Discrete optimization, SPSA and Markov Chain Monte Carlo methods

László Gerencsér, Stacy D. Hill, Zsuzsanna Vágó and Zoltán Vincze

***Abstract*— The minimization of a convex function defined over the grid $\mathbf{Z}^p$ is considered. A few relevant mathematical devices such as integer convexity, Markov Chain Monte Carlo (MCMC) methods, including stochastic comparison (SC), and simultaneous perturbation stochastic approximation (SPSA)are summarized. A truncated fixed gain SPSA method is proposed and investigated in combination with devices borrowed from the MCMC literature. The main contribution of the paper is the development and testing a number of devices that may eventually improve the convergence properties of the algorithm, such as various truncation techniques, averaging and choices of acceptance probabilities. The basis for comparison of performances is accuracy vs. number of function evaluations. We present experimental evidence for the superiority of an SC method allowing moves in wrong directions with small probability, where the underlying method is an SPSA method using averaging and adaptive truncation.**

I. Introduction

The basic problem to be considered in this paper is the minimization of a real valued "convex function" $L(\theta)$ over $\mathbf{Z}^p$, achieving its minimum in $\theta^* \epsilon \mathbf{Z}^p$, subject to the condition that only noise corrupted values of L are available: for any fixed θ at any time n we measure

$$Y_n(\theta) = L(\theta) + v_n,$$

where v_n is a state-independent noise. A *benchmark* problem that we will consider in detail is the quadratic problem when

$$L(\theta) = \frac{1}{2}(\theta - x^*)^T A(\theta - x^*) \tag{1}$$

where $x^* \epsilon \mathbb{R}^p$ is such that not all of its coordinates are integers, and A is a symmetric positive definite matrix. The notion of convexity for functions defined over $\mathbf{Z}^p$ will be discussed in Section 2. We would like to emphasize that this is a hard problem due to the restriction $\theta \epsilon \mathbf{Z}^p$.

An auxiliary problem that is closely related to our basic problem is the following: construct an ergodic Markov-chain with an invariant distribution π such that π attains its minimum at θ^*.

An important extension of the above problem is the following basic problem of resource allocation (see [2]): minimize a separable quadratic function

$$L(\theta) = \sum_{i=1}^{n} L_i(\theta_i), \qquad \theta_i \epsilon \mathbb{R}^m$$

subject to multiple capacity constraint

$$\sum_{i=1}^{n} \theta_i = K, \qquad K \epsilon \mathbb{R}^m$$

under the conditions that for any fixed θ_i and time n we can measure $L_i(\theta_i)$ with noise.

In discrete optimization it is usually assumed that noise-free values of L are available (see eg. [5]) In the noise free case a natural approach is to use Markov Chain Monte Carlo (MCMC) method, such as Metropolis' method [12], or Hasting's method [11]. Assume, that L has a unique minimum θ^* over D. To minimize L over D define a probability distribution

$$\pi_i = \frac{e^{-cL(i)}}{\sum_{j \epsilon D} e^{-cL(j)}}, \qquad c > 0.$$

Then obviously $\pi(\theta^*) > \pi(\theta)$ for all $\theta \neq \theta^*$ and for large c the probability assigned to θ^* will be close to 1.

In both the Metropolis and the Hastings method we construct a Markov-chain with invariant probability π. In both cases an irreducible and aperiodic Markov chain is generated with the property that its invariant probability distribution is maximal at the optimum-point, and in both cases an unnormalized invariant distribution is a priori given. To get this Markov-chain we start with an initial Markov chain, which will be called the q-chain, that is easy to realize and modify its dynamics so that each move is accepted with a probability that depends on consecutive values of the objective function.

For the noisy case a similar method has been developed in [9] using very different technical arguments. This so-called stochastic comparison method (SC) has been further developed by Andradottir, [1]. In both papers the initial chain is special, its transition probabilities are defined as $q_{ij} = 1/(N-1)$, for $i \neq j$, where N is the cardinality of the state-space. An extension of the SC method to a minimization problem on $\mathbf{Z}_+^1$ has been given in [1]. Finally, in [9] a SC method is developed with a general initial q-chain using repeated comparisons for every potential move. In stochastic comparison the invariant distribution of the resulting Markov-chain is not given a priori, but it is known to be maximal at optimum θ^*.

