

**OPTIMAL SAMPLING LAWS FOR CONSTRAINED SIMULATION OPTIMIZATION ON
FINITE SETS: THE BIVARIATE NORMAL CASE**

Susan R. Hunter

Operations Research and Information Engineering
Cornell University
Ithaca, NY USA

Nugroho Artadi Pujowidianto

Industrial and Systems Engineering
National University of Singapore
SINGAPORE

Chun-Hung Chen

Department of Electrical Engineering
National Taiwan University
Taipei, TAIWAN

Loo Hay Lee

Industrial and Systems Engineering
National University of Singapore
SINGAPORE

Raghu Pasupathy

Industrial and Systems Engineering
Virginia Tech
Blacksburg, VA, USA

Chee Meng Yap

Industrial and Systems Engineering
National University of Singapore
SINGAPORE

ABSTRACT

Consider the context of selecting an optimal system from amongst a finite set of competing systems, based on a “stochastic” objective function and subject to a single “stochastic” constraint. In this setting, and assuming the objective and constraint performance measures have a bivariate normal distribution, we present a characterization of the optimal sampling allocation across systems. Unlike previous work on this topic, the characterized optimal allocations are asymptotically exact and expressed explicitly as a function of the correlation between the performance measures.

1 INTRODUCTION

The simulation-optimization (SO) problem is a nonlinear optimization problem where the objective and constraint functions, defined on a set of candidate solutions, are observable only through consistent estimators. The consistent estimators can be defined implicitly, e.g., through a stochastic simulation model. The focus of this paper is a variation of the SO problem where there exist multiple performance measures defined on a finite set of candidate solutions, or “systems,” one of which is primary and called the objective function, while the others are secondary and called the constraint functions. Given that these objective and constraint function values are observable only as consistent estimators from a stochastic simulation, the constrained SO problem is to identify the “best feasible system” using only the simulation output, where the “best feasible system” is the system with the best objective function value amongst those systems whose constraint values satisfy predetermined threshold levels. While we limit our discussion here to the constrained SO problem, entry points into the corresponding objective-only and constraint-only problems can be found in Kim and Nelson (2006), Batur and Kim (2010), Chen et al. (2000), Glynn and Juneja (2004), and Szechtman and Yücesan (2008).

There exist two broad problem statements in the literature on “stochastically” constrained SO on finite sets. The first problem statement is to identify the best feasible system in finite-time with a pre-specified probabilistic guarantee. Andradóttir and Kim (2010) propose constrained Ranking and Selection (R&S) to solve the finite-time problem, after assuming the performance measures are independent and identically distributed (iid) replications from a bivariate normal distribution. The second problem statement in the literature, and the focus of this paper, is to find a computing budget allocation that minimizes the probability of selecting any system other than the best feasible system. This problem statement yields results that are asymptotic in nature, and solutions have been proposed by Lee et al. (2011) for normally distributed performance measures and by Hunter and Pasupathy (2011) for performance measures from general light-tailed distributions.

Within the context of the second problem statement, the effect of correlation between the objective and constraint functions on the asymptotically optimal sampling allocation has not been explored. In Lee et al. (2011), Bonferroni bounds dissolve the effect of dependence in the solution, and hence correlation is not explicitly taken into account in the allocation. In Hunter and Pasupathy (2011), the objective and constraints are assumed independent. The effect of correlation on the optimal allocation thus remains an important open question. Towards answering this question, we present a series of results that precisely characterize the effect of correlation on the optimal allocation. The characterized optimal allocations are asymptotically exact and expressed explicitly as a function of the correlation between the performance measures. Like Andradóttir and Kim (2010), we assume the performance measures are iid replications from a bivariate normal distribution. For brevity, we present results without proofs.

1.1 Problem Statement

Our problem statement is a one-constraint version of the problem statement presented in Hunter and Pasupathy (2011). We consider a finite set $i = 1, \dots, r$ of systems, each with an unknown objective value $h_i \in \mathbb{R}$ and unknown constraint value $g_i \in \mathbb{R}$. Given a constant $\gamma \in \mathbb{R}$, we wish to select the system with the lowest objective value h_i , subject to the constraint $g_i \leq \gamma$. That is, we consider

$$\begin{aligned} \text{Problem } P: \quad & \text{Find } \arg \min_i h_i \\ & \text{s.t. } g_i \leq \gamma, \end{aligned}$$

where h_i and g_i are expectations, estimates of h_i and g_i are observed together through simulation as sample means, and a unique solution to Problem P is assumed to exist.

