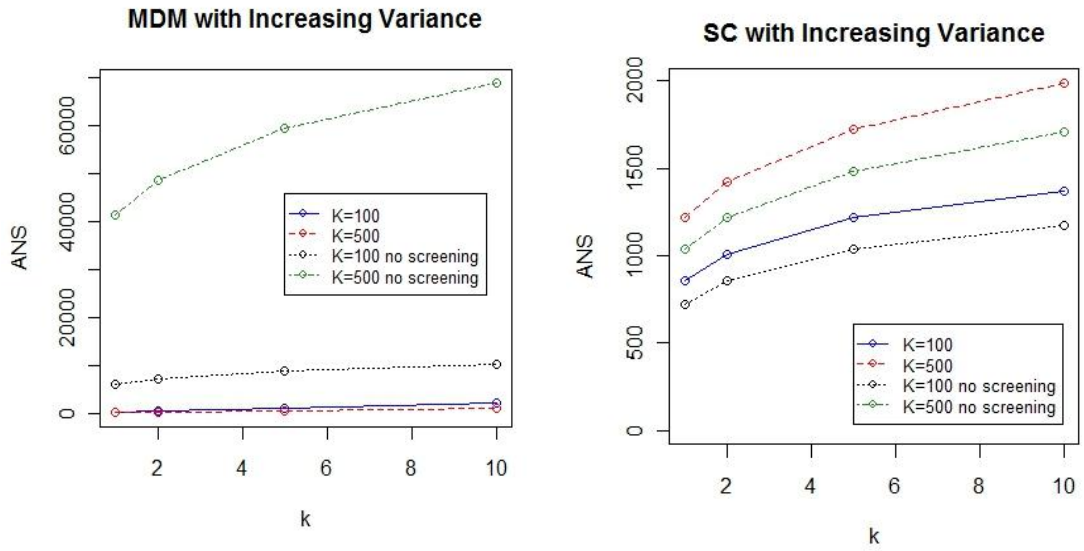

Figure 4.1: Graphical comparison of Average Number of Samples (ANS) vs Number of Best-Systems-to-be-Selected (k) for all increasing variance configurations.

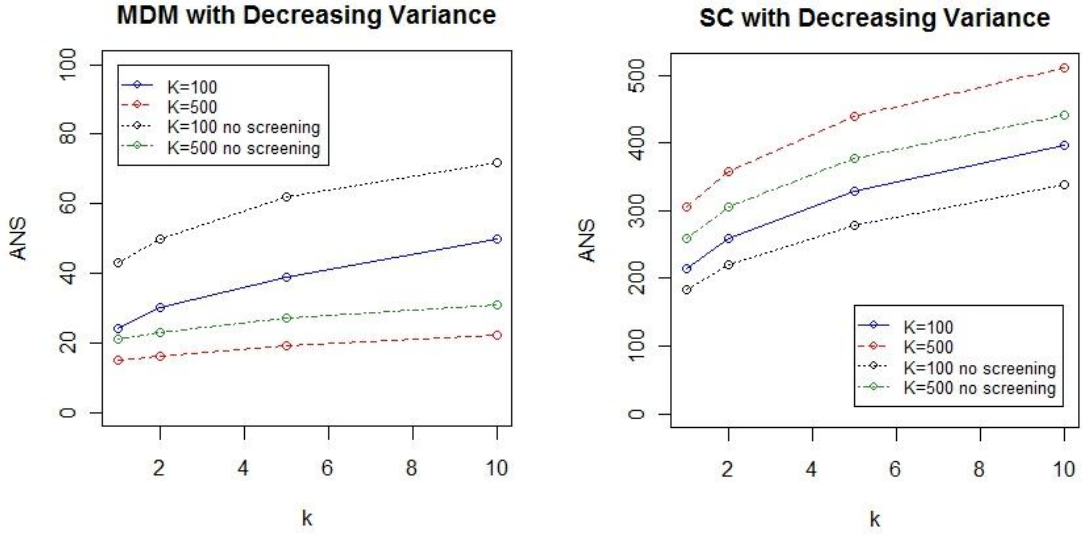

Figure 4.2: Graphical comparison of Average Number of Samples (ANS) vs Number of Best-Systems-to-be-Selected (k) for all decreasing variance configurations.

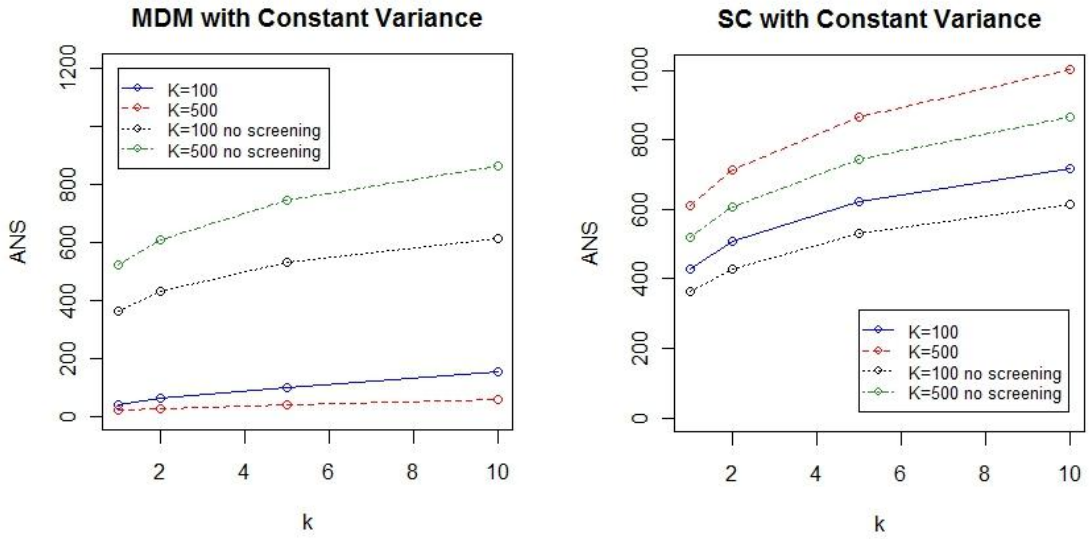

Figure 4.3: Graphical comparison of Average Number of Samples (ANS) vs Number of Best-Systems-to-be-Selected (k) for all constant variance configurations.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

