Proceedings of the 1997 Winter Simulation Conference
ed. S. Andradóttir, K. J. Healy, D. H. Withers, and B. L. Nelson

ACCELERATING THE CONVERGENCE OF THE STOCHASTIC RULER METHOD FOR DISCRETE STOCHASTIC OPTIMIZATION

Mahmoud H. Alrefaei

Department of Mathematical Sciences
Jordan University of Science & Technology
P.O. Box 3030, Irbid 22110, JORDAN

Sigrún Andradóttir

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332, U.S.A.

ABSTRACT

We present two new variants of the stochastic ruler method for solving discrete stochastic optimization problems. These two variants use the same mechanism for moving around the state space as the modified stochastic ruler method we have proposed earlier. However, the new variants use different approaches for estimating the optimal solution. In particular, the modified stochastic ruler method uses the number of visits to each state by the Markov chain generated by the algorithm to estimate the optimal solution. On the other hand, one of our new methods uses the number of visits to each state by the embedded chain of the Markov chain generated by the algorithm to estimate the optimal solution, and our other new method uses the feasible solution with the best average estimated objective function value to estimate the optimal solution. Like our earlier modification of the stochastic ruler method, these two new methods are guaranteed to converge almost surely to the set of global optimal solutions. We present theoretical and numerical results that indicate that our new approaches tend to lead to the set of global optimal solutions faster.

1 INTRODUCTION

Optimizing a stochastic system over a discrete set of decision parameters (discrete stochastic optimization) is an important and active area of research. If the number of decision parameters is small, then methods of ranking and selection, and multiple comparisons procedures can be used to select the best system with high probability. For more details, see Bechhofer, Santner, and Goldsman (1995) and Hsu (1996). Recently Norkin, Ermoliev, and Ruszczyński (1997) and Norkin, Pflug, and Ruszczyński (1996) have considered using a version of the Branch-and-Bound method for solving discrete stochastic optimization problems. Some other work in this area includes Gelfand and Mitter (1989), Alrefaei and Andradóttir (1995), Fox and Heine (1995), and Gutjahr and Pflug (1996) who consider using the simulated annealing algorithm for solving discrete stochastic optimization problems.

Yan and Mukai (1992) have proposed a random search method called the *stochastic ruler method* for solving discrete stochastic optimization problems. Their method requires an increasing number of observations of the objective function values per iteration as the number of iterations grows. One of the difficulties in using the original stochastic ruler method is that its convergence is very sensitive to the rate at which the number of observations per iteration increases. To avoid this difficulty, Alrefaei and Andradóttir (1996, 1997a) have proposed a modification of the stochastic ruler method that uses only a fixed number of observations per iteration. This method estimates the optimal solution using an approach that resembles the approach proposed by Andradóttir (1995, 1996); i.e., the number of visits the algorithm makes to the different states is used to estimate the optimal solution. Alrefaei and Andradóttir (1996, 1997a) present numerical results suggesting that their approach tends to perform better than the original stochastic ruler algorithm of Yan and Mukai (1992).

In this paper we propose two new variants of the modified stochastic ruler method of Alrefaei and Andradóttir (1996, 1997a) that use the same mechanism for moving around the state space, but they use different approaches for estimating the optimal solution. The first method uses the number of visits the embedded chain of the Markov chain generated by the modified algorithm makes to the different states to estimate the optimal solution, whereas, in the second method, we let the state that has the best average estimated objective function value obtained from all the previous observations of the objective function values be the estimate of the optimal solution. This latter approach for estimating the optimal solution has

been suggested by Andradóttir (1997). We discuss under what conditions these approaches converge almost surely and when they can be expected to accelerate the convergence to the set of global optimal solutions. Finally, we provide numerical results for our new methods and compare their performance with that of the earlier modified stochastic ruler method.

This paper is organized as follows: In Section 2 we give some background on the stochastic ruler method. In Sections 3 and 4 we present our new variants of the stochastic ruler method and discuss under what conditions they are expected to accelerate the convergence. In Section 5 we provide some numerical results. Finally, in Section 6 we give some concluding remarks.

