A MODEL REFERENCE ADAPTIVE SEARCH METHOD FOR STOCHASTIC GLOBAL OPTIMIZATION*

JIAQIAO HU†, MICHAEL C. FU‡, AND STEVEN I. MARCUS§

Abstract. We propose a randomized search method called Stochastic Model Reference Adaptive Search (SMRAS) for solving stochastic optimization problems in situations where the objective functions cannot be evaluated exactly, but can be estimated with some noise (or uncertainty), e.g., via simulation. The method generalizes the recently proposed Model Reference Adaptive Search (MRAS) for deterministic optimization, which is motivated by the well-known Cross-Entropy (CE) method. We prove global convergence of SMRAS in a general stochastic setting, and carry out numerical studies to illustrate its performance. An emphasis of this paper is on the application of SMRAS for solving static stochastic optimization problems; its various applications for solving dynamic decision making problems can be found in [7].

Keywords: stochastic optimization, global optimization, combinatorial optimization.

1. Introduction. Stochastic optimization problems arise in a wide range of areas such as manufacturing, communication networks, system design, and financial engineering. These problems are typically much more difficult to solve than their deterministic counterparts, either because an explicit relation between the objective function and the underlying decision variables is unavailable or because the cost of a precise evaluation of the objective function is too prohibitive. Oftentimes, one has to use simulation or real-time observations to evaluate the objective function. In such situations, all the objective function evaluations will contain some noise, so special techniques are generally used (as opposed to the deterministic optimization methods) to filter out the noisy components. Roger Brockett has made many important contributions in stochastic systems and optimization, having introduced the two senior authors to various topics in these areas during their graduate studies decades ago. In his paper [6], he discussed new issues in the mathematics of control; some of these issues were in adaptation and learning, and in particular in using scheduled randomization in stochastic optimization. It is in this spirit that we approach stochastic optimization problems.

A well-known class of methods for solving stochastic optimization problems with continuous decision variables is stochastic approximation [20], [23]. These methods mimic the classical gradient-based search method in deterministic optimization, and

---

†Department of Applied Mathematics and Statistics, State University of New York, Stony Brook, NY 11794. E-mail: jqhu@ams.sunysb.edu

‡Robert H. Smith School of Business & Institute for Systems Research, University of Maryland, College Park, MD 20742. E-mail: mfu@rhsmith.umd.edu

§Department of Electrical and Computer Engineering & Institute for Systems Research, University of Maryland, College Park, MD 20742. E-mail:marcus@umd.edu

rely on the estimation of the gradient of the objective function. Because they are gradient-based, these methods generally find local optimal solutions. In terms of the different gradient estimation techniques employed, the stochastic approximation algorithms can be generally divided into two categories: algorithms that are based on direct gradient estimation techniques, the best-known of which are perturbation analysis [18], [13], the likelihood ratio/score function methods [14], [27], the weak derivative methods [22], and algorithms that are based on indirect gradient estimation techniques like finite difference and its variations [29]. A detailed review of various gradient estimation techniques can be found in [21] and [12]. Also of relevance to our work is the low-dispersion point sets method of [33], which uses the idea of quasirandom search for continuous global optimization and large deviation principle to choose the evaluation points within the decision domain and to (adaptively) determine the number of simulation observations to be allocated to these points.

For discrete stochastic optimization problems, one popular approach is to use random search. This has given rise to many different stochastic discrete optimization algorithms, including the stochastic ruler method and its modification [34], [2], the random search methods of [4], modified simulated annealing [3], and the nested partitions method of [28]. The main idea throughout is to show that the algorithm induces a Markov chain over the solution space, and the Markov chain eventually settles down on the set of (possibly local) optimal solutions.

From an algorithmic point of view, there is another class of randomized search techniques, which [35] have termed the model-based methods, that can also be applied to stochastic *discrete* optimization problems. Unlike the aforementioned approaches, where the search for new candidate solutions depends directly on previously generated solutions, the model-based search methods are based on sampling from an intermediate probabilistic model on the solution space, which is updated iteratively after evaluating the performance of the samples at each iteration. Most of the algorithms that fall in this category are iterative methods involving the following two steps:

1) Generate candidate solutions (e.g., random samples) according to a specified probabilistic model on the solution space.
2) Update the probabilistic model based on the candidate solutions generated in the previous step, to bias the future search toward the region containing high quality solutions.

A well-established model-based method for stochastic discrete optimization is Stochastic Ant Colony Optimization (S-ACO) [15]. S-ACO is the extension of the original Ant Colony Optimization (ACO) algorithm [10] for deterministic problems. The method uses Monte-Carlo sampling to estimate the objective and is shown (under some regularity assumptions) to converge with probability one to the global optimal solution for stochastic combinatorial problems. Another method that is closely related to the work of this paper is the Cross-Entropy (CE) method [26]. The method was originally

motivated by the problem of estimating probabilities of rare events in simulation, before it was discovered that it could be modified to solving deterministic optimization problems (cf. e.g., [24]). The key idea of CE is to use a family of parameterized distributions to successively approximate an optimal (importance sampling) distribution concentrated only on the set of (near) optimal solutions, which is carried out by iteratively estimating the optimal parameter that minimizes the Kullback-Leibler (KL) distance between the parameterized distribution and the target optimal distribution. More recently, [25] shows that the method is also capable of handling stochastic network combinatorial optimization problems, and in that particular context, establishes the probability one convergence of the algorithm. However, to the authors' best knowledge, there is no version of CE intended for general (continuous) stochastic optimization problems, and the existing convergence analysis of CE is limited to very specific settings.

In this paper, we propose a model-based method called stochastic model reference adaptive search (SMRAS) for solving both continuous and discrete stochastic optimization problems. The method is a generalization of the recently proposed MRAS method for deterministic optimization [19]. MRAS is motivated by the CE method, and retains the advantages of CE, in that it also uses a family of parameterized distributions as sampling distributions to generate candidate solutions and update the parameters by minimizing the KL distance. However, instead of targeting the optimal importance sampling distribution as in CE, the method uses a *sequence* of user-specified *reference* distributions to facilitate and guide the updating of the parameters associated with the parameterized distribution family. In SMRAS, a major modification from the original MRAS method is in the way the sequence of reference distributions is constructed. In MRAS, reference distributions are *idealized* probabilistic models constructed based on the exact performance of the candidate solutions. In the stochastic case, however, the objective function cannot be evaluated deterministically, so the sample average approximations of the (idealized) reference distributions are used in SMRAS to guide the parameter updating. A similar "iterative focusing" approach is also proposed in [32] for finding bounded rational equilibria of common-interest games, but that approach is introduced in a deterministic context, and no attempt has been made to show its theoretical convergence. We establish general global convergence properties of SMRAS for a class of parameterized distributions called the natural exponential family (NEF), where in many cases of interest (e.g., normal distributions are used as sampling distributions), our convergence results imply that the sequence of sampling distributions will converge with probability one to a degenerate distribution with all probability mass at the global optimum.

The rest of the paper is structured as follows. In Section 2, we give a detailed description of the SMRAS method. In Section 3, we establish the global convergence properties of the method. Supporting numerical studies on both continuous and

combinatorial optimization problems are given in Section 4. Finally some future research topics are outlined in Section 5.

2. The Stochastic Model Reference Adaptive Search Method. We consider the following optimization problem:

$$(1) \qquad x^* \in \arg\max_{x \in \mathbb{X}} E_\psi[\mathcal{J}(x, \psi)],$$

where the solution space $\mathbb{X}$ is a non-empty set in $\Re^n$, which can be either continuous or discrete, $\mathcal{J}(\cdot, \cdot)$ is a deterministic, real-valued function, and ψ is a random variable (possibly depending on x) representing the stochastic effects of the system. We assume that $\mathcal{J}(x, \psi)$ is measurable and integrable with respect to the distribution of ψ for all $x \in \mathbb{X}$. We let $J(x) := E_\psi[\mathcal{J}(x, \psi)]$. Note that in many cases, $J(x)$ cannot be obtained easily, but the random variable $\mathcal{J}(x, \psi)$ can be observed, e.g., via simulation or real-time observation. Each time when $\mathcal{J}(x, \psi)$ is observed, we use $\mathcal{J}_i(x)$ to denote the ith i.i.d. random observation at x. Throughout this paper, we assume that (1) has a unique global optimal solution, i.e., $\exists\, x^* \in \mathbb{X}$ such that $J(x) < J(x^*) \,\forall\, x \neq x^*,\ x \in \mathbb{X}$. We also make the following assumption about $\mathcal{J}_i(x)$.

Assumption $L1$. *For any given $\epsilon > 0$, these exists a positive number n^* such that for all $n \geq n^*$,*

$$\sup_{x \in \mathbb{X}} P\Big(\Big|\frac{1}{n}\sum_{i=1}^{n} \mathcal{J}_i(x) - J(x)\Big| \geq \epsilon\Big) \leq \phi(n, \epsilon),$$

where $\phi(\cdot, \cdot)$ is strictly decreasing in its first argument and non-increasing in its second argument. Moreover, $\phi(n, \epsilon) \to 0$ as $n \to \infty$.

Assumption $L1$ is satisfied by many random sequences, e.g., the sequence of i.i.d. random variables with (asymptotically) uniformly bounded variance, or a class of random variables (not necessarily i.i.d.) that satisfy the large deviations principle (cf. e.g., [17], [33]). We remark that Assumption $L1$ also implies that for any given $\epsilon > 0$, there exist positive numbers m^* and n^* such that for all $m \geq m^*$ and $n \geq n^*$,

$$(2)\quad \sup_{x,y \in \mathbb{X}} P\Big(\Big|\frac{1}{m}\sum_{i=1}^{m}\mathcal{J}_i(x) - \frac{1}{n}\sum_{i=1}^{n}\mathcal{J}_i(y) - J(x) + J(y)\Big| \geq \epsilon\Big) \leq 2\phi\left(\min\{m,n\}, \epsilon/2\right).$$

To see this, note that

$$\sup_{x,y \in \mathbb{X}} P\Big(\Big|\frac{1}{m}\sum_{i=1}^{m}\mathcal{J}_i(x) - \frac{1}{n}\sum_{i=1}^{n}\mathcal{J}_i(y) - J(x) + J(y)\Big| \geq \epsilon\Big)$$

$$\leq \sup_{x,y \in \mathbb{X}} P\Big(\Big|\frac{1}{m}\sum_{i=1}^{m}\mathcal{J}_i(x) - J(x)\Big| + \Big|\frac{1}{n}\sum_{i=1}^{n}\mathcal{J}_i(y) - J(y)\Big| \geq \epsilon\Big)$$

$$\leq \sup_{x,y \in \mathbb{X}} \Big[P\Big(\Big|\frac{1}{m}\sum_{i=1}^{m}\mathcal{J}_i(x) - J(x)\Big| \geq \frac{\epsilon}{2}\Big) + P\Big(\Big|\frac{1}{n}\sum_{i=1}^{n}\mathcal{J}_i(y) - J(y)\Big| \geq \frac{\epsilon}{2}\Big)\Big]$$

$$\leq \phi(m, \epsilon/2) + \phi(n, \epsilon/2) \quad \text{for all } m \text{ and } n \text{ large enough by Assumption } L1$$

$$\leq 2\phi(\min\{m, n\}, \epsilon/2),$$

where the last inequality follows because $\phi(\cdot, \cdot)$ is strictly decreasing in its first argument.

2.1. Overview of the MRAS Method. The theoretical properties and practical performance of model-based methods are primarily determined by the two key issues of how to efficiently construct or represent the probabilistic models and how to efficiently sample from them over the solution space $\mathbb{X}$. As in CE, in the MRAS method, the probabilistic models are specified in terms of a family of parameterized distributions $\{f(\cdot, \theta), \; \theta \in \Theta\}$ on $\mathbb{X}$, where Θ is the parameter space. The idea is that once the parameter θ is determined, sampling from $f(\cdot, \theta)$ can be performed relatively efficiently. An additional advantage by using parameterized family is that the task of constructing the entire probabilistic model now translates to the simpler task of updating the associated parameter θ, which is carried out in MRAS by minimizing the KL distance between the parameterized family and a sequence of intermediate distributions called *reference* distributions. The sequence of reference distributions is used to express the desired properties (e.g., convergence) of the method, and is often selected such that it can be shown to converge to a degenerate distribution concentrated only on the set of optimal solutions. Thus, the key steps of MRAS are the following:

1) Selecting a sequence of reference distributions $\{g_k(\cdot)\}$ with the desired convergence properties (e.g., the limit distribution being concentrated on the set of optimal solutions).

2) Working with a parameterized family of distributions $\{f(\cdot, \theta), \; \theta \in \Theta\}$, where Θ is the parameter space.