L. Gerencsér is with MTA SZTAKI, 13-17 Kende u., Budapest, 1111, Hungary gerencser@sztaki.hu

S.D. Hill is with Applied Physics Lab. John Hopkins Univ., John Hopkins Rd., Laurel, MD 20723-6099, HillSD1@central.SSD.JHUAPL.edu

Zs. Vágó is with MTA SZTAKI, 13-17 Kende u., Budapest, 1111, Hungary, and also with Pázmány Péter Catholic University, Budapest, Hungary. vago@itk.ppke.hu

Z. Vincze is with Hungarian Gaming Corporation Directorate for Game Organisation vincze.zoltan@szerencsejatek.hu

Taking advantage of the structure of the set $\mathbf{Z}^p$ we propose to use a q-chain that imitates the move of a stochastic gradient or stochastic Newton-method. Since the gradient of L is not available, and L itself is evaluated with noise we propose to use a discrete version of the SPSA method (simultaneous perturbation stochastic approximation). For minimization problems over $\mathbb{R}^p$ this has been developed in [13] and [14], see also [4] and [6]. A discrete SPSA method has been first presented in [8]. We use a truncated fixed gain SPSA method and the stability of the method is ensured by a resetting mechanism.

The main *contribution* of the paper is the development and testing a number of devices that may eventually improve the convergence properties of the algorithm. These devices include: various truncation techniques, averaging, and choices of acceptance probabilities. The basis for comparison of performances is accuracy vs. number of function evaluations. These results will be summarized in Section 5.

Convexity over $\mathbf{Z}^p$ is an important practical issue to ensure that a local minimum of L is in fact a global minimum (see [5]. We cite only the following results:

Proposition 1: Let f be an integrally convex function on a discrete rectangle X. If x is a local minimum point for f over X, then x is a global minimum point.

Proposition 2: Let

$$f(x) = x^T C x + d^T x.$$

Assume, that C is symmetric, positive definite and diagonally dominant, i.e.$\sum_{j=1,j\neq i}^{n} |c_{ij}| \leq |c_{ii}|$, $i = 1, ..., n$.. Then f is integrally convex.

II. Markov Chain Monte Carlo methods

Let $L(\theta)$ be a function defined over some abstract, finite or countable set D, the elements of which are sometimes identified with natural numbers $1, 2, \ldots$. Assume, that L has a unique minimum θ^* over D. To minimize L over D define a probability distribution

$$\pi_i = \frac{e^{-cL(i)}}{\sum_{j\epsilon D} e^{-cL(j)}}, \qquad c > 0.$$

Then obviously $\pi(\theta^*) > \pi(\theta)$ for all $\theta \neq \theta^*$ and for large c the probability assigned to θ^* will be close to 1.

In both the Metropolis and the Hastings method we construct a Markov-chain with invariant probability π. Both methods rely on the following observation the proof of which is straightforward (see [10]):

Lemma 1: Let $P = (p_{ij})$ be a transition matrix of a Markov-chain over a finite or countable state space Θ and let $\pi = (\pi_i)$ be a probability distribution written as a row vector over Θ. Then

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i \neq j \tag{2}$$

implies that π is invariant for P, i.e. $\pi P = \pi$.

In both methods a Markov-chain satisfying (2), called a p-chain, will be constructed so that we start with an initial Markov-chain with transition matrix $Q = (q_{ij})$, called the q-chain, and its dynamics is modified by accepting a move from i to j, $i \neq j$with a pre-computed acceptance probability τ_{ij}. The p-chain will thus be defined by

$$p_{ij} = q_{ij}\tau_i j.$$

In the Metropolis-method the q-chain is symmetric, i.e. $q_{ij} = q_{ji}, \qquad \forall i, j$. In the more general Hastings method the q-chain is non-symmetric, and the acceptance probabilities τ_{ij} are given by

$$\tau_{ij} = \min\left(\frac{\pi_j q_{ji}}{\pi_i q_{ij}}, 1\right),$$

if $q_{ij} \neq 0$ and $\tau_{ij} = 0$ otherwise.