Let $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r)$ be a vector denoting the proportion of the total sampling budget n given to each system, so that $\sum_{i=1}^r \alpha_i = 1$ and $\alpha_i \geq 0$ for all $i = 1, \dots, r$. Let the system having the smallest estimated objective value amongst the estimated-feasible systems be selected as the estimated solution to Problem P . Then we ask, what vector of proportions α maximizes the rate of decay of the probability that this procedure returns a suboptimal solution to Problem P ?

1.2 Organization

We present notation and assumptions in Section 2. We derive the rate function of the probability of false selection in Section 3. In Section 4 we provide an optimal allocation strategy that maximizes the rate of decay of the probability of false selection, and in Section 5 we provide an approximate solution to the maximization problem when it is reasonable to assume that the optimal allocation of the best feasible system is much higher than the optimal allocation of other systems. Section 6 contains concluding remarks.

2 PRELIMINARIES

Let (H_i, G_i) be a random output vector from the simulation. We make the following assumption throughout the paper.

Assumption 1 We assume we may obtain iid replicates of the bivariate normal random vector (H_i, G_i) , with correlation ρ_i , $|\rho_i| < 1$. That is,

$$\begin{bmatrix} H_i \\ G_i \end{bmatrix} \sim BVN \left(\begin{bmatrix} h_i \\ g_i \end{bmatrix}, \Sigma_i \right),$$

where Σ_i is the variance-covariance matrix

$$\Sigma_i = \begin{pmatrix} \text{Var}(H_i) & \text{Cov}(H_i, G_i) \\ \text{Cov}(G_i, H_i) & \text{Var}(G_i) \end{pmatrix} = \begin{pmatrix} \sigma_{h_i}^2 & \rho_i \sigma_{h_i} \sigma_{g_i} \\ \rho_i \sigma_{h_i} \sigma_{g_i} & \sigma_{g_i}^2 \end{pmatrix},$$

and $\sigma_{h_i}^2 < \infty, \sigma_{g_i}^2 < \infty$. We denote the vector of correlations between the objective and constraint functions across systems as $\rho = (\rho_1, \rho_2, \dots, \rho_r)$.

We also assume the systems are simulated independent of each other.

Assumption 2 The output random vectors (H_i, G_i) are mutually independent for all $i = 1, \dots, r$.

Under Assumption 1, by the Gärtner-Ellis theorem (see, e.g., Dembo and Zeitouni 1998), $(\hat{H}_i, \hat{G}_i) = (\frac{1}{\alpha_i n} \sum_{k=1}^{\alpha_i n} H_{ik}, \frac{1}{\alpha_i n} \sum_{k=1}^{\alpha_i n} G_{ik})$ satisfies the large deviations principle (LDP) with good rate function

$$\begin{aligned} \alpha_i I_i(x, y) &= \frac{\alpha_i}{2} [x - h_i \quad y - g_i] \Sigma_i^{-1} \begin{bmatrix} x - h_i \\ y - g_i \end{bmatrix} \\ &= \frac{\alpha_i}{2(1 - \rho_i^2)} \left(\frac{(x - h_i)^2}{\sigma_{h_i}^2} - \frac{2\rho_i(x - h_i)(y - g_i)}{\sigma_{h_i} \sigma_{g_i}} + \frac{(y - g_i)^2}{\sigma_{g_i}^2} \right). \end{aligned}$$

Since the marginal distributions for each H_i and G_i are also normal, \hat{H}_i and \hat{G}_i likewise satisfy the LDP with good rate functions

$$\alpha_i I_i(x) = \frac{\alpha_i(x - h_i)^2}{2\sigma_{h_i}^2} \quad \text{and} \quad \alpha_i J_i(y) = \frac{\alpha_i(y - g_i)^2}{2\sigma_{g_i}^2}.$$

For notational convenience, let system 1 be the best feasible system. We also make the following assumption.

Assumption 3 We assume $h_i \neq h_1$ for all $i = 2, \dots, r$ and $g_i \neq g_1$ for all $i = 1, \dots, r$.