3) Optimizing the parameters $\{\theta_k\}$ iteratively by minimizing the KL distance between the parameterized distribution $f(\cdot, \theta)$ and the reference distributions $\{g_k(\cdot)\}$

$$\mathcal{D}(g_k, f(\cdot, \theta)) := E_{g_k}\left[\ln \frac{g_k(X)}{f(X, \theta)}\right] = \int_{\mathbb{X}} \ln \frac{g_k(x)}{f(x, \theta)} g_k(x) \nu(dx),$$

where ν is the Lebesgue/counting measure defined on $\mathbb{X}$, X is an n-dimensional random vector taking values in $\mathbb{X}$ with distribution $g_k(\cdot)$, and $E_{g_k}[\cdot]$ is the expectation taken with respect to the distribution $g_k(\cdot)$. Intuitively speaking, the sampling distributions $\{f(\cdot, \theta_k)\}$ can be viewed as compact approximations of the reference distributions (i.e., the projection of the reference distributions on the parameterized family $\{f(\cdot, \theta)\}$), and may hopefully retain some nice properties of $\{g_k(\cdot)\}$. Thus, as $\{g_k(\cdot)\}$ converges, the sequence of samples generated from their compact approximations $\{f(\cdot, \theta_k)\}$ should also converge to the optimal solution.

One primary difference between MRAS and CE has to do with the (implicit) distribution with respect to which the KL-distance is minimized. In CE, the target distribution is the optimal importance sampling distribution for estimating the

probability $P(J(X) \geq \gamma)$ for some γ close to $J(x^*)$, given by

$$g^*(x) = \frac{I\{J(x) \geq \gamma\}f(x, \theta)}{\int_{\mathbb{X}} I\{J(x) \geq \gamma\}f(dx, \theta)},$$

where $I\{\cdot\}$ is the indicator function. On the other hand, MRAS uses a *sequence* of user-specified *reference* distributions $\{g_k(\cdot)\}$. In particular, in the algorithm instantiation considered in [19], the sequence $\{g_k(x)\}$ is given by:

$$(3) \qquad g_k(x) = \frac{S(J(x))g_{k-1}(x)}{\int_{\mathbb{X}} S(J(x))g_{k-1}(dx)}, \quad \forall\, x \in \mathbb{X},$$

with $g_0(x) > 0$, $\forall\, x \in \mathbb{X}$ being an initial probability density/mass function (pdf/pmf) on $\mathbb{X}$, where $S(\cdot)$ is a non-negative increasing function to prevent negative probabilities. Thus, by assigning greater weight to solutions having larger values for J, the sequence $\{g_k(x)\}$ has the property that each iteration of (3) improves the expected performance, so that $\lim_{k \to \infty} E_{g_k}[S(J(X))] = S(J(x^*))$ regardless of the initial $g_0(x)$ used. In fact, the CE method can also be cast in the MRAS framework as a particular instantiation, in which $g_k(x)$ depends on the parameterized family $\{f(\cdot, \theta)\}$, not a natural choice a priori in MRAS, where the reference distributions would generally be chosen separately and independently from the choice of parameterized distributions. Therefore, it turns out that the theoretical convergence of CE depends heavily on the choices of $\{f(\cdot, \theta)\}$, and even within a given parameterized family, an inappropriate choice of the initial parameter can also lead to non-global convergence of the algorithm (cf. [19]).

2.2. Algorithm Description. We now focus on the sequence of reference distributions $\{g_k(\cdot)\}$ given by (3), and generalize the MRAS method to stochastic settings where the objective function $J(x)$ in (1) can only be estimated, e.g., via simulation or real-time observation. In the SMRAS method we propose below, one key modification of the original deterministic algorithm is to use approximations $\{\widetilde{g}_k(\cdot)\}$ of $\{g_k(\cdot)\}$ as the sequence of reference distributions, which is constructed based on the sample average approximation of the objective function $J(x)$.

A high-level description of the SMRAS method is presented in Figure 1. We provide a detailed discussion in the following subsections.

2.2.1. Initialization. In SMRAS, there are two allocation rules. The first one, denoted by $\{N_k, k = 0, 1 \ldots\}$, is called the sampling allocation rule, where each N_k determines the number of samples (candidate solutions) to be generated from the current sampling distribution at the kth iteration. The second is the observation allocation rule $\{M_k, k = 0, 1, \ldots\}$, which allocates M_k simulation observations to each of the candidate solutions generated at the kth iteration. We require both N_k and M_k to increase as the number of iterations grows for convergence, but other than that, there is considerable flexibility in their choices. To fix ideas, we use a parameter

Stochastic Model Reference Adaptive Search (SMRAS)

- **Initialization:** Specify an initial sample size $N_0 > 1$, $\alpha > 1$, $\rho \in (0, 1]$, $\varepsilon > 0$, a simulation allocation rule $\{M_k\}$, a strictly increasing performance function $S(\cdot) : \Re \to \Re^+$, a mixing coefficient $\lambda \in (0, 1)$, and an initial pdf/pmf $f(x, \theta_0) > 0$ $\forall x \in \mathbb{X}$. Set the iteration counter $k \leftarrow 0$.
- **Repeat until a specified stopping rule is satisfied:**
    1. **Sampling Candidate Solutions:** Generate N_k candidate solutions $X_k^1, \ldots, X_k^{N_k}$ by sampling from $\mathbf{f}(\cdot, \theta_k) := (1 - \lambda) f(\cdot, \theta_k) + \lambda f(\cdot, \theta_0)$. Let $\Lambda_k := \{X_k^1, \ldots, X_k^{N_k}\}$.
    2. **Performance Evaluation:** Take M_k observations for every $x \in \Lambda_k$, and calculate the sample average $\bar{\mathcal{J}}_k(x) = \frac{1}{M_k} \sum_{i=1}^{M_k} \mathcal{J}_i(x)$ for each $x \in \Lambda_k$, where $\mathcal{J}_i(x)$ is the ith i.i.d. observation of $\mathcal{J}(x, \psi)$.
    3. **Selecting Elite Samples:** Select the elite candidate solutions among Λ_k by calculating a threshold $\bar{\gamma}_k$; determine the sample size N_{k+1} to be used in the next iteration.
    4. **Parameter Updating:** Update the new parameter θ_{k+1} based on the set of elite solutions by solving

    $$(4) \quad \theta_{k+1} = \arg\max_{\theta \in \Theta} \frac{1}{N_k} \sum_{x \in \Lambda_k} \frac{[S(\bar{\mathcal{J}}_k(x))]^k}{\mathbf{f}(x, \theta_k)} \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) \ln f(x, \theta),$$

    where $\chi(y, \gamma) := \begin{cases} 0 & \text{if } y \leq \gamma - \varepsilon, \\ (y - \gamma + \varepsilon)/\varepsilon & \text{if } \gamma - \varepsilon < y < \gamma, \\ 1 & \text{if } y \geq \gamma. \end{cases}$
    5. **Updating Counters:** Set $k \leftarrow k + 1$.

FIG. 1. *Stochastic Model Reference Adaptive Search*

$\alpha > 1$, specified initially, to control the rate of increase in $\{N_k, k = 0, 1 \ldots\}$, and leave the sequence $\{M_k, k = 0, 1, \ldots\}$ as user-specified. We impose the following regularity condition on the observation allocation rule. The issue of how to determine the sample size N_k is discussed in Section 2.2.3.

Assumption L2. *The observation allocation rule $\{M_k, k = 0, 1, \ldots\}$ satisfies $M_k \geq M_{k-1}$ $\forall k = 1, 2, \ldots$, and $M_k \to \infty$ as $k \to \infty$. Moreover, for any $\epsilon > 0$, there exist $\delta_\epsilon \in (0, 1)$ and $\mathcal{K}_\epsilon > 0$ such that $\alpha^{2k} \phi(M_{k-1}, \epsilon) \leq (\delta_\epsilon)^k$, $\forall k \geq \mathcal{K}_\epsilon$, where $\phi(\cdot, \cdot)$ is defined as in L1.*

L2 is a mild condition and is very easy to verify. For instance, if $\phi(n, \epsilon)$ takes the form $\phi(n, \epsilon) = \frac{\mathcal{C}(\epsilon)}{n}$, where $\mathcal{C}(\epsilon)$ is a constant depending on ϵ, then the condition on M_{k-1} becomes $M_{k-1} \geq \mathcal{C}(\epsilon)(\frac{\alpha^2}{\delta_\epsilon})^k$ $\forall k \geq \mathcal{K}_\epsilon$. As another example, if $\mathcal{J}_i(x), i =$

$1, 2 \ldots$ satisfies the large deviations principle and $\phi(n, \epsilon) = e^{-n\mathcal{C}(\epsilon)}$, then the condition becomes $M_{k-1} \geq \left[\ln(\frac{\alpha^2}{\delta_\epsilon})/\mathcal{C}(\epsilon)\right]k, \ \forall k \geq \mathcal{K}_\epsilon$.

During the initialization step of SMRAS, a small positive number ε, a parameter $\rho \in (0, 1]$, and a strictly increasing function $S(\cdot) : \Re \to \Re^+$ are also specified. The parameters ε and ρ will be used to determine the set of promising solutions to be used in parameter updating (see sections 2.2.3 and 2.2.4), and as discussed earlier, the function $S(\cdot)$ is used to account for cases where $\bar{\mathcal{J}}_k(x)$ is negative for some x, and thus to prevent negative probabilities.

2.2.2. Sampling Candidate Solutions and Performance Evaluation. At each iteration k, N_k i.i.d. random samples are drawn from the density/mass function $\mathbf{f}(\cdot, \theta_k)$, which is a mixture of the initial density $f(\cdot, \theta_0)$ and the density calculated from the previous iteration $f(\cdot, \theta_k)$. See for example, [5] for a similar idea in the context of multiarmed bandit problems. The initial density $f(\cdot, \theta_0)$ can be chosen according to some prior knowledge of the problem structure; however, if nothing is known about where the good solutions are, this density should be chosen in such a way that each region in the solution space will have an (approximately) equal probability of being sampled. For example, if the solution space $\mathbb{X}$ is compact, then one possible choice of $f(\cdot, \theta_0)$ is the uniform distribution. Intuitively, mixing in the initial density enables the algorithm to explore the entire solution space and thus maintain a global perspective during the search process.

Given N_k generated solutions $\Lambda_k = \{X_k^1, \ldots, X_k^{N_k}\}$, their true performances $J(X_k^1), \ldots, J(X_k^{N_k})$ are estimated in Step 2 by allocating M_k observations to each $x \in \Lambda_k$, and then taking the sample average $\bar{\mathcal{J}}_k(x) = \frac{1}{M_k} \sum_{i=1}^{M_k} \mathcal{J}_i(x) \ \forall x \in \Lambda_k$.

2.2.3. Selecting Elite Samples. In SMRAS, as in CE and many population-based approaches such as genetic algorithms (GAs) (cf. e.g., [30]), only a portion of the samples – the set of "elite" samples – are used to update the probability model. This is achieved primarily by constructing a convergent sequence of thresholds $\{\bar{\gamma}_k\}$ based on quantile estimates of the performances of the current samples, and then taking those samples that have performances better than these thresholds. The idea is to concentrate the computational effort on the set of promising samples.

Another important issue is the choices of N_k, since both the theoretical convergence and practical performance of the algorithm will depend on the number of samples to be used at each iteration, but it is difficult to determine in advance the appropriate number of samples. A sample size that is too small could lead to nonconvergence of the algorithm and result in poor quality solutions, whereas a sample size that is too large may result in a waste of computational resources, especially in settings where the sampling and/or simulation cost is relatively expensive.

We now provide a detailed implementation of Step 3, which generates a convergent sequence of thresholds $\{\bar{\gamma}_k\}$ and adaptively determines the number of samples N_{k+1}

to be used in the next iteration.

Calculating Quantile Estimates: Order the sample performances $\bar{\mathcal{J}}_k(x)$, $\forall x \in \Lambda_k$ from the smallest to largest, $\bar{\mathcal{J}}_{k,(1)} \leq \bar{\mathcal{J}}_{k,(2)} \leq \cdots \leq \bar{\mathcal{J}}_{k,(N_k)}$, where $\bar{\mathcal{J}}_{k,(i)}$ represents the ith order statistic of the sequence $\{\bar{\mathcal{J}}_k(x), \ \forall x \in \Lambda_k\}$. Take $\kappa(\rho, N_k) := \bar{\mathcal{J}}_{k,(\lceil(1-\rho)N_k\rceil)}$ as an estimate of the true $(1-\rho)$-quantile of $J(X)$ with respect to $\mathbf{f}(\cdot, \theta_k)$, where $\lceil a \rceil$ is the smallest integer greater than a. Here we use the notation $\kappa(\rho, N_k)$ to emphasize the dependencies of the sample quantile estimates on both ρ and the sample size N_k.

Step 3a. If $k = 0$ (i.e., the first iteration) **or** $\kappa(\rho, N_k) \geq \bar{\gamma}_{k-1} + \varepsilon$, **then** use the quantile estimate as the current threshold $\bar{\gamma}_k = \kappa(\rho, N_k)$. Use an additional random variable X_k^* to record the sample that achieves the threshold $\bar{\gamma}_k$ by setting $X_k^* = \{x \in \Lambda_k : \bar{\mathcal{J}}_k(x) = \kappa(\rho, N_k)\}$. If more than one sample achieves the threshold value, ties are broken arbitrarily.

Step 3b. Else find the largest $\bar{\rho} \in (0, \rho)$ such that sample $(1-\bar{\rho})$-quantile estimate of the sequence $\{\bar{\mathcal{J}}_k(x), \ x \in \Lambda_k\}$ satisfies $\kappa(\bar{\rho}, N_k) \geq \bar{\gamma}_{k-1} + \varepsilon$. Use $\kappa(\bar{\rho}, N_k)$ as the current threshold, and set $\bar{\gamma}_k = \kappa(\bar{\rho}, N_k)$. Record the sample that achieves the current threshold $X_k^* = \{x \in \Lambda_k : \bar{\mathcal{J}}_k(x) = \kappa(\bar{\rho}, N_k)\}$. Set $\rho = \bar{\rho}$.

