

## SELECTION OF THE BEST WITH STOCHASTIC CONSTRAINTS

Alireza Kabirian

College of Business and Public Policy

3211 Providence Drive  
University of Alaska-Anchorage  
Anchorage, AK 99508, USA

Sigurdur Olafsson

Dept. of Industrial and Manufacturing Systems

Engineering  
Black Engineering 3004  
Iowa State University  
Ames, IA 5001, USA

### ABSTRACT

When selecting the best design of a system among a finite set of possible designs, there may be multiple selection criterion. One formulation of such a multi-criteria problem is minimization (or maximization) of one of the criteria while constraining the others. In this paper, we assume the criteria are unobservable mean values of stochastic outputs of simulation. We propose a new heuristic iterative algorithm for finding the best in this situation and use a number of experiments to demonstrate the performance of the algorithm.

### 1 INTRODUCTION

Selecting-the-Best design of a system among a finite set of designs when there are stochastic constraints on the performance of the desired system has wide applications, but is subject to various computational challenges. The applications range from toy examples of simulation-based optimization of a bank in which the bank manager wants the average waiting time of customers to be less than a target upper bound, to sophisticated manufacturing design models in which throughput should satisfy average frequency requirements while the maximum work in process ought to be less than available physical inventory capacity.

However, the application of Selection of the Best problems/methods in general and with expensive (stochastic) constraints in particular are not limited to the use of discrete event simulation as the expensive model for evaluating the objective function/constraints. Wherever an expensive model/experiment is needed to know the feasibility or quality of the designs, there is a potential need for such methods. In fact, the discrete-event simulation context studied here may be substituted with a fluids-mechanic simulation experiment in mechanical engineering, system dynamics simulation in business and economic modeling, war gaming in military, surveying experiment in statistics/sociology or the test of a new vaccine on animals in veterinary or medicine. The experiments in these cases are monetarily expensive, time consuming or computationally intensive, practically dangerous, or ethically unjustifiable.

Essentially, the Selection of the Best problem we consider in this paper has these properties:

1. The number of available designs is finite and each design is characterized only by a code assigned to it.
2. The selection is based on evaluating each design on a number of decision criteria. One of these criteria is called objective. A design is called feasible if the values of its non-objective criteria are within desired limits. The problem is about finding the feasible design with the best value of the objective.
3. For each design, the vector of criteria is the unobservable mean vector of a multivariate normal distribution. The outputs of running a discrete-event simulation model of the design is a random sample of this distribution. Hence, estimates of the mean values obtained by simulation sampling can be used to approximate the criteria.
4. Simulation is computationally expensive; hence evaluating the objective or the feasibility of each design using estimates of criteria is costly.
5. Simulation outputs are noisy. Hence the estimates of objective and non-objective criteria are subject to noise. This is why the limits put on non-objective criteria are called stochastic constraints. Moreover, unless infinite simulation effort is spent on each design which is practically impossible, no optimization algorithm could “surely” guarantee finding the optimum design or even determining if a design is feasible.

In the simulation literature, Ranking and Selection (RS) and Multiple Comparison (MC) procedures which include Selection of the Best methods have been the focus of research for many decades. A review of such methods can be found in Kim and Nelson (2006b), Bechhofer et. al. (1995), Goldsman and Nelson (1998).

There are two major paradigms for developing RS methods. Some authors try to bound the probability of correct selection and provide the number of simulation replications required to reach this probability. Bechhofer (1954) introduces the so-called indifference zone method that finds the best with a minimum probability when the difference between the best and the rest of the designs are more than indifference parameter. Since then, many more efficient procedures have been proposed. For examples, see KN procedure by Kim and Nelson (2001), and the KN++ by Kim and Nelson (2006a). The next paradigm is RS procedures that bound the total simulation replications consumed and allocating the simulation budget to the designs such that (approximated) probability of correct selection is maximized. An example of this approach is Optimal Computing Budget Allocation introduced by Chen et. al. (2000).

Although a lot of attention has been focused in the literature on Selection of the Best problems without stochastic constraints, the literature is surprisingly poor in proposing efficient methods for handling such constraints (Kim and Nelson 2007). Butler et. al (2001) combine multiple measures together using utility theory and apply an RS procedure to them. Andradóttir et. al (2005) proposes a procedure for one stochastic constraint. Batur and Kim(2005) provide a procedure for detecting feasibility in multiple stochastic constraint cases.

In this paper, we bridge the research gap by presenting an iterative heuristic algorithm which could be used with a limitation on the simulation budget or probability of correct selection as the terminating condition of the algorithm. The method is useful for Selection of the Best problems with stochastic constraints, though it could also be used for unconstrained problems as well.

In section 2, we define the problem mathematically. In section 3.1, the intuition behind our new method and its overall procedure is provided. In section 3.2, the detailed algorithm is outlined. Section 4 provides several numerical results and section 5 concludes the paper.