2- THE STOCHASTIC RULER METHOD

Consider the following optimization problem

$$\min_{x\in\mathcal{S}} f(x) = E[h(x, Y_x)], \psi \tag{1}$$

where $f : \mathcal{S} \to \mathbb{R}$, $\mathcal{S}$ is a discrete set, and it is assumed to be *finite* throughout this paper, h is a deterministic, real-valued function, and Y_x is a random variable that depends on the parameter $x \in \mathcal{S}$. Let $H(x) = h(x, Y_x)$ for all $x \in \mathcal{S}$, and let $\mathcal{S}^* = \{x^* \in \mathcal{S} : f(x^*) \leq f(x), \forall x \in \mathcal{S}\}$ denote the set of global optimal solutions to the optimization problem (1).

Yan and Mukai (1992) have proposed a method called the stochastic ruler method for solving the optimization problem (1). In their method, Yan and Mukai solve the following maximization problem

$$\max_{x\in\mathcal{S}} P(x, a, b) = P\{H(x) \leq \Theta(a, b)\}, \psi \tag{2}$$

where the stochastic ruler $\Theta(a, b)$ is a uniform random variable defined on the range (a, b) of the observed objective function values. Yan and Mukai show that a solution of the maximization problem (2) is also a solution to the original minimization problem (1) under certain conditions. The stochastic ruler method involves comparing the observed objective function values $H(x)$, where $x \in \mathcal{S}$, with observations of the stochastic ruler $\Theta(a, b)$. The maximum number of such comparisons in iteration k of the algorithm is M_k, where $M_k \to \infty$ as $k \to \infty$. One of the difficulties in implementing the stochastic ruler method is that its convergence depends on the rate at which the sequence $\{M_k\}$ is increased. If the sequence $\{M_k\}$ is increased rapidly, then the algorithm may end up at a local solution; on the other hand, if the sequence $\{M_k\}$ is increased slowly, then the algorithm tends to take a long time to converge. Alrefaei and Andradóttir (1996, 1997a) propose a modification of the stochastic ruler method that uses at most M comparisons per iteration, where $M > 0$ is a constant integer. Moreover, their method uses a different approach for estimating the optimal solution than the original stochastic ruler method of Yan and Mukai (1992); this approach resembles the approach proposed by Andradóttir (1995, 1996) that uses the state that is visited most often by the algorithm as the estimate of the optimal solution.

We need the following definitions and assumptions:

Definition 1 *For each $x \in \mathcal{S}$, there exists a subset $N(x)$ of $\mathcal{S} \setminus \{x\}$ which is called the set of neighbors of x.*

Assumption 1 *For any $x, x' \in \mathcal{S}$, x' is reachable from x; i.e., there exists a finite sequence $\{n_i\}_{i=0}^{l}$ for some l, such that $x_{n_0} = x, x_{n_l} = x'$, and $x_{n_{i+1}} \in N(x_{n_i})$, $i = 0, 1, 2, \ldots, l-1$.*

Definition 2 *A function $R : \mathcal{S} \times \mathcal{S} \to [0, 1]$ is said to be a transition probability for $\mathcal{S}$ and N if*

1. *$R(x, x') > 0 \Leftrightarrow x' \in N(x)$, and*
2. *$\sum_{x'\in\mathcal{S}} R(x, x') = 1$.*

Assumption 2 *Let*

$$R(x, x') = \frac{R'(x, x')}{D(x)}, \psi \tag{3}$$

where $R' : \mathcal{S}\times\mathcal{S} \to \mathbb{R}$ is a function such that $R'(x, x') > 0 \Leftrightarrow x' \in N(x)$ and $D : \mathcal{S} \to \mathbb{R}$ is defined by $D(x) = \sum_{x'\in\mathcal{S}} R'(x, x')$, $\forall x \in \mathcal{S}$. Then we assume that

1. *$x' \in N(x) \Leftrightarrow x \in N(x')$, and*
2. *$R'(x, x') = R'(x', x)$, $\forall x, x' \in \mathcal{S}$.*

Assumption 3 *The parameters $a, b \in \mathbb{R}$ satisfy*

1. *if $P(x, a, b) \geq P(x', a, b)$, then $f(x) \leq f(x')$, and*
2. *$0 < P(x, a, b) < 1$, for all $x \in \mathcal{S}$.*

Note that Assumption 3 implies that a solution to the optimization problem (2) is also a solution to the optimization problem (1). Yan and Mukai (1992) show that parameters a, b satisfying Assumption 3 exist when $E\{H(x)^2\} < \infty$ for all $x \in \mathcal{S}$.