Step 3c. Else if no such $\bar{\rho}$ exists, **then** take M_k i.i.d. observations for X_{k-1}^* (the sample that achieves the previous threshold value), and set the current threshold $\bar{\gamma}_k = \bar{\mathcal{J}}_k(X_{k-1}^*) = \frac{1}{M_k}\sum_{i=1}^{M_k}\mathcal{J}_i(X_{k-1}^*)$. Set $X_k^* = X_{k-1}^*$, and increase the sample size by a factor α, $N_{k+1} = \lceil \alpha N_k \rceil$.

Thus, by construction, the random sequence $\{\bar{\gamma}_k\}$ is (approximately) improving, and each increment in the sequence (at steps 3a and 3b) is lower bounded by a quantity ε. Intuitively, the primary reason for using the thresholds $\{\bar{\gamma}_k\}$ in SMRAS is that such an approach to selecting the elite samples will quickly direct the search of the algorithm towards a sequence of "improving" regions, which could be more efficient than simply using the sequence of quantile estimates or even a fixed threshold to determine the elite samples (see also Remark 3). Note that in a deterministic setting, i.e., $J(x)$ can be evaluated exactly, Step 3 of SMRAS coincides with that of the MRAS algorithm introduced in [19]. The convergence of the sequence $\{\bar{\gamma}_k\}$ is a direct consequence of the following lemma.

LEMMA 2.1. *If Assumptions L1 and L2 are satisfied, then the sequence of random variables* $\{X_k^*, k = 0, 1, \ldots\}$ *generated by SMRAS converges w.p.1 as* $k \to \infty$.

Proof. Let $\mathcal{A}_k$ be the event that Step 3a or 3b is visited at the kth iteration and $\mathcal{B}_k = \{J(X_k^*) - J(X_{k-1}^*) \leq \frac{\varepsilon}{2}\}$. Since each time Step 3a or 3b is visited, we have $\bar{\mathcal{J}}_k(X_k^*) - \bar{\mathcal{J}}_{k-1}(X_{k-1}^*) \geq \varepsilon$, $X_k^* \in \Lambda_k$, and $X_{k-1}^* \in \Lambda_i$ for some $0 \leq i \leq k-1$, it follows that

$$P(\mathcal{A}_k \cap \mathcal{B}_k) \leq P\left(\left\{\bar{\mathcal{J}}_k(X_k^*) - \bar{\mathcal{J}}_{k-1}(X_{k-1}^*) \geq \varepsilon\right\} \cap \left\{J(X_k^*) - J(X_{k-1}^*) \leq \frac{\varepsilon}{2}\right\}\right)$$

$$\leq P\left(\bigcup_{x \in \Lambda_k, y \in \Lambda_i} \left\{\bar{\mathcal{J}}_k(x) - \bar{\mathcal{J}}_{k-1}(y) \geq \varepsilon\right\} \cap \left\{J(x) - J(y) \leq \frac{\varepsilon}{2}\right\}\right)$$

$$\leq \sum_{x \in \Lambda_k, y \in \Lambda_i} P\left(\left\{\bar{\mathcal{J}}_k(x) - \bar{\mathcal{J}}_{k-1}(y) \geq \varepsilon\right\} \cap \left\{J(x) - J(y) \leq \frac{\varepsilon}{2}\right\}\right)$$

$$\leq |\Lambda_k||\Lambda_i| \sup_{x,y \in \mathbb{X}} P\left(\left\{\bar{\mathcal{J}}_k(x) - \bar{\mathcal{J}}_{k-1}(y) \geq \varepsilon\right\} \cap \left\{J(x) - J(y) \leq \frac{\varepsilon}{2}\right\}\right)$$

$$\leq |\Lambda_k||\Lambda_i| \sup_{x,y \in \mathbb{X}} P\left(\bar{\mathcal{J}}_k(x) - \bar{\mathcal{J}}_{k-1}(y) - J(x) + J(y) \geq \frac{\varepsilon}{2}\right)$$

$$\leq 2|\Lambda_k||\Lambda_i|\phi\left(\min\left\{M_k, M_{k-1}\right\}, \frac{\varepsilon}{4}\right) \quad \text{by Equation (2)}$$

$$\leq 2\alpha^{2k} N_0^2 \phi\left(M_{k-1}, \frac{\varepsilon}{4}\right)$$

$$\leq 2N_0^2(\delta_{\varepsilon/4})^k, \quad \forall k \geq \mathcal{K}_{\varepsilon/4} \quad \text{by Assumption } L2.$$

Therefore,

$$\sum_{k=1}^{\infty} P\left(\mathcal{A}_k \cap \mathcal{B}_k\right) \leq \mathcal{K}_{\varepsilon/4} + 2N_0^2 \sum_{k=\mathcal{K}_{\varepsilon/4}}^{\infty} (\delta_{\varepsilon/4})^k \leq \infty.$$

By the Borel-Cantelli lemma, we have

$$P\left(\mathcal{A}_k \cap \mathcal{B}_k \ i.o.\right) = 0.$$

It follows that if $\mathcal{A}_k$ occurs infinitely often, then w.p.1, $\mathcal{B}_k^c$ will also occur infinitely often. Thus,

$$\sum_{k=1}^{\infty} \left[J(X_k^*) - J(X_{k-1}^*)\right] = \sum_{k:\ \mathcal{A}_k \text{ occurs}} \left[J(X_k^*) - J(X_{k-1}^*)\right]$$

$$+ \sum_{k:\ \mathcal{A}_k^c \text{ occurs}} \left[J(X_k^*) - J(X_{k-1}^*)\right],$$

$$= \sum_{k:\ \mathcal{A}_k \text{ occurs}} \left[J(X_k^*) - J(X_{k-1}^*)\right]$$

$$\text{since } X_k^* = X_{k-1}^* \text{ if Step 3c is visited,}$$

$$= \sum_{k:\ \mathcal{A}_k \cap \mathcal{B}_k \text{ occurs}} \left[J(X_k^*) - J(X_{k-1}^*)\right]$$

$$+ \sum_{k:\ \mathcal{A}_k \cap \mathcal{B}_k^c \text{ occurs}} \left[J(X_k^*) - J(X_{k-1}^*)\right],$$

$$= \infty \quad w.p.1 \quad \text{since } \varepsilon > 0.$$

However, this is a contradiction, since $J(x)$ is bounded from above by $J(x^*)$. Therefore, w.p.1, $\mathcal{A}_k$ can only occur a finite number of times, which implies that the sequence $\{X_k^*, k = 0, 1, \ldots\}$ converges w.p.1. $\qquad\square$

Now define $\gamma_k = J(X_k^*)$, i.e., the true performance of the random sample X_k^*. Lemma 2.1 implies that the sequence $\{\gamma_k\}$ converges. And it is easy to see that the sequence of stochastic thresholds $\{\bar{\gamma}_k\}$ is just a sample average approximation of the sequence $\{\gamma_k\}$. As we will see, by using a slightly stronger condition than $L2$, we can show that $\bar{\gamma}_k$ not only converges to γ_k, but also does so at an exponential rate.

2.2.4. Parameter Updating. Given the performance threshold $\bar{\gamma}_k$, we can now concentrate the computational effort of parameter updating (4) on the set of elite samples. This is carried out via the use of a filter-like function $\chi(\cdot, \bar{\gamma}_k)$, which eliminates from consideration those obviously inferior solutions having performance worse than $\bar{\gamma}_k - \varepsilon$. Since all performance evaluations will contain some noise, $\chi(\cdot, \bar{\gamma}_k)$ is chosen to be a continuous function instead of the straightforward indicator function. The hope is that such a continuous threshold function will provide some robustness, in the sense that those solutions with true performance better than $\bar{\gamma}_k$ but whose current estimates are slightly worse than $\bar{\gamma}_k$ (between $\bar{\gamma}_k - \varepsilon$ and $\bar{\gamma}_k$) will still be included in parameter updating. Thus, in the long run, as more precise performance estimates are obtained, $\chi(\cdot, \bar{\gamma}_k)$ ensures (with probability one) that all solutions with true performance better than $\bar{\gamma}_k$ will be used to calculate the new parameter θ_{k+1}.

It is important to note that in Step 4, the set $\{x \in \Lambda_k : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\}$ could be empty, since it could happen that all the samples generated at the current iteration are worse than those generated at the previous iteration. If this is the case, then by the definition of $\chi(\cdot, \bar{\gamma}_k)$, the right hand side of equation (4) will be equal to zero, so any $\theta \in \Theta$ is a maximizer; we define $\theta_{k+1} = \theta_k$ in this case.

We now show that there is a sequence of reference models $\{\tilde{g}_k(\cdot)\}$ implicit in SMRAS, and the parameter θ_{k+1} computed at Step 4 indeed minimizes the KL-divergence $\mathcal{D}(\tilde{g}_{k+1}(\cdot), f(\cdot, \theta))$.

LEMMA 2.2. *The parameter θ_{k+1} computed at the kth iteration of SMRAS minimizes the KL-distance $\mathcal{D}(\tilde{g}_{k+1}(\cdot), f(\cdot, \theta))$, where*

$$\tilde{g}_{k+1}(x) := \begin{cases} \dfrac{\left([S(\bar{\mathcal{J}}_k(x))]^k / \mathbf{f}(x, \theta_k)\right) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)}{\sum_{x \in \Lambda_k} \left([S(\bar{\mathcal{J}}_k(x))]^k / \mathbf{f}(x, \theta_k)\right) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)} & if \; \{x \in \Lambda_k : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\} \neq \emptyset, \\ \tilde{g}_k(x) & otherwise, \end{cases}$$

$$(5) \qquad \forall \, k = 0, 1, \ldots .$$

Proof. We only need to consider the case where $\{x \in \Lambda_k : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\} \neq \emptyset$, since if this is not the case, then we can always backtrack and find a $\tilde{g}_k(\cdot)$ with non-empty support.

For brevity, we define $S_k(\bar{\mathcal{J}}_k(x)) := \frac{[S(\bar{\mathcal{J}}_k(x))]^k}{\mathbf{f}(x, \theta_k)}$. Note that at the kth iteration,

the KL distance between $\widetilde{g}_{k+1}(\cdot)$ and $f(\cdot, \theta)$ can be written as

$$
\begin{aligned}
\mathcal{D}\left(\widetilde{g}_{k+1}(\cdot), f(\cdot, \theta)\right) &= E_{\widetilde{g}_{k+1}}\left[\ln \widetilde{g}_{k+1}(X)\right] - E_{\widetilde{g}_{k+1}}\left[\ln f(X, \theta)\right] \\
&= E_{\widetilde{g}_{k+1}}\left[\ln \widetilde{g}_{k+1}(X)\right] - \frac{\frac{1}{N_k}\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x))\chi\left(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k\right)\ln f(x, \theta)}{\frac{1}{N_k}\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x))\chi\left(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k\right)},
\end{aligned}
$$

where X is a random variable with distribution $\widetilde{g}_{k+1}(\cdot)$, and $E_{\widetilde{g}_{k+1}}[\cdot]$ is the expectation taken with respect to $\widetilde{g}_{k+1}(\cdot)$. Thus the proof is completed by observing that minimizing $\mathcal{D}\left(\widetilde{g}_{k+1}(\cdot), f(\cdot, \theta)\right)$ is equivalent to maximizing the quantity $\frac{1}{N_k}\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x))$ $\chi\left(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k\right)\ln f(x, \theta)$. □

The distribution $\widetilde{g}_{k+1}(\cdot)$ is constructed based on the empirical performance of the sampled solutions in Λ_k. Intuitively, it can be view as a "model" of our best guess of the promising region $\{x \in \Lambda_k : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\}$. However, such a model is expensive to build and impractical to implement, because sampling from which would always result in solutions from the same set of points Λ_k. In contrast, by minimizing $\mathcal{D}(\widetilde{g}_{k+1}(\cdot), f(\cdot, \theta))$, we approximate $\widetilde{g}_{k+1}(\cdot)$ with a smooth distribution function $f(\cdot, \theta_{k+1})$, which can be efficiently sampled from and represented relatively compactly by its parameter. Moreover, there is no need to build $\widetilde{g}_{k+1}(\cdot)$ explicitly.

Remark 1: When the solution space is finite, it is often helpful to make efficient use of the past sampling information. This can be achieved by maintaining a list of all sampled candidate solutions (along with the number of observations made at each of these solutions), and then check if a newly generated solution is in that list. If a new solution at iteration k has already been sampled and, say M_l, observations have been made, then we only need to take $M_k - M_l$ additional observations from that point. This procedure is often effective when the solution space is relatively small. However, when the solution space is large, the storage and checking cost could be quite expensive. In SMRAS, we propose an alternative approach: at each iteration k of the method, instead of remembering all past samples, we only keep track of those samples that fall in the region $\{x : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\}$. As we will see, the sampling process will become more and more concentrated on these regions; thus the probability of getting repeated samples typically increases.

Remark 2: We have not provided a stopping rule for SMRAS; the discussion of this issue is deferred to the end of the next section.