## 2 PROBLEM DEFINITION

Assume there are  $\omega$  designs ,indexed by  $j \in J = \{1,2, \dots, \omega\}$ , of a stochastic system among which the best design must be selected. There are  $\zeta + 1$  decision criteria (performance measures) indexed by  $i$  for selecting the best design. Mathematically, the problem is:

$$\begin{aligned} & \min_{j \in J} \mu_{0j} & (1) \\ \text{s.t. } & \mu_{ij} < 0 \quad \forall i \in I = \{1,2, \dots, \zeta\} \end{aligned}$$

In this formulation, the objective function is to minimize one decision criterion, namely the expected value  $\mu_{0j}$ , while constraining the expected value of the other criteria,  $\mu_{ij}$ , to be less than an upper bound. Note that a typical constraint  $\mu_{ij} < 0$  in this formulation is broad enough to cover constraints such as  $\mu'_{ij} > 0$  (using  $\mu_{ij} = -\mu'_{ij} < 0$ ) or  $\mu'_{ij} < b$  where  $b \in \mathbb{R}$  (using  $\mu_{ij} = \mu'_{ij} - b < 0$ ). We assume the set of feasible designs  $\Phi = \{j \in J | \mu_{ij} < 0 \quad \forall i \in I\}$  is nonempty. Also we let  $\Phi' = J \setminus \Phi$  be the set of infeasible designs.

The vector  $\boldsymbol{\mu}_j = (\mu_{0j}, \mu_{1j}, \dots, \mu_{\zeta j})$  is unknown and unobservable. But we assume there exists a simulation model of design  $j$  of the stochastic system; the output of running this model in  $r$ -th replication is the observed vector  $\mathbf{y}_{jr} = (y_{0jr}, y_{1jr}, \dots, y_{\zeta jr})$  which is an independent sample of random vector  $\mathbf{Y}_j = (Y_{0j}, Y_{1j}, \dots, Y_{\zeta j})$ . We assume  $\mathbf{Y}_j$  follows a multivariate normal distribution with mean vector  $\boldsymbol{\mu}_j$  and covariance matrix  $\mathbf{C}_j = [C_{ii'j}]_{i,i' \in \{0,1,\dots,\zeta\}}$ .

## 3 THE NEW METHODOLOGY

In this section, we present our new algorithm for selection of the best with stochastic constraints. We first outline the basic ideas and intuition behind this methodology (section 3.1), and then provide a detailed outline of the algorithm (section 3.2).

### 3.1 Overview

Our heuristic algorithm has an iterative procedure that continues until a terminating condition (e.g. the probability of correct selection or budget) is satisfied.

To initialize the algorithm, each design is simulated with a number of replications denoted by  $\eta \geq 2$ . The idea here is to get initial rough estimates of the expected values of criteria and their variability. In each iteration, the algorithm then screens the eligibility of all designs for more simulation replications. The intuitive idea is that a design should be simulated further if there is reason to believe that the design is both feasible and of sufficiently high quality. Thus, the algorithm employs two measures, which we term the quality and feasibility indicators. Using these indicators, there are then two major rules:

1. A design must not be simulated more if there is sufficient evidence (feasibility indicator) that it is infeasible.
2. A design must not be simulated further if there is sufficient evidence (quality indicator) that its objective function is inferior to that of another design which is believed to be feasible.

Before describing the method further, we need to introduce some notations. Let  $n$  be the counter of the total number of simulation replications spent for all the designs. Denote the total number of simulation replications of design  $j$  right after  $n$ -th simulation replication by  $v_{jn}$ . Obviously  $\sum_{j \in J} v_{jn} = n$ . The sample mean vector  $\bar{y}_{jn} = (\bar{y}_{0jn}, \bar{y}_{1jn}, \dots, \bar{y}_{\zeta jn})$  and sample covariance matrix  $c_{jn} = [c_{ii'jn}]_{i,i' \in \{0,1,\dots,\zeta\}}$  of design  $j \in J$  up to replication  $n$  could then be obtained using:

$$\bar{y}_{ijn} = v_{jn}^{-1} \sum_{r=1}^{v_{jn}} y_{ijr} \quad \forall i \in \{0,1, \dots, \zeta\} \tag{2}$$

$$c_{ii'jn} = (v_{jn} - 1)^{-1} \left( \sum_{r=1}^{v_{jn}} (y_{ijr} y_{i'jr}) - v_{jn} \bar{y}_{ijn} \bar{y}_{i'jn} \right) \quad \forall i, i' \in \{0,1, \dots, \zeta\} \tag{3}$$

After  $n$ -th simulation replication, design  $j$  is labeled as feasible if, when estimated means are used in lieu of the expected values, all the constraints of the problem are satisfied for the design. Specifically, a feasibility label is defined as follows:

$$\hat{F}_{jn} = \begin{cases} 1 & \text{if } \bar{y}_{ijn} < 0 \quad \forall i \in I \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The algorithm then divides the set of designs into two parts called (approximated) feasible and infeasible sets denoted by  $\hat{\Phi}_n$  and  $\hat{\Phi}'_n$  respectively:

$$\hat{\Phi}_n = \{j \in J \mid \hat{F}_{jn} = 1\} \tag{5}$$

$$\hat{\Phi}'_n = \{j \in J \mid \hat{F}_{jn} = 0\} \tag{6}$$

Obviously,  $\hat{\Phi}_n \cup \hat{\Phi}'_n = J$  and  $\hat{\Phi}_n \cap \hat{\Phi}'_n = \{ \}$  for all  $n$ .