That is, Assumption 3 states that no system sits exactly on the the constraint value or has an objective value exactly equal to that of system 1. Assumption 3 ensures that the relevant rate functions are nonzero.

3 RATE FUNCTION OF THE PROBABILITY OF FALSE SELECTION

Let us formulate the probability of false selection as the probability that system 1 is declared infeasible or that system 1 is beaten in estimated objective value by another estimated-feasible system. That is,

$$P\{FS\} = P\{(\hat{G}_1 > \gamma) \cup (\cup_{i=2}^r (\hat{G}_i \leq \gamma) \cap (\hat{H}_i \leq \hat{H}_1))\} \tag{1}$$

The $P\{FS\}$ in equation (1) has the lower bound

$$\max \left(P\{\hat{G}_1 > \gamma\}, \max_{2 \leq i \leq r} P\{(\hat{G}_i \leq \gamma) \cap (\hat{H}_i \leq \hat{H}_1)\} \right)$$

and the upper bound

$$r \times \max \left(P\{\hat{G}_1 > \gamma\}, \max_{2 \leq i \leq r} P\{(\hat{G}_i \leq \gamma) \cap (\hat{H}_i \leq \hat{H}_1)\} \right)$$

such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log P\{FS\} = \lim_{n \rightarrow \infty} \frac{1}{n} \log \max \left(P\{\hat{G}_1 > \gamma\}, \max_{2 \leq i \leq r} P\{(\hat{G}_i \leq \gamma) \cap (\hat{H}_i \leq \hat{H}_1)\} \right). \quad (2)$$

We note that as in Hunter and Pasupathy (2011), there are two main terms: one for the feasibility of system 1, and a second term representing the event that another estimated-feasible system “beats” system 1 in estimated objective value. The following proposition states that the overall rate function is the minimum rate function of the probability of each of these two events.

Proposition 1 The rate function for $P\{FS\}$ is

$$-\lim_{n \rightarrow \infty} \frac{1}{n} \log P\{FS\} = \min \left(\alpha_1 J_1(\gamma), \min_{2 \leq i \leq r} \left(\inf_{x_i \leq x_1, y_i \leq \gamma} (\alpha_1 I_1(x_1) + \alpha_i I_i(x_i, y_i)) \right) \right).$$

For notational simplicity, let the rate function $\inf_{x_i \leq x_1, y_i \leq \gamma} (\alpha_1 I_1(x_1) + \alpha_i I_i(x_i, y_i))$ be denoted by $K_i(\alpha_1, \alpha_i)$. Then expanding the normal and bivariate normal rate functions, we have

$$K_i(\alpha_1, \alpha_i) = \inf_{\substack{x_i \leq x_1 \\ y_i \leq \gamma}} \left(\frac{\alpha_1 (x_1 - h_1)^2}{2\sigma_{h_1}^2} + \frac{\alpha_i}{2(1 - \rho_i^2)} \left(\frac{(x_i - h_i)^2}{\sigma_{h_i}^2} - \frac{2\rho_i (x_i - h_i)(y_i - g_i)}{\sigma_{h_i} \sigma_{g_i}} + \frac{(y_i - g_i)^2}{\sigma_{g_i}^2} \right) \right).$$

The following lemma provides the location of the infimum in $K_i(\alpha_1, \alpha_i)$.

Lemma 1 Under Assumption 1, the infimum in $K_i(\alpha_1, \alpha_i)$ is achieved at

$$x_1^* = x_i^* = \frac{(\alpha_1/\sigma_{h_1}^2)h_1 + (\alpha_i/\sigma_{h_i}^2)h_i}{(\alpha_1/\sigma_{h_1}^2) + (\alpha_i/\sigma_{h_i}^2)}, y_i^* = g_i + \rho_i \sigma_{g_i} \frac{(\alpha_1/\sigma_{h_1}^2)}{(\alpha_1/\sigma_{h_1}^2) + (\alpha_i/\sigma_{h_i}^2)} \frac{(h_1 - h_i)}{\sigma_{h_i}}, \quad i \in \Gamma(\rho, \alpha) \quad (3)$$

$$x_1^* = h_1, x_i^* = h_i + \rho_i \sigma_{h_i} \frac{(\gamma - g_i)}{\sigma_{g_i}}, y_i^* = \gamma, \quad i \in \mathcal{S}_b(\rho) \quad (4)$$