3. Convergence Analysis. Global convergence and computational efficiency of SMRAS clearly depend on the choice of the parameterized family of distributions. Throughout this paper, we restrict our discussion to the natural exponential family (NEF), which works well in practice, and for which convergence properties can be established.

DEFINITION 1. *A parameterized family of pdfs $\{f(\cdot, \theta),\ \theta \in \Theta \subseteq \Re^m\}$ on $\mathbb{X}$ is said to belong to the natural exponential family (NEF) if there exist functions $\ell(\cdot) : \Re^n \to \Re$, $\Gamma(\cdot) : \Re^n \to \Re^m$, and $K(\cdot) : \Re^m \to \Re$ such that*

(6)
$$f(x, \theta) = \exp\left\{\theta^T \Gamma(x) - K(\theta)\right\} \ell(x), \quad \forall\, \theta \in \Theta,$$

where $K(\theta) = \ln \int_{x \in \mathbb{X}} \exp\left\{\theta^T \Gamma(x)\right\} \ell(x) \nu(dx)$, and "T" denotes vector transposition. For the case where $f(\cdot, \theta)$ is a pdf, we assume that $\Gamma(\cdot)$ is a continuous mapping.

Many pdfs/pmfs can be put into the form of NEFs; some typical examples are Gaussian, Poisson, binomial, geometric, and certain multivariate forms of them.

To establish the global convergence of SMRAS, we make the following additional assumptions.

Assumptions:

A1. *There exists a compact set Π such that for the sequence of random variables $\{X_k^*, k = 0, 1, \ldots\}$ generated by SMRAS, $\exists \mathcal{N} < \infty$ w.p.1 such that $\{x : J(x) \geq J(X_k^*) - \varepsilon\} \cap \mathbb{X} \subseteq \Pi\ \forall\, k \geq \mathcal{N}$.*

A2. *For any constant $\xi < J(x^*)$, the set $\{x : J(x) \geq \xi\} \cap \mathbb{X}$ has a strictly positive Lebesgue or discrete measure.*

A3. *For any given constant $\delta > 0$, $\sup_{x \in A_\delta} J(x) < J(x^*)$, where $A_\delta := \{x : \|x - x^*\| > \delta\} \cap \mathbb{X}$, and we define the supremum over the empty set to be $-\infty$.*

A4. *For each point $z \leq J(x^*)$, there exist $\Delta_k > 0$ and $L_k > 0$, such that $\frac{|(S(z))^k - (S(\bar{z}))^k|}{|(S(z))^k|} \leq L_k |z - \bar{z}|$ for all $\bar{z} \in (z - \Delta_k, z + \Delta_k)$.*

A5. *The maximizer of equation (4) is an interior point of Θ for all k.*

A6. *$\sup_{\theta \in \Theta} \| \exp\left\{\theta^T \Gamma(x)\right\} \Gamma(x) \ell(x) \|$ is integrable/summable with respect to x, where θ, $\Gamma(\cdot)$, and $\ell(\cdot)$ are defined in Definition 1.*

A7. *$f(x, \theta_0) > 0\ \forall\, x \in \mathbb{X}$ and $f_* := \inf_{x \in \Pi} f(x, \theta_0) > 0$, where Π is defined in A1.*

Since the sequence $\{X_k^*\}$ generated by SMRAS converges (see Lemma 2.1), *A*1 requires that the search of SMRAS will eventually end up in a compact set. The assumption is trivially satisfied if the solution space $\mathbb{X}$ is compact. Assumption *A*2 ensures that the neighborhood of the optimal solution x^* will be sampled with a strictly positive probability. Since x^* is the unique global optimizer of $J(\cdot)$, *A*3 is satisfied by many functions encountered in practice. *A*4 can be understood as a locally Lipschitz condition on $[S(\cdot)]^k$; its suitability will be discussed later. In actual implementation of the algorithm, (4) is often posed as an unconstrained optimization problem, i.e., $\Theta = \Re^m$, in which case *A*5 is automatically satisfied. It is also easy to verify that *A*6 and *A*7 are satisfied by most NEFs.

The following lemma shows the connection between the sequence of empirical reference models $\{\widetilde{g}_k(\cdot)\}$ (see equation (5)) and their compact approximations $\{f(\cdot, \theta_k)\}$.

LEMMA 3.1. *If assumptions A5 and A6 hold, then we have*

$$E_{\theta_{k+1}}\left[\Gamma(X)\right] = E_{\widetilde{g}_{k+1}}\left[\Gamma(X)\right], \quad \forall\, k = 0, 1, \ldots,$$

where $E_{\theta_{k+1}}(\cdot)$ and $E_{\widetilde{g}_{k+1}}(\cdot)$ are the expectations taken with respect to the pdf/pmf $f(\cdot, \theta_{k+1})$ and $\widetilde{g}_{k+1}(\cdot)$, respectively.

Proof. For the same reason as discussed in the proof of Lemma 2.2, we only need to consider the case where $\{x \in \Lambda_k : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\} \neq \emptyset$. Define

$$J_k(\theta) = \frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) \ln f(x, \theta), \text{ where } S_k(\bar{\mathcal{J}}_k(x)) := \frac{[S(\bar{\mathcal{J}}_k(x))]^k}{\mathbf{f}(x, \theta_k)}.$$

Since $f(\cdot, \theta)$ belongs to the NEF, we can write

$$\begin{aligned}
J_k(\theta) = & \frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) \ln \ell(x) \\
& + \frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) \theta^T \Gamma(x) \\
& - \frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) \ln \int_{y \in \mathbb{X}} e^{\theta^T \Gamma(y)} \ell(y) \nu(dy).
\end{aligned}$$

Thus the gradient of $J_k(\theta)$ with respect to θ can be expressed as

$$\begin{aligned}
\nabla_\theta J_k(\theta) = & \frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) \Gamma(x) \\
& - \frac{\int e^{\theta^T \Gamma(y)} \Gamma(y) \ell(y) \nu(dy)}{\int e^{\theta^T \Gamma(y)} \ell(y) \nu(dy)} \frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k),
\end{aligned}$$

where the validity of the interchange of derivative and integral above is guaranteed by Assumption A6 and the dominated convergence theorem. By setting $\nabla_\theta J_k(\theta) = 0$, it follows that

$$\frac{\frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) \Gamma(x)}{\frac{1}{N_k} \sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)} = \frac{\int e^{\theta^T \Gamma(y)} \Gamma(y) \ell(y) \nu(dy)}{\int e^{\theta^T \Gamma(y)} \ell(y) \nu(dy)},$$

which implies that $E_{\widetilde{g}_{k+1}}[\Gamma(X)] = E_\theta[\Gamma(X)]$ by the definitions of $g_k(\cdot)$ (cf. (5)) and $f(\cdot, \theta)$.

Since θ_{k+1} is the optimal solution of the problem

$$\arg\max_{\theta \in \Theta} J_k(\theta),$$

we conclude that $E_{\widetilde{g}_{k+1}}[\Gamma(X)] = E_{\theta_{k+1}}[\Gamma(X)]$, $\forall k = 0, 1, \ldots$, by A5. $\square$

Remark 3: Roughly speaking, the sequence of regions $\{x : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\}$, $k = 0, 1, 2 \ldots$ tends to get smaller and smaller during the search process of SMRAS (since $\{\bar{\gamma}_k\}$ is approximately increasing). Lemma 3.1 shows that the sequence of sampling distributions $f(\cdot, \theta_{k+1})$ is adapted to this sequence of shrinking regions. For example,

consider the special case where $\{x : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\}$ is convex and $\Gamma(x) = x$. Since $E_{\widetilde{g}_{k+1}}[X]$ is a convex combination of $X_k^1, \ldots, X_k^{N_k}$, the lemma implies that $E_{\theta_{k+1}}[X] \in \{x : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\}$. Thus, it is natural to expect that the random samples generated at the next iteration will fall in the region $\{x : \bar{\mathcal{J}}_k(x) > \bar{\gamma}_k - \varepsilon\}$ with large probabilities (e.g., consider the normal distribution where its mean $\mu_{k+1} = E_{\theta_{k+1}}[X]$ is equal to its mode value). In contrast, if we use a fixed sampling distribution for all iterations, then sampling from this sequence of shrinking regions could be a substantially difficult problem in practice.

We now define a sequence of (idealized) pdfs/pmfs $\{g_k(\cdot)\}$ as

$$(7) \qquad g_{k+1}(x) = \frac{[S(J(x))]^k \chi(J(x), \gamma_{k-1})}{\int_{x \in \mathbb{X}} [S(J(x))]^k \chi(J(x), \gamma_{k-1}) \nu(dx)} \quad \forall k = 1, 2, \ldots,$$

where recall that $\gamma_{k-1} = J(X_{k-1}^*)$. Note that since X_{k-1}^* is a random variable, $g_{k+1}(x)$ is also random.

The outline of the convergence proof is as follows: first we establish the convergence of the sequence of idealized distributions $\{g_k(\cdot)\}$, then we claim that the reference models $\{\widetilde{g}_k(\cdot)\}$ are in fact the sample average approximations of the sequence $\{g_k(\cdot)\}$ by showing that $E_{\widetilde{g}_k}[\Gamma(X)] \to E_{g_k}[\Gamma(X)]$ w.p.1 as $k \to \infty$. Thus, the convergence of the sequence $\{f(\cdot, \theta_k)\}$ follows immediately from Lemma 3.1.

The convergence of the sequence $\{g_k(\cdot)\}$ is formalized in the following lemma.

LEMMA 3.2. *If Assumptions* $L1-L2$, $A1-A3$ *are satisfied, then*

$$\lim_{k \to \infty} E_{g_k}[\Gamma(X)] = \Gamma(x^*) \quad w.p.1.$$

Proof. Our proof is an extension of the proof of Theorem 1 in [19]. Let Ω_1 be the set of all sample paths such that Step 3a or 3b of SMRAS is visited finitely often, and let Ω_2 be the set of sample paths such that $\lim_{k \to \infty} \{x : J(x) \geq \gamma_k - \varepsilon\} \cap \mathbb{X} \subseteq \Pi$. By Lemma 2.1, we have $P(\Omega_1) = 1$, and for each $\omega \in \Omega_1$, there exists a finite $\mathcal{N}(\omega) > 0$ such that

$$X_k^*(\omega) = X_{k-1}^*(\omega) \quad \forall k \geq \mathcal{N}(\omega),$$

which implies that $\gamma_k(\omega) = \gamma_{k-1}(\omega) \; \forall k \geq \mathcal{N}(\omega)$. Furthermore, by $A1$, we have $P(\Omega_2) = 1$ and $\{x : J(x) \geq \gamma_{k-1}(\omega) - \varepsilon\} \cap \mathbb{X} \subseteq \Pi$, $\forall k \geq \mathcal{N}(\omega) \; \forall \omega \in \Omega_1 \cap \Omega_2$.

Thus, for each $\omega \in \Omega_1 \cap \Omega_2$, it is not difficult to see from equation (7) that $g_{k+1}(\cdot)$ can be expressed recursively as

$$g_{k+1}(x) = \frac{S(J(x)) g_k(x)}{E_{g_k}[S(J(X))]}, \quad \forall k > \mathcal{N}(\omega),$$

where we have used $g_k(\cdot)$ instead of $g_k(\omega)(\cdot)$ to simplify the notation. It follows that

$$(8) \qquad E_{g_{k+1}}[S(J(X))] = \frac{E_{g_k}[S^2(J(X))]}{E_{g_k}[S(J(X))]} \geq E_{g_k}[S(J(X))], \quad \forall k > \mathcal{N}(\omega),$$

which implies that the sequence $\{E_{g_k}[J(X)], k = 2, 3, \ldots\}$ converges (note that $E_{g_k}[J(X)]$ is bounded from above by $J(x^*)$).

Now we show that the limit of the above sequence is $S(J(x^*))$. To show this, we proceed by contradiction and assume that

$$\lim_{k \to \infty} E_{g_k}[S(J(X))] = S_* < S^* := S(J(x^*)).$$

Define the set $\mathcal{C} := \{x : J(x) \geq \gamma_{\mathcal{N}(\omega)} - \varepsilon\} \cap \{x : S(J(x)) \geq \frac{S^* + S_*}{2}\} \cap \mathbb{X}$. Since $S(\cdot)$ is strictly increasing, its inverse $S^{-1}(\cdot)$ exists, thus $\mathcal{C}$ can be formulated as $\mathcal{C} = \{x : J(x) \geq \max\{\gamma_{\mathcal{N}(\omega)} - \varepsilon, S^{-1}(\frac{S^* + S_*}{2})\}\} \cap \mathbb{X}$. By $A2$, $\mathcal{C}$ has a strictly positive Lebesgue/discrete measure.

