Stochastics: An International Journal of Probability and Stochastic Processes
Vol. 77, No. 1, February 2005, pp. 61–79

Numerical solutions for jump-diffusions with regime switching

G. YIN†*‡, Q. S. SONG†¶ and Z. ZHANG†§

†Department of Mathematics, Wayne State University, Detroit, MI 48202, USA

(Received 30 October 2004; in final form 30 November 2004)

This paper is devoted to numerical solutions for a class of jump-diffusions with regime switching. After briefly reviewing the notion of jump-diffusions with regime switching, finite-difference procedures are constructed. Under simple conditions, it is proved that the algorithm converges to the desired limit by means of a martingale problem formulation. Numerical experiments are carried out to demonstrate the performance of the algorithm.

Keywords: Numerical method; Discretization; Markov chain; Weak convergence; Martingale problem

2000 Mathematics Subject Classification: 60J27; 60J60; 60J75; 91B28; 34F05; 45J05

1. Introduction

We consider a class of jump-diffusion models with regime switching, which consists of the usual jump-diffusion processes together with a modulating continuous-time Markov chain. Suppose that there is a finite set $\mathcal{M} = \{1, \ldots, m_0\}$, representing the possible regimes of the environment. We work with a finite-time horizon $[0, T]$ for some $T > 0$. Let Γ be a compact subset of $\mathbb{R}$, $f(\cdot,\cdot) : \mathcal{M} \times \mathbb{R}^r \mapsto \mathbb{R}^r$, $\sigma(\cdot,\cdot) : \mathcal{M} \times \mathbb{R}^r \mapsto \mathbb{R}^r \times \mathbb{R}^r$, $g(\cdot,\cdot,\cdot) : \Gamma \times \mathcal{M} \times \mathbb{R}^r \mapsto \mathbb{R}^r$ and $\alpha(\cdot)$ be a continuous-time Markov chain having state space $\mathcal{M}$. Consider the dynamic system given by

$$\begin{aligned} \mathrm{d}x(t) &= f(\alpha(t), x(t))\,\mathrm{d}t + \sigma(\alpha(t), \mathrm{x}(t)\,\mathrm{d}w(t) + \mathrm{d}J(t), \\ J(t) &= \int_0^t \int_\Gamma g(\gamma, \alpha(s), x(s^-))N(\mathrm{d}s, \mathrm{d}\gamma), \\ x(0) &= x, \alpha(0) = \alpha, \end{aligned} \tag{1.1}$$

*Corresponding author. E-mail: gyin@math.wayne.edu

‡Research of this author was supported in part by the National Science Foundation under grant DMS-0304928, and in part by Wayne State University Research Enhancement Program.

¶Research of this author was supported in part by Wayne State University Graduate Research Assistantship. E-mail: song@math.wayne.edu

§Research of this author was supported in part by the National Science Foundation under grant DMS-0311807, and in part by Michigan Life Science Corridor. E-mail: zzhang@math.wayne.edu

Stochastics: An International Journal of Probability and Stochastic Processes
ISSN 1744-2508 print/ISSN 1744-2516 online © 2005 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/17442500512331341068

or in its integral form

$$x(t) = x + \int_0^t f(\alpha(s), x(s))\,\mathrm{d}s + \int_0^t \sigma(\alpha(s), x(s))\,\mathrm{d}w(s),$$
$$+ \int_0^t \int_\Gamma g(\gamma, \alpha(s), x(s^-))N(\mathrm{d}s, \mathrm{d}\gamma), \tag{1.2}$$

where $w(\cdot)$ is an r-dimensional standard Brownian motion, and $N(\cdot)$ is a Poisson measure. Throughout the paper, we assume that the Brownian motion $w(\cdot)$, the jump process $N(\cdot)$, and the Markov chain $\alpha(\cdot)$ are independent. This work focuses on the development of a numerical scheme for approximating the solutions of equation (1.2).

Our motivation stems from applications in risk theory and ruin probabilities arising in insurance and finance. In the classical insurance risk model, the surplus $X(t)$ of an insurance company at $t \geq 0$ is given by $X(t) = x + ct - S(t)$, where x is the initial surplus, $c > 0$ is the rate at which the premiums are received, and $S(t)$, a compound Poisson process, is the total claims in the duration $[0, t]$. In [7], the authors extended the classical risk model by adding an independent diffusion process so that the surplus is given by

$$X(t) = x + ct - S(t) + \sigma w(t), \tag{1.3}$$

where $w(t)$ is a standard Brownian motion that represents uncertainty (often referred to as oscillations) of premium incomes and claims. Since then, much work has been devoted to such jump-diffusion models; see also variations of the models in [15,18] and the references therein. Recently, there has been a resurgence of attention drawn to regime switching models in finance and insurance industry. For instance, taking the opportunity provided by using a switching process, the European options under the Black–Scholes formulation of the stock market, where the underlying economy switches among a finite number of states, were considered in [6]; the American options were dealt with in [4]; to make liquidation decision using solutions of two-point boundary value problems can be found in [25] and using a stochastic optimization approach is in [23]. The rationale is that for different environments, the system's behavior is markedly different. Therefore, we may introduce an "economic factor" process in the model. For instance, in a stock market, the regimes can be roughly divided into two states, bull market and bare market. The market sentiment and reaction to the two states are in stark contrast. Normally, a bare market is more volatile than that of a bull market. It is thus sensible and necessary to take such regime changes into consideration. To capture the features of insurance policies that are subject to the economic or political environment changes, we introduced a hybrid risk model and obtained the corresponding bounds for ruin probabilities in [20]. A more general class of jump-diffusion processes with regime switching takes the form (1.2), which is a nonlinear model, and a closed form solution is virtually impossible to obtain. As a viable alternative, one has to find feasible numerical schemes.

Numerical solutions of stochastic differential equations have been the focus of enormous research. An extensive treatment of the subject can be found in [9,14]; see also the recent survey [16] and the references therein. Solutions of differential equations with Levy driven processes can be found in [17]; finite element and difference methods for linear stochastic partial differential equations can be found in [1]; numerical methods for systems arising in stochastic control are in [12].

In this paper, we propose a numerical algorithm for (1.2), which is of finite difference type. We examine the convergence properties of the algorithm. Different from the usual convergence proofs, we use martingale problem formulation and weak convergence method to obtain the desired result. One of the advantages of this approach is that it brings out the profile and dynamic behavior of the process rather than dealing with the iterations directly.

The rest of the paper is arranged as follows. The second section begins with the review of certain notions for jump diffusions with regime switching. The third section proceeds with the description of the algorithm. We concentrate on a constant-step-size algorithm. The fourth section presents the convergence analysis. Using ideas from stochastic approximation methods [13], and taking continuous time interpolation, we show that the interpolated sequence converges to the desired jump-diffusion limit with switching by means of weak convergence and martingale averaging techniques. The main idea is to first verify the tightness of the sequence and then identify the limit process by examining a martingale problem with appropriate operator. The fifth section is devoted to several examples. Finally, the sixth section concludes the paper with a few more remarks.

2. Preliminary

In this section, we briefly review the notion of hybrid jump-diffusion processes, which is a generalization of the usual jump-diffusion processes. Since the jump diffusion is modulated by an additional continuous-time Markov chain, in lieu of one jump diffusion, we have a system of jump diffusions. The description to follow is a modification of that of [12] due to the appearance of the switching process. Suppose that $\alpha(\cdot)$ is a continuous-time Markov chain with state space $\mathcal{M} = \{1, \dots, m_0\}$ and generator $Q = (q_{ij})$ [21, Sections 2.3–2.5]. Let $\{v_n\}$ be an increasing sequence of stopping times independent of $\alpha(t)$ and the Brownian motion $w(\cdot)$ such that $v_{n+1} - v_n$ are independent and identically distributed exponential random variables. Let $\rho(\cdot)$ be a random process defined by

$$\rho(t) = \begin{cases} \rho_n & \text{if} \quad t = v_n, \\ 0, & \text{otherwise}, \end{cases}$$

where $\{\rho_n\}$ is a sequence of independent and identically distributed random variables, known as impulses. The process is termed a point process if $\{v_n\}$ has no finite accumulation point.