$$x_1^* = x_i^* = \frac{\frac{\alpha_1 h_1}{\sigma_{h_1}^2} + \frac{\alpha_i}{\sigma_{h_i}^2 (1 - \rho_i^2)} \left(h_i + \rho_i \sigma_{h_i} \frac{(\gamma - g_i)}{\sigma_{g_i}} \right)}{\left(\frac{\alpha_1}{\sigma_{h_1}^2} + \frac{\alpha_i}{\sigma_{h_i}^2 (1 - \rho_i^2)} \right)}, y_i^* = \gamma, \quad i \in \mathcal{S}_w(\rho, \alpha) \quad (5)$$

where

$$\Gamma(\rho, \alpha) = \left\{ i : h_i > h_1 \text{ and } g_i \leq \gamma + \rho_i \sigma_{g_i} \frac{(\sigma_{h_i}^2/\alpha_i)}{(\sigma_{h_1}^2/\alpha_1) + (\sigma_{h_i}^2/\alpha_i)} \frac{(h_i - h_1)}{\sigma_{h_i}} \right\},$$

$$\mathcal{S}_b(\rho) = \left\{ i : h_i \leq h_1 + \rho_i \sigma_{h_i} \frac{(g_i - \gamma)}{\sigma_{g_i}} \text{ and } g_i > \gamma \right\},$$

$$\mathcal{S}_w(\rho, \alpha) = \left\{ i : h_i > h_1 + \rho_i \sigma_{h_i} \frac{(g_i - \gamma)}{\sigma_{g_i}} \text{ and } g_i > \gamma + \rho_i \sigma_{g_i} \frac{(\sigma_{h_i}^2/\alpha_i)}{(\sigma_{h_1}^2/\alpha_1) + (\sigma_{h_i}^2/\alpha_i)} \frac{(h_i - h_1)}{\sigma_{h_i}} \right\}.$$

As in Hunter and Pasupathy (2011), let us define the sets for zero correlation as feasible and worse (Γ), infeasible and better (\mathcal{S}_b), and infeasible and worse (\mathcal{S}_w). Formally,

$$\Gamma = \{i : h_i > h_1 \text{ and } g_i \leq \gamma, i \neq 1\},$$

$$\mathcal{S}_b = \{i : h_i \leq h_1 \text{ and } g_i > \gamma, i \neq 1\},$$

$$\mathcal{S}_w = \{i : h_i > h_1 \text{ and } g_i > \gamma, i \neq 1\},$$

respectively.

In the set definitions given in Lemma 1, for positive ρ_i , $\Gamma(\rho, \alpha) \supseteq \Gamma$ and $\mathcal{S}_b(\rho) \supseteq \mathcal{S}_b$, while $\mathcal{S}_w(\rho, \alpha) \subseteq \mathcal{S}_w$. This result is intuitive: since the performance measures are positively correlated, a system from \mathcal{S}_w that is “close enough” to h_1 in objective value or “close enough” to γ in constraint value and succeeds in crossing one of these barriers will likely succeed in crossing the other. Therefore some systems in \mathcal{S}_w may “pretend” to be in either Γ or \mathcal{S}_b when there is positive correlation. When there is negative correlation, the opposite occurs. Thus $\Gamma(\rho, \alpha) \subseteq \Gamma$ and $\mathcal{S}_b(\rho) \subseteq \mathcal{S}_b$, while $\mathcal{S}_w(\rho, \alpha) \supseteq \mathcal{S}_w$. Now systems in Γ and \mathcal{S}_b “pretend” to be in \mathcal{S}_w . For example, due to negative correlation, a system in Γ that is “close enough” to the constraint might be declared infeasible when it “beats” system 1 in objective value. Of the two types of correlation, intuitively it seems that negative correlation is “better” since negative correlation may result in an increased rate of decay of the probability of false selection. However it is worth noting that this result depends on the formulation of the problem in Section 1.1.

For each set $\Gamma(\rho, \alpha)$, $\mathcal{S}_b(\rho)$, and $\mathcal{S}_w(\rho, \alpha)$, given the result from Lemma 1, we can solve for the rate functions. These rate functions are given in the following proposition.