Let the best design up to  $n$ -th simulation replication be denoted by  $j_n^*$  which is the design in  $\hat{\Phi}_n$  with least estimated objective function. When there is a tie between the estimated objective functions of two or more designs in  $\hat{\Phi}_n$ , we for simplicity introduce the design with lowest index among these designs as the best. Since  $\hat{\Phi}_n$  might potentially be empty in our algorithm, a “dummy” design 0 with objective value of  $+\infty$  is returned as the optimum:

$$j_n^* = \begin{cases} \min \{j' \mid j' \in \arg \min_{j \in \hat{\Phi}_n} \bar{y}_{0jn}\} & \text{if } |\hat{\Phi}_n| \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

The feasibility and quality indicators used in the proposed algorithm are both probabilistic terms defined on the distribution of the estimated criteria. We assume the estimated criteria of a design follow a multivariate normal distribution. Then the estimated means and co-variances are used for the multivariate normal distribution of each design. We hope that as the number of simulation replications of each design increases, the estimates get more accurate by the laws of large numbers, and that the normality assumption is better justified by central limit theorem.

For the feasibility indicator, denoted by  $\varphi_{jn}$ , the algorithm computes the probability of feasibility as follows:

$$\varphi_{jn} = \Pr \left\{ Z_{\zeta} \left( (\bar{y}_{1jn}, \bar{y}_{2jn}, \dots, \bar{y}_{\zeta jn}), \begin{bmatrix} c_{ii'jn} \\ v_{jn} \end{bmatrix}_{i,i' \in \{1,2,\dots,\zeta\}} \right) < (0,0, \dots, 0) \right\} \tag{8}$$

where  $Z_d(a, b)$  is a  $d$ -dimensional multivariate normal random vector with mean vector  $a$  and covariance matrix  $b$ .

To compute the quality indicator of design  $j \neq j_n^*$ , denoted by  $\tau_{jn}$ , first an indifference zone  $\delta_k > 0$  is defined for iteration  $k$  of the algorithm; then quality indicator would be the probability that the objective function of design  $j$  is at least  $\delta_k$  units better than that of design  $j_n^*$  as the following equation shows:

$$\tau_{jn} = \begin{cases} \Pr \left\{ Z_1 \left( \bar{y}_{0jn} - \bar{y}_{0j_n^*n}, \frac{c_{00jn}}{v_{jn}} + \frac{c_{00j_n^*n}}{v_{j_n^*n}} \right) < -\delta_k \right\} & \text{if } j_n^* \neq 0 \\ 1 & \text{if } j_n^* = 0 \end{cases} \quad (9)$$

With those preliminaries, we can now describe the overall procedure of the algorithm in a typical iteration  $k$ . Each iteration has three phases as follows:

1. Screen the feasibility of the best design in the set  $\hat{\Phi}_n$  which consists of the designs that are estimated to be feasible.
2. Screen the quality of the designs other than the best in  $\hat{\Phi}_n$ .
3. Screen the feasibility and quality of the designs in  $\hat{\Phi}'_n$  which consists of the designs that are estimated to be infeasible.

The algorithm is fully sequential and simulates a single replication of a design at a time.

In phase 1, the set  $\hat{\Phi}_n$  is ordered based on increasing estimated objective function. Then, the best design thus far ( $j_n^*$ ) must provide more evidence of being feasible. Therefore it is simulated as many simulation replications as necessary until one of the below conditions are satisfied:

1. The evidence is obtained, that is the feasibility indicator of the best is at least  $1 - \beta_k$  where  $\{\beta_k\}_{k=1}^\infty$  is a sequence of feasibility thresholds that gradually converges to zero as the iterations of the algorithm increases.
2. The best turns out to be infeasible, that is the feasibility label of the best becomes zero. In this case, the “best” is no longer the best because it is labeled infeasible. This old best is moved to the approximated infeasible set. Then a new best, if approximated feasible set is not empty, is determined and phase one is repeated for the new best.
3. The total number of simulation replications of the best design performed in the current iteration reaches a maximum allowed.
4. The “best” turns out to be actually worse than another design in the approximated feasible set. Then, the phase is repeated after a new best is determined.