This thesis focuses on the development of simulation ranking and selection procedures and applications to the NRD problem. This thesis thus makes contributions to both the simulation optimization and system reliability design literature.

In Chapter 2 of this thesis, we present a new objective and new methodology for solving the sNRD optimization problem. This methodology focused on two criteria: increasing the sampling efficiency of the sNRD problem and quantifying the probability of making a correct selection. We propose using a state-of-the-art R&S procedure to accomplish these goals, and develop a rule of thumb to approximate the minimum batch size that will maintain an acceptable normal assumption in this procedure as well as extend some justifications of this procedures to the sNRD population size. We then provide numerical examples to verify that the procedure performs accordingly, as well as provide numerical examples to showcase the benefits of the proposed procedure in contrast to the conventional approach. These examples show the proposed procedure achieving much higher probability, than the conventional method, of correctly selecting the optimal solution, given the same number of total Monte Carlo samples, as well as being more robust in terms of the selection of the input parameters and the ability of the user to efficiently make a correct selection.

In Chapter 3 of this thesis, we propose a new optimization framework for solving the large NRD problem. We call this framework CE+SSBS+DG as it is composed of the cross-entropy method and two statistical procedures that we develop, based on the concepts of R&S, to improve the sampling efficiency of the Monte Carlo simulation: sequential selection of the best subset (SSBS) and duplicate generation (DG). The purpose of this framework is to change the focus of the Monte Carlo simulation from accurate evaluation of all the systems searched to one concerned only with the relative

rankings of the system reliability estimates. Altogether this framework shows shorter runtimes and smaller sample sizes in comparison with the standard cross entropy approach, on the order of 500%-1500% over the three randomly generated example networks. We also show CE+SSBS+DG to be more robust in terms of the selection of the maximum sample size n, meaning that selection of n has less influence on total samples for the proposed method than it does for the standard cross entropy approach.

In Chapter 4 of this thesis, we present three new R&S procedures for the problem associated with the general simulation optimization problems concerning selection of the k-best alternatives. These procedures are novel in that they accounted for systems with unknown and potentially unequal mean and variances as well as an indifference parameter, which to our knowledge has not been studied within the R&S literature. For each procedure, we provide theoretical proofs regarding the probability of correct selection as well as an empirical analysis on its actual performance. Although these procedures are more conservative than the single best equivalent in general, we show that they only require a small number of extra samples, as compared to the single best procedures, to simultaneously select the multiple best systems.

For future work, we have plans for further applications of R&S in both system reliability design as well as simulation optimization. For the NRD problem, we plan on employing the use of our individual and combined selection of the k-best R&S procedures to develop better solution methodology for both small NRD and the large NRD. For R&S, we plan on extending the proof for our current IZP for selection of the k-best to a less conservative theoretical guarantee as well as publish a table of critical values associated with this new procedure for quicker reference. This will require the development of a new executable program to numerically evaluate a very large and complex integral, which will ideally be published as the companion to this procedure. We will also investigate the use of common random numbers in our proposed procedure for selection of the k-best. Finally, we

may also look at expanding the proof for our combined R&S procedure from a 2-stage procedure to a more efficient fully sequential procedure.

APPENDIX A

PROCEDURE R

1) Specify b_0, δ, α.

2) Take b_0 batch means from each design.

3) Determine $s_i^2 = \frac{\sum_{i=1}^{b_0} (\phi(y_{ij}) - \bar{R}_{i\cdot})^2}{b_0 - 1}$ the unbiased estimate of σ_i^2.

4) Calculate h, where h is the solution to

$$\int_0^\infty \int_0^\infty \left[\Phi\left(\frac{h}{\sqrt{(b_0-1)\left(\frac{1}{x}+\frac{1}{y}\right)}} \right) f_{b_0-1}(x) \right]^{K-1} f_{b_0-1}(y)\, dy\, dx = 1 - \alpha$$

where $\Phi(\cdot)$ denotes the s-normal cdf, and $f_v(\cdot)$ denotes the chi-squared pdf with v degrees of freedom. Tables for h are provided in [14].

5) Compute $b_i = \max\left\{ b_0, \left\lceil \left(\frac{h}{\delta}\right)^2 s_i^2 \right\rceil \right\} \quad \forall i \in (1,2,\dots,K)$.

6) Take $b_i - b_0$ additional batch means from each design.

7) Find the estimated reliability of each design, $\bar{R}_{i\cdot} = \frac{1}{b_i} \sum_{j=1}^{b_i} \phi(y_{ij}) \quad \forall i$.

8) Select the system with the largest sample mean $\bar{R}_{i\cdot}$ as the best.