The advantage of the Metropolis method over the Hastings method is that the explicit knowledge of the transition probabilities q_{ij} is not required. On the other hand the restriction that (q_{ij}) must be symmetric is quite severe.

Rate of convergence. If the p-chain is irreducible and aperiodic the n-step transition kernel $P^n(x, A)$ converges to the invariant measure π in total variation. The rate of convergence depends on Q and the acceptance probabilities τ. The estimation of the rate of convergence is not trivial, but it can be estimated in the case of a so-called independent q-chain, i.e. when $q_{ij} = q_j$ for all i, j. The following result is given in [16].

Theorem 1: Let $q_{ij} = q_j$ for all i, j and assume that $q_j \geq \beta\pi_j$ for all j, with $1 > \beta > 0$. Then for all A

$$|P^n(x, A) - \pi(A)| \leq (1 - \beta)^n.$$

Note that the acceptance probability for an independent q-chain is defined by

$$\tau_{ij} = \min\left(\frac{\pi_j q_i}{\pi_i q_j}, 1\right).$$

The limitations of the Metropolis-Hastings method is that the q-chain must be either symmetric or explicitly known. Now assume a q-chain is constructed by exploiting the special structure of the problem, such as a discrete SPSA method to be described in Section 4. Then Q is neither symmetric nor explicitly known. Therefore it can not be used as the initial q-chain for a Metropolis-Hastings method.

Since we have a large degree of freedom in choosing τ_{ij} by using different strictly monotone transformations of L we may try to find an appropriate τ_{ij} directly. A simple heuristic leads to the following choice:

$$\tau_{ij} = \begin{cases} \varepsilon > 0 & \text{if } \ L(j) > L(i) \\ 1 & \text{otherwise} \end{cases},$$

where $\varepsilon > 0$ is fixed.

Conjecture. For small ε the unique invariant probability of the p-chain with $p_{ij} = q_{ij}\tau_{ij}$ is maximal at θ^*.

The stochastic comparison method. The Metropolis-Hastings method is not applicable when the function values are evaluated with noise. In this case we follow a different path: our objective now is to construct an irreducible, aperiodic Markov-chain such that its invariant probability

distribution is maximal at θ^*. This objective is achieved by the stochastic comparison method (SC) originally proposed by Andradottir [1]. In contrast to the Metropolis-Hastings method, in the original formulation of the SC method the q-chain is not arbitrary. Let $\Theta = \{1, \dots, N\}$ be the state space. In our problem (1) we simply enumerate the gridpoint of a bounded domain $D \subset \mathbf{Z}^p$. Let $\theta^* = 1$ be the minimizing point. The q-chain is defined by

$$q_{ij} = \frac{1}{(N-1)} \qquad \text{for } i \neq j.$$

If a move from i to j is generated by the q-chain at time n then it is accepted if

$$y_{n+1}(j) < y_n(i).$$

Thus the acceptance probability is

$$\tau_{ij} = P(y_n(i) - y_{n+1}(j) > 0) = P(v_n - v_{n+1} > L(j) - L(i)) \tag{3}$$

It is easy to see that the p-chain with $p_{ij} = q_{ij}\tau_{ij}$ satisfies the following criterion:

Condition 1: The transition probability matrix P is strictly adapted to L in the following sense:

(i) for any $i, j \neq 1$ $p_{i1} > p_{ij}$
(ii) for any $i, j \neq 1$ $p_{1j} \leq p_{ij}$
(iii) for any $i \neq j$ $p_{i1} \geq p_{1i}$

The following proposition is essentially given in [1]:

Proposition 3: Let P be the transition matrix of an irreducible, aperiodic Markov-chain and let P be strictly adapted to L. Then $\pi_1 > \pi_i$ for all $i \neq 1$.

The realization of the initial q-chain with $q_{ij} = 1/(N-1)$ for $i \neq j$ may be impractical for large N-s. If the state-space has some neighborhood structure then it would be desirable to work with a Q-matrix allowing only local connections.

Definition 1: A q-chain is adapted to L if $L(i) \leq L(j)$ implies $q_{ij} \leq q_{ji}$.

Taking a symmetric Q we have the following result:

Proposition 4: Let Q be symmetric. Then the p-chain generated by the SC method is adapted to L.