Let Γ, a compact set not including the origin in $\mathbb{R}$, be the range space of $\rho(\cdot)$. Denote the σ-algebra of Borel sets of Γ by $\mathcal{B}(\Gamma)$. Suppose that the impulse time $v_n \to \infty$ as $n \to \infty$. For each $H \in \mathcal{B}(\Gamma)$, define

$$N(t, H) = \{\#\,\text{of impulses of}\,\rho(\cdot)\,\text{on}\,[0, t]\,\text{with values in}\,H\},$$

which is a counting process or counting measure that counts the number of impulses up to time t. Suppose that $EN(t, \Gamma) < \infty$, that $\mathcal{F}_t$ is a filtration such that $N(\cdot, H)$ is $\mathcal{F}_t$-adapted for each $H \in \mathcal{B}(\Gamma)$, that $N(\cdot, \cdot)$ is an $\mathcal{F}_t$-Poisson measure, and that $\rho(\cdot)$ is an $\mathcal{F}_t$-Poisson point process (i.e. $\{N(t + \cdot, H) - N(t, H) : H \in \mathcal{B}(\Gamma)\}$ is independent of $\mathcal{F}_t$). If $\rho(\cdot)$ is a Poisson process and the distribution of $\{N(t + s, H) - N(t, H) : H \in \mathcal{B}(\Gamma)\}$ is independent of t, then $\rho(\cdot)$ is a stationary Poisson point process. Then it is known that there exists a $\lambda > 0$ and probability measure $\pi(\cdot)$ on $\mathcal{B}(\Gamma)$ such that

$$E[N(t + s, H) - N(t, H) | \mathcal{F}_t] = s\pi(H)\lambda,$$

where λ is known as impulse rate of $\rho(\cdot)$ and/or the jump rate of $N(\cdot, \Gamma)$, and $\pi(H)$ is the jump distribution in the sense that

$$P(\rho(t) \in H | \rho(t) \neq 0, \quad \rho(u), \quad u < t) = \pi(H).$$

It follows from the properties of Poisson processes,

$$P(\text{there is one } \rho(s) \neq 0 \text{ on } [t, t+\delta) | \rho(u), u < t) = \lambda\delta + o(\delta),$$

$$P(\text{there are more than one } \rho(s) \neq 0 \text{ on } [t, t+\delta) | \rho(u), u < t) = 0(\delta).$$

The values and time of the impulses can be recovered from the integral

$$\varpi(t) = \int_0^t \int_\Gamma \gamma N(\mathrm{d}s, \mathrm{d}\gamma) = \sum_{s \leq t} \rho(s).$$

With the notation given above, we can construct hybrid jump-diffusion processes similar to [12, p. 28]. In particular, we note that for the equation given in (1.2),

$$J(t) = \sum_{v_n \leq t} g(\rho_n, \alpha(v_n), x(v_n^-)). \tag{2.1}$$

Next, we recall the existence and uniqueness of equation (2.1) [12, p. 30].

- We say that a solution of equation (1.2) exists in the weak sense if given any probability measure μ on $\mathbb{R}^r$, there is a probability space $(\Omega, \mathcal{F}, P)$, a filtration $\mathcal{F}_t$, an $\mathcal{F}_t$-Brownian motion $w(\cdot)$, an $\mathcal{F}_t$-Poisson measure $N(\cdot)$, a Markov chain $\alpha(\cdot)$, and an $\mathcal{F}_t$-adapted process $x(\cdot)$ satisfying equation (1.2) for all $t \geq 0$ and $P(x(0) \in A_0) = \mu(A_0)$.
- Suppose that $\{(\Omega_\ell, \mathcal{F}_\ell, P_\ell), \mathcal{F}_{\ell,t}, w_\ell(\cdot), N_\ell(\cdot), \alpha_\ell(\cdot), x_\ell(\cdot)\}$ for $\ell = 1, 2$ are two weak sense solutions to equation (1.2). The solution to equation (1.2) is unique in the weak sense if equality of the distribution induced on $\mathbb{R}^r$ by $x_\ell(0)$ under P_ℓ for $\ell = 1, 2$ implies equality of the distributions induced on $D([0, T]; \mathbb{R}^r)$ by $x_\ell(\cdot)$ under P_ℓ for $\ell = 1, 2$, where $D([0, T]; \mathbb{R}^r)$ is the space of $\mathbb{R}^r$-valued functions defined on [0, T] that are right continuous and have left limits endowed with the Skorohod topology.

Remark 2.1. In what follows, the asymptotic analysis of the algorithm is carried out by weak convergence method. As a result, we mainly concern ourselves with the weak sense solution. Note that sufficient conditions guaranteeing the existence and uniqueness can be devised; see [12, p. 30]. For example, if we assume that the coefficients $f(i,\cdot)$ and $\sigma(i,\cdot)$ satisfy the usual linear growth and Lipschitz conditions, then the existence and uniqueness of solution of equation (1.2) can be obtained not only in the weak sense but also in the strong sense. We shall not dwell on it here.

For future use, define a centered Poisson measure as

$$\tilde{N}(t, H) = N(t, H) - \lambda t \pi(H).$$

With this centered Poisson measure, we can rewrite $J(t)$ as

$$J(t) = \int_0^t \int_\Gamma g(\gamma, \alpha(s), x(s^-)) \tilde{N}(\mathrm{d}s, \mathrm{d}\gamma) + \lambda \int_0^t \int_\Gamma g(\gamma, \alpha(s), x(s^-)) \pi(\mathrm{d}\gamma)\, \mathrm{d}s, \tag{2.2}$$

which is the sum of a martingale and an absolute continuous process provided certain conditions are satisfied for the function $g(\cdot)$.

Define an operator $\mathcal{L}$ by

$$\mathcal{L}\varphi(i,x) = \varphi_x'(i,x)f(i,x) + \frac{1}{2}\mathrm{tr}[\varphi_{xx}(i,x)\sigma(i,x)\sigma'(i,x)] \\ + \lambda\int_\Gamma [\varphi(i,x+g(\gamma,i,x)) - \varphi(i,x)]\pi(\mathrm{d}\gamma) + Q\varphi(\cdot,x)(i), \quad i \in \mathcal{M}, \tag{2.3}$$

for a suitable function $\varphi(\cdot)$, where z' denotes the transpose of $z \in \mathbb{R}^{i\times j}$ for $i, j \geq 1$, and

$$Q\varphi(\cdot,x)(i) = \sum_{j=1}^{m} q_{ij}\varphi(j,x) = \sum_{j\neq i} q_{ij}(\varphi(j,x) - \varphi(i,x)).$$

Then a generalized Ito's formula reads: for a suitable function $\varphi(\cdot,\cdot)$,

$$\varphi(\alpha(t),x(t)) - \varphi(\alpha,x) = \int_0^t \mathcal{L}\varphi(\alpha(s),x(s))\,\mathrm{d}s + \int_0^t \varphi_x'(\alpha(s),x(s))\sigma(\alpha(s),x(s))\,\mathrm{d}w(s) \\ + \int_0^t\int_\Gamma [\varphi(\alpha(s),x(s^-) + g(\gamma,\alpha(s),x(s^-))) \\ - \varphi(\alpha(s),x(s^-))]\tilde{N}(\mathrm{d}s,\mathrm{d}\gamma). \tag{2.4}$$