Proposition 2 Under Assumption 1, the relevant rate functions are

$$\begin{aligned}
 K_{i \in \Gamma(\rho, \alpha)}(\alpha_1, \alpha_i) &= \frac{(h_1 - h_i)^2}{2(\sigma_{h_1}^2/\alpha_1 + \sigma_{h_i}^2/\alpha_i)}, \\
 K_{i \in \mathcal{S}_b(\rho)}(\alpha_1, \alpha_i) &= \frac{\alpha_i(\gamma - g_i)^2}{2\sigma_{g_i}^2}, \\
 K_{i \in \mathcal{S}_w(\rho, \alpha)}(\alpha_1, \alpha_i) &= \frac{(h_1 - h_i)^2 - 2\rho_i\sigma_{h_i}(h_1 - h_i)\frac{(\gamma - g_i)}{\sigma_{g_i}} + \frac{\alpha_i(\gamma - g_i)^2}{\sigma_{g_i}^2}(\sigma_{h_1}^2/\alpha_1 + \sigma_{h_i}^2/\alpha_i)}{2(\sigma_{h_1}^2/\alpha_1 + (\sigma_{h_i}^2/\alpha_i)(1 - \rho_i^2))}.
 \end{aligned}$$

The rate functions for sets $\Gamma(\rho, \alpha)$ and $\mathcal{S}_b(\rho)$ are identical to the rate functions of Γ and \mathcal{S}_b in the independent case. That is, while the sets depend on ρ , rate functions for systems in these set do not depend on ρ . In the independent case, the rate function for systems in \mathcal{S}_w comprises two added terms: one for “optimality” that is identical to the rate function for systems in Γ , and one for “feasibility” that is identical to the rate function for systems in \mathcal{S}_b . In the correlated case, the rate function for systems in $\mathcal{S}_w(\rho, \alpha)$ comprises three added terms: one for “optimality” and one for “feasibility” that are similar to the rate functions for systems in $\Gamma(\rho, \alpha)$ and $\mathcal{S}_b(\rho)$, plus an interaction term.

The following theorem states the probability of false selection for the bivariate normal case as a function of the correlation ρ .

Theorem 1 Under Assumption 1, the rate function of the probability of false selection is

$$\begin{aligned}
 -\lim_{n \rightarrow \infty} \frac{1}{n} \log P\{FS\} &= \min \left(\frac{\alpha_1(\gamma - g_1)^2}{2\sigma_{g_1}^2}, \min_{i \in \Gamma(\rho, \alpha)} \frac{(h_1 - h_i)^2}{2(\sigma_{h_1}^2/\alpha_1 + \sigma_{h_i}^2/\alpha_i)}, \min_{i \in \mathcal{S}_b(\rho)} \frac{\alpha_i(\gamma - g_i)^2}{2\sigma_{g_i}^2}, \right. \\
 &\quad \left. \min_{i \in \mathcal{S}_w(\rho, \alpha)} \frac{(h_1 - h_i)^2 - 2\rho_i\sigma_{h_i}(h_1 - h_i)\frac{(\gamma - g_i)}{\sigma_{g_i}} + \frac{\alpha_i(\gamma - g_i)^2}{\sigma_{g_i}^2}(\sigma_{h_1}^2/\alpha_1 + \sigma_{h_i}^2/\alpha_i)}{2(\sigma_{h_1}^2/\alpha_1 + (\sigma_{h_i}^2/\alpha_i)(1 - \rho_i^2))} \right).
 \end{aligned}$$

As in Hunter and Pasupathy (2011), we find that the overall rate function will be determined by the most likely of four events. In the correlated case, these events are: (i) system 1 is incorrectly declared infeasible; (ii) a system from $\Gamma(\rho, \alpha)$ “pretends” to be optimal; (iii) a system from $\mathcal{S}_b(\rho)$ “pretends” to be feasible; (iv) a system from $\mathcal{S}_w(\rho, \alpha)$ “pretends” to be optimal and feasible.