APPENDIX B

PROCEDURE CY

1) Take b_0 batch means from each design i.
2) Determine $s_{ij}^2 = \frac{1}{b_0-1}\sum_{l=1}^{b_0}\left(\phi(y_{il}) - \phi(y_{jl}) - (\bar{R}_{i\cdot} - \bar{R}_{j\cdot})\right)^2$, the sample variance of the difference between design i and design j.
3) $t = t_{1-\left[\frac{\alpha}{(K-1)}\right],b_0-1}$ the $1-\left[\frac{\alpha}{(K-1)}\right]$-quantile of the t-distribution with $b_0 - 1$ degrees of freedom.
4) Compute $b = \max\left\{b_0, \left\lceil \max_{j\neq i}\left(\frac{ts_{ij}}{\delta}\right)^2 \right\rceil\right\}$.
5) Take $b - b_0$ additional batch means from each design.
6) Find the estimated reliability of each design, $\bar{R}_{i\cdot} = \frac{1}{b}\sum_{j=1}^{b} y_{ij} \ \forall i$.
7) Select the system with the largest estimated reliability $\bar{R}_{i\cdot}$ as the best.

APPENDIX C

PROCEDURE 2SP

Step 1) Select $\alpha_0, \delta, n_0, \alpha_1 \;\; s.t. \alpha_0 + \alpha_1 = \alpha.$

Step 2) Calculate $t = t_{(1-\alpha_0)^{\frac{1}{K-1}}, b_0 - 1}$ for independent replications, or $t = t_{1-\left[\frac{\alpha}{(K-1)}\right], b_0 - 1}$ for CRN.

Step 3) Collect batch means $y_{ij}\ \forall i = 1, \dots, K\ \&\ j = 1, \dots, b_0$ and calculate $\bar{R}_{i\cdot}^{(1)}$ and s_i^2 for all i where $\bar{R}_{i\cdot}^{(1)} = \frac{1}{b_0}\sum_{j=1}^{b_0}\phi(y_{ij})$ and $s_i^2 = \frac{1}{b_0 - 1}\sum_{j=1}^{b_0}\left(\phi(y_{ij}) - \bar{R}_{i\cdot}^2\right)\ \forall i$ for IR, or $s_{ij}^2 = \frac{1}{b_0 - 1}\sum_{l=1}^{b_0}\left(\phi(y_{il}) - y_{jl} - \left(\bar{R}_{i\cdot} - \bar{R}_{j\cdot}\right)\right)^2$ for CRN.

Step 4) Set $W_{ij} = t\left(\frac{s_i^2}{b_0} + \frac{s_j^2}{b_0}\right)^{\frac{1}{2}}\ \forall i \neq j.$

Step 5) If employing CRN, replace $\frac{s_i^2}{b_0} + \frac{s_j^2}{b_0}$ with $\frac{s_{ij}^2}{b_0}$ where $s_{ij}^2 = \frac{1}{b-1}\sum_{l=1}^{b_0}\left(\phi(y_{il}) - \phi(y_{jl}) - \left(\bar{R}_{i\cdot} - \bar{R}_{j\cdot}\right)\right)^2$ and $t = t_{1-\frac{\alpha_0}{K-1}, b_0 - 1}$.

Step 6) $I = \left\{i : 1 \leq i \leq K \; and \; \bar{R}_{i\cdot} \geq \bar{R}_{j\cdot} - \left(W_{ij} - \delta\right)^+\ \forall j \neq i\right\}.$

Step 7) If $|I| = 1$ stop and select the design in I as the best, else compute $b_i = \max\left\{b_0, \left\lceil\left(\frac{hs_i}{\delta}\right)^2\right\rceil\right\}\ \forall i$ where h is Rinott's constant.

Step 8) Collect $b_i - b_0$ additional batch means $y_{ij}\ \forall i = 1, \dots, K\ \&\ j = b_0, \dots, b_i$ and calculate $\bar{R}_{i\cdot}^{(2)}\ \forall i$ where $\bar{R}_{i\cdot}^{(2)} = \frac{1}{b_i}\sum_{j=1}^{b_i}\phi(y_{ij})$.

Step 9) Select the design with the largest $\bar{R}_{i\cdot}^{(2)}$ as the best.

APPENDIX D

PROCEDURE KN PARAMETERS

$$g(\eta) \equiv \sum_{l=1}^{c}(-1)^{l+1}\left(1-\frac{1}{2}\psi(l=c)\right)\left(1+\frac{2\eta(2c-l)l}{c}\right)^{-(n_0-1)/2}$$

$\beta = \frac{\alpha}{(K-1)}$ for CRN and $\beta = 1-(1-\alpha)^{\frac{1}{K-1}}$ for independent replications.

APPENDIX E

BATCH SIZE RULE OF THUMB EXPERIMENT

Equation (2.12) was constructed as follows:

First, holding all other parameter levels constant at their middle range value, we determined the individual relationship between each of K, b_0, δ, R' and the critical batch size value, for the parameter levels: $\delta \in \{0.001,0.005,0.01,0.02,0.03,0.04,0.05\}, K \in [30,100,200,300,400,500,750,1000]$, $R \in [0.5,0.6,0.7,0.8,0.9,0.95,0.99,0.995,0.999]$, $b_0 \in [10,15,20]$. The best fit individual variable relationships with B^* were: $-\ln \delta \sim B^*$, $e^R \sim B^*$, $\frac{1}{b_0} \sim B^*$, $\ln K \sim B^*$.

Second, we determine the critical batch size for the fractional factorial experiment of δ, K, $\tilde{R}$, and b_0 for the 108 treatments shown in Table 2.1. Using this data, we then constructed various nonlinear models using the individual relationships of these four parameters with B^* as building blocks. These models were then tested against the same data set until we were able to represent a similar nonlinear relationship through a relatively simple model that also maintained a sufficient buffer zone between the RoT value and the actual B^* so as to never underestimate it.

The reason Equation (2.12) was not constructed using an additive linear or nonlinear regression model was because we wanted (2.12) to be simple and compact, and the afore mentioned techniques could not produce a model as compact as (2.12).