Note that equation (2.4) can be obtained by a combined approach of the extended Ito's formula in [3] and that of [11, p. 39]. When the switching process $\alpha(t)$ is missing, it reduces to the Ito's formula for jump-diffusion processes; when the Poisson jump is missing, it reduces to that for diffusions with regime switching; when both the switching process and the Poisson jumps are missing, it reduces to the Ito's formula for usual diffusion processes. If

$$M_\varphi(t) \overset{\text{def}}{=} \varphi(\alpha(t),x(t)) - \varphi(\alpha,x) - \int_0^t \mathcal{L}\varphi(\alpha(s),x(s))\,\mathrm{d}s \tag{2.5}$$

is an $\mathcal{F}_t$-martingale for each $\varphi(i,\cdot) \in C_0^2$ with $i \in \mathcal{M}$, where C_0^2 denotes the class of functions with compact support that are twice continuously differentiable, then $(\alpha(\cdot), x(\cdot))$ is said to be a solution of the martingale problem for the operator $\mathcal{L}$. To proceed, we need the following conditions.

(A1) For each $i \in \mathcal{M}$, the functions $f(i,\cdot)$ and $\sigma(i,\cdot)$ are continuous and grow at most linearly, i.e. $|f(i,x)| + |\sigma(i,x)| \leq K(1+|x|)$.

(A2) For each $i \in \mathcal{M}$, $g(\cdot,i,\cdot)$ is a bounded and continuous, $\mathbb{R}^r$-valued function on $\mathbb{R}\times\mathbb{R}^r$ satisfying $g(0,i,0) = 0$ for each $i \in \mathcal{M}$; for each x, the value of γ can be uniquely determined from the value of $g(\gamma,i,x)$.

LEMMA 2.2. Assume that (A1) and (A2) hold. Suppose that $(\alpha(\cdot), x(\cdot))$ is a solution of the martingale problem with $x(\cdot)$ taking values in $D^r[0,T]$. Then there exists a standard $\mathbb{R}^r$-valued Brownian motion $w(\cdot)$, and a Poisson measure $N(\cdot)$ such that

$$P(\text{one jump occurs on } (t,t+\delta]|\mathcal{F}_t) = \sum_{i=1}^{m} \delta\lambda\pi\{\gamma : g(\gamma,i,x(t)) \neq 0\}I_{\{\alpha(t)=i\}} + o(\delta),$$

$$P(\text{more than one jumps occur on } (t,t+\delta]|\mathcal{F}_t) = o(\delta),$$

$$P(\text{jump of magnitude in } H \text{ occurs at } t|\text{jump at } t, \mathcal{F}_{t^-}) = \sum_{i=1}^{m} \pi\{\gamma : g(\gamma,i,x(t^-)) \in H\}I_{\{\alpha(t)=i\}},$$

and that equation (1.2) holds.

Proof The proof is a slight modification of [11, Theorem 8.1, p. 42]. □

LEMMA 2.3. Suppose that (A1) and (A2) are satisfied and that $f(i,\cdot)$ and $\sigma(i,\cdot)$ are Lipschitz continuous for each $i \in \mathcal{M}$. Then there is a solution to the martingale problem with operator $\mathcal{L}$. Moreover, the solution is unique in the weak sense.

Proof The existence follows from the argument in [12]; see also [11]. The uniqueness can be proved by using a characteristic function as in [21, Lemma 7.18]. □

3. Algorithm

We describe the algorithm as follows:

1. Choose $\varepsilon > 0$, a small parameter, as the step size.
2. Construct a discrete-time Markov chain α_n with transition probability matrix $P^{\varepsilon} = I + \varepsilon Q$, where Q is the generator of the continuous-time Markov chain $\alpha(t)$, and I is an m-dimensional identity matrix.
3. For the Brownian motion $w(\cdot)$, define $\Delta w_k = w(\varepsilon(k+1)) - w(\varepsilon k)$.
4. Let $\{\tau_n\}$ and $\{\rho_n\}$ be sequences of independent and identically distributed random variables such that τ_n has an exponential distribution with parameter λ for some $\lambda > 0$, and ρ_n has distribution $\pi(\cdot)$. Define

$$v_{n+1} = v_n + \tau_n, \quad \text{with} \quad v_0 = 0.$$

5. Define the event $A_n^J = \{\varepsilon(n-1) < v_j \leq \varepsilon n \ \text{ for some } j\}$, and write the corresponding indicator function as $1_{A_n^J}$. Construct the approximation x_{n+1} by

$$x_{n+1} = x_n + \varepsilon f(\alpha_n, x_n) + \sigma(\alpha_n, x_n)\Delta w_n + g(\rho_j, \alpha_n, x_n) 1_{A_n^J}. \tag{3.1}$$

Remark 3.1. For construction of a discrete-time Markov chain, see [22, pp. 315–316]. The discrete-time Markov chain constructed is an approximation of a discretization obtained from $\alpha(t)$. In fact, we could define $\beta_n = \alpha(\varepsilon n)$ for any positive integer n. It is easily verified that the process so defined is a discrete-time Markov chain, whose transition probability matrix is given by $\exp(\varepsilon Q)$. This Markov chain has stationary transition probabilities or it is a time-homogeneous chain. The process β_n is known as a skeleton process in the literature [5]. One of the advantages of using a constant step size for the numerical procedure is that the skeleton process has stationary transition probabilities not depending on time, so it is easier to generate than that of a non-stationary process.

In the algorithm, we have used another fold of approximation, namely, using $I + \varepsilon Q$ in lieu of $\exp(\varepsilon Q)$ for the transition matrix. This further simplifies the computation and reduces the complexity in calculating $\exp(\varepsilon Q)$. Intuitively, the discrete-time Markov chain we are constructing can be considered as one whose transition probability matrix is obtained from that of β_n by a truncated Taylor expansion. This approximation makes sense since we can invoke the results in [24] to show that an interpolated process of α_n with interpolation interval $[\varepsilon n, \varepsilon n + \varepsilon)$ converges weakly to $\alpha(\cdot)$ generated by Q.

An equivalent way of writing equation (3.1) is:

$$x_{n+1} = x_0 + \varepsilon \sum_{k=0}^{n} f(\alpha_k, x_k) + \sum_{k=0}^{n} \sigma(\alpha_k, x_k)\Delta w_k + \sum_{v_j \le \varepsilon n} g\Big(\rho_j, \alpha_{\lfloor v_j/\varepsilon \rfloor}, x_{\lfloor v_j/\varepsilon \rfloor}\Big), \tag{3.2}$$

where $\lfloor y \rfloor$ denotes the integer part of a real number y. Since $w(\cdot)$ has independent increments, $\{\Delta w_k\}$ is a sequence of independent and identically distributed random variables with mean 0 and covariance εI. It follows from the independence assumption, $\{\Delta w_k\}$, $\{\alpha_k\}$, $\{\tau_n\}$ and $\{\rho_n\}$ are also independent.

Note that the sequences $\{\tau_n\} = \{v_{n+1} - v_n\}$ and $\{\rho_n\}$ are as in those discussed in the last session. The process v_n, in fact, represents the jump times of the underlying process. For convenience, with a slight abuse of notion, we often omit the floor function notation and write, for instance, $\lfloor v_j/\varepsilon \rfloor$ simply as v_j/ε henceforth.