The proof is trivial. Unfortunately, it does not in general follow that the invariant distribution for P is maximal at θ^*.

A straightforward extension of the SC method to general Q-s can be obtained using repeated comparisons, thus ensuring that the acceptance probabilities τ_{ij} will be close to 0 for $L(j) > L(i)$ and close to 1 for $L(j) < L(i)$. Then assuming the validity of our we would conclude that the invariant probability of the p-chain will be maximal at θ^*. However a large number of small entries in the p-matrix imply slow rate of convergence.

In the following definition we impose a much stronger condition on the invariant distribution, than being maximal at θ^*.

Definition 2: A probability distribution π on Θ is adapted to L if $L(i) > L(j)$ implies $\pi_i < \pi_j$.

Lemma 2: Let P be the transition probability of an irreducible, aperiodic Markov-chain and let $\overline{P}$ denote the transition probability of the reverse chain. Assume that both P and $\overline{P}$ are adapted to L. Then the invariant probability of P, denoted by π, is adapted to L.

Proof: We have for all i, j

$$\pi_i \overline{P}_{ij} = \pi_j P_{ji} \quad \text{and} \quad \pi_j \overline{P}_{ji} = \pi_i P_{ij}.$$

From here

$$\overline{P}_{ij} = \frac{\pi_j}{\pi_i} p_{ji} \quad \text{and} \quad \overline{P}_{ji} = \frac{\pi_i}{\pi_j} p_{ij}.$$

Let $i < j$ and assume that $\pi_j > \pi_i$. Then $P_{ji} > P_{ij}$ implies $\overline{P}_{ij} > \overline{P}_{ji}$, a contradiction.

III. SPSA OVER $\mathbf{Z}^p$

To approximate the gradient of L we use simultaneous random perturbations (see [13]). The advantage of SPSA is that only two function values are needed to get an unbiased estimator of the gradient, irrespective of the dimension. Let θ be the current approximation of θ^* and let k be the iteration time. We take a random vector

$$\Delta = \Delta_k = (\Delta_{k1}, ..., \Delta_{kp})^T,$$

where Δ_{ki} is a double sequence of i.i.d. random variables. A standard choice is to take a Bernoulli-sequence, taking values $+1$ or -1 with equal probability $1/2$. We take two measurements at $L(\theta + \Delta)$ and $L(\theta - \Delta)$. Note that the perturbation is *not scaled*, since we have to stay on the grid. In fact, for quadratic functions scaling is not needed: its role is to keep down the effect of the third error terms in the Taylor series expansion of L. Let $L_\theta(theta)$ denote the gradient of L at θ. Thus we get the following direction of move:

$$H(\theta) = \Delta^{-1}\Delta^T L_\theta(\theta),$$

where $\Delta^{-1} = (\Delta_1^{-1}, ..., \Delta_p^{-1})^T$. Now we are ready to define a stochastic gradient step, which is a fixed gain version of the simultaneous perturbation stochastic approximation or SPSA method constrained to $\mathbf{Z}^p$:

$$\theta_{n+1} = \theta_n - [aH(n, \theta_n)],$$

where [ ] is the integer part coordinatewise, we also write $[\theta] = \mathrm{round}(\theta)$. In general we use a mapping

$$\pi_{\mathbf{Z}}(\theta) : \mathbb{R}^p \to \mathbf{Z}^p$$

and use the correction term $\pi_{\mathbf{Z}}(aH_n(\theta_n)))$. An alternative correction term is obtained if by using an extension of the sign-function. For any real number y define

$$\mathrm{sgn}(y) = \begin{cases} 1 & \text{if } y \geq 1/2 \\ 0 & \text{if } |y| < 1/2, \\ -1 & \text{if } y \leq -1/2 \end{cases}$$

and for any p-vector x set $\mathrm{sgn}(x) = (\mathrm{sgn}(x_1), ..., \mathrm{sgn}(x_p))$. Then set

$$\pi_{\mathbf{Z}}(\theta) = \mathrm{sgn}(\theta).$$

Thus we get the signed, fixed gain SPSA-method. Note that for sufficiently small gains a the correction becomes 0. The

purpose of letting $\pi_{\mathbb{Z}}(y) = 0$ for $|y| < 1/2$ is to allow small moves with lots of 0 coordinates.

