Random-Direction Optimization Algorithms with Applications to Threshold Controls[1]

G. Yin,[2] Q. Zhang,[3] H. M. Yan,[4] and E. K. Boukas[5]

Communicated by C. T. Leondes

Abstract. This work develops a class of stochastic optimization algorithms. It aims to provide numerical procedures for solving threshold-type optimal control problems. The main motivation stems from applications involving optimal or suboptimal hedging policies, for example, production planning of manufacturing systems including random demand and stochastic machine capacity. The proposed algorithm is a constrained stochastic approximation procedure that uses random-direction finite-difference gradient estimates. Under fairly general conditions, the convergence of the algorithm is established and the rate of convergence is also derived. A numerical example is reported to demonstrate the performance of the algorithm.

Key Words. Gradient estimates, random directions, stochastic approximations, manufacturing.

1. Introduction

This work develops a class of stochastic optimization algorithms. Our primary motivation stems from systems involving optimal or suboptimal

hedging policies such as marketing-production manufacturing systems in which a firm aims to maximize its overall profit or to minimize its overall cost through adequate choice of the production rates. The underlying problem is to find optimal controls of models involving Markovian jump processes. By focusing on threshold-type controls, the optimal control problem is converted to an optimization problem that requires finding the optimal threshold values. To solve such problems and other problems that can be characterized by threshold-type controls, we develop in this paper feasible and easily implementable numerical algorithms. Based on recent progress in stochastic approximation, we suggest a class of algorithms using random-direction finite-difference gradient estimates, investigate the convergence, derive estimation error bounds, and obtain rates of convergence.

Before proceeding further, first we review some of the recent developments in production planning of manufacturing. Owing to the rapid progress in management science and computing technology, the research in production planning has received much needed attention; see Ref. 1 and the references therein. In Ref. 2, Akella and Kumar formulated a one-machine, one-part production system as a stochastic optimal control problem, in which the part demand is assumed to be a constant, the state of the machine is a two-state continuous-time Markov chain, and the objective function is a discounted inventory/shortage cost over an infinite-time horizon. It was shown that the optimal control is given by a single threshold inventory level. Bielecki and Kumar treated the long-run average-cost counterpart in Ref. 3; an optimal hedging policy was obtained. In these papers, the objective is to minimize an expected cost over a class of production control variables. For a discussion of the hedging policies or threshold-type controls, see Gershwin (Ref. 4) among others. In Ref. 5, a minimax production planning model with Markov capacity process and deterministic but unknown demand rate was considered by Boukas, Yang, and Zhang; an optimal control under discounted cost criterion was obtained. In lieu of concentrating on infinite-horizon problems with discounted cost function, optimal policies for finite-horizon problems were obtained in the paper of Zhang and Yin (Ref. 6); here, the corresponding optimal controls were obtained in terms of the time-dependent turnpike sets under traceability conditions. To be able to treat large-dimensional manufacturing systems, the framework of hierarchical decomposition and appropriate aggregation was dealt with in Sethi and Zhang (Ref. 7). The related singularly perturbed Markovian systems were studied in depth in the papers of Yin and Zhang (Ref. 8) and references therein. In a recent paper (Ref. 9), Zhang, Yin, and Boukas proposed a class of marketing-production models. The underlying problem is a stochastic control problem with the demand being a Poisson process. Assuming that the system consists of a single machine and that the machine

is reliable, analytic solutions were provided together with illuminating descriptions on various control regions. These results yield managerial insight for many applications. Certainly, it is helpful to model the machine capacity also by a continuous-time Markov chain. Unfortunately, under such a setup, the problem becomes extremely difficult; one cannot expect any closed-form solutions. In addition, for many systems in applications, it is important to treat a many-machine system in lieu of a single-machine system. To resolve the problem, we seek numerical solutions and efficient algorithms. One of the possibilities is to solve the optimal control problems numerically. Nevertheless, it is well known that such an approach may suffer from the curse of dimensionality for a large-scale system. This leads us to seek alternative approaches via stochastic optimization.

In Yan, Yin, and Lou (Ref. 10), we proposed a combined approach of stochastic approximation with gradient estimation techniques. In that paper, we concentrated on threshold-type controls, and converted the optimal control problem to an optimization problem, in which the gradient estimator was constructed with the help of infinitesimal perturbation analysis [IPA, see Ho and Cao (Ref. 11) and references therein]; promising results were obtained. For related work, see also Liberopoulos and Caramanis (Ref. 12) and Yan, Zhou, and Yin (Ref. 13).

In view of these developments, by focusing on threshold-type controls, we are concerned with alternative procedures in this paper. We use also a combined approach of stochastic approximation and gradient estimates. Nevertheless, in lieu of the IPA approach, we use random-direction finite-difference methods; see Refs. 14–16. Such an idea is quite efficient and is convenient especially for large-dimensional problems. Using such an approach, the computation is rather systematic, and one need not figure out the form of the gradient estimates as in the IPA approach.

One of the features of this paper is that we do not even require the capacity and demand to be Markovian, since we are not using a dynamic programming approach. We assume only that the processes are stationary satisfying appropriate moment conditions. From an application point of view, it is appealing particularly without the Markovian assumption. Not only the algorithm suggested is applicable to the manufacturing systems, but is useful also for many other optimal control problems requiring finding the optimal threshold values.

To proceed, the rest of the paper is arranged as follows. Section 2 gives the setup of the problem together with the proposed algorithm. Section 3 begins with a modified algorithm using projection, states the conditions needed, establishes the convergence of the recursive algorithm, and ascertains the convergence rate via a local analysis. Section 4 presents an example

to illustrate the performance of the algorithm. Finally, Section 5 concludes the paper with additional remarks.

2. Formulation

We formulate the problem in a general setting. The main idea is to convert an optimal control problem to an optimization problem by focusing our attention to a particular class of controls, namely, controls of the threshold type. Then, the underlying problem becomes to find the optimal threshold values.

By a threshold-type control in the control space Γ, we mean a policy $u(t) \in \Gamma \subset \mathbb{R}^l$ with constant threshold levels such that, for $A_i \subset \mathbb{R}^l$ and constants $c_i \in \Gamma \subset \mathbb{R}^l$, $i = 1, \ldots, l_1$, such that, for some integer j,

$$u(t) = \sum_{i=1}^{j} c_i I_{\{x(t) \in A_i\}},$$

where $x(t)$ represents the state.