4. Convergence

In this section, we demonstrate the convergence of the algorithm. To obtain the desired convergence, we take a continuous-time interpolation. We then focus on the interpolated process. The work is naturally divided into two parts. In the first part, we show that the interpolated process is tight, and in the second part, we characterize the limit process via martingale problem formulation [8,11]. For future use, define

$$J_n = \sum_{v_j \le \varepsilon n} g(\rho_j, \alpha_{v_j/\varepsilon}, x_{v_j/\varepsilon}). \tag{4.1}$$

To proceed, let us define the interpolated processes via piecewise constant interpolations as:

$$\left.\begin{aligned} x^{\varepsilon}(t) &= x_n, \\ \alpha^{\varepsilon}(t) &= \alpha_n, \\ w^{\varepsilon}(t) &= \textstyle\sum_{k=0}^{t/\varepsilon - 1} \Delta w_k, \\ J^{\varepsilon}(t) &= J_n, \end{aligned}\right\} t \in [\varepsilon n, \varepsilon n + \varepsilon). \tag{4.2}$$

Again we use the convention that t/ε denotes the integer part of t/ε here and hereafter. Before proceeding further, we state a lemma.

Lemma 4.1. The following results hold.
(a) $\alpha^{\varepsilon}(\cdot)$ converges weakly to $\alpha(\cdot)$, the Markov chain with generator Q.
(b) $w^{\varepsilon}(\cdot)$ converges weakly to $w(\cdot)$, the standard $\mathbb{R}^r$-valued Brownian motion.

Proof The proof of part (a) can be obtained by an application of [24, Theorem 3.1], whereas part (b) is a standard functional central limit theorem (see [2,8,13] among others). □

4.1 Compactness

It is well known that in a complete separable metric space, tightness is equivalent to sequential compactness. Our objective is to establish the desired compactness. We will show that the pair $\{\alpha^{\varepsilon}(\cdot), x^{\varepsilon}(\cdot)\}$ is tight in $D([0,T]; \mathcal{M} \times \mathbb{R}^r)$, the space of functions that are defined on $[0,T]$ taking values in $\mathcal{M}\times\mathbb{R}^r$, and that are right continuous and have left limits endowed with the Skorohod topology. The result is recorded next followed by the proof.

THEOREM 4.2. Suppose that all the conditions of Lemma 2.3 are satisfied. Then $\{\alpha^{\varepsilon}(\cdot), x^{\varepsilon}(\cdot)\}$ is tight in $D([0,T];\mathcal{M}\times\mathbb{R}^r)$.

Proof We first note that in view of Lemma 4.1 (a), $\{\alpha^{\varepsilon}(\cdot)\}$ is tight. Thus owing to [2, Theorem 7.7, p. 48], it suffices to prove the tightness of $\{x^{\varepsilon}(\cdot)\}$.

Next, we claim that the following *a priori* estimate

$$\sup_{0\le n\le T/\varepsilon} E|x_n|^2 \le K < \infty \tag{4.3}$$

holds. Here and henceforth, K represents a generic positive constant, whose values may be different for different usages. Thus, the convention $K+K=K$ and $KK=K$ is used for notational simplicity. Note that the remark in [11, p. 39] yields that

$$E|J_n|^2 \le K, \quad \text{and} \quad E|J_{t/\varepsilon}|^2 \le K.$$

Denote the expectation conditioned on $\mathcal{F}_k$ by E_k. Therefore, (3.2) leads to

$$\begin{aligned}
E|x_{n+1}|^2 &\le K\left[E|x_0|^2 + E\left|\varepsilon\sum_{k=0}^{n} f(\alpha_k,x_k)\right|^2 + E\left|\sum_{k=0}^{n}\sigma(\alpha_k,x_k)\Delta w_k\right|^2 + E|J_n|^2\right]\\
&\le K + K\varepsilon^2\sum_{k=0}^{n}\sum_{k_1=0}^{n} Ef'(\alpha_k,x_k)f(\alpha_{k_1},x_{k_1}) + K\sum_{k=0}^{n} E\sigma'(\alpha_k,x_k)\sigma(\alpha_x,x_k)E_k\Delta w_k'\Delta w_k\\
&\le K + K\varepsilon^2\sum_{k=0}^{n}\sum_{k_1=0}^{n}(1+E|x_k||x_{k_1}|) + K\varepsilon\sum_{k=0}^{n} E|\sigma(\alpha_k,x_k)|^2\\
&\le K + K\varepsilon\sum_{k=0}^{n} E|x_k|^2.
\end{aligned} \tag{4.4}$$

In the above, from the second line to the third line, we used the linear growth of $f(i,\cdot)$ and $\sigma(i,\cdot)$, and the familiar inequality $ab\le(a^2+b^2)/2$ for two real numbers a and b. An application of the Gronwall's inequality to equation (4.4) then yields

$$E|x_{n+1}|^2 \le K\exp(K\varepsilon n) \le K.$$

Moreover, the above inequality also holds uniformly in n, so the desired second moment bound equation (4.3) follows.

To proceed, in view of equations (3.2) and (4.2), for any $\delta>0$, $t>0$, $0<s<\delta$, and $t+s\le T$, a direct calculation leads to

$$x^{\varepsilon}(t+s) - x^{\varepsilon}(t) = \varepsilon\sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1} f(\alpha_k,x_k) + \sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1}\sigma(\alpha_k,x_k)\Delta w_k + [J_{(t+s)/\varepsilon} - J_{t/\varepsilon}]. \tag{4.5}$$

As a result,

$$
\begin{aligned}
&E|x^{\varepsilon}(t+s)-x^{\varepsilon}(t)|^2\\
&\leq K\left[E\left|\varepsilon\sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1}f(\alpha_k,x_k)\right|^2+E\left|\sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1}\sigma(\alpha_k,x_k)\Delta w_k\right|^2+E|J_{(t+s)/\varepsilon}-J_{t/\varepsilon}|^2\right]\\
&\leq K\varepsilon\sum_{k=t/\varepsilon}^{(t+s)/\varepsilon-1}(1+E|x_k|^2)+KE|J_{(t+s)/\varepsilon}-J_{t/\varepsilon}|^2\\
&\leq K\delta.
\end{aligned}
\tag{4.6}
$$

In the above, we have used equation (4.3), the independence of α_k and Δw_k, and the properties of the Brownian motion together with the estimate (see [11, p. 39])

$$E\left|\int_t^{t+s}\int_{\Gamma}g(\gamma,\alpha^{\varepsilon}(u),x^{\varepsilon}(u^-))\pi(\mathrm{d}\gamma)\,\mathrm{d}u\right|=O(\delta),$$

so

$$E|J_{(t+s)/\varepsilon}-J_{t/\varepsilon}|^2=O(\delta).$$

It follows from equation (4.6),

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}E|x^{\varepsilon}(t+s)-x^{\varepsilon}(t)|^2=0.$$

Thus, the tightness criterion implies that $\{x^{\varepsilon}(\cdot)\}$ is tight (see [13, Chapter 7]). □

Since $\{\alpha^{\varepsilon}(\cdot),\ x^{\varepsilon}(\cdot)\}$ is tight, by the Prohorov's Theorem, we can extract convergent subsequences. Select such a subsequence and still denote it by $\{\alpha^{\varepsilon}(\cdot),\ x^{\varepsilon}(\cdot)\}$ for notational simplicity. Denote the limit by $(\alpha(\cdot),\ x(\cdot))$. By Skorohod representation, we may assume without loss of generality that $(\alpha^{\varepsilon}(\cdot),\ x^{\varepsilon}(\cdot))$ converges to $(\alpha(\cdot),\ x(\cdot))$ w.p.1, and the convergence is uniform on any bounded interval. We proceed to characterize the limit process.

4.2 Martingale problem formulation

The aim of this section is to characterize $(\alpha(\cdot),\ x(\cdot))$, the limit of $(\alpha^{\varepsilon}(\cdot),x^{\varepsilon}(\cdot))$. We record the result as follows.