In phase 2, the algorithm seeks more evidence that all the non-best members of the set  $\hat{\Phi}_n$  are actually not better than the best. Hence more simulation replications are spent to reach the desired evidences. We define a quality threshold sequence  $\{\alpha_k\}_{k=1}^\infty$  that gradually converges to zero as the iterations of the algorithm increases. In this phase, first the eligibility of the second best design in the ordered set  $\hat{\Phi}_n$  is studied for further simulation by computing the quality indicator of this design. If the quality indicator of the second best is more than  $\alpha_k$ , then it means that the second best might actually be better than the already known best, hence either the best or the second best (whichever has a higher estimated objective function variance and whose number of simulation replications in the current iteration has not reached the maximum allowed) is simulated with one more replication to get more accurate estimates of objective function. After setting  $n = n + 1$ , some estimates are updated with the data from the new simulation replication. If the previous second best design is still feasible and is not the new best after the new simulation replication, the quality indicator of this design is computed again and the eligibility of this design or the best for more simulation is examined at this stage. If the quality is less than the threshold, the algorithm moves to the third design in the ordered set  $\hat{\Phi}_n$ . This process continues to the last design in  $\hat{\Phi}_n$ .

After screening the quality of the approximated feasible designs, the algorithm in phase 3 computes both quality and feasibility of the designs in  $\hat{\Phi}'_n$  one at a time. A design in  $\hat{\Phi}'_n$  is simulated one more replication if its quality and feasibility indicators are more than  $\alpha_k$  and  $\beta_k$  respectively and the number of simulation replications of the design in this iteration has not reached a predefined maximum. However, after simulating a design in  $\hat{\Phi}'_n$ , if the design turns out to be feasible, then the iteration is immediately restarted (going back to phase 1). This is done because a design that is recently believed to be feasible, might actually be the new best introduced design and hence the quality indicators of all the other designs and consequently their eligibility for further simulation must be defined based on the quality of this design, not the old best.

The three parameters  $\alpha_k$ ,  $\beta_k$  and  $\delta_k$  are reduced for each iteration based on a user defined plan. An example plan is the following:

$$\begin{aligned} \alpha_k &= (c_\alpha)^{k-1} \alpha_1 \\ \beta_k &= (c_\beta)^{k-1} \beta_1 \\ \delta_k &= (c_\delta)^{k-1} \delta_1 \end{aligned} \quad (10)$$

where  $\alpha_1 \in (0,1)$ ,  $\beta_1 \in (0,1)$ ,  $\delta_1 \in (0, +\infty)$  are iteration-one parameters and  $c_\alpha \in (0,1)$ ,  $c_\beta \in (0,1)$ ,  $c_\delta \in (0,1)$  are decrease coefficients.

As we said earlier, the total number of simulation replications spent on each design in each iteration must not exceed an upper bound denoted by  $\gamma$ . This is because the estimated objective functions of a non-best design in  $\hat{\Phi}_n$  and the best ( $j_n^*$ ) might get arbitrarily close. Hence, when unknown to us, none of these two designs are actually the optimum design of the

problem, clarifying which design among these two is better is just futile, though very costly. Therefore, the algorithm postpones the “fight” between these two designs until the actual best hopefully comes out of shadow and overwhelms both or the algorithm continues the fight partly in the next iteration(s).

### 3.2 The Algorithm

After providing an overall picture of the algorithm of this paper in previous section, we present detailed steps here.

#### I: Initialization:

- Step 1: Simulate each design with  $\eta \geq 2$  replications and let  $n = \eta\omega$ . Set  $v_{jn} = \eta$  for  $j \in J$ .  
 Step 2: For  $j \in J$ , compute the sample means and sample co-variances using Equations 2 and 3.  
 Step 3: For  $j \in J$ , compute the feasibility label using Equation 4.  
 Step 4: Create the approximated feasible and infeasible sets  $\widehat{\Phi}_n, \widehat{\Phi}'_n$  using Equations 5 and 6.  
 Step 5: Determine the introduced optimum  $j_n^*$  using Equation 7.  
 Step 6: Set iteration counter  $k = 1$ .

#### II: Iterative Phase:

- Step 7: Let  $\xi_j$  denote the counter of simulation replications of design  $j$  performed in the current iteration. Set  $\xi_j = 0$  for  $j \in J$ .  
 Step 8: If  $|\widehat{\Phi}_n| \leq 1$ , then go to step 13.  
 Step 9: Let  $j_h$  be the  $h$ -th best design of the set  $\widehat{\Phi}_n$  in terms of estimated objective function. If there is tie between estimated objective functions of two or more designs, consider the lower-indexed design as a better design. Obviously  $j_1 = j_n^*$  and  $\bar{y}_{0j_1n} \leq \bar{y}_{0j_2n} \leq \dots \leq \bar{y}_{0j_{|\widehat{\Phi}_n|}n}$  and if  $\bar{y}_{0j_hn} = \bar{y}_{0j_{h+1}n}$  for some  $h \in \{1, 2, \dots, |\widehat{\Phi}_n| - 1\}$ , then  $j_h < j_{h+1}$ . Let  $S = |\widehat{\Phi}_n|$ .