4 OPTIMAL ALLOCATION STRATEGY

We wish to allocate the α_i 's to solve the following optimization problem:

$$\begin{aligned} \max \quad & \min \left(\alpha_1 J_1(\gamma), \min_{2 \leq i \leq r} K_i(\alpha_1, \alpha_i) \right) \\ \text{s.t.} \quad & \sum_{i=1}^r \alpha_i = 1, \quad \alpha_i \geq 0. \end{aligned} \tag{6}$$

This problem can be re-expressed as

$$\begin{aligned} \text{Problem } Q: \quad & \max \quad z \quad \text{s.t.} \\ & \alpha_1 J_1(\gamma) \geq z, \\ & K_i(\alpha_1, \alpha_i) \geq z, \quad 2 \leq i \leq r, \\ & \sum_{i=1}^r \alpha_i = 1, \quad \alpha_i \geq 0. \end{aligned}$$

Proposition 3 The rate function $K_i(\alpha_1, \alpha_i)$ is a concave function of (α_1, α_i) .

Since $\alpha_1 J_1(\gamma)$ is also a concave function of α , Problem Q is a concave maximization problem.

Let Problem Q_{BVN}^* be as Problem Q except with the inequality constraints on $K_i(\alpha_1, \alpha_i)$ replaced by equality constraints and forcing $\alpha_i > 0$. Following a proof similar to that given in Hunter and Pasupathy (2011), it can be shown that Problems Q and Q_{BVN}^* are equivalent, that is, they have an identical optimal solution and optimal value. Problem Q_{BVN}^* is written as

$$\begin{aligned} \text{Problem } Q_{BVN}^*: \quad & \max \quad z \quad \text{s.t.} \\ & \frac{\alpha_1(\gamma - g_1)^2}{2\sigma_{g_1}^2} \geq z, \\ & \frac{(h_1 - h_i)^2}{2(\sigma_{h_1}^2/\alpha_1 + \sigma_{h_i}^2/\alpha_i)} = z, \quad i \in \Gamma(\rho, \alpha) \\ & \frac{\alpha_i(\gamma - g_i)^2}{2\sigma_{g_i}^2} = z, \quad i \in \mathcal{S}_b(\rho) \\ & \frac{(h_1 - h_i)^2 - 2\rho_i\sigma_{h_i}(h_1 - h_i)\frac{(\gamma - g_i)}{\sigma_{g_i}} + \frac{\alpha_i(\gamma - g_i)^2}{\sigma_{g_i}^2} (\sigma_{h_1}^2/\alpha_1 + \sigma_{h_i}^2/\alpha_i)}{2(\sigma_{h_1}^2/\alpha_1 + (\sigma_{h_i}^2/\alpha_i)(1 - \rho_i^2))} = z, \quad i \in \mathcal{S}_w(\rho, \alpha) \\ & \sum_{i=1}^r \alpha_i = 1, \quad \alpha_i > 0. \end{aligned}$$

We propose that the solution to Problem Q_{BVN}^* be obtained using a solver.

5 AN APPROXIMATE CLOSED-FORM SOLUTION

This section presents a closed-form solution to Problem Q_{BVN}^* under the assumption that $\alpha_1^* \gg \alpha_i^*$. Thus the solutions in this section are only appropriate when such an assumption is reasonable. For example, such a scenario may arise when the number of systems in $\Gamma(\rho, \alpha)$ and $\mathcal{S}_w(\rho, \alpha)$ is large.

The $\alpha_1^* \gg \alpha_i^*$ assumption appears often in previous OCBA literature (see, e.g., Lee et al. (2011)). Notationally, we remove the “stars” from α in this section to emphasize that the α value presented here

is an approximation to the true optimal solution. First, let us use the $\alpha_1 \gg \alpha_i$ assumption to simplify the characterization of the sets $\Gamma(\rho, \alpha)$ and $\mathcal{S}_w(\rho, \alpha)$. We define the sets $\Gamma^{\alpha_1 \gg \alpha_i}(\rho)$ and $\mathcal{S}_w^{\alpha_1 \gg \alpha_i}(\rho)$ as

$$\Gamma^{\alpha_1 \gg \alpha_i}(\rho) = \left\{ i : h_i > h_1 \text{ and } g_i < \gamma + \rho_i \sigma_{g_i} \frac{(h_i - h_1)}{\sigma_{h_i}}, i \neq 1 \right\},$$

$$\mathcal{S}_w^{\alpha_1 \gg \alpha_i}(\rho) = \left\{ i : h_i > h_1 + \rho_i \sigma_{h_i} \frac{(g_i - \gamma)}{\sigma_{g_i}} \text{ and } g_i > \gamma + \rho_i \sigma_{g_i} \frac{(h_i - h_1)}{\sigma_{h_i}}, i \neq 1 \right\}.$$