THEOREM 4.3. Under the conditions of Theorem 4.2, $(\alpha^{\varepsilon}(\cdot),x^{\varepsilon}(\cdot))$ converges weakly to $(\alpha(\cdot),\ x(\cdot))$ as $\varepsilon\to 0$ such that $(\alpha(\cdot),\ x(\cdot))$ is the solution of the martingale problem with operator $\mathcal{L}$.

The main line of work in what follows is devoted to showing that for each $i\in\mathcal{M}$, and for any $\varphi(i,\cdot)\in C_0^2$,

$$M_{\varphi}(t)=\varphi(\alpha(t),x(t))-\varphi(\alpha,x)-\int_0^t\mathcal{L}\varphi(\alpha(s),x(s))\mathrm{d}s \text{ is a martingale.} \tag{4.7}$$

This is realized via a series of approximations that directly averaged out the unwanted terms.

Proof The claim (4.7) is verified if for each $i \in \mathcal{M}$, any bounded and continuous function $h_j(i,\cdot)$, any positive integer κ, any $0 < t_j \le t$ with $j \le \kappa$, $s > 0$, and $t+s \le T$,

$$E\prod_{j=1}^{\kappa} h_j(\alpha(t_j),x(t_j))\left[\varphi(\alpha(t+s),x(t+s)) - \varphi(\alpha(t),x(t)) - \int_t^{t+s} \mathcal{L}\varphi(\alpha(u),x(u))\mathrm{d}u\right] = 0. \quad (4.8)$$

The underlying problem now reduces to verify equation (4.8). We proceed to verify this via the use of the pair $(\alpha^{\varepsilon}(\cdot), x^{\varepsilon}(\cdot))$.

Step 1: We begin by choosing a sequence of integers $\{m_\varepsilon\}$ satisfying $m_\varepsilon \to \infty$ as $\varepsilon \to 0$ but $\Delta_\varepsilon = \varepsilon m_\varepsilon \to 0$. Again, to avoid the use of the floor function notation, assume that t and s are both constant multiples of Δ_ε. Define the index set

$$\chi_l^{\varepsilon} = \{k : lm_\varepsilon \le k \le lm_\varepsilon + m_\varepsilon - 1\}.$$

Then

$$\begin{aligned}
&\varphi(\alpha^{\varepsilon}(t+s), x^{\varepsilon}(t+s)) - \varphi(\alpha^{\varepsilon}(t), x^{\varepsilon}(t)) \\
&\quad = \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \left[\varphi\big(\alpha_{lm_\varepsilon+m_\varepsilon}, x_{lm_\varepsilon+m_\varepsilon}\big) - \varphi\big(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon+m_\varepsilon}\big)\right] \\
&\qquad + \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \left[\varphi\big(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon+m_\varepsilon}\big) - \varphi\big(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon}\big)\right].
\end{aligned} \quad (4.9)$$

Step 2: For $k \in \chi_l^{\varepsilon}$, denote $\hat{I}_k = I_{\{\varepsilon(k-1) < v_j \le \varepsilon k \quad \text{with} \quad lm_\varepsilon \le k < lm_\varepsilon + m_\varepsilon\}}$. Using equation (3.1), note that

$$\begin{aligned}
&\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^{\varepsilon}} \varphi_x^{\prime}(\alpha_{lm_\varepsilon}, x_k) g(\rho_j, \alpha_k, x_k) 1_{A_k^J} \\
&\quad = \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^{\varepsilon}} \left[\varphi(\alpha_{lm_\varepsilon}, x_k + g(\rho_j, \alpha_k, x_k)) - \varphi(\alpha_{lm_\varepsilon}, x_k)\right]\hat{I}_k + o(1),
\end{aligned}$$

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

$$\begin{aligned}
&\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \left[\varphi(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon+m_\varepsilon}) - \varphi(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon})\right] \\
&\quad = \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^{\varepsilon}} \left[\varphi(\alpha_{lm_\varepsilon}, x_{k+1}) - \varphi(\alpha_{lm_\varepsilon}, x_k)\right],
\end{aligned}$$

we obtain

$$
\begin{aligned}
&\lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \left[\varphi(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon+m_\varepsilon}) - \varphi(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon})\right] \\
&= \lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \Bigg[\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \varepsilon \sum_{k:k\in\chi_l^\varepsilon} \varphi_x'(\alpha_{lm_\varepsilon}, x_k) f(\alpha_k, x_k) \\
&+ \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varphi_x'(\alpha_{lm_\varepsilon}, x_k)\sigma(\alpha_k, x_k)\Delta w_k \\
&+ \frac{1}{2}\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varepsilon^2 f'(\alpha_k, x_k)\varphi_{xx}(\alpha_{lm_\varepsilon}, x_k) f(\alpha_k, x_k) \\
&+ \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varepsilon f'(\alpha_k, x_k)\varphi_{xx}(\alpha_{lm_\varepsilon}, x_k)\sigma(\alpha_k, x_k)\Delta w_k \\
&+ \frac{1}{2}\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \Delta w_k'\sigma'(\alpha_k, x_k)\varphi_{xx}(\alpha_{lm_\varepsilon}, x_k)\sigma(\alpha_k, x_k)\Delta w_k \\
&+ \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \int_0^1 \widetilde{\Delta x}_k'[\varphi_{xx}(\alpha_{lm_\varepsilon}, x_k + u\widetilde{\Delta x_k}) - \varphi_{xx}(\alpha_{lm_\varepsilon}, x_k)]\widetilde{\Delta x}_k \,\mathrm{d}u \\
&+ \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} [\varphi(\alpha_{lm_\varepsilon}, x_k + g(\rho_j, \alpha_k, x_k)) - \varphi(\alpha_{lm_\varepsilon}, x_k)]\hat{I}_k\Bigg],
\end{aligned}
$$

where

$$\widetilde{\Delta x}_k = \varepsilon f(\alpha_k, x_k) + \sigma(\alpha_k, x_k)\Delta w_k.$$

Since $\phi_{xx}(i,\cdot)$ is continuous with compact support, the weak convergence of $(\alpha^{\varepsilon}(\cdot), x^{\varepsilon}(\cdot))$, the Skorohod representation, and the boundedness of $h_j(\cdot)$ then yield

$$
\begin{aligned}
&E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \int_0^1 \widetilde{\Delta x}_k'[\varphi_{xx}(\alpha_{lm_\varepsilon}, x_k + u\widetilde{\Delta x}_k) \\
&\quad - \varphi_{xx}(\alpha_{lm_\varepsilon}, x_k)]\widetilde{\Delta x}_k \mathrm{d}u \to 0 \text{ as } \varepsilon\to 0. \qquad (4.10)
\end{aligned}
$$

Thus, it suffices to consider the terms involving $\varphi_x(\alpha_{lm_\varepsilon}, x_k)$, $\varphi_{xx}(\alpha_{lm_\varepsilon}, x_k)$, and the jumps in the expression.