Now we state the modified stochastic ruler algorithm of Alrefaei and Andradóttir (1996, 1997a). Note that for all $x \in \mathcal{S}$ and $k \in \mathbb{N}$, $V_k(x)$ is the number of times that the Markov chain $\{X_k\}$ has visited state x in the first k iterations, and X_k^* is the state that maximizes the value of $V_k(x)/D(x)$, where $x \in \mathcal{S}$.

Algorithm 1

Step 0: *Select a starting point $X_0 \in \mathcal{S}$. Let $V_0(X_0) = 1$, and $V_0(x) = 0$, for all $x \in \mathcal{S}$, $x \neq X_0$. Let $k = 0$ and $X_k^* = X_0$.*

Step 1: *Given $X_k = x$, choose a candidate Z_k from $N(x)$ with probability distribution*

$$P\{Z_k = z | X_k = x\} = \frac{R'(x,z)}{D(x)},$$

where $z \in N(x)$, and $R'(x,z)$ and $D(x)$ are defined in equation (3).

Step 2: *Given $Z_k = z$, draw a sample $h(z)$ from $H(z)$. Then draw a sample θ from $\Theta(a,b)$. If $h(z) > \theta$, then let $X_{k+1} = X_k$ and go to Step 3. Otherwise, draw another sample $h(z)$ from $H(z)$ and draw another sample θ from $\Theta(a,b)$ that are independent of the previous samples. If $h(z) > \theta$, then let $X_{k+1} = X_k$ and go to Step 3. Otherwise, continue to draw and compare M times. If all M tests, $h(z) > \theta$, fail, then accept the candidate Z_k and set $X_{k+1} = Z_k = z$.*

Step 3: *Let $k = k+1, V_k(X_k) = V_{k-1}(X_k) + 1$, and $V_k(x) = V_{k-1}(x)$, for all $x \in \mathcal{S}$, $x \neq X_k$. If*

$$\frac{V_k(X_k)}{D(X_k)} > \frac{V_k(X_{k-1}^*)}{D(X_{k-1}^*)},$$

then let $X_k^* = X_k$; otherwise let $X_k^* = X_{k-1}^*$. Go to Step 1.

Under Assumptions 1 through 3, Alrefaei and Andradóttir (1997a) show that the sequence $\{X_k^*\}$ generated by Algorithm 1 converges almost surely to the set $\mathcal{S}^*$.

In some cases Algorithm 1 may spend a lot of time at a bad state x, especially if the neighboring states $x' \in N(x)$ are also bad. This is because moving from a state x to a state x' does not depend on the distribution of $H(x)$, but only on that of $H(x')$. This means that the state x could remain an estimate of the optimal solution for a long time. To avoid this, we propose two methods in Sections 3 and 4 that are expected to accelerate the convergence of the algorithm.

3- THE EMBEDDED CHAIN APPROACH

This approach involves modifying Algorithm 1 by focusing on the embedded chain of the Markov chain $\{X_k\}$ generated by Algorithm 1. In this approach, instead of counting how many visits the Markov chain $\{X_k\}$ makes to each state, including the time it spends at that state before moving to another state, we just count how many times this Markov chain $\{X_k\}$ enters each state without counting the time it spends at that state. This yields the following algorithm:

Algorithm 2

Step 0: *Identical to Step 0 of Algorithm 1.*

Step 1: *Identical to Step 1 of Algorithm 1.*

Step 2: *Identical to Step 2 of Algorithm 1.*

Step 3: *Let $k = k+1$. If $X_k = Z_{k-1}$, then let $V_k(X_k) = V_{k-1}(X_k) + 1$; otherwise let $V_k(X_k) = V_{k-1}(X_k)$. Let $V_k(x) = V_{k-1}(x)$, for all $x \in \mathcal{S}$, $x \neq X_k$. If*

$$\frac{V_k(X_k)}{D(X_k)} > \frac{V_k(X_{k-1}^*)}{D(X_{k-1}^*)},$$

then let $X_k^* = X_k$; otherwise let $X_k^* = X_{k-1}^*$. Go to Step 1.