Note that $g_{k+1}(\cdot)$ can be written as

$$g_{k+1}(x) = \left[\prod_{i=\mathcal{N}(\omega)+1}^{k} \frac{S(J(x))}{E_{g_i}[S(J(X))]} \right] \cdot g_{\mathcal{N}(\omega)+1}(x), \quad \forall k > \mathcal{N}(\omega).$$

Since $\lim_{k \to \infty} \frac{S(J(x))}{E_{g_k}[S(J(X))]} = \frac{S(J(x))}{S_*} > 1, \ \forall x \in \mathcal{C}$, we conclude that

$$\liminf_{k \to \infty} g_k(x) = \infty, \quad \forall x \in \mathcal{C}.$$

We have, by Fatou's lemma,

$$1 = \liminf_{k \to \infty} \int_{\mathbb{X}} g_{k+1}(x) \nu(dx) \geq \liminf_{k \to \infty} \int_{\mathcal{C}} g_{k+1}(x) \nu(dx) \geq \int_{\mathcal{C}} \liminf_{k \to \infty} g_{k+1}(x) \nu(dx) = \infty,$$

which is a contradiction. Hence, it follows that

$$(9) \qquad \lim_{k \to \infty} E_{g_k}[S(J(X))] = S^*, \ \forall \omega \in \Omega_1 \cap \Omega_2.$$

We now bound the difference between $E_{g_{k+1}}[\Gamma(X)]$ and $\Gamma(x^*)$. We have

$$\|E_{g_{k+1}}[\Gamma(X)] - \Gamma(x^*)\| \leq \int_{x \in \mathbb{X}} \|\Gamma(x) - \Gamma(x^*)\| g_{k+1}(x) \nu(dx)$$

$$(10) \qquad = \int_{\mathcal{G}} \|\Gamma(x) - \Gamma(x^*)\| g_{k+1}(x) \nu(dx),$$

where $\mathcal{G} := \{x : J(x) \geq \gamma_{\mathcal{N}(\omega)} - \varepsilon\} \cap \mathbb{X}$ is the support of $g_{k+1}(\cdot), \ \forall k > \mathcal{N}(\omega)$.

By the assumption on $\Gamma(\cdot)$ in Definition 1, for any given $\zeta > 0$, there exists a $\delta > 0$ such that $\|x - x^*\| \leq \delta$ implies $\|\Gamma(x) - \Gamma(x^*)\| \leq \zeta$. Let A_δ be defined as in $A3$; then we have from (10)

$$\|E_{g_{k+1}}[\Gamma(X)] - \Gamma(x^*)\| \leq \int_{A_\delta^c \cap \mathcal{G}} \|\Gamma(x) - \Gamma(x^*)\| g_{k+1}(x) \nu(dx)$$

$$+ \int_{A_\delta \cap \mathcal{G}} \|\Gamma(x) - \Gamma(x^*)\| g_{k+1}(x) \nu(dx)$$

$$(11) \qquad \leq \zeta + \int_{A_\delta \cap \mathcal{G}} \|\Gamma(x) - \Gamma(x^*)\| g_{k+1}(x) \nu(dx), \quad \forall k > \mathcal{N}(\omega).$$

The rest of the proof amounts to showing that the second term in (11) is also bounded. Clearly by $A1$, the term $\|\Gamma(x) - \Gamma(x^*)\|$ is bounded on the set $A_\delta \cap \mathcal{G}$. We only need to find a bound for $g_{k+1}(x)$.

By $A3$, we have

$$\sup_{x \in A_\delta \cap \mathcal{G}} J(x) \leq \sup_{x \in A_\delta} J(x) < J(x^*).$$

Define $S_\delta := S^* - S(\sup_{x \in A_\delta} J(x))$. And by the monotonicity of $S(\cdot)$, we have $S_\delta > 0$. It is easy to see that

$$(12) \qquad S(J(x)) \leq S^* - S_\delta, \quad \forall\, x \in A_\delta \cap \mathcal{G}.$$

From (8) and (9), there exists $\bar{\mathcal{N}}(\omega) \geq \mathcal{N}(\omega)$ such that for all $k \geq \bar{\mathcal{N}}(\omega)$

$$(13) \qquad E_{g_{k+1}}[S(J(X))] \geq S^* - \frac{1}{2}S_\delta.$$

Observe that $g_{k+1}(x)$ can be rewritten as

$$g_{k+1}(x) = \left[\prod_{i=\bar{\mathcal{N}}}^{k} \frac{S(J(x))}{E_{g_i}[S(J(X))]} \right] \cdot g_{\bar{\mathcal{N}}}(x), \quad \forall\, k \geq \bar{\mathcal{N}}(\omega).$$

Thus, it follows from (12) and (13) that

$$g_{k+1}(x) \leq \left(\frac{S^* - S_\delta}{S^* - \frac{1}{2}S_\delta} \right)^{k - \bar{\mathcal{N}} + 1} \cdot g_{\bar{\mathcal{N}}}(x), \quad \forall\, x \in A_\delta \cap \mathcal{G}, \quad \forall\, k \geq \bar{\mathcal{N}}(\omega).$$

Therefore,

$$\|E_{g_{k+1}}[\Gamma(X)] - \Gamma(x^*)\| \leq \zeta + \sup_{x \in A_\delta \cap \mathcal{G}} \|\Gamma(x) - \Gamma(x^*)\| \int_{A_\delta \cap \mathcal{G}} g_{k+1}(x)\nu(dx)$$

$$\leq \zeta + \sup_{x \in A_\delta \cap \mathcal{G}} \|\Gamma(x) - \Gamma(x^*)\| \left(\frac{S^* - S_\delta}{S^* - \frac{1}{2}S_\delta} \right)^{k - \bar{\mathcal{N}} + 1}, \quad \forall\, k \geq \bar{\mathcal{N}}(\omega)$$

$$\leq \left(1 + \sup_{x \in A_\delta \cap \mathcal{G}} \|\Gamma(x) - \Gamma(x^*)\| \right)\zeta, \quad \forall\, k \geq \widehat{\mathcal{N}}(\omega),$$

where $\widehat{\mathcal{N}}(\omega)$ is given by $\widehat{\mathcal{N}}(\omega) := \max\left\{ \bar{\mathcal{N}}(\omega),\ \lceil \bar{\mathcal{N}}(\omega) - 1 + \ln \zeta / \ln\left(\frac{S^* - S_\delta}{S^* - \frac{1}{2}S_\delta} \right) \rceil \right\}$.

Since ζ is arbitrary, we have

$$\lim_{k \to \infty} E_{g_k}[\Gamma(X)] = \Gamma(x^*), \quad \forall\, \omega \in \Omega_1 \cap \Omega_2.$$

And since $P(\Omega_1 \cap \Omega_2) = 1$, the proof is thus completed.

$$\square$$

The rest of the convergence proof now amounts to showing that $E_{\widetilde{g}_k}[\Gamma(X)] \to E_{g_k}[\Gamma(X)]$ w.p.1 as $k \to \infty$. However, there is one more complication: Since $S(\cdot)$ is an increasing function and is raised to the kth power in both $\widetilde{g}_{k+1}(\cdot)$ and $g_{k+1}(\cdot)$

(see equations (5), (7)), the associated estimation error between $\bar{\mathcal{J}}_k(x)$ and $J(x)$ is exaggerated. Thus, even though we have $\lim_{k\to\infty}\bar{\mathcal{J}}_k(x) = J(x)$ w.p.1, the quantities $[S(\bar{\mathcal{J}}_k(x))]^k$ and $[S(J(x))]^k$ may still differ considerably as k gets large. Therefore, the sequence $\{\bar{\mathcal{J}}_k(x)\}$ not only has to converge to $J(x)$, but it should also do so at a fast enough rate in order to keep the resultant approximation error between $S^k(\bar{\mathcal{J}}_k(x))$ and $S^k(J(x))$ at a manageable level. This requirement is summarized in the following assumption.

Assumption L3. *For any given $\zeta > 0$, there exist $\delta^* \in (0,1)$ and $\mathcal{K} > 0$ such that the observation allocation rule $\{M_k, k = 1, 2 \ldots\}$ satisfies*

$$\alpha^k \phi\Big(M_k, \min\big\{\Delta_k, \frac{\zeta}{\alpha^{k/2}}, \frac{\zeta}{\alpha^{k/2}L_k}\big\}\Big) \leq (\delta^*)^k \ \forall k \geq \mathcal{K},$$

where $\phi(\cdot,\cdot)$ is defined as in L1, Δ_k and L_k are defined as in A4.

Let $S(z) = e^{\tau z}$, for some positive constant τ. We have $S^k(z) = e^{\tau k z}$ and $[S^k(z)]' = k\tau e^{\tau k z}$. It is easy to verify that $\frac{|S^k(z)-S^k(\bar{z})|}{S^k(z)} \leq k\tau e^{\tau k \Delta_k}|z - \bar{z}| \ \forall \bar{z} \in (z - \Delta_k, z + \Delta_k)$, and A4 is satisfied for $\Delta_k = 1/k$ and $L_k = \tau e^\tau k$. Thus, the condition in L3 becomes $\alpha^k \phi(M_k, \bar{\zeta}/\alpha^{k/2}k) \leq (\delta^*)^k \ \forall k \geq \mathcal{K}$, where $\bar{\zeta} = \zeta/\tau e^\tau$. We consider the following two special cases of L3. Let $\mathcal{J}_i(x)$ be i.i.d. with $E(\mathcal{J}_i(x)) = J(x)$ and uniformly bounded variance $\sup_{x\in\mathbb{X}} \sigma^2(x) \leq \sigma^2$. By Chebyshev's inequality

$$P\Big(\big|\bar{\mathcal{J}}_k(x) - J(x)\big| \geq \frac{\bar{\zeta}}{\alpha^{k/2}k}\Big) \leq \frac{\sigma^2 \alpha^k k^2}{M_k \bar{\zeta}^2}.$$

Thus, it is easy to check that L3 is satisfied by $M_k = (\eta\alpha^2)^k$ for any constant $\eta > 1$.

As a second example, consider the case where $\mathcal{J}_1(x), \ldots, \mathcal{J}_{N_k}(x)$ are i.i.d. with $E(\mathcal{J}_i(x)) = J(x)$ and bounded support $[a, b]$. By the Hoeffding inequality [16]

$$P\Big(\big|\bar{\mathcal{J}}_k(x) - J(x)\big| \geq \frac{\bar{\zeta}}{\alpha^{k/2}k}\Big) \leq 2\exp\Big(\frac{-2M_k\bar{\zeta}^2}{(b-a)^2\alpha^k k^2}\Big).$$

In this case, L3 is satisfied by $M_k = (\eta\alpha)^k$ for any constant $\eta > 1$.

We note that Assumption L3 can be replaced by the weaker condition

$$\sum_{k=1}^{\infty} \phi\Big(M_k, \min\big\{\Delta_k, \frac{\zeta}{\alpha^{k/2}}, \frac{\zeta}{\alpha^{k/2}L_k}\big\}\Big) < \infty$$

when the solution space $\mathbb{X}$ is finite.

The following result shows that under Assumption L3, the stochastic threshold $\bar{\gamma}_k$ converges to γ_k exponentially fast, whose proof can be found in [7].

PROPOSITION 3.1. *If Assumptions L1−L3 are satisfied, then*

$$\lim_{k\to\infty} \alpha^{k/2}\big|\bar{\gamma}_k - \gamma_k\big| = 0 \ \ w.p.1.$$

We are now ready to state the main theorem.

THEOREM 1. *Let φ be a positive constant satisfying the condition that the set $\{x : S(J(x)) \geq \frac{1}{\varphi}\}$ has a strictly positive Lebesgue/counting measure. If assumptions $L1-L3$, $A1-A7$ are satisfied and $\alpha > (\varphi S^*)^2$, where $S^* = S(J(x^*))$, then*

$$(14) \qquad \lim_{k \to \infty} E_{\theta_k}[\Gamma(X)] = \Gamma(x^*) \quad w.p.1,$$

where the limit above is component-wise.

Remark 4: By the monotonicity of $S(\cdot)$ and Assumption $A2$, it is easy to see that such a positive constant φ in Theorem 1 always exists. Moreover, for continuous problems, φ can be chosen such that $\varphi S^* \approx 1$; for discrete problems, if the counting measure is used, then we can choose $\varphi = 1/S^*$.

Remark 5: Note that when $\Gamma(x)$ is a one-to-one function (which is the case for many NEFs used in practice), the above result can be equivalently written as

$$\Gamma^{-1}\left(\lim_{k \to \infty} E_{\theta_k}[\Gamma(X)]\right) = x^*.$$

Also note that for some particular pdfs/pmfs, the solution vector x itself will be a component of $\Gamma(x)$ (e.g., multivariate normal pdf), in which case (14) is equivalent to $\lim_{k \to \infty} E_{\theta_k}[X] = x^*$, i.e., the mean of the sampling distribution converges to the optimal solution x^*. Another special case of particular interest is when the components of the random vector $X = (X_1, \ldots, X_n)$ are independent, and each has a univariate pdf/pmf of the form

$$f(x_i, \vartheta_i) = \exp(x_i \vartheta_i - K(\vartheta_i))\ell(x_i), \ \vartheta_i \subset \Re, \forall\, i = 1, \ldots, n.$$

In this case, since the distribution of the random vector X is simply the product of the marginal distributions, we have $\Gamma(x) = x$. Thus, (14) is again equivalent to $\lim_{k \to \infty} E_{\theta_k}[X] = x^*$, where $\theta_k := (\vartheta_1^k, \ldots, \vartheta_n^k)$, and ϑ_i^k is the value of ϑ_i at the kth iteration of the algorithm.

Proof. As mentioned earlier, the convergence proof amounts to showing that $E_{\widetilde{g}_{k+1}}[\Gamma(X)] \to E_{g_{k+1}}[\Gamma(X)]$ w.p.1 as $k \to \infty$, and then directly applying Lemma 3.1 and 3.2. The proof of $E_{\widetilde{g}_{k+1}}[\Gamma(X)] \to E_{g_{k+1}}[\Gamma(X)]$ can be found in [7]. □

We now address some of the special cases discussed in Remark 5; the proofs are straightforward and hence omitted.