#### II.I: Screening the feasibility of the best

- Step 10: If  $\xi_{j_1} < \gamma$ , then do 10.1-10.2 .  
 10.1. Compute  $\varphi_{j_1n}$ , the feasibility probability of design  $j_1$ , using Equation 8.  
 10.2. If  $\varphi_{j_1n} < 1 - \beta_k$ , then do 10.2.1-10.2.7 :  
 10.2.1. Simulate  $j_1$  for one more independent replication and let  $\xi_{j_1} = \xi_{j_1} + 1$  and  $n = n + 1$ .  
 10.2.2. Set  $\tilde{j} = j_1$  and then set:

$$v_{jn} = \begin{cases} v_{j,n-1} + 1 & \text{if } j = \tilde{j} \\ v_{j,n-1} & \text{otherwise} \end{cases} \quad (11)$$

$$\widehat{\Phi}_n = \widehat{\Phi}_{n-1}; \widehat{\Phi}'_n = \widehat{\Phi}'_{n-1} \quad (12)$$

- 10.2.3. For  $j = j_1$ , calculate parameters in Equations 2-4.  
 10.2.4. For  $j \in J \setminus \{j_1\}$ , set  $\bar{\mathbf{y}}_{jn} = \bar{\mathbf{y}}_{j,n-1}$ ,  $\mathbf{c}_{jn} = \mathbf{c}_{j,n-1}$  and  $\widehat{F}_{jn} = \widehat{F}_{j,n-1}$ .  
 10.2.5. If  $\widehat{F}_{j_1n} = 0$ , then move  $j_1$  from  $\widehat{\Phi}_n$  to  $\widehat{\Phi}'_n$ , update the best  $j_n^*$  in Equation 7 and go to step 8.  
 10.2.6. If  $\widehat{F}_{j_1n} = 1$  and  $\bar{y}_{0j_1n} \leq \bar{y}_{0j_2n}$ , then set  $j_n^* = j_{n-1}^*$  and go to step 10.  
 10.2.7. If  $\widehat{F}_{j_1n} = 1$  and  $\bar{y}_{0j_1n} > \bar{y}_{0j_2n}$ , let  $j_n^* = j_2$  and go to step 9.

#### II.II: Screening the quality of estimated feasible set

- Step 11: For  $h = 2, 3, \dots, S$ , do 11.1.  
 11.1. If  $\xi_{j_h} < \gamma$  or  $\xi_{j_h^*} < \gamma$ , then do 11.1.1-11.1.2.  
 11.1.1. Compute  $\tau_{j_hn}$ , the quality probability of design  $j_h$ , using Equation 9.  
 11.1.2. If  $\tau_{j_hn} > \alpha_k$ , then do 11.1.2.1-11.1.2.2.  
 11.1.2.1. If  $\frac{c_{00j_hn}}{v_{j_hn}} > \frac{c_{00j_n^*n}}{v_{j_n^*n}}$  and  $\xi_{j_h} < \gamma$ , then do 11.1.2.1.1-11.1.2.1.7:  
 11.1.2.1.1. Simulate  $j_h$  for one more replication and let  $\xi_{j_h} = \xi_{j_h} + 1$  and  $n = n + 1$ .  
 11.1.2.1.2. Set  $\tilde{j} = j_h$  and update Equations 11-12.  
 11.1.2.1.3. For  $j = j_h$ , calculate parameters in Equations 2-4.  
 11.1.2.1.4. For  $j \in J \setminus \{j_h\}$ , set  $\bar{\mathbf{y}}_{jn} = \bar{\mathbf{y}}_{j,n-1}$ ,  $\mathbf{c}_{jn} = \mathbf{c}_{j,n-1}$  and  $\widehat{F}_{jn} = \widehat{F}_{j,n-1}$ .

- 11.1.2.1.5. If  $\hat{F}_{j_h n} = 0$ , then move  $j_h$  from  $\hat{\Phi}_n$  to  $\hat{\Phi}'_n$  and set  $j_n^* = j_{n-1}^*$ .
- 11.1.2.1.6. If  $\hat{F}_{j_h n} = 1$  and  $\bar{y}_{0j_h n} < \bar{y}_{0j_{n-1}^* n}$ , then set  $j_n^* = j_h$  and go to step 9.
- 11.1.2.1.7. If  $\hat{F}_{j_1 n} = 1$  and  $\bar{y}_{0j_h n} \geq \bar{y}_{0j_{n-1}^* n}$ , then set  $j_n^* = j_{n-1}^*$  and go to step 11.1.
- 11.1.2.2. If  $\frac{c_{00j_h n}}{v_{j_h n}} \leq \frac{c_{00j_n^* n}}{v_{j_n^* n}}$  and  $\xi_{j_n^*} < \gamma$ , then do 11.1.2.2.1-11.1.2.2.7:
  - 11.1.2.2.1. Simulate  $j_n^*$  for one more replication and let  $\xi_{j_n^*} = \xi_{j_n^*} + 1$  and  $n = n + 1$ .
  - 11.1.2.2.2. Set  $\tilde{j} = j_{n-1}^*$  and update Equations 11-12.
  - 11.1.2.2.3. For  $j = j_{n-1}^*$ , calculate parameters in Equations 2-4.
  - 11.1.2.2.4. For  $j \in J \setminus \{j_{n-1}^*\}$ , set  $\bar{y}_{jn} = \bar{y}_{j,n-1}$ ,  $\mathbf{c}_{jn} = \mathbf{c}_{j,n-1}$  and  $\hat{F}_{jn} = \hat{F}_{j,n-1}$ .
  - 11.1.2.2.5. If  $\hat{F}_{j_{n-1}^* n} = 0$ , then move  $j_{n-1}^*$  from  $\hat{\Phi}_n$  to  $\hat{\Phi}'_n$ , update  $j_n^*$  in Equation 7 and go to step 8.
  - 11.1.2.2.6. If  $\hat{F}_{j_{n-1}^* n} = 1$  and  $\bar{y}_{0j_{n-1}^* n} \neq \min_{j \in \hat{\Phi}_n} \bar{y}_{0jn}$ , then update  $j_n^*$  in Equation 7 and go to step 9.
  - 11.1.2.2.7. If  $\hat{F}_{j_{n-1}^* n} = 1$  and  $\bar{y}_{0j_{n-1}^* n} = \min_{j \in \hat{\Phi}_n} \bar{y}_{0jn}$ , then set  $j_n^* = j_{n-1}^*$  and go to step 11.1.