For all $x \in \mathcal{S}$, define

$$u(x) = \sum_{z \in N(x)} R'(x,z) \left[P(z,a,b)\right]^M, \psi \qquad (4)$$

where $R'(x,z)$ is given in equation (3) and $P(z,a,b)$ is defined in equation (2). The following two theorems have been proved by Alrefaei and Andradóttir (1997b). Note that the notation $|A|$ refers to the number of elements in the set A.

Theorem 1
Suppose that Assumption 3 holds. If for all $x, y \in \mathcal{S}, x \neq y$, $N(x) = \mathcal{S} \setminus \{x\}$ and $R'(x,y) = 1/(|\mathcal{S}| - 1)$, then the sequence $\{X_k^*\}$ generated by Algorithm 2 converges almost surely to the set $\mathcal{S}^*$.

Theorem 2
Suppose that Assumptions 1 through 3 hold, and that at least for one global optimal solution $x^* \in \mathcal{S}^*$, we have

$$\frac{u(x^*)}{D(x^*)} \geq \frac{u(x)}{D(x)}, \forall x \in \mathcal{S}, \psi \qquad (5)$$

where $D(\cdot)$ is defined in equation (3) and $u(\cdot)$ is defined in equation (4). Then the sequence $\{X_k^*\}$ generated by Algorithm 2 converges almost surely to the set $\mathcal{S}^*$.

Suppose that π is the stationary distribution for the Markov chain $\{X_k\}$ generated by Algorithms 1 and 2 and that π' is the stationary distribution of the embedded chain of the Markov chain $\{X_k\}$. (Alrefaei and Andradóttir (1997a, 1997b) have shown that π and π' exist.) Then the following proposition was proved by Alrefaei and Andradóttir (1997b).

Proposition 3

1. *Suppose that Assumption 3 holds and that for all* $x, y \in \mathcal{S}, x \neq y$, $N(x) = \mathcal{S}\backslash\{x\}$ *and* $R'(x,y) = 1/(|\mathcal{S}|-1)$. *Then* $\pi'_{x^*} < \pi_{x^*}$ *for all* $x^* \in \mathcal{S}^*$.

2. *Suppose that Assumptions 1 through 3 and equation* (5) *hold. Then* $\pi'_{x^*} \geq \pi_{x^*}$ *for the global optimal solution* $x^* \in \mathcal{S}^*$ *satisfying equation* (5).

Alrefaei and Andradóttir (1997a, 1997b) show that Algorithms 1 and 2 are in fact maximizing $\pi_x/D(x)$ and $\pi'_x/D(x)$, respectively. So if $\pi'_{x^*} > \pi_{x^*}$ then Algorithm 2 would be expected to perform better than Algorithm 1; the reverse is true when $\pi'_{x^*} < \pi_{x^*}$. Therefore, by Proposition 3, Algorithm 2 is expected to perform better than Algorithm 1 under the conditions of Theorem 2, but not under those of Theorem 1.

4- THE BEST AVERAGE ESTIMATE APPROACH

In this approach, we use the state that has the best (lowest since we are minimizing) average estimated objective function value as the estimate of the optimal solution. This approach for estimating the global optimal solution was originally proposed by Andradóttir (1997). We have the following algorithm:

Algorithm 3

Step 0: *Select a starting point* $X_0 \in \mathcal{S}$. *For all* $x \in \mathcal{S}$, *let* $A_0(x) = 0$ *and* $C_0(x) = 0$. *Let* $k = 0$ *and* $X_k^* = X_0$.

Step 1: *Identical to Step 1 of Algorithm 1.*

Step 2: *Given* $Z_k = z$, *draw a sample* $h(z)$ *from* $H(z)$, *and let* $A_k(z) = A_k(z) + h(z)$ *and* $C_k(z) = C_k(z) + 1$. *Then draw a sample* θ *from* $\Theta(a,b)$. *If* $h(z) > \theta$, *then let* $X_{k+1} = X_k$ *and go to Step 3. Otherwise, draw another sample* $h(z)$ *from* $H(z)$ *that is independent of the previous samples and update* $A_k(z)$ *and* $C_k(z)$ *as before, and draw another sample* θ *from* $\Theta(a,b)$ *that is independent of the previous samples. If* $h(z) > \theta$, *then let* $X_{k+1} = X_k$ *and go to Step 3. Otherwise, continue to draw and compare* M *times, each time updating* $A_k(z)$ *and* $C_k(z)$ *as before. If all* M *tests,* $h(z) > \theta$, *fail, then accept the candidate* Z_k *and set* $X_{k+1} = Z_k = z$.