COROLLARY 1. **(Multivariate Normal)** *For continuous optimization problems in $\Re^n$, if multivariate normal pdfs are used in SMRAS, i.e.,*

$$f(x, \theta_k) = \frac{1}{\sqrt{(2\pi)^n |\Sigma_k|}} \exp\left(-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k)\right),$$

where $\theta_k := (\mu_k;\ \Sigma_k)$, assumptions $L1 - L3$, $A1 - A5$ are satisfied, and $\alpha > (\varphi S^*)^2$, then

$$\lim_{k \to \infty} \mu_k = x^*, \quad and \quad \lim_{k \to \infty} \Sigma_k = 0_{n \times n} \quad w.p.1,$$

where $0_{n \times n}$ represents an n-by-n zero matrix.

COROLLARY 2. **(Independent Univariate)** *If the components of the random vector $X = (X_1, \ldots, X_n)$ are independent, each has a univariate pdf/pmf of the form*

$$f(x_i, \vartheta_i) = \exp(x_i \vartheta_i - K(\vartheta_i))\ell(x_i), \;\; \vartheta_i \subset \Re, \forall\, i = 1, \ldots, n,$$

assumptions $L1 - L3$, $A1 - A7$ are satisfied, and $\alpha > (\varphi S^*)^2$, then

$$\lim_{k \to \infty} E_{\theta_k}[X] = x^* \;\; w.p.1, \quad where \;\; \theta_k := (\vartheta_1^k, \ldots, \vartheta_n^k).$$

Remark 6 **(Stopping Rule):** We now return to the issue of designing a valid stopping rule for SMRAS. In practice, this can be achieved in many different ways. The simplest method is to stop the algorithm when the total computational budget is exhausted or when a prescribed maximum number of iterations is reached. Since Proposition 3.1 indicates that the sequence $\{\bar{\gamma}_k, \; k = 0, 1, \ldots\}$ generated by SMRAS converges, an alternative stopping criteria could be based on identifying whether the sequence has settled down to its limit value. To do so, we consider the moving average process $\{\Upsilon_k^{(l)}\}$ defined as follows

$$\Upsilon_k^{(l)} := \frac{1}{l} \sum_{i=k-l+1}^{k} \bar{\gamma}_i, \;\; \forall\, k \geq l - 1,$$

where $l \geq 1$ is a predefined constant. It is easy to see that an unbiased estimator of the sample variance of $\Upsilon_k^{(l)}$ is

$$\widetilde{\mathbf{var}}(\Upsilon_k^{(l)}) := \frac{\sum_{i=k-l+1}^{k} [\bar{\gamma}_i - \Upsilon_k^{(l)}]^2}{l(l-1)},$$

which approaches zero as the sequence $\{\bar{\gamma}_k\}$ approaches its limit. Thus, a reasonable approach in practice is to stop the algorithm when the value of $\widetilde{\mathbf{var}}(\Upsilon_k^{(l)})$ falls below some pre-specified tolerance level, i.e., $\exists\, k > 0$ such that $\widetilde{\mathbf{var}}(\Upsilon_k^{(l)}) \leq \tau$, where $\tau > 0$ is the tolerance level.

4. Numerical Examples. In this section, we test the performance of SMRAS on several continuous and combinatorial optimization problems. In the former case, we first illustrate the global convergence of SMRAS by testing the algorithm on four multi-extremal functions; then we apply the algorithm to an inventory control problem. In the latter case, we consider the problem of optimizing the buffer allocations in a tandem queue with unreliable servers, which has been previously studied in e.g., [1] and [31] .

We begin with some implementation issues of SMRAS.
  1. Since SMRAS was presented in a maximization context, the following slight modifications are required before it can be applied to minimization problems:

(i) $S(\cdot)$ needs to be initialized as a strictly decreasing function instead of strictly increasing. Throughout this section, we take $S(z) := e^{rz}$ for maximization problems and $S(z) := e^{-rz}$ for minimization problems, where $r > 0$ is some predefined constant. (ii) The sample $(1-\rho)$-quantile $\kappa(\rho, N_k)$ will now be calculated by first ordering the sample performances $\bar{\mathcal{J}}_k(X_k^i)$, $i = 1, \ldots, N_k$ from largest to smallest, and then taking the $\lceil (1 - \rho)N_k \rceil$th order statistic. ($iii$) The threshold function should now be modified as

$$
\chi(y, \gamma) = \begin{cases} 0 & \text{if } y \geq \gamma + \varepsilon, \\ (\gamma + \varepsilon - y)/\varepsilon & \text{if } \gamma < y < \gamma + \varepsilon, \\ 1 & \text{if } y \leq \gamma. \end{cases}
$$

(iv) The inequalities at steps 3a and 3b need to be replaced with $\kappa(\rho, N_k) \leq \bar{\gamma}_{k-1} - \varepsilon$ and $\kappa(\bar{\rho}, N_k) \leq \bar{\gamma}_{k-1} - \varepsilon$, respectively.

2. In actual implementation of SMRAS, a smoothed parameter updating procedure (cf. e.g. [9], [24]) is used, i.e., first a smoothed parameter vector $\widehat{\theta}_{k+1}$ is computed at each iteration k according to

$$
\widehat{\theta}_{k+1} := \upsilon\, \theta_{k+1} + (1 - \upsilon)\widehat{\theta}_k, \quad \forall k = 0, 1, \ldots, \text{ and } \widehat{\theta}_0 := \theta_0,
$$

where θ_{k+1} is the parameter vector derived at Step 4 of SMRAS, and $\upsilon \in (0, 1]$ is the smoothing parameter, then $f(x, \widehat{\theta}_{k+1})$ (instead of $f(x, \theta_{k+1})$) is used in Step 1 to generate new samples. Although this modification will not affect the theoretical convergence results, it may improve the empirical performance of the algorithm.

4.1. Continuous Optimization. For continuous problems, we use multivariate normal pdfs as the parameterized probabilistic model. Initially, a mean vector μ_0 and a covariance matrix Σ_0 are specified; then at each iteration of the algorithm, it is easy to see that the new parameters μ_{k+1} and Σ_{k+1} are updated according to the following formula:

$$
\mu_{k+1} = \frac{\frac{1}{N_k}\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x))\chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)x}{\frac{1}{N_k}\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x))\chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)},
$$

and

$$
\Sigma_{k+1} = \frac{\frac{1}{N_k}\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x))\chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)(x - \mu_{k+1})(x - \mu_{k+1})^T}{\frac{1}{N_k}\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x))\chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)}.
$$

By Corollary 1, the sequence of mean vectors $\{\mu_k\}$ will converge to the optimal solution x^* and the sequence of covariance matrices $\{\Sigma_k\}$ to the zero matrix. In subsequent numerical experiments, we will use μ_{k+1} to represent the current best solution found at iteration k.

4.1.1. Global Convergence. To demonstrate the global convergence of SM-RAS, we consider four multi-extremal test functions with additive noise, where the noise ψ is normally distributed with mean 0 and variance 100. The graphical representations of the deterministic versions of these functions in two dimensions are given in Figure 2.

(1) Goldstein-Price function with additive noise

$$\mathcal{J}_1(x, \psi) = (1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2))$$
$$(30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)) + \psi,$$

where $-3 \leq x_i \leq 3$, $i = 1, 2$. The function $J_1(x) = E_\psi[\mathcal{J}_1(x, \psi)]$ has four local minima and a global minimum at $x^* = (0, -1)^T$, $J_1(x^*) = 3$.

(2) Rosenbrock function with additive noise $(n = 5)$

$$\mathcal{J}_2(x, \psi) = \sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2 + 1 + \psi,$$

where $-10 \leq x_i \leq 10$, $i = 1, \ldots, n$. Its deterministic counterpart $J_2(x) = E_\psi[\mathcal{J}_2(x, \psi)]$ is widely used to test the performance of different global optimization algorithms. The function has a global minimum at $x^* = (1, \ldots, 1)^T$, $J_2(x^*) = 1$.

(3) Pintér's function with additive noise $(n = 5)$

$$\mathcal{J}_3(x, \psi) = \sum_{i=1}^{n} ix_i^2 + \sum_{i=1}^{n} 20i \sin^2\left(x_{i-1}\sin x_i - x_i + \sin x_{i+1}\right)$$
$$+ \sum_{i=1}^{n} i \log_{10}\left(1 + i(x_{i-1}^2 - 2x_i + 3x_{i+1} - \cos x_i + 1)^2\right) + 1 + \psi,$$

where $x_0 = x_n$, $x_{n+1} = x_1$, $-10 \leq x_i \leq 10$, $i = 1, \ldots, n$, $x^* = (0, \ldots, 0)^T$, $J_3(x^*) = 1$.

(4) Griewank function with additive noise $(n = 10)$

$$\mathcal{J}_4(x) = \frac{1}{40}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 2 + \psi,$$

where $-10 \leq x_i \leq 10$, $i = 1, \ldots, n$, $x^* = (0, \ldots, 0)^T$, $J_4(x^*) = 1$.

For all four problems, the same set of parameters are used to test SMRAS: $r = 0.01$, $\varepsilon = 0.01$, mixing coefficient $\lambda = 0.01$, initial sample size $N_0 = 500$, initial $\rho = 0.1$, $\alpha = 1.04$, and the observation allocation rule is $M_k = \lceil 1.05M_{k-1} \rceil$ with $M_0 = 10$, the smoothing parameter $\upsilon = 0.5$, the initial mean vector μ_0 is an n-by-1 vector with each component uniformly selected from the interval $[-3, 3]$ for J_1 and from $[-10, 10]$ for J_2, J_3, and J_4, and Σ_0 is initialized as a n-by-n diagonal matrix with all diagonal elements equal to 100, where n is the dimension of the problem.
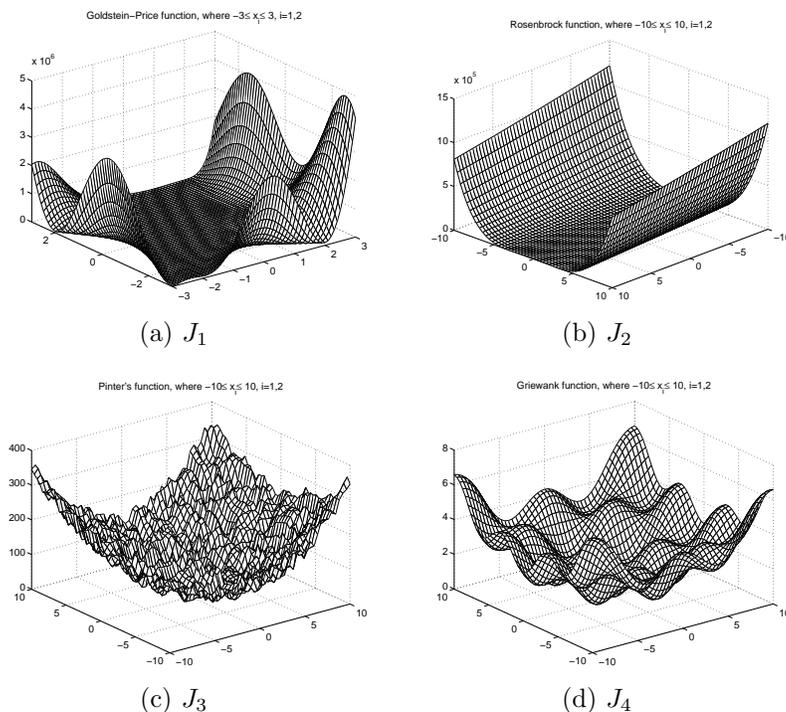
Fig. 2. *Test functions (deterministic version) in two dimensions, (a) J_1: Goldstein-Price; (b) J_2: Rosenbrock; (c) J_3: Pintér; (d) J_4: Griewank.*

Note that if the random observations obey the large deviations principle, then α and M_k satisfy the relevant conditions in Theorem 1. We refer the readers to [19] for a discussion of general guidelines for selecting r and υ.

For numerical comparison purposes, we also applied the simultaneous perturbation stochastic approximation (SPSA) algorithm [29] and the simulated annealing (SA) algorithm [8] to all test cases. In SPSA, the gradient is estimated by averaging q independent simultaneous perturbation approximations at each iteration, and whenever the update results in a solution that violates the constraint, we simply project the solution back into the feasible region. In our experiments, we choose $q = 1$, i.e., the algorithm requires only two measurements of the objective function at each iteration. We have also used a standard gain sequence $a_k = 1/k$ and a step-size sequence $c_k = 1/(c + k)^{0.25}$, where $k = 1, 2, \ldots$ is the iteration counter, $c = 5 \times 10^4$ for J_1 and J_2, and $c = 5 \times 10^2$ for J_3 and J_4. The SA algorithm we have considered is a simple stochastic version of standard SA, where each time the algorithm visits a solution, we allocate L independent simulation observations to that solution, estimate the performance of that solution by averaging over L replications, and then use standard SA to solve the underlying problem. We have used the following set of parameters: $L = 50$, initial temperature $T = 5 \times 10^4$, annealing factor $r_T = 0.85$,

and the neighborhood of a solution x to be searched at each iteration is taken to be $\mathcal{N}(x) = \{y : \max_{1 \leq i \leq n} |x - y| \leq 0.5\}$. For both SPSA and SA, the initial solution is uniformly selected from $[-3, 3]^n$ for J_1, and from $[-10, 10]^n$ for J_2, J_3, and J_4.