### II.III: Screening the quality and feasibility of estimated infeasible set

*Step 12:* If  $\hat{\Phi}'_n$  is empty, go to step 15.

*Step 13:* Let  $j'_h$  be the  $h$ -th best design of the set  $\hat{\Phi}'_n$  in terms of estimated objective function. If there is tie between estimated objective functions of two or more designs, consider the lower-indexed design as a better design. Let  $S' = |\hat{\Phi}'_n|$ .

*Step 14:* For  $h = 1, 2, \dots, S'$ , do 14.1.

14.1. If  $\xi_{j'_h} < \gamma$ , then do 14.1.1-14.1.2.

14.1.1. Compute  $\tau_{j'_h n}$ , the quality probability of design  $j'_h$ , using Equation 9.

14.1.2. If  $\tau_{j'_h n} > \alpha_k$ , then do 14.1.2.1-14.1.2.2.

14.1.2.1. Compute  $\varphi_{j'_h n}$ , the feasibility probability of design  $j'_h$ , using Equation 8.

14.1.2.2. If  $\varphi_{j'_h n} > \beta_k$ , then do 14.1.2.2.1-14.1.2.2.6:

14.1.2.2.1. Simulate  $j'_h$  for one more replication and let  $\xi_{j'_h} = \xi_{j'_h} + 1$  and  $n = n + 1$ .

14.1.2.2.2. Set  $\tilde{j} = j'_h$  and update Equations 11-12.

14.1.2.2.3. For  $j = j'_h$ , calculate parameters in Equations 2-4.

14.1.2.2.4. For  $j \in J \setminus \{j'_h\}$ , set  $\bar{y}_{jn} = \bar{y}_{j,n-1}$ ,  $\mathbf{c}_{jn} = \mathbf{c}_{j,n-1}$  and  $\hat{F}_{jn} = \hat{F}_{j,n-1}$ .

14.1.2.2.5. If  $\hat{F}_{j'_h n} = 1$ , then move  $j'_h$  from  $\hat{\Phi}'_n$  to  $\hat{\Phi}_n$ , update  $j_n^*$  in Equation 7 and go to step 8.

14.1.2.2.6. If  $\hat{F}_{j'_h n} = 0$ , then set  $j_n^* = j_{n-1}^*$  and go to step 14.1.

*Step 15:* If a termination condition of the algorithm holds, return  $j_n^*$  as the optimum and exit; otherwise set  $k = k + 1$  and go to step 7.

## 4 NUMERICAL RESULTS

In this section we analyze the performance of the new algorithm through a set of experiments. In the first section we consider how the correct selection probability advances as the number of simulation replications is increased. In the second experiment we consider how the problem size affects the algorithm performance. In the third section we illustrate how various types of infeasibility settings for the problem affect the algorithm, and finally, in the fourth section we show how the number of constraints influences performance.

### 4.1 Average Correct Selection Curve

The empirical probability of correct selection is a key performance measure of the algorithm. Intuitively we expect this probability to increase as the number of simulation runs increases, but the rate at which this occurs is of great interest. Thus, in our first experiment we will generate a average correct selection curve for a single problem, as a function of the number of simulation replications. The properties of the problem are given in Table 1.

We set the parameters of the algorithm as follows:  $\eta = 5$ ,  $\gamma = 10$ ,  $c_\alpha = 0.95$ ,  $c_\beta = 0.95$ ,  $c_\delta = 0.95$ ,  $\alpha_1 = 0.5$ ,  $\beta_1 = 0.5$ , and  $\delta_1 = 10$ . These are our educated guesses and are not guaranteed to be the best parameter settings. More experimentation is needed to determine guidelines for setting these parameters. The terminating condition of the algorithm was set to a budget of 10000 simulation replications.