APPENDIX F

RAW DATA FOR EXAMPLE 1 OF SECTION 3.4

Table F.1: Example 1 Raw Data.

CE+SSBS+DG using (3.8)			CE+SSBS			CE+DG			CE		
Iterations	Time	Total Samples	Iterations	Time	Total Samples	Iterations	Time	Total Samples	Iterations	Time	Total Samples
8	9,598	20,150,550	7	11,250	28,593,950	7	28,474	63,500,000	7	39970	90,000,000
8	8,181	18,404,450	6	9,950	22,341,550	8	32,251	71,460,000	7	40295	90,000,000
7	9,360	19,430,150	7	10,374	28,509,050	7	30,086	67,120,000	7	40511	90,000,000
6	7,060	14,677,200	8	15,525	35,320,450	6	26,531	62,500,000	6	33857	75,000,000
8	9,660	19,907,000	7	10,213	27,684,950	7	30,540	67,940,000	6	33680	75,000,000
7	8,375	17,315,800	7	10,222	28,897,650	7	29,928	68,040,000	6	33583	75,000,000
6	7,361	15,390,300	6	9,484	224,637,00	7	27,899	61,440,000	6	33634	75,000,000
8	8,645	18,098,300	6	9,036	21,842,100	7	27,878	61,240,000	6	33648	75,000,000
6	7,971	16,700,700	7	11,174	28,958,800	7	29,010	63,960,000	7	40463	90,000,000
6	7,067	17,796,850	7	11678	28,901,950	7	28,813	64,120,000	7	40604	90,000,000

APPENDIX G

RAW DATA FOR EXAMPLE 2 OF SECTION 3.4

Table G.1: Example 2 Raw Data.

CE+SSBS+DG using (3.8)			CE+SSBS			CE+DG			CE		
Iterations	Time	Total Samples	Iterations	Time	Total Samples	Iterations	Time	Total Samples	Iterations	Time	Total Samples
14	1,747	3,540,440	11	3,482	8,654,940	12	15,008	31,044,000	10	19,469	36,000,000
15	1,872	4,263,060	9	2,065	5,502,180	15	14,456	29,680,000	11	21,933	40,000,000
9	1,678	2,990,720	9	2,113	5,696,260	13	14,521	29,616,000	10	15,626	36,000,000
10	1,654	2,946,640	10	2,376	6,219,840	9	11,526	24,180,000	8	11,881	28,000,000
12	1,735	3,295,460	9	1,905	5,029,960	9	11,938	25,252,000	10	15,902	36,000,000
9	1,673	2,910,720	9	2,083	5,635,660	15	13,810	28,392,000	9	13,880	32,000,000
17	1,988	4,539,980	9	2,111	5,645,840	11	14,352	29,560,000	9	13,951	32,000,000
10	1,722	3,273,500	9	2,133	5,803,900	13	16,092	30,160,000	8	11,938	28,000,000
13	1,857	4,135,120	10	2,498	6,338,080	10	12,581	26,096,000	9	13,833	32,000,000
12	1,764	3,409,580	11	3,148	7,730,720	9	11,398	23,872,000	8	11,831	28,000,000

APPENDIX H

RAW DATA FOR EXAMPLE 3 OF SECTION 3.4

Table H.1: Example 3 Raw Data.

CE+SSBS+DG using (3.8)			CE+SSBS			CE+DG			CE		
Iterations	Time	Total Samples	Iterations	Time	Total Samples	Iterations	Time	Total Samples	Iterations	Time	Total Samples
10	1,720	4,043,880	11	4,497	13,331,760	10	22,738	55,160,000	10	28,438	72,000,000
11	1,999	4,641,880	9	2,527	8,660,340	14	26,396	63,744,000	9	23,681	64,000,000
11	2,153	4,937,540	10	2,756	9,339,560	12	25,781	60,420,000	9	23,813	64,000,000
10	1,765	4,147,920	11	3,245	11,042,800	12	26,081	62,686,000	9	23,905	64,000,000
11	1,984	4,672,840	9	2,100	7,944,160	11	23,969	57,492,000	11	30,587	80,000,000
15	2,458	5,642,400	10	2,841	9,708,300	11	26,655	62,624,000	9	23,539	64,000,000
12	2,266	5,267,860	9	2,489	9,424,460	11	28,982	58,644,000	10	27,482	72,000,000
10	1,758	4,105,140	9	2,002	7,326,420	10	24,579	60,008,000	10	27,422	72,000,000
10	1,756	4,159,680	10	2,653	8,772,420	12	25,430	61,472,000	10	26,905	72,000,000
10	1,757	4,117,460	10	3,140	11,050,080	10	22,860	54,272,000	11	30,437	80,000,000

APPENDIX I

BINOMIAL CONFIDENCE LIMITS EXPERIMENT

Table I.1: Example 1
Raw Data

CE+SSBS+DG using (3.9)		
Iterations	Time	Total Samples
7	9,667	20,266,300
7	8,898	18,612,250
7	9,817	20,588,500
6	7,859	16,732,250
7	9,267	19,314,650
6	9,957	20,987,350
6	7,761	16,424,900
6	8,584	18,337,850
7	9,057	19,039,500
7	10,135	20,863,050

Table I.2: Example 2
Raw Data

CE+SSBS+DG using (3.9)		
Iterations	Time	Total Samples
14	1,928	3,891,180
10	1,526	3,191,520
10	1,503	3,118,300
10	1,592	3,302,780
13	1,938	4,057,860
13	2,134	4,363,460
11	1,934	4,021,920
16	2,081	4,217,500
12	1,914	3,919,880
12	1,704	3,516,340

Table I.3: Example 3
Raw Data

CE+SSBS+DG using (3.9)		
Iterations	Time	Total Samples
11	2,076	4,835,760
13	2,120	4,797,360
12	2,107	4,791,340
9	1,772	4,134,580
10	1,802	4,097,380
11	2,125	4,818,500
10	1,871	4,255,080
16	3,073	6,974,440
11	2,004	4,753,200
10	2,033	4,624,500

APPENDIX J

RAW DATA FOR α SENSITIVITY IN SECTION 3.6.1

Table J.1: α Sensitivity Raw Data Example 2.