For each function, we performed 100 independent simulation runs of all three algorithms. The numerical results are reported in Table 1, where $\bar{J}_i$, $i = 1, 2, 3, 4$ are the averaged function values J_i evaluated at the final solutions found by the algorithms, and the corresponding optimal values J_i^* are also included for reference. Our performance comparison is based on the same amount of simulation effort, where the total number of allowed function measurements is set to 3×10^5 for J_1 and J_3, 2×10^6 for J_2, and 10^6 for J_4. In Figure 3, we also plotted the average function values of the current solutions as a function of the number of function measurements.

Numerical results indicate convergence of all three algorithms. Since SPSA uses gradient information, it may quickly locate a local optimum by using only a small number of function measurements. However, depending on the initial solutions used, the algorithm may converge to solutions that are far from optimal. Since SA permits uphill moves, it has the capability to escape from local optimal solutions. Therefore, except for the J_2 case (which is often considered as a unimodal function), SA generally yields better performance than SPSA does in the long run. SMRAS consistently outperforms SA in all test cases, and finds better solutions than SPSA does when the number of function measurements is large enough.

TABLE 1

Performance of SMRAS, SPSA, and SA on benchmark problems $J_1 - J_4$, based on 100 independent simulation runs. The standard errors are in parentheses.

Algorithm	$\bar{J}_1$(std err)	J_1^*	$\bar{J}_2$(std err)	J_2^*	$\bar{J}_3$(std err)	J_3^*	$\bar{J}_4$(std err)	J_4^*
SMRAS	3.12(0.01)	3	1.37(0.02)	1	1.60(0.03)	1	1.75(0.03)	1
SPSA $q = 1$	31.2(11.7)	3	2.02(0.20)	1	116.4(6.1)	1	6.54(0.15)	1
SA	18.8(3.2)	3	4.21(0.22)	1	14.2(1.5)	1	3.89(0.09)	1

4.1.2. An Inventory Control Example. To further illustrate the algorithm, we consider an (s, S) inventory control problem with i.i.d. exponentially distributed continuous demands, zero order lead times, full backlogging of orders, and linear ordering, holding and shortage costs. The inventory level is periodically reviewed, and an order is placed when the inventory position (on hand plus that on order) falls below the level s, and the amount of the order is the difference between S and the current inventory position. Formally, we let D_t denote the demand in period t, X_t the inventory position in period t, p the per period per unit demand shortage penalty cost, h the per period per unit inventory holding cost, c the per unit ordering cost, and K the set-up cost per order. The inventory position $\{X_t\}$ evolves according to
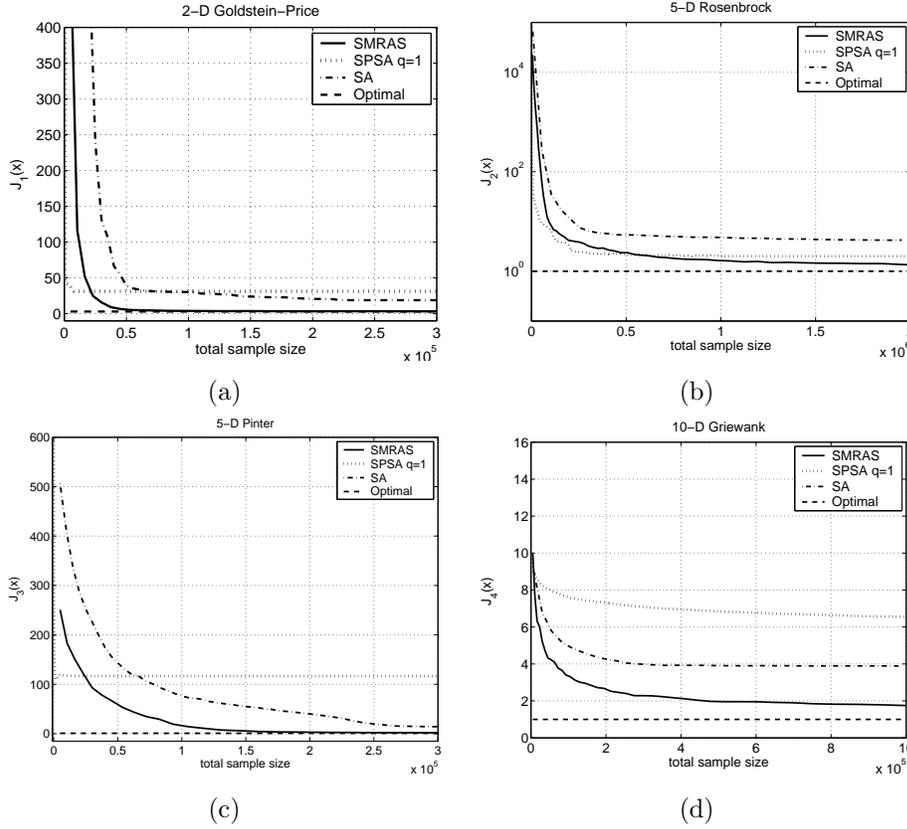
FIG. 3. *Performance of SMRAS, SPSA, and SA on (a) Goldstein-price function; (b) 5-D Rosenbrock function; (c) 5-D Pintér function; (d) 10-D Griewank function.*

the following dynamics

$$X_{t+1} = \begin{cases} S - D_{t+1} & X_t < s, \\ X_t - D_{t+1} & X_t \geq s. \end{cases}$$

The goal is to choose the thresholds s and S such that the long-run average cost per period is minimized, i.e.,

$$(s^*, S^*) = \arg\min J(s, S) := \arg\min \lim_{t \to \infty} J_t(s, S),$$

where $J_t(s, S) := \frac{1}{t} \sum_{i=1}^{t} \left[I\{X_i < s\}(K + c(S - X_i)) + hX_i^+ + pX_i^- \right]$, $x^+ = \max(0, x)$, and $x^- = \max(0, -x)$. Note that the above objective function is convex; thus a natural choice of comparison algorithm is a gradient-based algorithm like SPSA. However, we will not exploit this structure in SMRAS.

The following four test cases, taken from [11], are used to test the performance of SMRAS and SPSA. The cost coefficients and the analytical optimal solutions are given in Table 2, each with $c = h = 1$ and exponentially distributed demands with mean $E[D]$.

TABLE 2

The four test cases.

Case	$E[D]$	p	K	J^*	s^*	S^*
1	200	10	100	740.9	341	541
2	200	10	10000	2200.0	0	2000
3	200	100	100	1184.4	784	984
4	200	100	10000	2643.4	443	2443

In our simulation experiments with SMRAS, the initial mean vector is uniformly selected from $[0, 2000] \times [0, 4000]$ for all four cases, and the covariance matrices are initialized as diagonal matrices with all diagonal elements equal to 10^6. The other parameters are: $r = 0.01$, $\varepsilon = 0.01$, $\lambda = 0.01$, $N_0 = 100$, initial $\rho = 0.1$, $\alpha = 1.04$, $M_k = \lceil 1.05 M_{k-1} \rceil$ with $M_0 = 10$, smoothing parameter $\upsilon = 0.5$. For SPSA, we have considered two cases: $q = 1$ and $q = 10$, where in both cases, the initial solutions are uniformly selected from $[0, 2000] \times [0, 4000]$, and a gain sequence $a_k = 200/k$ and a step-size sequence $c_k = 200/k^{0.25}$ are used, which give reasonable performance for different starting values. In both SMRAS and SPSA, the average cost per period is estimated by averaging the accumulated cost over 50 periods after a warm-up length of 50 periods.

The average performances of both algorithms for all test cases are given in Table 3, where N_p indicates the total number of periods (including the warm-up periods) simulated, and the entries represent the averaged function values J of the final sample solutions obtained for different choices of N_p, each based on 30 independent simulation replications. Since SPSA ($q = 10$) uses more precise gradient estimates, it can generally produce better solutions than SPSA ($q = 1$) within the same number of algorithm iterations. However, the performance gain of $q = 10$ over $q = 1$ is compromised by the additional simulation effort required in estimating the gradient, in the sense that for a fixed simulation budget, the solutions found by SPSA ($q = 10$) are actually worse than those found by SPSA ($q = 1$). SMRAS does not explicitly exploit the gradient structure, however, the algorithm still does very well as compared to an algorithm designed for convex problems and that utilizes the gradient information. We see that in all four cases, SMRAS provides superior empirical performance over SPSA, and finds solutions that are reasonably close to the optimal solution. Moreover, the algorithm also shows a significant variance reduction over SPSA.

4.2. Combinatorial Optimization. To illustrate the performance of SMRAS on discrete problems, we consider the buffer allocation problem in a service facility with unreliable servers. The system consists of m servers in series, which are separated by $m - 1$ buffer locations. Each job enters the system from the first server, goes through all intermediate servers and buffer locations in a sequential order, and

TABLE 3

Performance of SMRAS and SPSA on four test cases, based on 30 *independent simulation runs. The standard errors are in parentheses.*

Algorithm	Case	$N_p = 10^3$	$N_p = 10^4$	$N_p = 10^5$	$N_p = 10^6$	J^*
SMRAS	1	–	996.2(26.2)	820.1(9.1)	747.3(1.0)	740.9
	2	–	2342.3(14.7)	2263.3(8.2)	2216.6(2.8)	2200.0
	3	–	1310.8(15.7)	1291.4(13.6)	1219.5(3.7)	1184.4
	4	–	2737.3(13.9)	2690.2(5.2)	2663.5(3.3)	2643.4
SPSA $q=1$	1	1399.7(85.0)	968.6(37.9)	836.8(17.1)	813.6(17.1)	740.9
	2	2903.7(101.3)	2447.1(63.8)	2312.9(35.4)	2242.2(14.6)	2200.0
	3	2204.8(189.8)	1774.6(161.7)	1559.3(135.1)	1378.3(87.8)	1184.4
	4	4358.2(418.7)	3123.0(144.6)	2943.7(119.5)	2819.2(85.3)	2643.4
SPSA $q=10$	1	–	1498.1(94.6)	950.5(33.3)	829.6(16.7)	740.9
	2	–	2956.7(97.3)	2402.2(34.8)	2265.9(11.8)	2200.0
	3	–	2979.1(500.8)	1640.6(117.8)	1468.1(87.3)	1184.4
	4	–	4337.6(518.3)	2987.1(101.0)	2845.9(76.8)	2643.4

finally exits from the last server. The service times at each server are independent exponentially distributed with service rate μ_i, $i = 1, \ldots, m$. The servers are assumed to be unreliable, and are subject to random failures. When a server fails, it has to be repaired. The time to failure and the time for repair are both i.i.d. exponentially distributed with respective rates f_i and r_i, $i = 1, \ldots, m$. A server is blocked when the buffer associated with the server coming next to it is full and is starved when no jobs are offered to it. Thus, the status of a server (busy/broken) will affect the status of all other servers in the system. We assume that the failure rate of each server remains the same, regardless of its current status. Given n limited buffer spaces, our goal is to find an optimal way of allocating these n spaces to the $m-1$ buffer locations such that the throughput (average production rate) is maximized.

When applying SMRAS, we have used the same technique as in [1] to generate admissible buffer allocations; the basic idea is to use an $(m-1)$-by-$(n+1)$ matrix P, whose (i, j)th entry specifies the probability of allocating $j-1$ buffer spaces to the ith buffer location, and then generate allocations according to P. Please refer to their paper for a detailed discussion on how to generate admissible allocations. We define an allocation scheme x as an $(m-1)$-by-1 vector, whose ith element $x(i)$ specifies the number of buffer spaces allocated to the ith location. Thus, when parameterized by P, the probability of generating x is

$$f(x, P) = \prod_{i=1}^{m-1} \prod_{j=1}^{n+1} (P_{i,j})^{I\{x \in \mathbf{X}_{i,j}\}} = \prod_{i=1}^{m-1} e^{(\theta^i)^T \Gamma^i(x)},$$

where $\mathbf{X}_{i,j}$ represents the set of allocation schemes in which $j-1$ buffer spaces are allocated to the ith buffer location, $\theta^i = [\ln P_{i,0}, \ldots, \ln P_{i,n}]^T$, and $\Gamma^i(x) = [I\{x \in$

$\mathbf{X}_{i,0}\}, \ldots, I\{x \in \mathbf{X}_{i,n}\}]^T$. Once the admissible allocations are generated, it is not difficult to see that the entries of the matrix P are updated at the kth iteration as

$$P_{i,j}^{k+1} = \frac{\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k) I\{x \in \mathbf{X}_{i,j}\}}{\sum_{x \in \Lambda_k} S_k(\bar{\mathcal{J}}_k(x)) \chi(\bar{\mathcal{J}}_k(x), \bar{\gamma}_k)},$$

where $\Lambda_k = \{X_k^1, \ldots, X_k^{N_k}\}$ is the set of N_k admissible buffer allocations generated, and $\bar{\mathcal{J}}_k(x)$ is the average throughput obtained via simulation when the allocation x is used. It is not difficult to see that a straightforward interpretation of Theorem 1 yields $\lim_{k \to \infty} P_{i,j}^k = I\{x^* \in \mathbf{X}_{i,j}\} \ \forall i = 1, \ldots, m-1, \ \forall j = 1, \ldots, n+1$, which indicates that the sequence of stochastic matrices $\{P^k\}$ will converge to a matrix P^* with all mass at the optimal allocation scheme x^*, i.e., $P_{i,j}^* = 1$ for $j = x^*(i) + 1$ and $P_{i,j}^* = 0$ for all $j \neq x^*(i) + 1$, $i = 1, \ldots, m-1$.