For this experiment we ran the algorithm 1000 times to solve the problem. We define the correct selection ( $CS_n$ ) as the following indicator function

$$CS_n = \begin{cases} 1 & \text{if } j_n^* = 1 \\ 0 & \text{otherwise} \end{cases}$$

This indicator is then averaged over the 1000 replications, and Figure 1 shows the resulting average correct selection function. We observe from those results that for this problem, to get an average 90% and 95% correct selection, one needs to spend an average simulation replication budget of 4496 and 5698 respectively. In other words, a budget of more than 4496 would find the optimum with more than 90% probability.

Table 1: Properties of the test problem

Features of the problem	Description
Number of all designs	100
Feasible Designs	50
Stochastic Constraints	5
Joint distribution of Criteria	6-dimensional multivariate normal
Mean Values of objective function (Criterion 0) of each design (except design 1)	Uniformly randomly from (0,100]
Mean Values of objective function (Criterion 0) of design 1	0
Mean Values of Criteria 1-5 for each feasible designs	Uniformly randomly from [-100,0)
Mean Values of Criteria 1-5 for each infeasible designs	Uniformly randomly from [-100,100] and guaranteeing that at least one of the mean values is positive.
Variances of the Criteria of each design	Uniformly randomly from (0,50]
Correlations between each two criteria of each design	Uniformly randomly from [-1,1] and guaranteeing that the covariance matrix of the distribution of criteria remains positive definite.

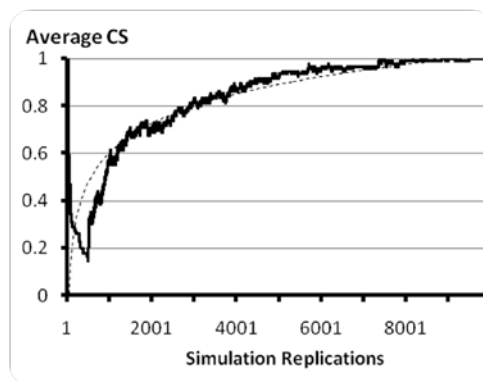


Figure 1: The average correct selection indicator function obtained from 1000 runs of the algorithm and a logarithmic trend line (dashed line)

Furthermore, at least from this single example, the rate of increase for the empirical probability of correct selection appears to be well modeled by a logarithmic function. That is, the increase is initially very rapid but when the probability approaches one improving it further requires a large increase in the average number of simulation replications.

## 4.2 Problem Size

In the second experiment we investigate the effect of problem size and the number of feasible designs on the performance of the algorithm. To this end we generated random problems with 20,60,200,500,1000 designs, each of which with 3 different levels of feasible designs: one with too many feasible designs (95% feasible), one with too many infeasible designs (5% feasible) and another with balanced number of feasible and infeasible designs (50% feasible). The specific settings are shown in the first two columns of Table 2.

For this experiment we generated 500 problems with given number of all designs and feasible designs. Except these two features, the other features of the generated problems were the same as Table 1. Moreover, the parameters of the algorithm were the same as those described in section 4.1. The algorithm was run 1 time on each problem. Table 2 summarizes the results obtained.

From Table 2 we observe the trend that when the fraction of feasible designs in a set of designs is increased, the problem gets more difficult. This is intuitive because more infeasible designs means more opportunity of quarantining them in the approximated infeasible set.

Another trend that can be observed from Table 2 is that as the total number of designs increases, the required budget increases nonlinearly. (From these results, the required budget appears to be growing exponentially.)

Table 2: Average Simulation Replications Budget for Different Correct Selections in 500 random problems with 1 single run of the algorithm on each problem

# of designs	# of feasible designs	Average # of Simulation Replications		
		CS=75%	CS=90%	CS=95%
20	1	<100	147	243
	10	250	408	507
	19	359	960	1499
60	3	<300	356	457
	30	372	565	1007
	57	1367	3108	3973
200	10	1069	1316	2294
	100	3110	5671	8848
	190	7501	14520	16819
500	25	2871	3462	4173
	250	11889	26275	36439
	475	20813	39936	63057
1000	50	6614	11590	15194
	500	39462	103703	179340
	950	102492	161834	219599

Table 3: Average Budget. In Configuration 1, the infeasible designs have a better objective function than the best, but in configuration 2, the infeasible designs are worse than the best.

Configuration	Average # of Simulation Replications		
	CS=75%	CS=90%	CS=95%
1	13670	28940	42980
2	7697	16940	27497

## 4.3 Infeasibility Challenge

The number of infeasible designs in a problem has a mixed effect on the difficulty of the problem. On the one hand, the more the number of infeasible designs the easier the work of the algorithm to eliminate them as part of the estimated infeasible list and spend the computation effort on the reduced feasible set. On the other hand, when the infeasible designs have very good objective function values (better than the best feasible design) and are only marginally infeasible, i.e. the mean value of all of



the criteria in the constraints are less than a small positive number, then the chance of wrongly labeling them as feasible and consequently announcing them as the best are high.

To illustrate this mixed effect of infeasible designs, we conducted an experiment with 500 designs with 250 feasible designs and 2 different configurations:

1. The mean objective function of all the designs are better than that of the best feasible design.
2. The mean objective function of all the designs are worse than that of the best design.

The other features of the problem were the same as section 4.1. We ran the algorithm on 1000 generated problems with above features. Table 3 reports the results that confirms our initial intuition.