To begin, let $x(t) \in \mathbb{R}^r$ be the state variable, let $u(t) \in \Gamma_u \subset \mathbb{R}^{r_1}$ be the control, where Γ_u is a compact set. Suppose that $\alpha(\cdot)$ and $\beta(\cdot)$ are stationary random processes taking values in compact sets $\mathcal{M}_\alpha$ and $\mathcal{M}_\beta$. For example, $\alpha(\cdot)$ and $\beta(\cdot)$ may be finite-state Markov chains with stationary transition probabilities, generators Q_α and Q_β, and state spaces $\mathcal{M}_\alpha$ and $\mathcal{M}_\beta$. Nevertheless, since we do not use the HJB (Hamilton–Jacobi–Bellman) equations, the explicit Markovian assumption is not needed. For appropriate functions $f(\cdot)$ and $G(\cdot)$ with

$$f: \mathbb{R}^r \times \Gamma_u \times \mathcal{M}_\alpha \times \mathcal{M}_\beta \mapsto \mathbb{R}^r,$$

$$G: \mathbb{R}^r \times \Gamma_u \times \mathcal{M}_\alpha \times \mathcal{M}_\beta \mapsto \mathbb{R},$$

the general problem setup is as follows: Choose u so as to maximize

$$J(u) = \lim_{\hat{T} \to \infty} (1/\hat{T}) E \int_0^{\hat{T}} G(x(t), u(t), \alpha(t), \beta(t)) \, dt, \tag{1a}$$

$$\text{s.t.} \quad \dot{x}(t) = f(x(t), u(t), \alpha(t), \beta(t)), \qquad x(0) = x. \tag{1b}$$

An interpretation of the above general setup in the manufacturing setting is now given: $x(t)$ is the inventory level of the manufacturing system, $u(t)$ represents the rate of production, $\alpha(\cdot)$ and $\beta(\cdot)$ model the machine capacity and the demand process, respectively. Our objective is to maximize the overall expected profit, or equivalently to minimize the overall expected cost.

In lieu of solving the optimal control problem directly, we formulate the problem as an optimization problem with the use of threshold-type controls so that the profit $J(\cdot)$ becomes a function of the threshold levels, i.e., $J = J(\theta)$. Threshold-type control policies have drawn renewed attention lately, since the idea is appealing and the principle is easy to implement. It is well known that such a control may lead only to suboptimality. Nevertheless, in applications, one is often content with suboptimality or near optimality, owing to the inherent uncertainty and other unknown disturbances. As was shown in Ref. 9, even for a seemingly not so complex problem, the solution can be rather involved. Frequently, a suboptimal or nearly-optimal control is as valuable as an optimal control, since within a vicinity of the optimal solution, any approximation is as good as the optimal one from a practical point of view.

Example 2.1. Let us consider a manufacturing firm facing a stochastic demand for its product over time. For $t \geq 0$, let $x(t), u(t), d(t)$ denote the inventory level, the production rate, and the demand rate. They evolve in accordance with the following dynamic system:

$$\dot{x}(t) = u(t) - d(t), \qquad x(0) = x. \tag{2}$$

Let

$$h(x) = c^+ x^+ + c^- x^- \tag{3}$$

be the inventory cost function, where c^+ and c^- are positive constants, and where

$$x^+ = \max\{0, x\}, \qquad x^- = \max\{0, -x\}.$$

Let $\pi(t) \in \{\pi_1, \pi_2\}$ denote the price for a unit product. In addition to the production rate $u(t)$, the price $\pi(t)$ is also a control variable, which in turn, depends on the state—the inventory level $x(t)$ and demand. Note that the demand depends typically on the price $\pi(t)$, since the higher the price, the lower the demand. Suppose that the demand rate is a two-state Markov chain with state space $\{d_1, d_2\}$ and generator

$$Q(\pi) = \begin{bmatrix} -\lambda(\pi) & \lambda(\pi) \\ \mu(\pi) & -\mu(\pi) \end{bmatrix},$$

where $\pi \in \{\pi_1, \pi_2\}$ is the price rate. Consider the cost function

$$J(x, d, u(\cdot), \pi(\cdot)) = \lim_{T \to \infty} (1/T) E \int_0^T [h(x(t)) - \pi(t)d(t)] \, dt. \tag{4}$$

The problem is to choose (u, π) so as to minimize J. Focusing on threshold-type controls, we convert this problem into an optimization problem. The threshold value $\theta = (\theta_1, \theta_2)$ is a 2-dimensional vector such that $\theta_1 \le \theta_2$. The first component θ_1 represents the lower limit of the inventory level that triggers a promotion action. The second component θ_2 controls the maximum inventory level, which is the same as the threshold control level considered in Ref. 3. In a production system controlled by a threshold control policy, the production discipline is regulated as follows:

if $x(t)$ is below θ_2, produce as much as one can;
if $x(t)$ is above θ_2, produce nothing;
if $x(t) = \theta_2$, produce exactly the same as the demand.

When the problem is converted to an optimization problem, the important task becomes to locate the optimal threshold levels. To proceed, we denote the threshold parameter by θ with $\theta \in \mathbb{R}^r$. Similar to Ref. 10, we consider a combined process

$$\xi(t) = (x(t), \alpha(t), \beta(t)).$$

In what follows, we refer to $\xi(t)$ as a noisy process, since it collects various random effects. It follows that $G(\cdot)$ can be written as a function of θ and the noise effect ξ. For each n, choose a random direction vector $\Delta_n = (\Delta_{n,1}, \ldots, \Delta_{n,r})'$, with independent components such that each component $\Delta_{n,i}$ is a Bernoulli random variable taking values ± 1, each with probability $1/2$. The stochastic optimization algorithm is of the form

$$\theta_{n+1} = \theta_n + \epsilon \Delta_n D\tilde{J}_n(\theta_n), \tag{5}$$

where

$$D\tilde{J}(\theta_n) = (1/2\delta T) \int_{nT}^{(n+1)T} [G(\theta_n + \delta\Delta_n, \xi^+(t)) - G(\theta_n - \delta\Delta_n, \xi^-(t))]\, dt, \tag{6}$$

where δ is the finite-difference stepsize and $\{\xi^+(t)\}$, $[\xi^-(t)]$ are observation noises. Note that the random-direction vectors are chosen as Bernoulli sequences following the suggestion of Ref. 15; see also Refs. 14, 16. The choice of the random-direction vector implies that

$$|\Delta_{n,i}|^2 = 1 \quad \text{and} \quad |\Delta_n|^2 = r.$$

We use the random-direction finite-difference estimates (with the random directions vector Δ_n)

$$(1/2\delta T)\begin{bmatrix} \int_{nT}^{(n+1)T} [G(\theta_n + \delta\Delta_n, \xi^+(t)) - G(\theta_n - \delta\Delta_n, \xi^-(t))]\, dt/\Delta_{n,1} \\ \vdots \\ \int_{nT}^{(n+1)T} [G(\theta_n + \delta\Delta_n, \xi^+(t)) - G(\theta_n - \delta\Delta_n, \xi^-(t))]\, dt/\Delta_{n,r+1} \end{bmatrix}.$$

Then, multiplying the component i by $\Delta_{n,i}/\Delta_{n,i}$ yields the increments $\Delta_n D\check{J}_n(\theta_n)$ in (3).

By virtue of the idea of using common random numbers, which is well known in the simulation culture, we use henceforth $\xi^+(t) = \xi^-(t)$. The advantages of this are explained for example in Ref. 16, p. 14. For all the components of the gradient estimate, the same observations are used. Only the components of the random-direction vector differ. Throughout the paper, we assume that $\delta = \delta_\epsilon$, such that $\epsilon/\delta_\epsilon \to 0$ as $\epsilon \to 0$. That is, $\epsilon \ll \delta_\epsilon$ or δ_ϵ goes to zero much slower than the stepsize ϵ. Different from Ref. 10, we will not assume $T = T_\epsilon$ such that $T_\epsilon \to \infty$ as $\epsilon \to 0$. This will be helpful in the actual computation.