α	Iterations	Total Samples	α	Iterations	Total Samples	α	Iterations	Total Samples	α	Iterations	Total Samples
0.01	10	3505460	0.10	11	2644580	0.30	10	1788700	0.50	11	1490080
0.01	12	3947880	0.10	15	2517380	0.30	11	1946040	0.50	13	1655120
0.01	16	5040000	0.10	10	2267080	0.30	11	1995360	0.50	12	1524920
0.01	12	3789360	0.10	9	2270100	0.30	11	2168120	0.50	18	1975700
0.01	11	3739320	0.10	13	2703860	0.30	14	2280940	0.50	12	1529760
0.01	13	4194100	0.10	17	4037600	0.30	13	2157020	0.50	9	1047100
0.01	11	3465840	0.10	12	2701760	0.30	14	2279140	0.50	11	1184760
0.01	9	2990720	0.10	11	2459260	0.30	12	2161500	0.50	17	1987380
0.01	10	3925900	0.10	12	2879720	0.30	11	2045840	0.50	14	1522800
0.01	12	4381820	0.10	9	2156840	0.30	10	1616380	0.50	12	1818540
0.05	14	3540440	0.20	11	2257780	0.40	17	2154580	0.60	12	1072580
0.05	15	4263060	0.20	19	2767720	0.40	9	1502220	0.60	9	1071180
0.05	9	2990720	0.20	11	1901200	0.40	15	1644620	0.60	11	1287280
0.05	10	2946640	0.20	21	3168900	0.40	13	1892320	0.60	16	1561860
0.05	12	3295460	0.20	10	1943340	0.40	14	2091080	0.60	10	1044180
0.05	9	2910720	0.20	21	2667240	0.40	13	1794880	0.60	12	915740
0.05	17	4539980	0.20	14	2442140	0.40	11	1599720	0.60	13	1234760
0.05	10	3273500	0.20	15	2723200	0.40	11	1580380	0.60	10	919360
0.05	13	4135120	0.20	13	2240120	0.40	9	1282840	0.60	12	1048980
0.05	12	3409580	0.20	13	2153780	0.40	13	1902480	0.60	13	1210120

APPENDIX K

RAW DATA FOR n SENSITIVITY IN SECTION 3.6.2

Table K.1: n Sensitivity Raw Data Example 2 for CE+SSBS+DG using $\alpha = 0.05$.

n	Iter	Total	n	Iter	Total	n	Iter	Total
3,000	10	2,821,940	5,000	11	3,940,040	8,000	10	5,280,400
3,000	11	2,699,000	5,000	10	4,012,020	8,000	11	5,615,440
3,000	19	3,707,220	5,000	19	4,685,100	8,000	12	6,507,540
3,000	10	2,640,320	5,000	12	4,249,840	8,000	11	5,447,000
3,000	12	2,950,580	5,000	12	4,244,780	8,000	10	4,637,440
3,000	12	3,157,460	5,000	11	3,641,840	8,000	11	5,863,520
3,000	13	2,921,440	5,000	10	3,731,240	8,000	10	5,232,480
3,000	13	3,076,900	5,000	12	3,689,800	8,000	11	5,259,740
3,000	11	2,515,860	5,000	10	3,588,400	8,000	10	4,845,280
3,000	11	2,579,220	5,000	11	3,709,260	8,000	13	5,846,720
4,000	14	3,540,440	6,000	11	4,367,940	10,000	9	5,535,680
4,000	15	4,263,060	6,000	14	5,790,720	10,000	10	6,378,920
4,000	9	2,990,720	6,000	9	3,863,420	10,000	9	5,586,300
4,000	10	2,946,640	6,000	10	3,825,140	10,000	9	5,697,320
4,000	12	3,295,460	6,000	10	4,143,940	10,000	10	5,895,780
4,000	9	2,910,720	6,000	9	3,685,840	10,000	10	5,987,120
4,000	17	4,539,980	6,000	12	5,061,280	10,000	13	7,293,160
4,000	10	3,273,500	6,000	10	4,029,220	10,000	10	6,046,740
4,000	13	4,135,120	6,000	10	3,837,100	10,000	10	6,086,540
4,000	12	3,409,580	6,000	9	3,826,120	10,000	10	5,235,440

Table K.2: n Sensitivity Raw Data Example 2 Sample for CE.