Using the defining relation (3.2) and the continuity of $f(i,\cdot)$ for each $i \in \mathcal{M}$,

$$
\begin{aligned}
&\lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \varepsilon \sum_{k:k\in\chi_l^\varepsilon} \varphi_x'(\alpha_{lm_\varepsilon}, x_k) f(\alpha_k, x_k) \\
&= \lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \varepsilon \sum_{k:k\in\chi_l^\varepsilon} E_{lm_\varepsilon} \sum_{i=1}^{m_0} \varphi_x'(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon}) f(i, x_{lm_\varepsilon}) I_{\{\alpha_k = i\}}. \qquad (4.11)
\end{aligned}
$$

Note that when $\varepsilon \to 0$ and $l\Delta_\varepsilon = \varepsilon l m_\varepsilon \to u$, $l\Delta_\varepsilon + \Delta_\varepsilon \to u$ and for any k satisfying $lm_\varepsilon \le k \le lm_\varepsilon + m_\varepsilon$, $\varepsilon k \to u$. In addition, The weak convergence of $\alpha^\varepsilon(\cdot)$ to $\alpha(\cdot)$ and the Skorohod representation (without changing notation), we may assume that $I_{\{\alpha^\varepsilon(u)=i\}} \to I_{\{\alpha(u)=i\}}$ w.p.1. As a result, by equation (4.11),

$$\begin{aligned}
&\lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \varepsilon \sum_{k:k\in\chi_l^\varepsilon} E_{lm_\varepsilon} \sum_{i=1}^{m_0} \varphi_x'(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon}) f(i, x_{lm_\varepsilon}) I_{\{\alpha_k = i\}} \\
&\quad = \lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j)) \sum_{i=1}^{m_0} \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \Delta_\varepsilon \varphi_x'(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon}) f(i, x_{lm_\varepsilon}) \\
&\qquad \times E_{lm_\varepsilon} \frac{1}{m_\varepsilon} \sum_{k:k\in\chi_l^\varepsilon} I_{\{\alpha^\varepsilon(\varepsilon k)=i\}} \\
&\quad = \lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j)) \sum_{i=1}^{m_0} \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \Delta_\varepsilon \varphi_x'(\alpha^\varepsilon(l\Delta_\varepsilon), x^\varepsilon(l\Delta_\varepsilon)) f(i, x^\varepsilon(l\Delta_\varepsilon)) \\
&\qquad \times E_{lm_\varepsilon} \frac{1}{m_\varepsilon} \sum_{k:k\in\chi_l^\varepsilon} I_{\{\alpha^\varepsilon(\varepsilon k)=i\}} \\
&\quad = E\prod_{j=1}^{\kappa} h(\alpha(t_j), x(t_j)) \sum_{i=1}^{m_0} \int_t^{t+s} \varphi_x'(\alpha(u), x(u)) f(i, x(u)) I_{\{\alpha(u)=i\}}\, \mathrm{d}u \\
&\quad = E\prod_{j=1}^{\kappa} h(\alpha(t_j), x(t_j)) \int_t^{t+s} \varphi_x'(\alpha(u), x(u)) f(\alpha(u), x(u))\, \mathrm{d}u.
\end{aligned}$$

By virtue of the independent of Δw_k with α_{lm_ε} and x_k, and noting that $E_k \Delta w_k = 0$,

$$\begin{aligned}
&E\prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varphi_x'(\alpha_{lm_\varepsilon}, x_k) \sigma(\alpha_{lm_\varepsilon}, x_k) \Delta w_k \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} E_{lm_\varepsilon} \varphi_x'(\alpha_{lm_\varepsilon}, x_k) \sigma(\alpha_{lm_\varepsilon}, x_k) E_k \Delta w_k = 0. \qquad (4.12)
\end{aligned}$$

Likewise,

$$E\prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varepsilon f'(\alpha_{lm_\varepsilon}, x_k) \varphi_{xx}(\alpha_{lm_\varepsilon}, x_k) \sigma(\alpha_{lm_\varepsilon}, x_k) \Delta w_k = 0. \qquad (4.13)$$

Using the linear growth of $f(i,\cdot)$, the boundedness of $\phi_{xx}(i,\cdot)$, and the *a priori* estimate (4.3), we have

$$\begin{aligned}
&E\left| \prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varepsilon^2 f'(\alpha_{lm_\varepsilon}, x_k) \varphi_{xx}(\alpha_{lm_\varepsilon}, x_k) f(\alpha_{lm_\varepsilon}, x_k) \right| \\
&\quad \le K \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varepsilon^2 E|f(\alpha_{lm_\varepsilon}, x_k)|^2 |\varphi_{xx}(\alpha_{lm_\varepsilon}, x_k)| \\
&\quad \le K \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \varepsilon^2 (1 + E|x_k|^2) \\
&\quad \le O(\varepsilon) \to 0 \ \text{ as } \ \varepsilon \to 0. \qquad (4.14)
\end{aligned}$$

Recall that K is a generic positive constant; its values may vary for different usage.

Next, using the independent increment property of the Brownian motion and inserting conditional expectations yield

$$
\begin{aligned}
&E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j),x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon-1} \sum_{k:k\in\chi_l^\varepsilon} \Delta w_k' \sigma'(\alpha_k,x_k)\varphi_{xx}(\alpha_{lm_\varepsilon},x_k)\sigma(\alpha_k,x_k)\Delta w_k \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j),x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon-1} \sum_{k:k\in\chi_l^\varepsilon} E_{lm_\varepsilon}\mathrm{tr}[\varphi_{xx}(\alpha_{lm_\varepsilon},x_k)\sigma(\alpha_k,x_k)\sigma'(\alpha_k,x_k)E_k\Delta w_k\Delta w_k'] \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j),x^{\varepsilon}(t_j))\varepsilon \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon-1} \sum_{k:k\in\chi_l^\varepsilon} E_{lm_\varepsilon}\mathrm{tr}[\varphi_{xx}(\alpha_{lm_\varepsilon},x_k)\sigma(\alpha_k,x_k)\sigma'(\alpha_k,x_k)] \\
&\quad \to E\prod_{j=1}^{\kappa} h_j(\alpha(t_j),x(t_j))\int_t^{t+s} \mathrm{tr}[\varphi_{xx}(\alpha(u),x(u))\sigma(\alpha(u),x(u))\sigma'(\alpha(u),x(u))]\mathrm{d}u, \text{as } \varepsilon\to 0. \quad (4.15)
\end{aligned}
$$

Step 3: Now, we come back to the term on the second line of equation (4.9). Using the continuity of $\varphi(i,\cdot)$ for each $i\in\mathcal{M}$, we can argue along the same line as in Step 2 that the limit of

$$
\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon-1} [\varphi(\alpha_{lm_\varepsilon+m_\varepsilon},x_{lm_\varepsilon+m_\varepsilon}) - \varphi(\alpha_{lm_\varepsilon},x_{lm_\varepsilon+m_\varepsilon})]
$$

is the same as that of

$$
\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon-1} [\varphi(\alpha_{lm_\varepsilon+m_\varepsilon},x_{lm_\varepsilon}) - \varphi(\alpha_{lm_\varepsilon},x_{lm_\varepsilon})] + \eta_\varepsilon,
$$

where $\eta_\varepsilon\to 0$ in probability uniformly in t. Moreover,

$$
\begin{aligned}
&\sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon-1} [\varphi(\alpha_{lm_\varepsilon+m_\varepsilon},x_{lm_\varepsilon}) - \varphi(\alpha_{lm_\varepsilon},x_{lm_\varepsilon})] \\
&\quad = \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon-1} \sum_{k:k\in\chi_l^\varepsilon} [\varphi(\alpha_{k+1},x_{lm_\varepsilon}) - \varphi(\alpha_k,x_{lm_\varepsilon})].
\end{aligned}
$$

Recall that for the Markov chain $\{\alpha_k\}$, the transition probability matrix is given by $P^\varepsilon = I+\varepsilon Q$. Define

$$
\Phi(x) = (\varphi(1,x),\ldots,\varphi(m_0,x))' \in \mathbb{R}^{m_0\times 1}, \quad \text{and} \quad \Psi(\alpha) = (I_{\{\alpha=1\}},\ldots,I_{\{\alpha=m_0\}}) \in \mathbb{R}^{1\times m_0}.
$$