Step 3: *Select* $X_k^* \in \arg\min_{x\in\mathcal{S}} A_k(x)/C_k(x)$. *Let* $k = k+1$, *and for all* $x \in \mathcal{S}$, *let* $A_k(x) = A_{k-1}(x)$ *and* $C_k(x) = C_{k-1}(x)$. *Go to Step 1.*

The following theorem has been proved by Alrefaei and Andradóttir (1997b).

Theorem 4
Under Assumptions 1 through 3, the sequence $\{X_k\}$ *generated by Algorithm 3 converges almost surely to the set* $\mathcal{S}^*$.

Theorem 5 in Alrefaei and Andradóttir (1997a) shows that the Markov chain $\{X_k\}$ generated by Algorithms 1, 2, and 3 is attracted to good states (states that have small objective function values) in the sense that the expected number of visits the Markov chain $\{X_k\}$ makes to good states is larger than the expected number of visits to bad states (states that have large objective function values). This means that good estimates of the objective function values at good states will be obtained quickly, suggesting that Algorithm 3 may converge rapidly to the set of global optimal solutions. Numerical results supporting this intuition on a particular example are presented in Section 5.

5- NUMERICAL APPLICATION

In this section, we present empirical results obtained by applying Algorithms 1, 2, and 3 to solve the following discrete stochastic optimization problem:

$$\min_{x\in\mathcal{S}} f(x) = E[W(x)],$$

where $\mathcal{S} = \{1, \ldots, 50\}$, $W_i(x)$ is the system time of customer i for all $x \in \mathcal{S}$, and $W(x) = \frac{1}{200}\sum_{i=1}^{200} W_i(x)$ is the average system time per customer of the first 200 customers in an M/M/1 queue with fixed arrival rate $\lambda = 1$ and service rate $\mu(x)$, $\forall x \in \mathcal{S}$. Figure 1 shows the estimated values of $f(x) = E[W(x)]$ for all $x \in \mathcal{S}$ obtained from separate long simulation runs.

For each $x \in \mathcal{S}$ and $i \in \mathbb{N}$, let $S_i(x)$ be the service time of customer i and let T_i be the inter-arrival time between customers $i-1$ and i. Assume that $T_0 = 0$ and $W_0(x) = 0$ for all $x \in \mathcal{S}$. Then we use the following recursive formula to generate the system time of customer i :

$$W_i(x) = \max\{S_i(x), W_{i-1}(x) + S_i(x) - T_i\}.$$

We apply Algorithms 1, 2, and 3 using two different neighborhood structures. The first neighborhood structure is given by

$$N_1(x) = \begin{cases} \{2\} \leftarrow & \text{if } x = 1, \\ \{x \pm 1\} & \text{if } 2 \leq x \leq 49, \\ \{49\} \leftarrow & \text{if } x = 50. \end{cases} \tag{6}$$

For this neighborhood structure, for all $x \in \mathcal{S}$, we take $R'_1(x,x') = 1$ for all $x' \in N_1(x)$, $R'_1(x,x') = 0$

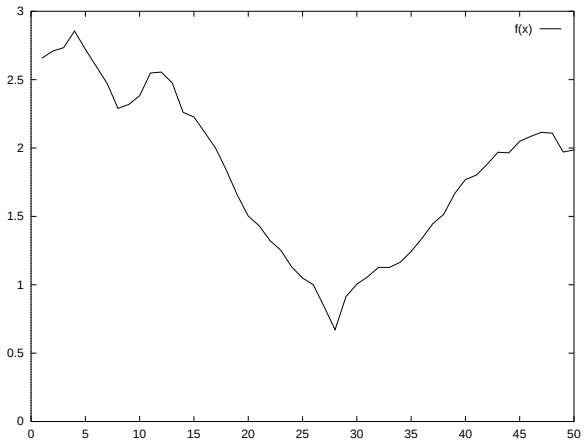

Figure 1: The estimated objective function values $f(x) = E[W(x)]$ for all $x \in \mathcal{S}$.