n	Iter	Total	n	Iter	Total	n	Iter	Total
3,000	11	30,000,000	5,000	10	45,000,000	8,000	9	64,000,000
3,000	9	24,000,000	5,000	8	35,000,000	8,000	10	72,000,000
3,000	10	27,000,000	5,000	8	35,000,000	8,000	9	64,000,000
3,000	9	24,000,000	5,000	10	45,000,000	8,000	10	72,000,000
3,000	14	39,000,000	5,000	9	40,000,000	8,000	9	64,000,000
3,000	11	30,000,000	5,000	11	50,000,000	8,000	8	56,000,000
3,000	9	24,000,000	5,000	8	35,000,000	8,000	9	64,000,000
3,000	9	24,000,000	5,000	11	50,000,000	8,000	9	64,000,000
3,000	8	21,000,000	5,000	10	45,000,000	8,000	11	80,000,000
3,000	9	24,000,000	5,000	8	35,000,000	8,000	10	72,000,000
4,000	10	36,000,000	6,000	9	48,000,000	10,000	10	90,000,000
4,000	11	40,000,000	6,000	10	54,000,000	10,000	9	80,000,000
4,000	10	36,000,000	6,000	10	54,000,000	10,000	11	100,000,000
4,000	8	28,000,000	6,000	10	54,000,000	10,000	9	80,000,000
4,000	10	36,000,000	6,000	9	48,000,000	10,000	9	80,000,000
4,000	9	32,000,000	6,000	9	48,000,000	10,000	10	90,000,000
4,000	9	32,000,000	6,000	8	42,000,000	10,000	8	70,000,000
4,000	8	28,000,000	6,000	11	60,000,000	10,000	10	90,000,000
4,000	9	32,000,000	6,000	9	48,000,000	10,000	9	80,000,000
4,000	8	28,000,000	6,000	9	48,000,000	10,000	9	80,000,000

REFERENCES

1. Banks, J., *Discrete-event system simulation*, Upper Saddle River, N.J.; Singapore: Prentice Hall.

2. Shannon, R.E., S.S. Long, and B.P. Buckles, *Operation Research Methodologies in Industrial Engineering: A Survey.* AIIE Transactions, 1980. **12**(4): p. 364 - 367.

3. Forgionne, G.A., *Corporate Management Science Activities: An Update.* INTERFACES, 1983. **13**(3): p. 20-23.

4. Harpell, J.L., M.S. Lane, and A.H. Mansour, *Operations Research in Practice: A Longitudinal Study.* INTERFACES, 1989. **19**(3): p. 65-74.

5. Lane, M.S., A.H. Mansour, and J.L. Harpell, *Operations Research Techniques: A Longitudinal Update 1973-1988.* INTERFACES, 1993. **23**(2): p. 63-68.

6. Thomas, G. and J.-A. DaCosta, *A Sample Survey of Corporate Operations Research.* INTERFACES, 1979. **9**(4): p. 102-111.

7. Goldsman, D. and B.L. Nelson. *Ranking, selection and multiple comparisons in computer simulation*. in *Simulation Conference Proceedings, 1994. Winter*. 1994.

8. Fu, M.C., *Optimization via simulation: a review*. 1994.

9. Fu, M.C., *A tutorial review of techniques for simulation optimization*, in *Proceedings of the 26th conference on Winter simulation*. 1994, Society for Computer Simulation International: Orlando, Florida, United States.

10. Law, A.M. and D.W. Kelton, *Simulation Modeling and Analysis*. 2000, New York: McGraw-Hill.

11. James, R.S., et al., *Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey.* ACM Trans. Model. Comput. Simul., 2003. **13**(2): p. 134-154.

12. Goldsman, D., *Ranking and selection in simulation*, in *Proceedings of the 15th conference on Winter Simulation - Volume 2*. 1983, IEEE Press: Arlington, Virginia, United States.

13. Goldsman, D. and B.L. Nelson, *Statistical screening, selection, and multiple comparison procedures in computer simulation*, in *Proceedings of the 30th conference on Winter simulation*. 1998, IEEE Computer Society Press: Washington, D.C., United States.

14. Bechhofer, R.E., T.J. Santner, and D.M. Goldsman, *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons*. Vol. 38. 1995, New York: John Wiley and Sons. 289.

15. Kim, S.-H. and B.L. Nelson. *Selecting the Best System*. in *Elsevier Handbooks in Operations Research and Management Science: Simulation*. forthcoming: Elsevier.

16. Gupta, S.S., *On a Decision Rule for a Problem in Ranking Means*. 1956, Institute of Statistics, University of North Carolina: Chapel Hill, North Carolina.

17. Gupta, S.S. and T.J. Santer. *On selection and ranking procedures - a restricted subset selection rule.* in *Proceedings of the 39th Session of the International Statistical Institute*. 1973.

18. Santner, T.J., *A Restricted Subset Selection Approach to Ranking and Selection Problems.* The Annals of Statistics, 1975. **3**(2): p. 334-349.

19. Sullivan, D.W. and J.R. Wilson, *Restricted Subset Selection Procedures for Simulation.* Operations Research, 1989. **37**(1): p. 52-71.

20. Nelson, B.L., et al., *Simple procedures for selecting the best simulated system when the number of alternatives is large.* Operations Research, 2001. **49**(6): p. 950-963.

21. Wilson, J.R., *A Multiplicative Decomposition Property of the Screening-and-Selection Procedures of Nelson et al.* Operations Research, 2001. **49**(6): p. 964-966.

22. Carroll, R.J., S.S. Gupta, and D.-Y. Huang, *On Selection Procedures for the t Best Populations and some Related Problems*. 1975.

23. Bofinger, E. and K. Mangersen, *Subset selection of the t best populations.* Communications in Statistics - Theory and Methods, 1986. **15**(10): p. 3145 - 3161.

24. Koenig, L.W. and A.M. Law, *A procedure for selecting a subset of size m containing the l best of k independent normal populations, with applications to simulation.* Communications in Statistics - Simulation and Computation, 1985. **14**(3): p. 719 - 734.

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

26. Paulson, E., *A sequential procedure for selecting the population with the largest mean from k normal populations.* Annals of Mathematical Statistics, 1964. **35**: p. 174-180.

27. Bechhofer, R.E., Kiefer, J., and Sobel, M., *Sequential Identification and Ranking Procedures.* The University of Chicago Press, Chicago, Ill., 1968.