A compromise between the above two procedures is obtained by defining

$$\pi_{\mathbb{Z}}(x) = \text{round}(h\frac{x}{||x||_\infty}) = \text{sig}_h(x) \quad h > 0$$

which allows more variability in the length of the truncated vector, but its L_∞-norm is bounded by a pre-fixed h.

Lemma 3: For strictly convex quadratic functions L the Markov-chain generated by a fixed gain truncated SPSA method using sig_h satisfies

$$p_{ij} \leq p_{ji} \qquad \text{for} \quad L(j) \geq L(i),$$

i.e. it is adapted to L.

A simple and useful variant of the fixed gain, truncated SPSA method is obtained if our starting point is a second order SPSA method developed in [15]. It has been shown there that using three function values per iteration a consistent estimator of the Hessian matrix of L can be obtained. This estimate can be used in a Newton-like procedure.

Following Spall [14] we estimate the Hessian of L as follows: assume that the gradient of L, denoted by g is available, then we can apply the SP procedure for each component of g and we would finally end up with the estimate

$$\widehat{A} = \frac{1}{2c}\Delta^{-1}\delta g^T$$

where $\delta g = g(\theta + c\Delta) - g(\theta - c\Delta)$. Now since g is not available we have to resort to its estimate. Using *one-sided* estimate of g we have for any θ' the estimate

$$\hat{g}(\theta') = \tilde{\Delta}^{-1}[L(\theta' + c\tilde{\Delta}) - L(\theta')]/c \ .$$

Substitute this in place of g. To ensure that our estimate of H is symmetric, we symmetrize it. Thus the final estimate will become

$$\hat{A} = \frac{1}{2}[\frac{1}{2c}\Delta^{-1}\delta\hat{g}^T + \frac{1}{2c}\delta\hat{g}\Delta^{-T}] \ .$$

Obviously $E\hat{H} = A$. Consider the special case when $\Delta = e_i = (0, \ldots, 1, 0, \ldots, 0)^T$ with 1 at the i-th position, $\tilde{\Delta} = e_j$ with i, j random, uniform over $\{1, \ldots, p\}$. Such a selection of Δ will be called a random coordinate vector (RCV) Then we get for the symmetrized $\hat{A}$

$$\hat{A} = \frac{1}{2}[A_{ij} + A_{ji}]$$

where A_{ij} is the matrix with 0-s in all positions except in i, j, where the entry is (a_{ij}). The argument indicates a potential shortcoming of using $\Delta = e_i$, namely in higher dimensions the matrix $\hat{H}$ will fill up very slowly. E.g. for $p = 200$ we need cca. 20000 experiments to make sure that all entries will be calculated.

For exploring the properties of fixed gain second order SPSA method we actually pretend that A and thus A^{-1} are known and define

$$H(n, \theta) = A^{-1}\Delta^{-1}\Delta^T L_\theta(\theta).$$

A second order SPSA method. Consider the following fixed gain version of a second order simultaneous perturbation stochastic approximation (SPSA) method (see also [7]):

$$\theta_{n+1} = \theta_n - [aH_n(\theta_n)] \tag{4}$$

where $a > 0$ is the gain, and

$$H_n(\theta) = A^{-1}\Delta_n^{-1}\Delta_n^T A(\theta - \theta^*). \tag{5}$$

IV. Resource allocation

Our interest in SPSA on grids is motivated by multiple discrete resource allocation problems, which we shortly describe. of discrete resource allocation is to distribute a finite amount of resources of different types to finitely many classes of users, where the amount of resources that can be allocated to any user class is discrete. Suppose there are n types of resources, and that the number of resources of type i is N_i. Resources of the same type are identical. The resources are allocated over M user classes: the number of resources of type i that are allocated to user class j is denoted by θ_{ij}. The matrix consisting of the θ_{ij}'s is denoted by Θ.

For each allocation the cost, such as performance or reliability is associated, which is denoted by $L(\Theta)$. We assume that the total cost is weakly separable in the following sense:

$$L(\Theta) = \sum_{j=1}^{M} L_j(\theta_j)$$

where $L_j(\theta_j)$ is the individual cost incurred by class j, $\theta_j = (\theta_{ij}, ..., \theta_{ij})$, i.e. the class j cost depends only on the resources that are allocated to class j. An important feature of resource allocation problems is that often the cost L_j is not given explicitly, but rather in the form of an expectation or in practical terms by simulation results.