for all $x \in \mathcal{S} \setminus N_1(x)$, and $D_1(x) = |N_1(x)|$. The second neighborhood structure is given by

$$N_2(x) = \mathcal{S} \setminus \{x\}, \ \forall x \in \mathcal{S}. \ \psi \tag{7}$$

For this neighborhood structure we take $R_2(x, x') = R_2'(x, x') = 1/(|\mathcal{S}| - 1)$ for all $x, x' \in \mathcal{S}, x \neq x'$, and $R_2(x, x) = R_2'(x, x) = 0$ for all $x \in \mathcal{S}$ (so that $D_2(x) = 1$ for all $x \in \mathcal{S}$). Note that Proposition 3 implies that Algorithm 2 is expected to perform better than Algorithm 1 when the first neighborhood structure N_1 given in equation (6) is used. The reverse is true when the second neighborhood structure N_2 given in equation (7) is used. We let $M = 2$, $a = 0.5$, and $b = 2.5$. We always select the initial state randomly and we run the program for 100 replications.

Figure 2 shows the average performance of the three algorithms over 100 replications when the first neighborhood structure N_1 given in equation (6) is used. The x-axis shows the number of customers that have been used in our simulations for estimating the objective function values and the y-axis shows the average estimated optimal objective function values over the 100 replications. It is clear from this figure that the performance of Algorithm 3 is superior to the performance of the other two algorithms. It is also clear that the performance of Algorithm 2 is better than the performance of Algorithm 1 as expected since the structure of this problem using the neighborhood structure N_1 satisfies equation (5). Figure 3 shows the results when the second neighborhood structure N_2 given in equation (7) is used. Again, we note that the performance of Algorithm 3 is superior to the performance of the other two algorithms. However, there is no significant difference between Algorithms 1 and 2 in this setting. This is because

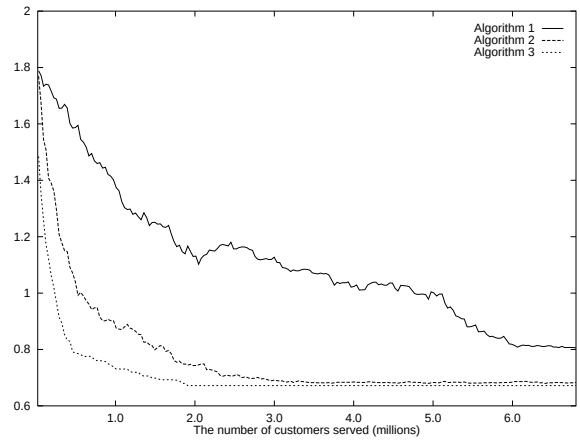

Figure 2: Performance of Algorithms 1, 2, and 3 using the neighborhood structure N_1. The estimated objective function value at the global optimal solution is $\hat{f}^* = 0.6717$.

equation (5) is not satisfied when the neighborhood structure N_2 is used.

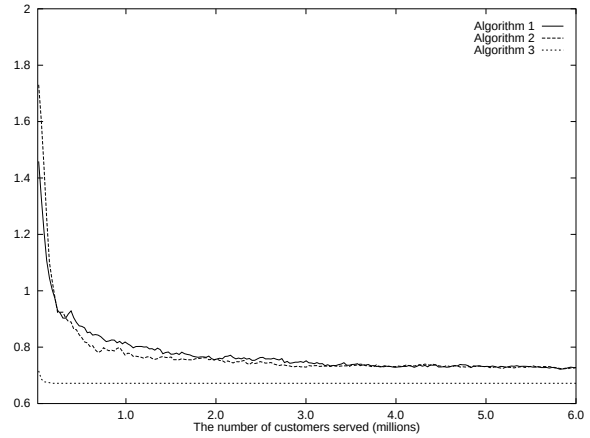

Figure 3: Performance of Algorithms 1, 2, and 3 using the neighborhood structure N_2. The estimated objective function value at the global optimal solution is $\hat{f}^* = 0.6717$.