28. Zinger, A. and J. St-Pierre, *On the Choice of the Best Amongst Three Normal Populations with Known Variances.* Biometrika, 1958. **45**(3/4): p. 436-446.

29. Dudewicz, E.J. and S.R. Dalal, *Allocation of observations in ranking and selection with unequal variances.* Indi J. Stat., 1975. **37B**(1): p. 27-78.

30. Rinott, Y., *On two-stage selection procedures and related probability-inequalities.* Communications in Statistics - Theory and Methods, 1978. **A7**(8): p. 799-811.

31. Clark, G.M. and W. Yang. *A Bonferroni selection procedure when using common random numbers with unknown variances*. 1986. Washington, DC, USA: SCS.

32. Nelson, B.L. and F.J. Matejcik, *Using common random numbers for indifference-zone selection and multiple comparisons in simulation.* Management Science, 1995. **41**(12): p. 1935-45.

33. Bechhofer, R.E., *A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with known Variances.* The Annals of Mathematical Statistics, 1954. **25**(1): p. 16-39.

34. Bechhofer, R.E., C.W. Dunnett, and M. Sobel, *A Two-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with a Common Unknown Variance.* Biometrika, 1954. **41**(1-2): p. 170-176.

35. Nelson, B.L. and D. Goldsman, *Comparisons with a standard in simulation experiments.* Management Science, 2001. **47**(3): p. 449-63.

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

37. Kim, S.-H. and B.L. Nelson, *A Fully Sequential Procedure for Indifference-Zone Selection in Simulation.* ACM Transactions on Modeling and Computer Simulation, 2001. **11**(3): p. 251-273.

38. Garey, M.R. and D.S. Johnson, *Computers and intractibility : A guide to the theory of NP- Completeness*. 1979: San Francisco : W. H. Freeman and Company.

39. Aggarwal, K.K., Y.C. Chopra, and J.S. Bajwa, *Topological layout of links for optimizing the s-t reliability in a computer communication system*. Microelectronics and Reliability, 1982. **22**(3): p. 341-5.

40. Aggarwal, K.K., Y.C. Chopra, and J.S. Bajwa, *Reliability evaluation by network decomposition.* IEEE Transactions on Reliability, 1982. **R-31**(4): p. 355-8.

41. Jan, R.-H., F.-J. Hwang, and S.-T. Cheng, *Topological optimization of a communication network subject to a reliability constraint.* IEEE Transactions on Reliability, 1993. **42**(1): p. 63-70.

42. Ball, M.O. and G.L. Nemhauser, *Matroids and a reliability analysis problem*. Mathematics of Operations Research, 1979. **4**(2): p. 132-43.

43. Colbourn, C.J., *Combinatorics of network reliability*. 1987, New York, NY: Oxford University Press, Inc. 176.

44. Page, L.B. and J.E. Perry, *A note on computing environments and network reliability.* Microelectronics and Reliability, 1991. **31**(1): p. 185-6.

45. Resende, L.I.P., *Implementation of a factoring algorithm for reliability evaluation of undirected networks.* IEEE Transactions on Reliability, 1988. **37**(5): p. 462-8.

46. Resende, M.G.C., *A program for reliability evaluation of undirected networks via polygon-to-chain reductions.* IEEE Transactions on Reliability, 1986. **R-35**(1): p. 24-9.

47. Satyanarayana, A. and M.K. Chang, *Network reliability and the factoring theorem.* Networks, 1983. **13**(1): p. 107-20.

48. Van Slyke, R. and H. Frank, *Network reliability analysis. I.* Networks, 1972. **1**(3): p. 279-90.

49. Ball, M.O. and J.S. Provan, *Calculating bounds on reachability and connectedness in stochastic networks.* Networks, 1983. **13**(2): p. 253-78.

50. Lomonosov, M.V. and V.P. Polesskii, *Lower bound of network reliability.* Problems of Information Transmission, 1972. **8**(2): p. 118-23.

51. Aboelfotoh, H.M. and C.J. Colbourn, *Series-parallel bounds for the two-terminal reliability problem.* ORSA Journal on Computing, 1989. **1**(4): p. 209-22.

52. Brecht, T.B. and C.J. Colbourn, *Lower bounds on two-terminal network reliability.* Discrete Applied Mathematics, 1988. **21**(3): p. 185-98.

53. Li, V.O.K. and J.A. Silvester, *Performance analysis of networks with unreliable components.* IEEE Transactions on Communications, 1984. **CM-32**(10): p. 1105-10.

54. Lam, Y.F. and V.O.K. Li, *An improved algorithm for performance analysis of networks with unreliable components.* IEEE Transactions on Communications, 1986. **CM-34**(5): p. 496-7.

55. Yang, C.L. and P. Kubat, *Efficient computation of most probably states for communication networks with multimode components.* IEEE Transactions on Communications, 1989. **37**(5): p. 535-8.

56. Srivaree-ratana, C., A. Konak, and A.E. Smith, *Estimation of all-terminal network reliability using an artificial neural network.* Computers and Operations Research, 2002. **29**(7): p. 849-868.

57. Kumamoto, H., et al., *Dagger-sampling Monte Carlo for system unavailability evaluation.* IEEE Transactions on Reliability, 1980. **R-29**(2): p. 122-5.

58. Kumamoto, H., K. Tanaka, and K. Inoue, *Efficient evaluation of system reliability by Monte Carlo method.* IEEE Transactions on Reliability, 1977. **R-26**(5): p. 311-15.

59. Elperin, T., I. Gertsbakh, and M. Lomonosov, *Estimation of network reliability using graph evolution models.* IEEE Transactions on Reliability, 1991. **40**(5): p. 572-81.