Then the discrete, multiple constrained resource allocation problem is to minimize $L(\Theta)$ subject to

$$\sum_{j=1}^{M} \theta_{ij} = N_i, \ \theta_{ij} \geq 0, \ 1 \leq i \leq n. \tag{6}$$

where the θ_{ij}'s are non-negative integers. We will assume that a solution exists with strictly positive components. Then the minimization problem is unconstrained on the linear manifold defined by the balance equations.

Problem (6) includes many problems of practical interest such as the scheduling of time slots for the transmission of messages over nodes in a radio network (cf.[3]). The above problem is a generalization of the single resource allocation problem with $m = 1$, see [2].

Cassandras et al. in [2] present a relaxation-type algorithm for the single resource, in which at any time the allocation is rebalanced between exactly two tasks. For multiple resource rebalancing between user j and k is based on comparing the gradients $\frac{\partial}{\partial\theta_j}L_j(\theta_j)$ and $\frac{\partial}{\partial\theta_k}L_j(\theta_k)$. The computation of these gradients may require time-consuming

simulation, hence we use their estimates obtained by simultaneous perturbation at time t, and denoted by $H_j(t, \theta_j)$. Thus we arrive to the following recursion: at time t select a pair (j, k) and then modify the allocation for this pair of tasks as follows:

$$\begin{aligned} \theta_{j,t+1} &= \theta_{j,t} + [a(H_k(t, \theta_k) - H_j(t, \theta_j))] \\ \theta_{k,t+1} &= \theta_{k,t} - [a(H_k(t, \theta_k) - H_j(t, \theta_j))], \end{aligned}$$

where a is a fixed gain. Obviously, the balance equations are not violated by the new allocation. The selection of the pair (j, k) can be done by a simple cyclic visiting schedule.

V. Simulation results

Simulations have been carried our for randomly generated quadratic function

$$L(\theta) = \frac{1}{2}(\theta - \theta^*)^T A(\theta - \theta^*) \tag{7}$$

in dimensions $p = 50, 100$. The eigenvalues of the symmetric matrix A have been generated by sampling from an exponential distribution with parameter $\mu = 0.5$, and shifting them by a positive amount $\delta = 0.5$. Then the resulting diagonal matrix has been multiplied by p randomly chosen rotations, where the pairs of indices defining the plane of the rotation and the angle of rotation have been both chosen uniformly. The coordinates of the continuous optimum θ^* have been chosen randomly according to a uniform distribution in $[0, 1]$.

The initial point for the iteration is chosen by taking the integer part of a sample from a standard normal distribution $\mathcal{N}(0, I_p)$. In all cases we add a *resetting* mechanism, which in some cases - depending on the value of the rejection probability - is dramatically active, to ensure stability of the procedure. The exact value of the discrete optimum is unknown, but a rough upper bound is obtained by evaluating L at $\theta^* = 0$. We carry out our experiments on two randomly generated test problems with dimensions 50 and 100, respectively, with various noise levels. The noise level is determined by fixing a signal to noise ratio, say 2, and then setting the variance of the noise, denoted by Σ, equal to 2 of the current function value. The simulation experiments are of increasing complexity and in each step we vary one aspect of the algorithm. All the plots show (noise free) function values against the number of function evaluations. Functions values are plotted on a logarithmic scale.

On Fig.1 the plain finite difference stochastic approximation (FDSA), and two SPSA methods are compared, with two different truncation functions, sig_1 and sig_3. On Fig.2 the same experiment is repeated but with higher degree of averaging: $q = 5$ as opposed to $q = 2$. We see that SPSA is superior to FDSA in both cases, and averaging improves the stability of the procedure. (Note the differences in the vertical scale).

It is seen that in the choice of the truncation method there is a trade-off: for large h we get faster initial descent and for small h we get better stability. This motivates the introduction of an adaptive procedure: fix $h_1 < h_2$ and perform a step with both, then choose one of the two resulting positions with smaller measured (noisy) function value. The performance of an adaptive SPSA is compared with two non-adaptive procedure on Fig.3.