The stochastic approximation algorithm given above can be thought of as a mixed-mode scheme. It uses information from continuous dynamic systems [the inclusion of $\xi(t)$], but the recursion is done at discrete instance. Thus, to some extent, we may call it a hybrid procedure.

3. Asymptotic Properties of the Algorithm

This section is devoted to the asymptotic properties of the stochastic approximation algorithm. The main technique is the method of weak convergence. We work with a modified version of the algorithm: Select θ_0 to be an arbitrary nonrandom initial approximation and generate

$$\theta_{n+1} = \Pi_H(\theta_n + \epsilon\Delta_n D\check{J}_n(\theta_n)), \tag{7}$$

where Π_H is the projection onto the constrained set H. That is, $\Pi_H(\theta)$ is the closest point in H to θ; in particular, if $\theta \in H$, $\Pi_H(\theta) = \theta$.

As illustrated in Ref. 16, for the recursive approximation, "allowing unboundedness can lead to needless mathematical complication because some sort of stability must be shown or otherwise assumed, ..., and it generally adds little to the understanding of practical algorithms." Note

that we could have chosen the initial vector to be an ϵ-dependent random variable θ_0^ϵ and assume that θ_0^ϵ converges weakly to θ_0. However, for the applications we are dealing with, the nonrandom initial condition appears to be sufficient.

Following the development of Ref. 16 and the references therein, introduce a correction or reflection term z_n by writing (5) as

$$\theta_{n+1} = \theta_n + \epsilon\Delta_n D\tilde{J}_n(\theta_n) + \epsilon z_n, \tag{8}$$

where z_n is the vector having the shortest Euclidean length required to bring $\theta_n + \epsilon D\tilde{J}_n(\theta_n)$ back to the constraint set H.

3.1. Convergence of the Recursive Algorithm.   To obtain the convergence of the algorithm, we make the following assumptions:

(A1) $\{\Delta_n\}$ is a sequence of i.i.d. Bernoulli random variables on the unit cube in $\mathbb{R}^r$; i.e., each component $\Delta_{n,i}$ takes the values ± 1 with probability $1/2$. H is a compact set, and satisfies any one of the three conditions below:

   (i) H is a hyperrectangle; i.e., there are real numbers $a_i < b_i$, $i = 1, \ldots, r$, such that $H = \{x : a_i \leq x_i \leq b_i\}$.
   (ii)   Let $q_i(\cdot), i = 1, \ldots, p$, be continuously differentiable real-valued functions on $\mathbb{R}^r$, with gradients $q_{i,x}(\cdot)$. Without loss of generality, let $q_{i,x}(x) \neq 0$, if $q_i(x) = 0$. $H = \{x : q_i(x) \leq 0, i = 1, \ldots, p\} \neq \varnothing$ is connected and compact.
   (iii) H is an $\mathbb{R}^r$-dimensional connected compact surface with a continuously differentiable outer normal.

(A2)   The functions $G(\cdot)$ and $J(\cdot)$ satisfy the following three conditions:
   (i) $J(\cdot)$ and $\nabla J(\cdot)$ are continuous.
   (ii)   For each $\xi, G(\cdot, \xi)$ and $G_\theta(\cdot, \xi)$ are continuous and bounded on bounded θ sets, where $G_\theta(\cdot, \xi)$ denotes the partial derivative of $G(\cdot)$ with respect to θ.
   (iii)   For each bounded θ, $\{G(\theta, \xi(t))\}$ and $\{G_\theta(\theta, \xi(t))\}$ are uniformly integrable.

(A3) $\epsilon > 0, \delta = \delta_\epsilon$ such that $\epsilon/\delta_\epsilon^2 \to 0$ as $\epsilon \to 0$.
(A4)   The process $\{\xi(\cdot)\}$ is stationary. For each θ,
   $E|G(\theta, \xi(t))|^2 < \infty, \ EG(\theta, \xi(t)) - J(\theta) = 0$,

   for each n, $EG_\theta(\theta, \xi(t)) - \nabla J(\theta) = 0$,

and as $n \to \infty$,

$$(1/n) \sum_{j=m}^{n+m-1} E_m(1/T)$$

$$\times \int_{jT}^{jT+T} [G_\theta(\theta, \xi(t)) \, dt - \nabla J(\theta)] \, dt \to 0, \qquad \text{in probability,}$$

where E_m denotes the conditional expectation with respect to $\mathscr{F}_m$, the σ-algebra generated by $\{\theta_0, \psi_j, \Delta_j, b_j : j < m\}$.

Remark 3.1. More general random directions can be used; see the setup in Ref. 16. However, the random directions on the unit cube seem to be more easily implementable.

The constrained set in (i) is a hypercube, and the sets in (ii) and (iii) are used often in nonlinear programming and optimization. The possibilities of the constrained sets in (A1) above cover a wide variety of pervasive practical cases. The conditions are just those of (A3.1), or (A3.2), or (A3.3) in Ref. 16, p. 77.

Regarding the second moment condition in (A4), consider for instance Example 2.1. Solving the equation leads to

$$|x(t)| \le |x| + K \int_0^t (|x(s)| + 1) \, ds.$$

Under Markovian capacity and demand,

$$E|x(t)|^2 \le K.$$

Normally, the capacity is modeled by a finite-state Markov chain. If the demand rate is changed to a periodic diffusion (a diffusion defined on a compact manifold), the boundedness mentioned above still hold.

In a more general setting, consider the dynamic system given by (1). From a practical point of view, the inventory level must be bounded, so are the machine capacity and the demand. Suppose that, for each (α, β), the function $f(\cdot, \cdot, \alpha, \beta)$ is smooth enough. Then, the moment bound also holds. This demonstrates that such a condition is not a strong restriction.

Assumption (A4) is an averaging condition. It is satisfied if $\xi(t)$ is a stationary ϕ-mixing process. An alternative condition is as in Ref. 3, where it is assumed that the underlying process is ergodic. However, the condition used here appears to be slightly weaker.