60. Fishman, G.S., *A Monte Carlo sampling plan for estimating network reliability.* Operations Research, 1986. **34**(4): p. 581-94.

61. Easton, M.C. and C.K. Wong, *Sequential destruction method for Monte Carlo evaluation of system reliability*. IBM Technical Disclosure Bulletin, 1980. **22**(12): p. 5543-6.

62. Fishman, G.S., *A comparison of four Monte Carlo methods for estimating the probability of s-t connectedness*. IEEE Transactions on Reliability, 1986. **R-35**(2): p. 145-55.

63. Belovich, S.G., *A design technique for reliable networks under a nonuniform traffic distribution*. IEEE Transactions on Reliability, 1995. **44**(3): p. 377-87.

64. Dengiz, B., F. Altiparmak, and A.E. Smith, *Efficient optimization of all-terminal reliable networks, using an evolutionary approach*. IEEE Transactions on Reliability, 1997. **46**(1): p. 18-26.

65. Kumar, A., et al., *A genetic algorithm for distributed system topology design*. Computers & Industrial Engineering, 1995. **28**(3): p. 659-70.

66. Altiparmak, F., B. Dengiz, and A.E. Smith. *Reliability optimization of computer communication networks using genetic algorithms*. 1998. San Diego, CA, USA: IEEE, Piscataway, NJ, USA.

67. Ruey-Shun, C., C. Chin Ching, and Y.S. Yeh. *A genetic algorithm for the reliability optimization of a distributed system*. 1998. Vienna, Austria: IEEE Comput. Soc.

68. Reichelt, D., F. Rothlauf, and P. Gmilkowsky, *Designing Reliable Communication Networks with a Genetic Algorithm Using a Repair Heuristic*, in *Evolutionary Computation in Combinatorial Optimization*. 2004, Springer. p. 177-187.

69. Dengiz, B., F. Altiparmak, and A.E. Smith, *Local search genetic algorithm for optimal design of reliable networks*. IEEE Transactions on Evolutionary Computation, 1997. **1**(3): p. 179-88.

70. Kumar, A., R.M. Pathak, and Y.P. Gupta, *Genetic-algorithm-based reliability optimization for computer network expansion*. IEEE Transactions on Reliability, 1995. **44**(1): p. 63-72.

71. Deeter, D.L. and A.E. Smith, *Economic design of reliable networks*. IIE Transactions (Institute of Industrial Engineers), 1998. **30**(12): p. 1161-1174.

72. Aboelfotoh, H.M.F. and L.S. Al-Sumait, *A neural approach to topological optimization of communication networks, with reliability constraints*. IEEE Transactions on Reliability, 2001. **50**(4): p. 397-408.

73. Atiqullah, M.M. and S.S. Rao, *Reliability optimization of communication networks using simulated annealing*. Microelectronics and Reliability, 1993. **33**(9): p. 1303-1319.

74. Ravi, V., B.S.N. Murty, and J. Reddy, *Nonequilibrium simulated-annealing algorithm applied to reliability optimization of complex systems*. IEEE Transactions on Reliability, 1997. **46**(2): p. 233-9.

75. Nariai, S., K.P. Hui, and D. Kroese, *Designing an Optimal Network Using the Cross-Entropy Method,* in *Intelligent Data Engineering and Automated Learning*. 2005, Springer.

76. Hui, K.P., et al., *The cross-entropy method for network reliability estimation.* Annals of Operations Research, 2005. **134**: p. 101-18.

77. Hui, K.-P., *Monte Carlo network reliability ranking estimation.* IEEE Transactions on Reliability, 2007. **56**(1): p. 50-7.

78. Kroese, D.P., K.-P. Hui, and S. Nariai, *Network reliability optimization via the cross-entropy method.* IEEE Transactions on Reliability, 2007. **56**(2): p. 275-287.

79. Nariai, S. and D.P. Kroese. *On the Design of Multi-type Networks via the Cross-Entropy Method*. in *Proceeding of the Fifth International Workshop on the Design of Reliable Communication Networks*. 2005: IEEE.

80. Altiparmak, F. and B. Dengiz, *A cross entropy approach to design of reliable networks.* European Journal of Operational Research, 2009. **199**(2): p. 542-552.

81. Rubinstein, R.Y., *Optimization of computer simulation models with rare events.* European Journal of Operational Research, 1997. **99**(1): p. 89-112.

82. S. Nariai, K.P.H., and D. Kroese, *Designing an Optimal Network Using the Cross-Entropy Method,* in *Intelligent Data Engineering and Automated Learning*. 2005, Springer.

83. Clopper, C.J. and E.S. Pearson, *The use of confidence or fiducial limits illustrated in the case of the binomial.* Biometrika, 1934. **26**: p. 404-413.

84. Wilson, E.B., *Probable inference, the law of succession, and statistical inference.* Journal of the American Statistical Association, 1927. **22**: p. 209-212.

85. Kim, S.-H., B.L. Nelson, and S.G. Henderson, *Chapter 17 Selecting the Best System*, in *Handbooks in Operations Research and Management Science*. 2006, Elsevier. p. 501-534.

86. Banerjee, S., *On Confidence Interval for Two-Means Problem Based on Separate Estimates of Variances and Tabulated Values of t-Table.* SankhyÄ • : The Indian Journal of Statistics, Series A, 1961. **23**(4): p. 359-378.

87. Slepian, D., *The one-sided barrier problem for Gaussian noise.* Bell System Tech. J., 1962. **41**: p. 463-501.

88. Tamhane, A.C., *Multiple comparisons in model i one-way anova with unequal variances.* Communications in Statistics - Theory and Methods, 1977. **6**(1): p. 15 - 32.