Then we have

$$
\begin{aligned}
&E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} [\varphi(\alpha_{lm_\varepsilon+m_\varepsilon}, x_{lm_\varepsilon}) - \varphi(\alpha_{lm_\varepsilon}, x_{lm_\varepsilon})] \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} [\varphi(\alpha_{k+1}, x_{lm_\varepsilon}) - \varphi(\alpha_k, x_{lm_\varepsilon})] \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \sum_{i=1}^{m_0} E_k[\varphi(\alpha_{k+1}, x_{lm_\varepsilon}) - \varphi(i, x_{lm_\varepsilon})] I_{\{\alpha_k = i\}} \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \sum_{i=1}^{m_0} \sum_{j=1}^{m_0} [\varphi(j, x_{lm_\varepsilon}) P(\alpha_{k+1} = j | \alpha_k = i) \\
&\qquad - \varphi(i, x_{lm_\varepsilon})] I_{\{\alpha_k = i\}} \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} \Psi(\alpha_k)(P^{\varepsilon} - I)\Phi(x_{lm_\varepsilon}) \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \varepsilon \sum_{k:k\in\chi_l^\varepsilon} \Psi(\alpha_k) Q \Phi(x_{lm_\varepsilon}) \\
&\quad \to E\prod_{j=1}^{\kappa} h_j(\alpha(t_j), x(t_j)) \int_t^{t+s} Q\varphi(\cdot, x(u))(\alpha(u))\mathrm{d}u \ \text{ as } \ \varepsilon \to 0.
\end{aligned}
\tag{4.16}
$$

Step 4: Work with the jump term. We have

$$
\begin{aligned}
&\lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \sum_{l=t/\Delta_\varepsilon}^{(t+s)/\Delta_\varepsilon - 1} \sum_{k:k\in\chi_l^\varepsilon} [\varphi(\alpha_{lm_\varepsilon}, x_k + g(\rho_j, \alpha_k, x_k)) - \varphi(\alpha_{lm_\varepsilon}, x_k)]\hat{I}_k \\
&\quad = \lim_{\varepsilon\to 0} E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j)) \int_t^{t+s} \lambda \int_\Gamma [\varphi(\alpha^{\varepsilon}(u), x^{\varepsilon}(u^-) + g(\gamma, \alpha^{\varepsilon}(u), x^{\varepsilon}(u^-))) \\
&\qquad - \varphi(\alpha^{\varepsilon}(u), x^{\varepsilon}(u^-))]\pi(\mathrm{d}\gamma)\mathrm{d}u \\
&\quad = E\prod_{j=1}^{\kappa} h_j(\alpha(t_j), x(t_j)) \int_t^{t+s} \lambda \int_\Gamma [\varphi(\alpha(u), x(u^-) + g(\gamma, \alpha(u), x(u^-))) \\
&\qquad - \varphi(\alpha(u), x(u^-))]\pi(\mathrm{d}\gamma)\mathrm{d}u.
\end{aligned}
\tag{4.17}
$$

Step 5: Now rework with the term

$$
E\prod_{j=1}^{\kappa} h_j(\alpha^{\varepsilon}(t_j), x^{\varepsilon}(t_j))[\varphi(\alpha^{\varepsilon}(t+s), x^{\varepsilon}(t+s)) - \varphi(\alpha^{\varepsilon}(t), x^{\varepsilon}(t))].
$$

By virtue of the weak convergence of $(\alpha^\varepsilon(\cdot), x^\varepsilon(\cdot))$, the Skorohod representation, and the continuity of $h_j(i, \cdot)$ and $\varphi(i, \cdot)$, we have

$$E\prod_{j=1}^{\kappa} h_j(\alpha^\varepsilon(t_j), x^\varepsilon(t_j))[\varphi(\alpha^\varepsilon(t+s), x^\varepsilon(t+s)) - \varphi(\alpha^\varepsilon(t), x^\varepsilon(t))]$$
$$\to \prod_{j=1}^{\kappa} h_j(\alpha(t_j), x(t_j))[\varphi(\alpha(t+s), x(t+s)) - \varphi(\alpha(t), x(t))] \text{ as } \varepsilon \to 0. \quad (4.18)$$

In view of Steps 1–5, using equations (4.9)–(4.18), equation (4.8) is verified. Thus the desired result follows and the proof of the theorem is completed. □

5. Examples

In this section, we provide several examples for demonstration. We shall work with a fixed time $T > 0$. In view of our algorithm (3.1) and the system given in equation (1.1), we compare the computed value $x_{T/\varepsilon}$ with $x(T)$ the solution of equation (1.1). We are mainly interested in the approximation errors and the variances. The numerical experiments were done using MATLAB on a desktop computer equipped with WinXP.

Example 5.1. We begin by examining the system

$$\mathrm{d}x(t) = \mu x(t)\mathrm{d}t + \sigma(\alpha(t))x(t)\mathrm{d}w(t) + \mathrm{d}J(t), \quad x(0) = x_0, \quad (5.1)$$

where x_0 is assumed to be non-random, $\mu \in \mathbb{R}$ is a constant does not depend on the Markov chain, and $J(t)$ is of the simple form $J(t) = \sum_{j=0}^{N(t)} \zeta_j$, with $N(t)$ being a Poisson counting process, and $\{\zeta_k\}$ being a sequence of i.i.d. random variables.

The motivation for examining a model of the form (5.1) stems from the so-called risk-neutral consideration. In fact, beginning with a regime-switching model, one may derive a suitable probability space upon which the expected rate of return of all securities is equal to the risk-free interest rate; see [19]. One of the advantages is that the excepted value $\mathrm{E}x(T)$ can be computed explicitly.

To compute the deviation and the variance, we take n_0 batches each of which has $\lfloor 1/\varepsilon \rfloor$ replications. Denoting by $x_{T/\varepsilon}^{\ell,j}$ the approximated value computed in accordance with equation (3.2) of jth replication within the ℓth batch, consider

$$\hat{x}^{\ell,\varepsilon} = \varepsilon \sum_{j=1}^{\lfloor 1/\varepsilon \rfloor} x_{T/\varepsilon}^{\ell,j} - \mathrm{E}x(T), \quad 1 \le \ell \le n_0, \quad \hat{x}^\varepsilon = \frac{1}{n_0}\sum_{\ell=1}^{n_0} \hat{x}^{\ell,\varepsilon}. \quad (5.2)$$

Next define $z^{\ell,\varepsilon} = \hat{x}^{\ell,\varepsilon}/\sqrt{\varepsilon}$. It is easily shown that as $\varepsilon \to 0$, $z^{\lambda,\varepsilon}$ converges in distribution to a normal random variable. As a result, for practical purposes, we may regard $z^\varepsilon = \hat{x}^\varepsilon/\sqrt{\varepsilon}$ as a random variable following the t-distribution. Based on this t-distribution, we can then construct confidence intervals for the estimate of the mean deviation of the trajectories of solutions, [10].