Note that

$$\Delta_n D\tilde{J}_n(\theta_n) = (\Delta_n/2\delta T) \int_{nT}^{(n+1)T} [G(\theta_n + \delta\Delta_n, \xi(t)) - G(\theta_n - \delta\Delta_n, \xi(t))] \, dt. \quad (9)$$

Define

$$\hat{G}^+(\theta_n + \delta\Delta_n) = (1/T) \int_{nT}^{(n+1)T} [G(\theta_n + \delta\Delta_n, \xi(t)) \, dt,$$

$$\hat{G}^-(\theta_n - \delta\Delta_n) = (1/T) \int_{nT}^{(n+1)T} [G(\theta_n - \delta\Delta_n, \xi(t)) \, dt.$$

Then,

$$\Delta_n D\tilde{J}_n(\theta_n) = (\Delta_n/2\delta)[\hat{G}^+(\theta_n + \delta\Delta_n) - \hat{G}^-(\theta_n - \delta\Delta_n)].$$

Define a piecewise constant interpolation of θ_n as follows:

$$\theta^\epsilon(t) = \begin{cases} \theta_0, & \text{for } t < 0, \\ \theta_n, & \text{for } 0 \le t \in [n\epsilon, n\epsilon + \epsilon). \end{cases}$$

Define the interpolation for the reflection term as

$$z^\epsilon(t) = \begin{cases} 0, & \text{for } t < 0, \\ \epsilon \sum_{j=0}^{t/\epsilon - 1} z_j, & \text{for } t \ge 0. \end{cases}$$

Then, for any $t, s > 0$,

$$\theta^\epsilon(t+s) - \theta^\epsilon(t) = \theta_{(t+s)/\epsilon} - \theta_{t/\epsilon}$$

$$= \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \Delta_j D\tilde{J}_j(\theta_j) + [z^\epsilon(t+s) - z^\epsilon(t)].$$

In the above, t/ϵ and $(t+s)/\epsilon$ are taken to be integers without loss of generality (we can always take the integral parts of the above quantities in any case). We could use the notation $\lfloor t/\epsilon \rfloor$, i.e., the integer part of t/ϵ. But, for saving notation, we simply use t/ϵ throughout.

Theorem 3.1.   Under Assumptions (A1)–(A4), $\{(\theta^\epsilon(\cdot), z^\epsilon(\cdot))\}$ is tight in $D^{2r}(-\infty, \infty)$.

Proof.   First, the projection implies that $\{\theta_n\}$ is bounded, so are $\{\theta_n + \delta\Delta_n\}$ and $\{\theta_n - \delta\Delta_n\}$. By virtue of a truncated Taylor expansion, we have

$$G(\theta_n + \delta\Delta_n, \xi(t)) = G(\theta_n, \xi(t)) + \delta\Delta_n' G_\theta(\theta_n + \tilde{s}\delta\Delta_n, \xi(t)), \quad (10)$$

for some $0 \leq \tilde{s} \leq 1$. Similarly, we have

$$G(\theta_n - \delta\Delta_n, \xi(t)) = G(\theta_n, \xi(t)) - \delta\Delta_n' G_\theta(\theta_n - \hat{s}\delta\Delta_n, \xi(t)), \qquad (11)$$

for some $0 \leq \hat{s} \leq 1$. As a result,

$\Delta_n D\tilde{J}_n(\theta_n)$

$$= (\Delta_n\Delta_n'/2T) \int_{nT}^{nT+T} [G_\theta(\theta_n + \tilde{s}\delta\Delta_n, \xi(t)) + G_\theta(\theta_n - \hat{s}\delta\Delta_n, \xi(t))] \, dt, \qquad (12)$$

and the recursion takes the form

$\theta_{n+1} = \theta_n + \epsilon(\Delta_n\Delta_n'/2T)$

$$\times \int_{nT}^{nT+T} [G_\theta(\theta_n + \tilde{s}\delta\Delta_n, \xi(t)) + G_\theta(\theta_n - \hat{s}\delta\Delta_n, \xi(t))] \, dt + \epsilon z_n.$$

Since z_n is the minimum force needed to keep θ_n in the constraint set H, it is readily seen that

$$\left| z^\epsilon(t+s) - z^\epsilon(t) \right| \leq \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \epsilon \left| D\tilde{J}_j(\theta_j) \right|. \qquad (13)$$

The uniform integrability in (A2) and Ref. 17, Lemma 3.7, p. 51 (see also Theorem 8.2 of Ref. 18) imply the tightness of $\{(\theta^\epsilon(\cdot), z^\epsilon(\cdot))\}$. Moreover, similar to Part 2 of the proof of Theorem 8.2.1 in Ref. 16, any limit of the underlying process has Lipschitz continuous sample paths w.p.1. This concludes the proof.   □

To proceed, define

$$\psi_n = [\hat{G}^+(\theta_n + \delta\Delta_n) - J(\theta_n + \delta\Delta_n)] - [\hat{G}^-(\theta_n - \delta\Delta_n) - J(\theta_n - \delta\Delta_n)].$$

Furthermore, note that

$$[J(\theta_n + \delta\Delta_n) - J(\theta_n - \delta\Delta_n)]/2\delta = \Delta_n' \nabla J(\theta_n) + b_n, \qquad (14)$$

where b_n denotes the bias of the finite-difference approximation of $\nabla J(\theta_n)$. Thus, (6) can be rewritten as

$$\theta_{n+1} = \theta_n + \epsilon\nabla J(\theta_n) + \epsilon\Delta_n b_n + \epsilon(\Delta_n\Delta_n' - I)\nabla J(\theta_n) + \epsilon(\Delta_n/2\delta)\psi_n + \epsilon z_n. \qquad (15)$$

The recursive form above indicates that the algorithm can be decomposed into the gradient of the objective function, the finite-difference bias, the observation noise, the random-direction noise, and the reflection term. In what follows, we show that, through averaging, this algorithm has a close

connection to a continuous-time projected ordinary differential equation (see Ref. 16, p. 77),

$$\dot{\theta} = \nabla J(\theta) + \tilde{z}, \qquad \tilde{z}(t) \in -C(\theta(t)), \qquad \theta(0) = \theta_0. \tag{16}$$

The form of $C(\theta(t))$ depends on which one of the conditions in (A1) is satisfied. Under (i) in (A1), the set $C(\theta)$ is as follows:

If $\theta \in H^0$, the interior of H, $C(\theta) = \{0\}$.

If $\theta \in \partial H$, the boundary of H, $C(\theta)$ is the infinite convex cone generated by the outer normals at θ of the faces on which θ lies. Under (ii) of (A1), $C(\theta)$ is the convex cone generated by the set of outward normal $\{y: y = q_{i,\theta}(\theta), i \in A(x)\}$. If there is no active constraints at θ, $C(\theta) = \{0\}$. Under (iii), $C(\theta)$ is the linear span of the outer normal at θ.

Next, we characterize the limit process. The result is stated below.

Theorem 3.2. Under the conditions of Theorem 3.1, extract a weakly convergent subsequence of $(\theta^\epsilon(\cdot), z^\epsilon(\cdot))$ with limit $(\theta(\cdot), z(\cdot))$. Then, the limit satisfies (14).

Proof. First, in view of (12), $b_n = O(\delta^2)$. Thus, the boundedness of Δ_j, $\delta = \delta_\epsilon \to 0$ as $\epsilon \to 0$, and the bounded convergence theorem imply that

$$E\epsilon \left| \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \Delta_j b_j \right| \le \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} E|\Delta_j| |b_j|$$

$$= \sqrt{r}\epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} O(\delta^2) \to 0, \qquad \text{as } \epsilon \to 0.$$

Furthermore, the limit is uniform in t. Thus, it suffices to consider

$$\theta^\epsilon(t+s) - \theta^\epsilon(t) = \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \nabla J(\theta_j) + \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j \Delta_j' - I) \nabla J(\theta_j)$$

$$+ \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \Delta_j \psi_j / 2\delta + [z^\epsilon(t+s) - z^\epsilon(t)] + o(1),$$

where $o(1) \to 0$ in probability uniformly in t.