For the numerical experiments, we consider two cases: (i) $m = 3$, $n = 1, \ldots, 10$, $\mu_1 = 1$, $\mu_2 = 1.2$ $\mu_3 = 1.4$, failure rates $f_i = 0.05$ and repair rates $r_i = 0.5$ for all $i = 1, 2, 3$; (ii) $m = 5$, $n = 1, \ldots, 10$, $\mu_1 = 1$, $\mu_2 = 1.1$, $\mu_3 = 1.2$, $\mu_4 = 1.3$, $\mu_5 = 1.4$, $f_i = 0.05$ and $r_i = 0.5$ for all $i = 1, \ldots, 5$.

Apart from their combinatorial nature, an additional difficulty in solving these problems is that different buffer allocation schemes (samples) have similar performances. Thus, when only noisy observations are available, it could be very difficult to identify the best allocation from a set of candidate allocation schemes. Because of this, in SMRAS we have used a relatively large $r = 2.3$. The other parameters are as follows: $\varepsilon = 0.001$, $\lambda = 0.01$, initial sample size $N_0 = 10$ for case (i) and $N_0 = 20$ for case (ii), initial $\rho = 0.1$, $\alpha = 1.2$, observation allocation rule $M_k = \lceil 1.5 M_{k-1} \rceil$ with $M_0 = 1$, the stopping control parameters $\tau = 1e-4$ and $l = 5$ (see Remark 6), smoothing parameter $\upsilon = 0.7$, and the initial P^0 is taken to be a uniform matrix with each column sum equal to one, i.e., $P_{i,j}^0 = \frac{1}{n+1} \ \forall i, j$. We start all simulation replications with the system empty. The steady-state throughputs are simulated after 100 warm-up periods, and then averaged over the subsequent 900 periods. Note that we have employed the sample reuse procedure (cf. Remark 1) in actual implementation of the algorithm.

Tables 4 and 5 give the performances of SMRAS for each of the respective cases (i) and (ii). In each table, N_{avg} is the averaged number of simulations over 16 independent trials, $Alloc$ is the true optimal allocation scheme and N_{A^*} is the number of times the optimal allocation was found out of 16 runs, $\bar{\mathcal{T}}$ is the averaged throughput value calculated by the algorithm, and $\mathcal{T}^*$ represents the exact optimal solution (cf. [31]). We see that in both cases, SMRAS produces very accurate solutions while using only a small number of simulations.

5. Conclusions and Future Research. We have proposed a Stochastic Model Reference Adaptive Search (SMRAS) for solving both continuous and discrete stochastic global optimization problems. The method generalizes the MRAS method for de-

TABLE 4

Performance of SMRAS on the buffer allocation problems case (i), based on 16 independent simulation runs. The standard errors are in parentheses.

n	$N_{avg}(std\ err)$	$Alloc\ (N_{A^*})$	$\bar{\mathcal{T}}(std\ err)$	$\mathcal{T}^*$
1	33.1(4.9)	[1,0] (16)	0.634(4.06e-4)	0.634
2	46.8(3.2)	[1,1] (16)	0.674(6.35e-4)	0.674
3	43.9(1.5)	[2,1] (16)	0.711(6.11e-4)	0.711
4	49.8(3.5)	[3,1] (14)	0.735(6.47e-4)	0.736
5	50.4(3.7)	[3,2] (13)	0.758(1.06e-3)	0.759
6	64.0(6.3)	[4,2] (12)	0.776(1.39e-3)	0.778
7	59.1(4.3)	[5,2] (14)	0.792(1.04e-3)	0.792
8	63.9(4.8)	[5,3] (10)	0.805(1.20e-3)	0.806
9	60.6(3.5)	[6,3] (10)	0.817(6.53e-4)	0.818
10	63.7(5.7)	[7,3] (12)	0.826(9.88e-4)	0.827

TABLE 5

Performance of SMRAS on the buffer allocation problem case (ii), based on 16 independent simulation runs. The standard errors are in parentheses.

n	$N_{avg}(std\ err)$	$Alloc\ (N_{A^*})$	$\bar{\mathcal{T}}(std\ err)$	$\mathcal{T}^*$
1	1.02e+2(7.49)	[0,1,0,0] (16)	0.523(6.79e-4)	0.521
2	1.29e+2(14.8)	[1,1,0,0] (16)	0.555(3.86e-4)	0.551
3	1.75e+2(15.7)	[1,1,1,0] (16)	0.587(4.57e-4)	0.582
4	2.51e+2(25.9)	[1,2,1,0] (11)	0.606(1.20e-3)	0.603
5	3.37e+2(42.0)	[2,2,1,0] (10)	0.626(6.57e-4)	0.621
6	4.69e+2(55.2)	[2,2,1,1] (8)	0.644(1.10e-3)	0.642
7	4.56e+2(58.2)	[2,2,2,1] (7)	0.659(1.10e-3)	0.659
8	4.45e+2(54.9)	[3,2,2,1] (7)	0.674(1.10e-3)	0.674
9	5.91e+2(56.1)	[3,3,2,1] (6)	0.689(1.39e-3)	0.689
10	5.29e+2(54.0)	[3,3,3,1] (8)	0.701(1.10e-3)	0.701

terministic optimization, which is motivated by the well-known Cross-Entropy method. SMRAS is general, requires only a few mild regularity conditions on the underlying problem; and thus can be applied to a wide range of problems with little modification. More importantly, we believe that the idea behind the method offers a general framework for stochastic global optimization, based on which one can possibly design and implement other efficient algorithms.

There are several input parameters in SMRAS. In our preliminary numerical experiments, the choices of these parameters are based on trial and error. For a given problem, how to determine a priori the most appropriate values of these parameters is an open issue. One research topic is to study the effects of these parameters on the

performance of the method, and possibly design an adaptive scheme to choose these parameters adaptively during the search process.

Our current numerical study with the algorithm shows that the objective function need not be evaluated very accurately during the initial search phase. Instead, it is sufficient to provide the algorithm with a rough idea of where the good solutions are located. This has motivated our research to use observation allocation rules with adaptive increasing rates during different search phases. For instance, during the initial search phase, we could increase M_k at a linear rate or even keep it at a constant value; and exponential rates will only be used during the later search phase when more accurate estimates of the objective function values are required.

Some other research topics include analyzing the convergence rate of the method for a class of optimization problems and incorporating local search techniques in the method to further enhance its performance.

REFERENCES

[1] G. ALLON, D. P. KROESE, T. RAVIV, AND R. Y. RUBINSTEIN, *Application of the cross-entropy method to the buffer allocation problem in a simulation-based environment*, Annals of Operations Research, 134(2005), pp. 137-151.

[2] M. H. ALREFAEI AND S. ANDRADÓTTIR, *A modification of the stochastic ruler method for discrete stochastic optimization*, European Journal of Operational Research, 133(1995), pp. 160-182.

[3] M. H. ALREFAEI AND S. ANDRADÓTTIR, *A simulated annealing algorithm with constant temperature for discrete stochastic optimization*, Management Science, 45(1999), pp. 748-764.

[4] S. ANDRADÓTTIR, *A method for discrete stochastic optimization*, Management Science, 41(1995), pp. 1946-1961.

[5] P. AUER, N. CESA-BIANCHI, AND P. FISHER, *Finite-time analysis of the multiarmed bandit problem*, Machine Learning, 47(2002), pp. 235-256.

[6] R. W. BROCKETT, *New issues in the mathematics of control*, in: Mathematics Unlimited − 2001 and Beyond (B. Engquist and W. Schmid, editors), pp. 189-219, Springer Verlag, Berlin, 2001.

[7] H. S. CHANG, M. C. FU, J. HU, AND S. I. MARCUS, *Simulation-based Algorithms for Markov Decision Processes.* Springer-Verlag, London, 2007.

[8] A. CORANA, M. MARCHESI, C. MARTINI, AND S. RIDELLA, *Minimizing multimodal functions of continuous variables with the simulated annealing algorithm*, ACM Trans. Mathematical Software, 13(1987), pp. 262-280.

[9] P. T. DE BOER, D. P. KROESE, S. MANNOR, AND R. Y. RUBINSTEIN, *A tutorial on the cross-entropy method*, Annals of Operations Research, 134(2005), pp. 19-67.

[10] M. DORIGO AND L. M. GAMBARDELLA, *Ant colony system: a cooperative learning approach to the traveling salesman problem*, IEEE Trans. on Evolutionary Computation, 1(1997), pp. 53-66.

[11] M. C. FU AND K. J. HEALY, *Techniques for simulation optimization: an experimental study on an (s, S) inventory system*, IIE Transactions, 29(1997), pp. 191-199.

[12] M. C. FU, *Stochastic gradient estimation.* Chapter 19 in Handbooks in Operations Research and Management Science: Simulation, S.G. Henderson and B.L. Nelson, eds., Elsevier, 2005.

[13] P. GLASSERMAN, *Gradient estimation via perturbation analysis.* Kluwer Academic Publisher, Boston, Massachusetts, 1991.

[14] P. W. GLYNN, *Likelihood ratio gradient estimation: an overview*, Proceedings of the 1987 Winter Simulation Conference, pp. 366-375.

[15] W. J. GUTJAHR, *A converging ACO algorithm for stochastic combinatorial optimization*, Proc. SAGA 2003 Stochastic Algorithms: Foundations and Applications, Hatfield (UK), A. Albrecht, K. Steinhoefl, eds., Springer LNCS 2827 $10 - 25$.

[16] W. HOEFFDING, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association, 58(1963), pp. 13-30.

[17] L. J. HONG AND B. L. NELSON, *Discrete optimization via simulation using COMPASS*, Operations Research, 54(2006), pp. 115-129.

[18] Y. C. HO AND X. R. CAO, *Perturbation analysis of discrete event dynamic systems.* Kluwer Academic Publisher, Norwell, Massachusetts, 1991.

[19] J. HU, M. C. FU, AND S. I. MARCUS, *A model reference adaptive search method for global optimization*, Operations Research, 55(2007), pp. 549-568.

[20] J. KIEFER AND J. WOLFOWITZ, *Stochastic estimation of the maximum of a regression function*, Annals of Mathematical Statistics, 23(1952), pp. 462-466.

[21] P. L'ECUYER, *An overview of derivative estimation*, Proceedings of the 1991 Winter Simulation Conference, pp. 207-217.

[22] G. C., PFLUG, *Sampling derivatives of probabilities*, Computing, 42(1989), pp. 315-328.

[23] H. ROBBINS AND S. MONRO, *A stochastic approximation method*, Annals of Mathematical Statistics, 22(1951), pp. 400-407.

[24] R. Y. RUBINSTEIN, *The cross-entropy method for combinatorial and continuous optimization*, Methodology and Computing in Applied Probability, 2(1999), pp. 127-190.

[25] R. Y. RUBINSTEIN, *Combinatorial optimization, ants and rare events*, Stochastic Optimization: Algorithms and Applications, $304 - 358$, S. Uryasev and P. M. Pardalos, eds., Kluwer, 2001.

[26] R. Y. RUBINSTEIN AND D. P. KROESE, The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation, and machine learning. Springer, New York, 2004.

[27] R. Y. RUBINSTEIN AND A. SHAPIRO, Discrete Event Systems: Sensitivity Analysis and Stochastic Optimization by the Score Function Method. John Wiley & Sons, 1993.

[28] L. SHI AND S. ÓLAFSSON, *Nested partitions method for stochastic optimization*, Methodology and Computing in Applied Probability, 2(2000), pp. 271-291.

[29] J. C. SPALL, *Multivariate stochastic approximation using simultaneous perturbation gradient approximation*, IEEE Transactions on Automatic Control, 37(1992), pp. 332-341.

[30] M. SRINIVAS AND L. M. PATNAIK, *Genetic Algorithms: A Survey*, IEEE Computer, 27(1994), pp. 17-26.

[31] G. A. VOUROS AND H. T. PAPADOPOULOS, *Buffer allocation in unreliable production lines using a knowledge based system*, Computer & Operations Research, 25(1998), pp. 1055-1067.

[32] D. H. WOLPERT, *Finding Bounded Rational Equilibria Part I: Iterative Focusing*, Proceedings of ISDG 04, T. Vincent (Ed.), 2004.

[33] S. YAKOWITZ, P. L'ECUYER, AND F. VÁZQUEZ-ABAD, *Global Stochastic Optimization with Low-Dispersion Point Sets*, Operations Research, 48(2000), pp. 939-950.

[34] D. YAN AND H. MUKAI, *Stochastic discrete optimization*, SIAM Journal on Control and Optimization, 30(1992), pp. 594-612.

[35] M. ZLOCHIN, M. BIRATTARI, N. MEULEAU, AND M. DORIGO, *Model-based search for combinatorial optimization: A Critical Survey*, Annals of Operations Research, 131(2004), pp. 373-395.