A further improvement in stability can be achieved by blocking moves in wrong directions following techniques from MCMC. For this we use a fixed acceptance probability τ: if $L(\theta_{n+1}) > L(\theta_n)$ then we accept θ_{n+1} with probability τ. We found that choosing τ very small slows down the algorithm considerably. A non-blocked procedure is compared with two blocked procedures on Fig.4. It is seen that $\tau = 0.04$ considerably improves the performance of the algorithm. Note that the number of function value evaluations now include the number of comparisons before an actual move is made. For $\tau = 0.3$ 40% of the moves were blocked, while for $\tau = 0.04$ the corresponding figure is 73%.

For theoretical analysis a potential tool is a discrete version of the ODE method. The quality of the method can be partially assessed by computing the expected value of the correction term H (whatever H means in different procedures) and see if its angle with the gradient of L is less than 90 degrees. It was found that for the best procedure shown on Fig.4 this angle less than 5 degrees.

In most experiment Bernoulli SPSA and random coordinate vectors (RCV) SPSA with $\Delta = \sqrt{p}e_i$ yield the same performance. This is not the case for an interesting application: resource allocation. We considered a problem with 50 users and 10 types of resources, and found that RCV SPSA is superior to Bernoulli SPSA. The results are given on Fig.5.

Acknowledgement

The first author expresses his thanks to James C. Spall of the Applied Physics Laboratory of Johns Hopkins University for cooperating in this research.

References

[1] S. Andradottir. A method for discrete stochastic optimization. *Management Science*, 41:1946–1961, 1995.

[2] C.G. Cassandras, L. Dai, and C.G. Panayiotou. Ordinal optimization for a class of deterministic and stochastic discrete resource allocation problems. *IEEE Trans. Auto. Contr.*, 43(7):881–900, 1998.

[3] C.G. Cassandras and V. Julka. Scheduling policies using marked/phantom slot algorithms. *Queueing Systems: Theory and Appl.*, 20:207–254, 1995.

[4] H.F. Chen, T.E. Duncan, and B. Pasik-Duncan. A stochastic approximation algorithm with random differences. In J.Gertler, J.B. Cruz, and M. Peshkin, editors, *Proceedings of the 13th Triennal IFAC World Congress, San Francisco, USA*, pages 493–496, 1996. Volume editors: R.Bitmead, J. Petersen, H.F. Chen and G. Picci.

[5] P. Favati and F. Tardella. Convexity in nonlinear integer programming. (53), 1990.

[6] L. Gerencsér. Rate of convergence of moments for a simultaneuous perturbation stochastic approximation method for function minimization. *IEEE Trans. Automat. Contr.*, 44:894–906, 1999.

[7] L. Gerencsér, S. D. Hill, and Zs. Vágó. Optimization over discrete sets via SPSA. In *Proceedings of the 38-th Conference on Decision and Control, CDC'99*, pages 1791–1794. IEEE, 1999.

[8] L. Gerencsér, S. D. Hill, and Zs. Vágó. Discrete optimization via spsa. In *Proceedings of the American Control Conference, ACC'01*, pages 1503–1504. IEEE, 2001.
[9] W.B. Gong, Y.C. Ho, and W. Zhai. Stochastic comparison algorithm for discrete optimization with estimation. *SIAM J. Optimization*, 10(2):384–404.
[10] J. M. Hammersley and D.C. Handscomb. *Monte Carlo methods*. Fletcher & Son Ltd. Norwich, 1967.
[11] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 1970.
[12] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem Phys.*, 21:1087–1091, 1970.
[13] J.C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Automat. Contr.*, 37:332–341, 1992.
[14] J.C. Spall. Adaptive stochastic approximation by the simultaneous perturbation method. In *Proceedings of the 1998 IEEE CDC*, pages 3872 – 3879, 1998.
[15] J.C. Spall. Adaptive stochastic approximation by the simultaneous perturbation method. *IEEE Trans. Automat. Contr.*, 45:1839–1853, 2000.
[16] R. L. Tweede. Markov chains: Structure and applications. Technical report, Colorado State University, October 1998.