To prove the desired result, define

$$M(t) = \theta(t) - \theta_0 - \int_0^t \nabla J(\theta(\tau)) \, d\tau - z(t).$$

All that is needed is to show that $M(t)$ is a continuous martingale. Once the martingale properties are verified, by virtue of Theorem 4.1.1 of Ref. 16,

$M(t) \equiv$ const, since $M(\cdot)$ is locally Lipschitz. To verify the martingale property, it is equivalent to show that, for any bounded and continuous function $h(\cdot)$, any integer $\kappa > 0$, and any $0 < t_i < t$, and $s > 0$,

$$Eh(\theta(t_i); i \le \kappa) \left[\theta(t+s) - \theta(t) - \int_t^{t+s} \nabla J(\theta(\tau))\, d\tau - (z(t+s) - z(t)) \right] = 0. \quad (17)$$

To this end, let us begin with the process indexed by the small parameter $\epsilon > 0$. In fact, by the weak convergence and the Skorohod representation,

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa)[\theta^\epsilon(t+s) - \theta^\epsilon(t)]$$

$$= Eh(\theta(t_i); i \le \kappa)[\theta(t+s) - \theta(t)].$$

On the other hand, in view of the equation for $\theta^\epsilon(t+s) - \theta^\epsilon(t)$, choosing a sequence $\{k_\epsilon\}$ such that $k_\epsilon \to \infty$ as $\epsilon \to 0$ and $\gamma_\epsilon = \epsilon k_\epsilon \to 0$, we have

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa)[\theta^\epsilon(t+s) - \theta^\epsilon(t)]$$

$$= \lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa)$$

$$\times \left[\sum_{lk_\epsilon = t/\epsilon}^{(t+s)/\epsilon} \gamma_\epsilon (1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} \nabla J(\theta_j) \right.$$

$$+ \sum_{lk_\epsilon = t/\epsilon}^{(t+s)/\epsilon} \gamma_\epsilon (1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} (\Delta_j \Delta_j' - I) \nabla J(\theta_j)$$

$$\left. + \sum_{lk_\epsilon = t/\epsilon}^{(t+s)/\epsilon} \gamma_\epsilon (1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} \Delta_j \psi_j / 2\delta + z^\epsilon(t+s) - z^\epsilon(t) \right].$$

Letting $\epsilon l k_\epsilon \to \tau$, by the continuity of $\nabla J(\cdot)$, the weak convergence of $\theta^\epsilon(\cdot)$, and the Skorohod representation, we have

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[\sum_{lk_\epsilon = t/\epsilon}^{(t+s)/\epsilon} \gamma_\epsilon (1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} \nabla J(\theta_j) \right]$$

$$= \lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[\sum_{lk_\epsilon = t/\epsilon}^{(t+s)/\epsilon} \gamma_\epsilon \nabla J(\theta_{lk_\epsilon}) \right]$$

$$= Eh(\theta(t_i); i \le \kappa) \left[\int_t^{t+s} \nabla J(\theta(\tau))\, d\tau \right].$$

Recall that $\mathscr{F}_n$ denotes the σ-algebra generated by $\{\theta_0, \psi_j, \Delta_j, b_j : j < n\}$ and that E_n denotes the corresponding conditional expectation. Then, θ_n is

$\mathscr{T}_n$-measurable, so is $\nabla J(\theta_n)$. Consequently, for the term involving random direction noise, we have

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon+k_\epsilon-1} (\Delta_j \Delta_j' - I) \nabla J(\theta_j) \right]$$

$$= \lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon+k_\epsilon-1} E_{lk_e}(\Delta_j \Delta_j' - I) \nabla J(\theta_j) \right]$$

$$= \lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon+k_\epsilon-1} E_{lk_\epsilon}[E_j(\Delta_j \Delta_j' - I)] \nabla J(\theta_j) \right],$$

$$= 0,$$

since θ_j is $\mathscr{T}_j$ measurable and

$$E_j \Delta_j \Delta_j' = I,$$

by the independence and the Bernoulli assumption on the random noise.

Subtracting $J(\theta_n + \delta \Delta_n)$ from (8) and $J(\theta_n - \delta \Delta_n)$ from (9), we obtain

$$\psi_n = (\delta \Delta_n'/T) \int_{nT}^{nT+T} [\tilde{G}_\theta(\theta_n + \tilde{s}\delta\Delta_n, \xi(t)) + \tilde{G}_\theta(\theta_n - \hat{s}\delta\Delta_n, \xi(t))] \, dt, \qquad (18)$$

where

$$\tilde{G}_\theta(\theta_n + \tilde{s}\delta\Delta_n, \xi(t)) = G_\theta(\theta_n + \tilde{s}\delta\Delta_n, \xi(t)) - \nabla J(\theta_n + \tilde{s}\delta\Delta_n),$$

$$\tilde{G}_\theta(\theta_n - \hat{s}\delta\Delta_n, \xi(t)) = G_\theta(\theta_n - \hat{s}\delta\Delta_n, \xi(t)) - \nabla J(\theta_n - \hat{s}\delta\Delta_n).$$

Again, using $\epsilon lk_\epsilon \to \tau$, by the continuity of the functions $g_\theta(\cdot)$ and $\nabla J(\cdot)$, the weak convergence of $\theta^\epsilon(\cdot)$, and the Skorohod representation, we have

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon+k_\epsilon-1} (\Delta_j \Delta_j'/2T) \int_{jT}^{jT+T} \tilde{G}_\theta(\theta_j + \tilde{s}\delta\Delta_j, \xi(t)) \, dt \right]$$

$$= \lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon+k_\epsilon-1} (\Delta_j \Delta_j'/2T) \int_{jT}^{jT+T} \tilde{G}_\theta(\theta_{lk_\epsilon}, \xi(t)) \, dt \right] + o(1)$$

$$= \lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \le \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon+k_\epsilon-1} (\Delta_j \Delta_j'/2T) \int_{jT}^{jT+T} \tilde{G}_\theta(\theta(\tau), \xi(t)) \, dt \right] + o(1),$$

where $o(1) \to 0$ in probability uniformly in t.