Set $\mathcal{S}_b(\rho)$ needs no modification, as it does not depend on (α_1, α_i) . The corresponding rate functions are

$$K_{i \in \Gamma^{\alpha_1 \gg \alpha_i}(\rho)}(\alpha_1, \alpha_i) = \frac{\alpha_i (h_1 - h_i)^2}{2\sigma_{h_i}^2} = \alpha_i I_i(h_1),$$

$$K_{i \in \mathcal{S}_b(\rho)}(\alpha_1, \alpha_i) = \frac{\alpha_i (\gamma - g_i)^2}{2\sigma_{g_i}^2} = \alpha_i J_i(\gamma),$$

$$K_{i \in \mathcal{S}_w^{\alpha_1 \gg \alpha_i}(\rho)}(\alpha_1, \alpha_i) = \frac{\alpha_i}{2(1 - \rho_i^2)} \left(\frac{(h_1 - h_i)^2}{\sigma_{h_i}^2} - 2\rho_i \frac{(h_1 - h_i)}{\sigma_{h_i}} \frac{(\gamma - g_i)}{\sigma_{g_i}} + \frac{(\gamma - g_i)^2}{\sigma_{g_i}^2} \right) = \alpha_i I_i(h_1, \gamma).$$

Let $\mathbb{I}_{(\cdot)}$ denote the indicator function. Under Assumption 1 and the assumption that $\alpha_1^* \gg \alpha_i^*$, it follows from the equality constraints in Problem Q_{BVN}^* that the sample allocated to systems other than system 1 should follow the proportion

$$\frac{\alpha_i}{\alpha_k} = \frac{I_k(h_1) \mathbb{I}_{(k \in \Gamma^{\alpha_1 \gg \alpha_i}(\rho))} + J_k(\gamma) \mathbb{I}_{(k \in \mathcal{S}_b(\rho))} + I_k(h_1, \gamma) \mathbb{I}_{(k \in \mathcal{S}_w^{\alpha_1 \gg \alpha_i}(\rho))}}{I_i(h_1) \mathbb{I}_{(i \in \Gamma^{\alpha_1 \gg \alpha_i}(\rho))} + J_i(\gamma) \mathbb{I}_{(i \in \mathcal{S}_b(\rho))} + I_i(h_1, \gamma) \mathbb{I}_{(i \in \mathcal{S}_w^{\alpha_1 \gg \alpha_i}(\rho))}}.$$

6 CONCLUDING REMARKS

Simulation optimization on a finite set and subject to stochastic constraints is a topic that remains largely unexplored. For the specific but useful setting where the objective and constraint performance measures can be jointly sampled from a multivariate normal distribution, there appears to be much inherent problem structure that can be exploited toward developing asymptotically optimal and easily implementable algorithms. For instance, suppose we use a simple fully-sequential algorithm that selects that system having the smallest estimated objective function value among the estimated-feasible systems, after expending a given simulation budget. Within the setting of such algorithms, our results characterize the nature of sampling plans that maximize the decay rate of the probability that the algorithm returns a suboptimal system, as the simulation budget tends to infinity. The characterization takes the form of a concave maximization problem.

We note that the theory presented in this paper assumes that the rate functions of the systems are known. In practice, the rate functions must be estimated, and hence our results may only be implemented approximately through estimators. As in Hunter and Pasupathy (2011), our results readily inspire an implementable sequential selection algorithm using sample means and covariances to estimate the rate functions. Ongoing research provides such an implementable sequential algorithm and extends our current results to the context of multiple constraints.

REFERENCES

- Andradóttir, S., and S.-H. Kim. 2010. "Fully Sequential Procedures for Comparing Constrained Systems via Simulation". *Naval Research Logistics* 57 (5): 403–421.
- Batur, D., and S.-H. Kim. 2010. "Finding Feasible Systems in the Presence of Constraints on Multiple Performance Measures". *ACM Transactions on Modeling and Computer Simulation* 20 (3): 13:1–26.
- Chen, C.-H., J. Lin, E. Yücesan, and S. E. Chick. 2000. "Simulation Budget Allocation for Further Enhancing the Efficiency of Ordinal Optimization". *Discrete Event Dynamic Systems* 10 (3): 251–270.