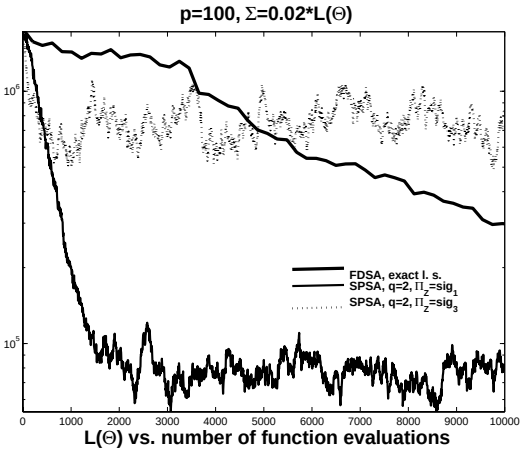

Fig. 1.

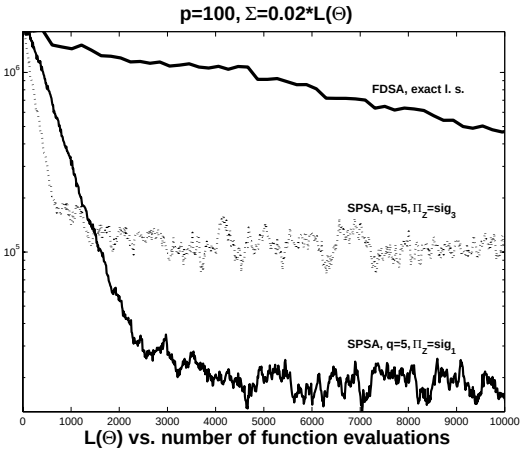

Fig. 2.

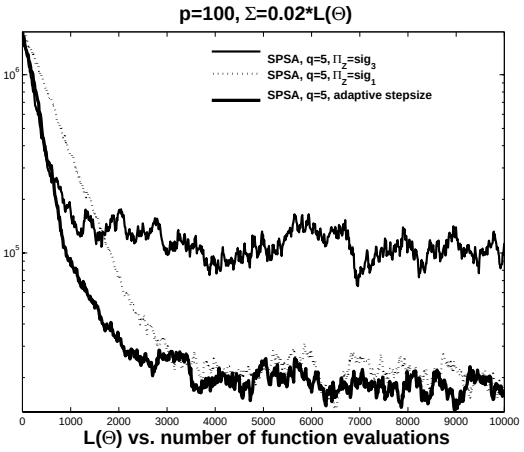

Fig. 3.

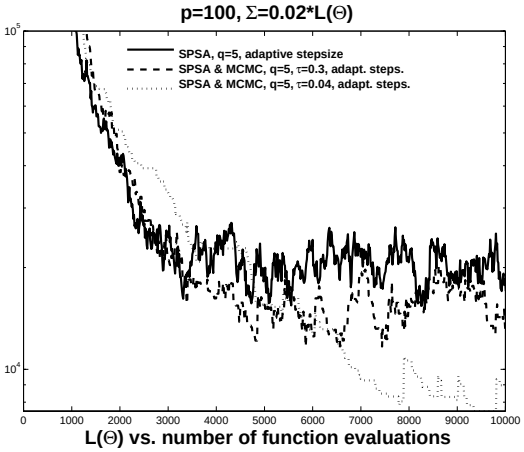

Fig. 4.

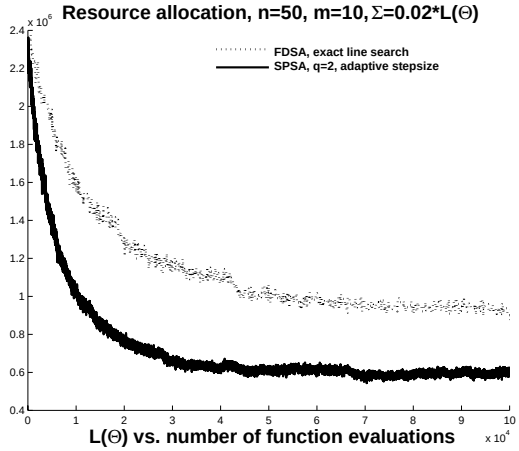

Fig. 5.