Suppose that $\theta(\tau)$ takes finitely many values, say $x_1, \ldots, x_\iota$ for some positive integer ι. Then, using nested expectation and (A4), we have

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \leq \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} (\Delta_j \Delta_j'/2T) \int_{jT}^{jT+T} E_j \tilde{G}_\theta(\theta(\tau), \xi(t)) \, dt \right]$$

$$= \lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \leq \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} (\Delta_j \Delta_j'/2T) \right.$$

$$\left. \times \sum_{k_1=1}^{\iota} \int_{jT}^{jT+T} E_j \tilde{G}_\theta(x_{k_1}, \xi(t)) I_{\{\theta(\tau) = x_{k_1}\}} \, dt \right]$$

$$= 0.$$

For more general $\theta(\tau)$, for each $\eta > 0$, choose a finite number of disjoint sets B_i^η, $k_1 = 1, \ldots, \iota$, which cover the range of $\{\theta_n\}$ such that

$$P(\theta(\tau) \in \partial B_{k_1}^\eta) = 0, \qquad \text{diam}(B_{k_1}^\eta) \leq \eta,$$

where $\partial B_{k_1}^\eta$ denotes the boundary of $B_{k_1}^\eta$ and $\text{diam}(B)$ denotes the diameter of the set B. Choose $\theta_{k_1}^\eta \in B_{k_1}^\eta$. Then, proceeding with the above argument, we obtain also that

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i \leq \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} (\Delta_j \Delta_j'/2T) \int_{jT}^{jT+T} \tilde{G}_\theta(\theta(\tau), \xi(t)) \, dt \right] = 0.$$

A few details are omitted. Similarly, we also have

$$\lim_{\epsilon \to 0} Eh(\theta^\epsilon(t_i); i < \kappa) \left[(1/k_\epsilon) \sum_{j=lk_\epsilon}^{lk_\epsilon + k_\epsilon - 1} (\Delta_j \Delta_j'/2T) \int_{jT}^{jT+T} \tilde{G}_\theta(\theta_j - \hat{s}\delta\Delta_j, \xi(t)) \, dt \right]$$

$$= 0.$$

Finally, note that the Lipschitz continuity of $z(\cdot)$ implies that it is absolutely continuous, and hence there is a $\tilde{z}(\cdot)$ such that

$$z(t+s) - z(t) = \int_t^{t+s} \tilde{z}(\tau) \, d\tau.$$

Piecing together the estimates obtained thus far, the desired theorem follows.   □

Corollary 3.1. Let $\{t_\epsilon\}$ be a sequence such that $t_\epsilon \to \infty$ and $\epsilon t_\epsilon \to \infty$ as $\epsilon \to 0$. Then, any weakly convergent subsequence of $(\theta^\epsilon(\epsilon t_\epsilon + \cdot), z^\epsilon(\epsilon t_\epsilon + \cdot))$ has a limit $(\theta(\cdot), z(\cdot))$ such that $\theta(\cdot)$ lies in an invariant set $\hat{A}$ of (14). Use L_H to denote the set of limit points of (14) over all initial conditions in H.

If L_H is asymptotically stable in the sense of Lyapounov, then the invariant set $\hat{A}$ is in L_H.

The proof of this corollary is similar to the previous theorem. Theorem 3.2 deals with large but still bounded t, whereas in Corollary 3.1 we let both $\epsilon \to 0$ and $n \to \infty$. Note that, if there is a unique stationary point θ^*, then $\theta(\cdot)$ is concentrated at θ^*.

3.2. Rate of Convergence. This section deals with the rate of convergence issues. Since our main concern here is the convergence rate of θ_n to θ^*, we assume that $\theta^* \in H^0$ (the interior of H) and the convergence of θ_n to θ^*. To study the convergence rate, using a local analysis, we bring the asymptotic features of $\theta_n - \delta^*$ to the foreground.

Since $\theta^* \in H^0$, $\theta_n \in H^0$ for sufficiently large n. Thus, without loss of generality, we assume that $\theta_n \in H^0$ for all n and henceforth drop the reflection term z_n from the recursion. Consider

$$\theta_{n+1} = \theta_n + \epsilon \nabla J(\theta_n) + \epsilon \Delta_n b_n + \epsilon(\Delta_n \Delta_n' - I)\nabla J(\theta_n) \qquad (19)$$
$$+ \epsilon(\Delta_n/2\delta)\psi_n.$$

In what follows, assume that $\delta = \epsilon^{1/6}$. This choice stems from the decreasing stepsize KW (Kiefer–Wolfowitz) algorithm; see Chapter 10 of Ref. 16. We will exploit the dependence of the estimation error $\theta_n - \theta^*$ on ϵ. Define a rescaled sequence

$$U_n = (\theta_n - \theta^*)/\epsilon^{1/3}.$$

Let the following conditions hold:

(B1)   As $\epsilon \to 0$, $\theta^\epsilon(t_\epsilon + \cdot)$ converges weakly to $\theta^* \in H^0$.

(B2)   For some $N_\epsilon \to \infty$ as $\epsilon \to 0$, $\{U_n : n \geq N_\epsilon\}$ is tight.

(B3) $J(\cdot)$ is three times continuously differentiable. There is an asymptotically stable matrix (i.e., all of its eigenvalues have negative real parts) $A \in \mathbb{R}^{r \times r}$ such that

$$\nabla J(\theta) = A(\theta - \theta^*) + O(|\theta - \theta^*|^2).$$

(B4)   As $\epsilon \to 0$, $\sqrt{\epsilon} \sum_{j=0}^{t/\epsilon} (1/T) \int_{jT}^{jT+T} G_\theta(\theta^*, \xi(t))\, dt$ converges weakly to Brownian motion $w(\cdot)$ with covariance Σt.

Remark 3.2. A few words and comments of the assumptions are in order. Since our main objective is to obtain the convergence rate, the convergence is assumed. The tightness in (B2) can be obtained by means of a

perturbed Lyapounov function approach (Chapter 10 of Ref. 16) and sufficient conditions can be provided. We choose to assume this tightness condition for simplicity. Condition (B3) indicates that $J(\cdot)$ is locally (near θ^*) linear. This allows a localization leading to the desired asymptotic normality. Finally, (B4) assumes the weak convergence to a Brownian motion. Conditions guaranteeing this can be found in the literature; see for example Ref. 16, Chapter 7, and Ref. 18 among others.

For $n \geq N_\epsilon$, it follows from (17) that

$$U_{n+1} = U_n + \epsilon A U_n + \epsilon(\Delta_n b_n / \epsilon^{1/3}) + \epsilon(\Delta_n \Delta_n' - I) A U_n$$
$$+ (\sqrt{\epsilon} \Delta_n / 2) \psi_n + \epsilon o(|U_n|). \tag{20}$$

Next, define the piecewise constant interpolation

$$u^\epsilon(0) = U_{N_\epsilon},$$

$$u^\epsilon(t) = U_n, \qquad \text{for } t \in [\epsilon(n - N_\epsilon), \epsilon(n - N_\epsilon) + \epsilon).$$

In view of the definition of the interpolation, the space we are working with is $D^r[0, \infty)$. Alternatively, one could work on $D^r(-\infty, \infty)$. An interested reader is referred to Chapter 10 of Ref. 16 for this aspect.

Since it is not known a priori that U_n is bounded, a pertinent analysis requires using a truncation device; see Ref. 16. The idea is as follows: For each $N > 0$, work on a truncated process $U_n^{\epsilon,N}$ taking values inside and on the sphere with radius N centered at the origin. Denote the corresponding interpolated process by $\{u^{\epsilon,N}(\cdot)\}$. Then, we show that it is tight and characterize its limit $u^N(\cdot)$. Finally, let the radius of the sphere grow to ∞ to complete the proof. However, for notational simplicity, we assume henceforth the boundedness of U_n.