#### 4.4 Number of Constraints

Finally, we address the issue of how the number of constraints affects the performance of the algorithm. This time, a new experiment was conducted on a problem with 20 designs, of which 10 are feasible. We used 1, 5, 10, and 20 stochastic constraints. The algorithm was run one time on 1000 random problems for each constraint-size configuration. The other features of the generated problems and the algorithm are the same as section 4.1.

Table 4: Average Budget for Different Number of Stochastic Constraints

# of Stochastic Constraints	Average # of Simulation Replications		
	CS=75%	CS=90%	CS=95%
1	104	234	386
5	246	396	519
10	316	728	1709
20	867	2067	3591

Table 4 reports the results of the above experiment. We note from these results that the increase in the number of constraints has mixed effects on the total computational time to solve the problem with our algorithm. On the one hand, as the number of constraints grows the chance of detecting the infeasibility of infeasible designs increases, and this facilitates the selection of the feasible best process by eliminating the infeasible designs. (This effect is further confirmed by this experiment.) On the other hand, when there is a large number of stochastic constraints, there is a greater chance of incorrectly listing a feasible design into the approximated infeasible set exists. Moreover, our algorithm has to find the feasibility indicator, whose computations exponentially grows with the increase in the number of stochastic constraints.

## 5 CONCLUSION

In this paper, we introduced a new methodology for selection of the best when there are some stochastic constraints. We presented the theory of the asymptotic almost-sure convergence of the method to optimal solution. We used many numerical examples to illustrate how the correct selection would behave in practice. The results show the effectiveness of the algorithm in dealing with different test problems.

Several research directions are open based on the new methodology presented here. One open question is the study of the correct selection probability analytically, that is, finding an expression (or a lower bound) for the probability of correct selection under general assumption.

We believe Bayesian prior distributions could be defined for the means and covariance matrices of the design and then, the posterior distribution of the means be used for calculating the quality and feasibility indicators.

We used educated guesses for the parameters of our algorithm, but more experiments are required to study the performance of the algorithm under different parameter settings.

## REFERENCES

- Andradóttir, S., D. Goldsman, and S. Kim. 2005. Finding the best in the presence of a stochastic constraint, In *Proceedings of the 2005 Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 732-738.
- Batur, D., and S. Kim. 2005. Procedures for feasibility detection in the presence of multiple constraints. In *Proceedings of the 2005 Winter Simulation Conference*. ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 692- 698.
- Bechhofer, R. E., T. J. Santner, and D. Goldsman. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening and Multiple Comparisons*. New York: John Wiley & Sons.

- Butler, J., D. J. Morrice, and P. W. Mullarkey. 2001. A multiple attribute utility theory approach to ranking and selection. *Management Science*, 47(6): 800-816.
- Chen, C., J. Lin, E. Yucesan, and S. E. Chick. 2000. Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization, *Discrete Event Dynamic Systems: Theory and Applications*, 10(3): 251–270.
- Goldman, D., and B. L. Nelson. 1998. Comparing systems via simulation. In *Handbook of Simulation*, ed. J. Banks, 273–306. New York: John Wiley.
- Kim, S., and B. L. Nelson. 2001. A fully sequential procedure for indifference zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation*, 11(3): 251-273.
- Kim, S., and B. L. Nelson. 2006a. On the asymptotic validity of fully sequential selection procedures for steady-state simulation. *Operations Research*, 54(3): 475-488.
- Kim, S., and B. L. Nelson. 2006b. Selecting the best system. In Henderson and Nelson (eds.), *Handbooks in operations research and management science*, 13, 501-534, Elsevier.
- Kim S., and B. L. Nelson. 2007. Recent advances in ranking and selection. In *Proceedings of the 2007 Winter Simulation Conference*. ed. S. G. Henderson, B. Biller, M.-H. Hsieh, J. Shortle, J. D. Tew, and R. R. Barton, 162- 172.

#### AUTHOR BIOGRAPHIES

**ALIREZA KABIRIAN** is an assistant professor in the College of Business and Public Policy at the University of Alaska-Anchorage. In 2009, he received his Ph.D. in Industrial Engineering from Iowa State University. He also obtained his degrees of M.Sc. and B.Sc. in Industrial Engineering respectively from Sharif University of Technology and Azad University of Najafabad in Iran. His specific research interests are simulation optimization and energy planning. He is a member of INFORMS and IIE. His e-mail address is <[a\\_kabirian@yahoo.com](mailto:a_kabirian@yahoo.com)>.

**SIGURDUR OLAFSSON** is an associate professor in the Department of Industrial and Manufacturing Systems Engineering at Iowa State University. He received a M.S. and Ph.D. from the Department of Industrial Engineering at the University of Wisconsin – Madison, and B.S. in mathematics from the University of Iceland. His research interests focus on discrete optimization, especially metaheuristics, and their uses in simulation optimization and data mining. His web page can be found at <[www.public.iastate.edu/~olafsson](http://www.public.iastate.edu/~olafsson)>.