- Dembo, A., and O. Zeitouni. 1998. *Large Deviations Techniques and Applications*. 2nd ed. New York: Springer.
- Glynn, P., and S. Juneja. 2004. "A large deviations perspective on ordinal optimization". In *Proceedings of the 2004 Winter Simulation Conference*, edited by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 577–585. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.
- Hunter, S. R., and R. Pasupathy. 2011. "Optimal sampling laws for stochastically constrained simulation optimization on finite sets". Under Review.
- Kim, S.-H., and B. L. Nelson. 2006. "Selecting the best system". In *Simulation*, edited by S. G. Henderson and B. L. Nelson, Handbooks in O.R. and M.S., Vol. 13, 501–534. Elsevier.
- Lee, L. H., N. A. Pujowidianto, L.-W. Li, C.-H. Chen, and C. M. Yap. 2011. "Asymptotic Simulation Budget Allocation for Selecting the Best Design in the Presence of Stochastic Constraints". Submitted to *IEEE Transactions on Automatic Control*.
- Szechtman, R., and E. Yücesan. 2008. "A New Perspective on Feasibility Determination". In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, and J. W. Fowler, 273–280. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

SUSAN R. HUNTER is a postdoctoral associate in the School of Operations Research and Information Engineering at Cornell University. Her research interests include Monte Carlo methods and simulation optimization. Her email address is hunter@cornell.edu.

NUGROHO ARTADI PUJOWIDIANTO is a Ph.D. candidate in the Department of Industrial and Systems Engineering, National University of Singapore. He received his B.Eng. (Mechanical Engineering) degree from Nanyang Technological University in 2006. His research interest is in simulation optimization and its application in health care. His email address is nugroho@nus.edu.sg.

CHUN-HUNG CHEN is a Professor in the Department of Electrical Engineering at National Taiwan University. He received his Ph.D. from Harvard University in 1994. His research interests are mainly in development of very efficient methodology for simulation and optimization and its applications. Dr. Chen has served as Co-Editor of the Proceedings of the 2002 Winter Simulation Conference and Program Co-Chair for 2007 Informs Simulation Society Workshop. He is currently an associate editor of *IEEE Transactions on Automatic Control*, area editor of *Journal of Simulation Modeling Practice and Theory*, associate editor of *International Journal of Simulation and Process Modeling*, and simulation department editor for *IIE Transactions*. His email address is cchen9@gmu.edu.

LOO HAY LEE is an Associate Professor and Deputy Head (Graduate Studies and Research) in the Department of Industrial and Systems Engineering, National University of Singapore. He received his B.S. (Electrical Engineering) degree from the National Taiwan University in 1992 and his Ph.D. degree in 1997 from Harvard University. He is currently a senior member of IEEE, a member of ORSS and INFORMS. He is the associate editor for *IIE Transactions*, *Flexible Services and Manufacturing Journal*, the *Asia Pacific Journal of Operational Research*, *Journal of Simulation* and is a member in the advisory board for *OR Spectrum*. His research interests include simulation-based optimization, and maritime Logistics. His email address is iseleelh@nus.edu.sg.

RAGHU PASUPATHY is an associate professor in the Industrial and Systems Engineering Department at Virginia Tech. His research interests lie broadly in Monte Carlo methods with a specific focus on simulation optimization and stochastic root finding. He is a member of INFORMS, IIE, and ASA, and serves as an Associate Editor for *ACM TOMACS* and *INFORMS Journal on Computing*. His email address

Hunter, Pujowidianto, Chen, Lee, Pasupathy, and Yap

is pasupath@vt.edu, and his webpage is <https://filebox.vt.edu/users/pasupath/pasupath.htm>

CHEE MENG YAP is a Senior Lecturer in the Department of Industrial and Systems Engineering and Associate Dean in the Faculty of Engineering, National University of Singapore. He received his B. Eng. (First Class Honours) in Civil Engineering from National University of Singapore in 1987 and his M.S. in Industrial Engineering and Ph.D. in Engineering Management in 1989 and 1992 from University of Pittsburgh. In addition, he holds the Chartered Financial Analyst designation. His research interests include R&D investment and firm value, and operations management in the healthcare industry. His email address is iseyapcm@nus.edu.sg.