Note that

$$u^\epsilon(t+s) - u^\epsilon(t) = \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} A U_j$$

$$+ \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j b_j / \epsilon^{1/3}) + \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j \Delta_j' - I) A U_j$$

$$+ \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j / 2) \psi_j + \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} o(|U_j|)$$

and that

$$\sqrt{\epsilon} \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j / 2) \psi_j = \sqrt{\epsilon} \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \Delta_j \Delta_j' \hat{G}_j + o(1), \tag{21}$$

where

$$\hat{G}_j = (1/(2T)) \int_{jT}^{jT+T} G_\theta(\theta^*, \xi(t)) \, dt$$

and $o(1) \to 0$ in probability uniformly in t. In addition,

$$E \left| \sqrt{\epsilon} \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j \Delta_j' - I)\hat{G}_j \right|^2$$

$$= 2\epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \sum_{k>j}^{(t+s)/\epsilon} E\hat{G}_j'(\Delta_j\Delta_j' - I)(\Delta_k\Delta_k' - I)\hat{G}_k$$

$$+ \epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} E\hat{G}_j'(\Delta_j\Delta_j' - I)(\Delta_j\Delta_j' - I)\hat{G}_j.$$

In view of the independence of $\{\Delta_j\}$, for $k > j$, we have

$$E\hat{G}_j'(\Delta_j\Delta_j' - I)(\Delta_k\Delta_k' - I)\hat{G}_k$$

$$= E\hat{G}_j'[E_j(\Delta_j\Delta_j' - I)(\Delta_k\Delta_k' - I)]\hat{G}_k$$

$$= E\hat{G}_j'[E_j(\Delta_j\Delta_j' - I)E_k(\Delta_k\Delta_k' - I)]\hat{G}_k = 0.$$

Similarly, since

$$\Delta_k'\Delta_k = 1 \quad \text{and} \quad E\Delta_k\Delta_k = E_k\Delta_k\Delta_k' = I,$$

we have

$$E\hat{G}_k'(\Delta_k\Delta_k' - I)(\Delta_k\Delta_k' - I)\hat{G}_k$$

$$= E\hat{G}_k'[E_k(\Delta_k\Delta_k'\Delta_k\Delta_k' - 2\Delta_k\Delta_k' + I)]\hat{G}_k$$

$$= E\hat{G}_k'[E_k(-\Delta_k\Delta_k' + I)]\hat{G}_k$$

$$= 0.$$

Therefore, the above arguments and (19) imply that

$$\sqrt{\epsilon} \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j/2)\psi_j = \sqrt{\epsilon} \sum_{j=t/\epsilon}^{(t+s)/\epsilon} \hat{G}_j + o(1),$$

where $o(1) \to 0$ in probability. As a result, Condition B4 yields that $\sqrt{\epsilon} \sum_{j=t/\epsilon}^{(t+s)/\epsilon}(\Delta_j/2)\psi_j$ converges weakly to $\int_t^{t+s} dw(\tau)$, where $w(\cdot)$ is the Brownian motion with covariance Σt.

Next, in view of the formula for $u^\epsilon(t+s) - u^\epsilon(t)$ above, similar to the last section, it can be shown that

$$\epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j \Delta'_j - I) A U_j \text{ converges weakly to } 0,$$

$$\epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} A U \text{ converges to } \int_t^{t+s} Au(\tau) \, d\tau.$$

As for the bias term, recall that $\delta = \epsilon^{1/6}$ and note that

$$b_n / \epsilon^{1/3} = (1/3! \epsilon^{1/3}) \left[\sum_{i,j,k} [\partial^3 J(\theta_n)/\partial\theta_i \partial\theta_j \partial\theta_k] \Delta_{n,i} \Delta_{n,j} \Delta_{n,k} \delta^2 + o(\delta^2) \right]$$

$$= (1/3!) \sum_{i,j,k} [\partial^3 J(\theta_n)/\partial\theta_i \partial\theta_j \partial\theta_k] \Delta_{n,i} \Delta_{n,j} \Delta_{n,k} + o(1),$$

where $o(1) \to 0$ as $\epsilon \to 0$. The weak convergence of $\theta^\epsilon(t_\epsilon + \cdot)$ to θ^* together with the Skorohod representation leads to

$$\epsilon \sum_{j=t/\epsilon}^{(t+s)/\epsilon} (\Delta_j b_j / \epsilon^{1/3}) \text{ converges weakly to } \bar{b}(\theta^*)s,$$

where

$$\bar{b}(\theta^*) = (1/3!) E \Delta_n \sum_{i,j,k} [\partial^3 J(\theta^*)/\partial\theta_i \partial\theta_j \partial\theta_k] \Delta_{n,i} \Delta_{n,j} \Delta_{n,k}. \tag{22}$$

We summarize the discussion above in the following theorem.

Theorem 3.3.   Assume Conditions (B1)–(B4). Suppose that U_{N_ϵ} converges weakly to u_0. Then, $\{u^\epsilon(\cdot)\}$ is tight in $D^r[0, \infty)$. Any weakly convergent subsequence of $u^\epsilon(\cdot)$ has a limit $u(\cdot)$ such that $u(\cdot)$ satisfies the stochastic differential equation

$$du = [Au + \bar{b}(\theta^*)] \, dt + dw, \qquad u(0) = u_0, \tag{23}$$

with $\bar{b}(\theta^*)$ given by (20) and $w(\cdot)$ being the Brownian motion given in (B4).

The stationary covariance of the limit diffusion is

$$\hat{\Sigma} = \int_0^\infty \exp(At) \Sigma \exp(A't) \, dt. \tag{24}$$

A consequence of Theorem 3.3 is that $(\theta_n - \theta^*)/\epsilon^{1/3}$ converges in distribution to a normal random variable with mean $\bar{b}(\theta^*)$ and covariance $\hat{\Sigma}$ given by (22).

If starting from the constant-step algorithm, we use a decreasing step-size algorithm by replacing ϵ by $\epsilon_n = O(1/n)$ and δ by $\delta_n = O(1/n^{1/6})$, then the convergence rate result will be that $n^{1/3}(\theta_n - \theta^*)$ converges in distribution to a normal random variable with mean $\bar{b}(\theta^*)$ and covariance $\hat{\Sigma}$ given by (22).

4. Numerical Example

A numerical study is carried out to demonstrate the performance of our algorithm. We use Example 2.1. Various quantities and variables are as defined there, and the system parameters are shown in Table 1. In the asymptotic analysis, it is necessary that the stepsizes be small. In the actual computation, we chose them large enough to speedup the computation (we used a stepsize $\epsilon = 10$ and a finite-difference stepsize $\delta = 1000$, respectively).

We consider the random processes to be Markov chains. More general random processes can be handled in essentially the same way as far as the simulation is concerned. This example illustrates that, when one uses the proposed algorithm, one may need to modify the algorithm as needed.

Since no analytical result is available for the optimal threshold values, to show the convergence of the algorithm for different initial data, we plot the learning curves that are the contour curves generated by the simulations for different sets of control parameters (threshold values) and the corresponding costs. In the numerical experiments, different initial data are chosen; we have used $\theta_0 = (900, 100)'$ and $\theta_0 = (3500, 6000)'$. The learning curves are shown in Fig. 1. The iterates appear to be converging quickly. After a few iterations, the average costs are already close to their optimal values.