6- CONCLUSION

In this paper, we have proposed two new variants of the modified stochastic ruler method of Alrefaei and Andradóttir (1996, 1997a). The first variant uses the number of visits the embedded chain of the Markov chain generated by the modified stochastic ruler method makes to the different states to estimate the optimal solution, whereas the second variant uses

an approach that has been proposed by Andradóttir (1997) that uses the state with the best average estimated objective function value as estimate of the optimal solution. Our methods are guaranteed to converge *almost surely* to the set of global optimal solutions under mild assumptions. From our numerical results, we conclude that in comparison with the other two variants of the stochastic ruler method, the variant that uses the state with the best average estimated objective function value as estimate of the optimal solution shows the best performance.

ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under grants DDM–9210679 and DMI–9523111.

REFERENCES

Alrefaei, M. H., and S. Andradóttir. 1995. A new search algorithm for discrete stochastic optimization. In *Proceedings of the 1995 Winter Simulation Conference*, ed. C. Alexopoulos, K. Kang, W. R. Lilegdon, and D. Goldsman, 236–241. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Alrefaei, M. H., and S. Andradóttir. 1996. Discrete stochastic optimization via a modification of the stochastic ruler method. In *Proceedings of the 1996 Winter Simulation Conference*, ed. J. M. Charnes, D. M. Morrice, D. T. Brunner, and J. J. Swain, 406–411. Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.

Alrefaei, M. H., and S. Andradóttir. 1997a. A modification of the stochastic ruler method for discrete stochastic optimization. Working paper.

Alrefaei, M. H., and S. Andradóttir. 1997b. Discrete stochastic optimization using variants of the stochastic ruler method. Working paper.

Andradóttir, S. 1995. A method for discrete stochastic optimization. *Management Science* 41:1946–1961.

Andradóttir, S. 1996. A global search method for discrete stochastic optimization. *SIAM Journal on Optimization* 6:513–530.

Andradóttir, S. 1997. Accelerating the convergence of random search methods for discrete stochastic optimization. Working paper.

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

Fox, B. L., and G. W. Heine. 1995. Probabilistic search with overrides. *The Annals of Applied Probability* 5:1087–1094.

Gelfand, S. B., and S. K. Mitter. 1989. Simulated annealing with noisy or imprecise energy measurements. *Journal of Optimization Theory and Applications* 62:49–62.

Gutjahr, W. J., and G. Ch. Pflug. 1996. Simulated annealing for noisy cost functions. *Journal of Global Optimization* 8:1–13.

Hsu, J. C. 1996. *Multiple Comparisons: Theory and Methods*. New York: Chapman & Hall.

Norkin, V., Y. Ermoliev, and A. Ruszczyński. 1997. On optimal allocation of individuals under uncertainty. To appear in *Operations Research*.

Norkin, V., G. Ch. Pflug, and A. Ruszczyński. 1996. A branch and bound method for stochastic global optimization. Working paper.

Yan, D., and H. Mukai. 1992. Stochastic discrete optimization. *SIAM Journal on Control and Optimization* 30:594–612.

AUTHOR BIOGRAPHIES

MAHMOUD HASAN ALREFAEI is an Assistant Professor in the Department of Mathematical Sciences at the Jordan University of Science and Technology. He received his B.S. and M.S. degrees in Mathematics from Yarmouk University in Jordan and a joint Ph.D. degree in Industrial Engineering and Mathematics from the University of Wisconsin – Madison in 1997. His research interests include stochastic optimization, simulation, and simulated annealing. He is presently a member of INFORMS and AMS.

SIGRÚN ANDRADÓTTIR is an Associate Professor in the School of Industrial and Systems Engineering at the Georgia Institute of Technology. Previously, she was an Assistant Professor and later an Associate Professor in the Departments of Industrial Engineering, Mathematics, and Computer Sciences at the University of Wisconsin – Madison. She received a B.S. in Mathematics from the University of Iceland in 1986, an M.S. in Statistics from Stanford University in 1989, and a Ph.D. in Operations Research from Stanford University in 1990. Her research interests include stochastic optimization, simulation, and stochastic processes. She presently serves as Associate Editor for *IIE Transactions*, *Stochastic Models*, and the *ACM Transactions on Modeling and Computer Simulation*. She is a member of INFORMS and Editor of the *Proceedings of the 1997 Winter Simulation Conference*.