The example treated concerns mainly the production planning taking the market action (setting up the price of the product) into consideration. In Ref. 9, a somewhat different model is considered, in which in addition to planning the production rate, a firm makes a decision on the amount of

Table 1.   Parameters of the numerical experiment.

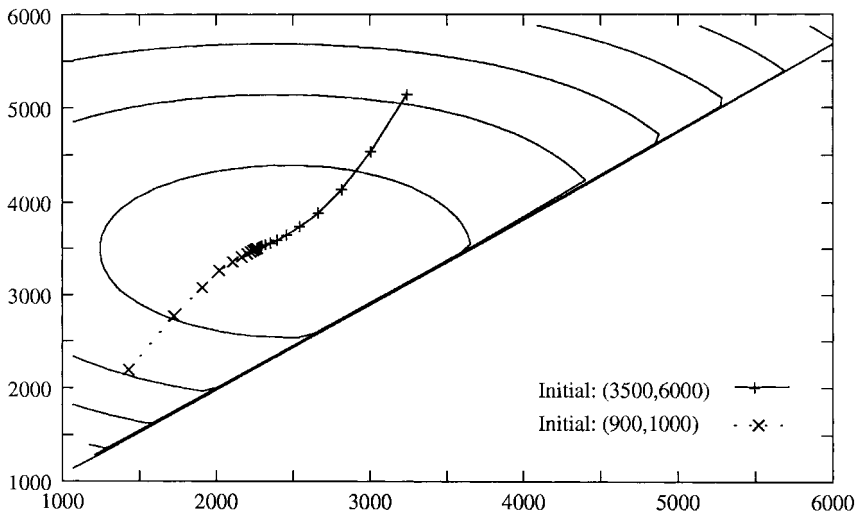| Demand rates	(d_1, d_2)	(150, 250)
Price rate	π_1	10
Promotion price rate	π_2	8
Maximum capacity	u	900
Inventory holding cost rate	c^+	0.1
Penalty cost rate	c^-	0.7
Mean machine up time	$1/\lambda$	20
Mean machine down time	$1/\mu$	8

Fig. 1.   Learning curves.

effort needed in the promotion activities; to obtain analytic solutions in that paper, it is necessary that one uses the conditions that the demand is a Poisson process and that the machine is reliable. The algorithm developed here can be used also to solve the problem as studied in Ref. 9, but allows for unreliable machines, non-Markovian processes, and multiple machine shops.

5. Concluding Remarks

In this paper, we have developed a class of stochastic approximation algorithms with applications to threshold-control problems. An example of a production planning in manufacturing has been examined. The algorithm developed has mixed modes, is recursive in form, and is easy to implement. It uses random-direction methods for the gradient estimates. One of the advantages of the algorithm is that, unlike the usual numerical schemes for solving optimal control problems that depend on the solutions of the corresponding HJB equations, it can deal with large-scale problems.

For the class of algorithms developed, normally the convergence is to the target of a local maximum. Algorithms for convergence to a global maximum can be designed. However, it is known (Ref. 19) that such algorithms have a very slow convergence rate. Since we are concentrating on threshold-type controls, the approximation that we get is in the sense of

a suboptimal solution. Thus, the local maximum is good enough for an approximation purpose.

In evaluating the convergence rates, the computational budget can be taken into consideration (Ref. 20). The basic idea is to treat both bias and noise as functions of the budget devoted to the computation, and then obtain the convergence rates accordingly.

References

1. YIN, G., and ZHANG, Q., Editors, *Mathematics of Stochastic Manufacturing Systems*, Proceedings of the 1996 AMS-SIAM Summer Seminar in Applied Mathematics, Lectures in Applied Mathematics, American Mathematical Society, Providence, Rhode Island, 1997.

2. AKELLA, R., and KUMAR, P. R., *Optimal Control of Production Rate in a Failure-Prone Manufacturing System*, IEEE Transactions on Automatic Control, Vol. 31, pp. 116–126, 1986.

3. BIELECKI, T. R., and KUMAR, P. R., *Optimality of Zero-Inventory Policies for Unreliable Manufacturing Systems*, Operations Research, Vol. 36, pp. 532–541, 1988.

4. GERSHWIN, S. B., *Manufacturing Systems Engineering*, Prentice-Hall, Englewood Cliffs, New Jersey, 1994.

5. BOUKAS, E. K., YANG, H., and ZHANG, Q., *Minimax Production Planning in Failure-Prone Manufacturing Systems*, Journal of Optimization Theory and Applications, Vol. 87, pp. 269–286, 1995.

6. ZHANG, Q., and YIN, G., *Turnpike Sets in Stochastic Manufacturing Systems with Finite-Time Horizon*, Stochastics and Stochastics Report, Vol. 51, pp. 11–40, 1994.

7. SETHI, S. P., and ZHANG, Q., *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, Massachusetts, 1994.

8. YIN, G., and ZHANG, Q., *Continuous-Time Markov Chains and Applications: A Singular Perturbations Approach*, Springer Verlag, New York, NY, 1998.

9. ZHANG, Q., YIN, G., and BOUKAS, E. K., *Optimal Control of a Marketing-Production System*, IEEE Transactions on Automatic Control (to appear).

10. YAN, H. M., YIN, G., and LOU, S. X. C., *Using Stochastic Optimization to Determine Threshold Values for Control of Unreliable Manufacturing Systems*, Journal of Optimization Theory and Applications, Vol. 83, pp. 511–539, 1994.

11. HO, Y. C., and CAO, X. R., *Perturbation Analysis of Discrete-Event Dynamic Systems*, Kluwer, Boston, Massachusetts, 1991.

12. LIBEROPOULOS, G., and CARAMANIS, M., *Infinitesimal Perturbation Analysis for Second-Derivative Estimation and Design of Manufacturing Flow Controllers*, Journal of Optimization Theory and Applications, Vol. 81, pp. 297–327, 1994.

13. YAN, H. M., ZHOU, X. Y., and YIN, G., *Approximating an Optimal Production Policy in a Continuous Flow Line: Recurrence and Asymptotic Properties,* Operations Research, Vol. 47, pp. 535–549, 1999.

14. KUSHNER, H. J., and CLARK, D. S., *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer Verlag, Berlin, Germany, 1978.
15. SPALL, J. C., *Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation*, IEEE Transactions on Automatic Control, Vol. 37, pp. 331–341, 1992.
16. KUSHNER, H. J., and YIN, G., *Stochastic Approximation Algorithms and Applications*, Springer Verlag, New York, NY, 1997.
17. KUSHNER, H. J., *Approximation and Weak Convergence Methods for Random Processes*, *with Applications to Stochastic Systems Theory*, MIT Press, Cambridge, Massachusetts, 1984.
18. BILLINGSLEY, P., *Convergence of Probability Measures*, J. Wiley, New York, NY, 1968.
19. YIN, G., *Rates of Convergence for a Class of Global Stochastic Optimization Algorithms*, SIAM Journal on Optimization, Vol. 10, pp. 99–120, 1999.
20. L'ÉCUYER, P., and YIN, G., *Budget-Dependent Convergence Rate of Stochastic Approximation*, SIAM Journal on Optimization, Vol. 8, pp. 217–247, 1998.