To be more specific, let $\mu = 1.5$, $\alpha(t) \in \{1, 2\}$, $\sigma(1) = 0.15$, $\sigma(2) = 0.05$, $\{\rho_k\}$ is a sequence of i.i.d. random variables, $J(t) = \sum_{k=1}^{N(t)} \zeta_k$ with $\zeta_k = g(\rho_k)$, and $g(\rho) = 0.01\rho$,

$$\pi(x) = \begin{cases} 0.25 & \text{if } x = 1 \\ 0.5 & \text{if } x = 2 \\ 0.25 & \text{if } x = 3 \end{cases},$$

$\alpha(t)$ is Markov Chain, with state space $\mathcal{M} = \{1, 2\}$, and generator

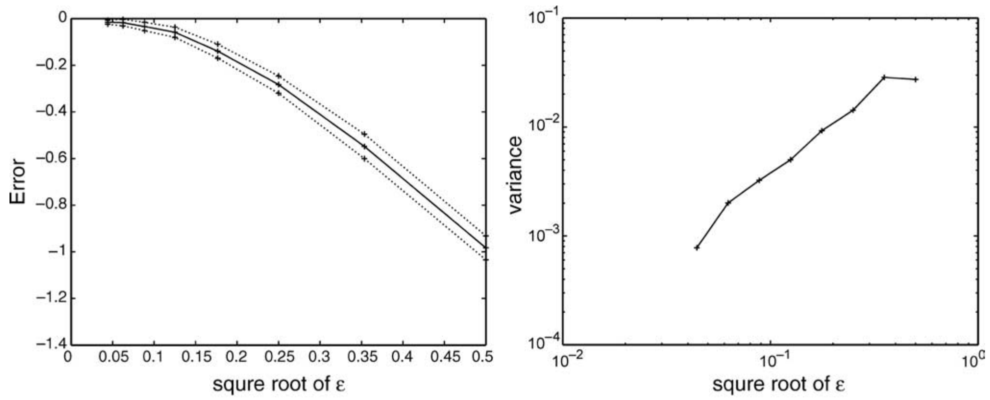
$$Q = \begin{pmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

Let $\lambda = 4$ and τ_n be exponentially distributed with mean $1/\lambda$. As can be seen from the computation results, the iterates behaves well. Both the mean deviation and variance of the errors decrease as ε decreases; see figure 1 and table 1.

Example 5.2. Consider

$$\mathrm{d}x(t) = \mu(\alpha(t))x(t)\mathrm{d}t + \sigma(\alpha(t))x(t)\mathrm{d}w(t) + \mathrm{d}J(t). \tag{5.3}$$

The added complexity is that the drift $\mu = \mu(\alpha)$ is also a function of the Markov chain. However, as in the previous example, the desired computation can be carried out. For the numerical experiment, we use $\mu(1) = 1.5968$, $\mu(2) = -1.4116$, $\sigma(1) = 0.44$, $\sigma(2) = 0.63$ and $g(\rho) = 0.1\rho$. Since, it is difficult to find the explicit expression of the mean, we use its sample mean instead. As can be seen from the numerical result, both the mean deviation and the variance decrease as ε gets smaller; see figure 2 and table 2.



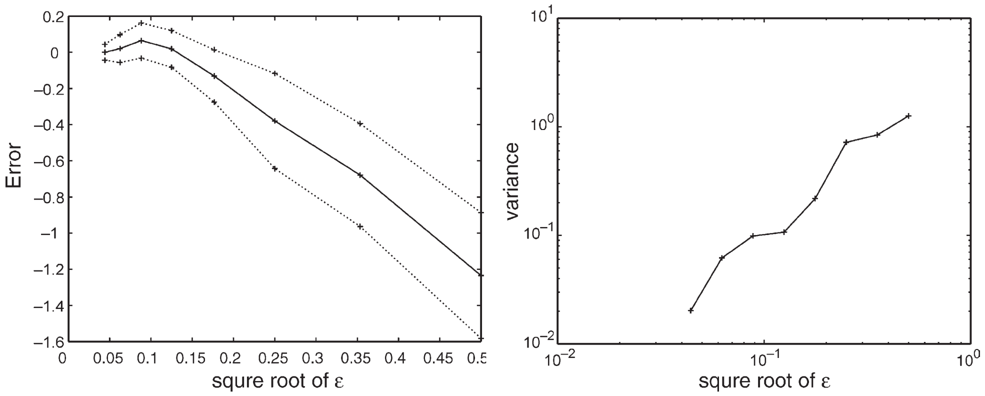
(a) Mean Error vs Step Size $\varepsilon > 0$, and 90% Confidence Interval

(b) log log Chart: Variance vs $\sqrt{\varepsilon}$

Figure 1. Linear system with drift independent of the Markov chain and Markov dependent variance.

Table 1. Variance of error vs. ε and $\sqrt{\varepsilon}$.

ε	$\sqrt{\varepsilon}$	*Variance of error*
0.2500	0.5000	0.0273
0.0156	0.1250	0.0050
0.0020	0.0442	0.0008



(a) Mean Error vs Step Size $\varepsilon > 0$, and 90% Confidence Interval

(b) log log Chart: Variance vs $\sqrt{\varepsilon}$

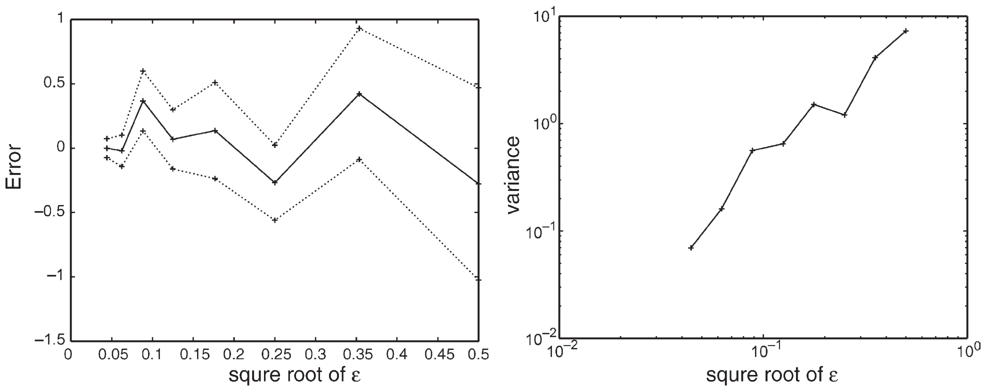
Figure 2. Linear system with Markovian dependent drift and variance.

Table 2. Variance of error vs. ε and $\sqrt{\varepsilon}$.

ε	$\sqrt{\varepsilon}$	*Variance of error*
0.2500	0.5000	1.2557
0.0156	0.1250	0.1069
0.0020	0.0442	0.0203

Example 5.3. In this example, we treat a nonlinear system as given in equation (1.1). We specify the functions as follows: $f(1,x) = -\sqrt{x}\ln(x)$, $f(2,x) = \sqrt[3]{x}\ln(x)$, $\sigma(1,x) = x$, $\sigma(2,x) = 2x$, and $g(\rho) = 0.5\rho$. Use the same Markov Chain $\alpha(t)$ and the same jump process as before. Due to the inherent nonlinearity, the numerical result is not as good as in the linear cases.

Nevertheless, the behavior of the mean deviation and variance still decrease as ε gets smaller; see figures 3 and table 3.



(a) Mean Error vs Step Size $\varepsilon > 0$, and 90% Confidence Interval

(b) log log Chart: Variance vs $\sqrt{\varepsilon}$

Figure 3. Nonlinear system with Markovian dependent drift and variance.

Table 3. Variance of error vs. ε and $\sqrt{\varepsilon}$.

ε	$\sqrt{\varepsilon}$	*Variance of error*
0.2500	0.5000	5.7999
0.0156	0.1250	0.5498
0.0020	0.0442	0.0560

6. Further remarks

One of the variations of the algorithm studied in the paper is to use a sequence of decreasing step sizes. The modifications are as follows. Let $\{\varepsilon_n\}$ be a sequence of nonnegative real numbers such that $\varepsilon_n \to 0$ and $\sum_{n=0}^{\infty}\varepsilon_n = \infty$. For example, we may take $\varepsilon_n = 1/n$, or $\varepsilon_n = 1/n^{\gamma}$ for some $0 < \gamma < 1$. Define

$$t_n = \sum_{l=0}^{n-1} \varepsilon_l, \quad \text{and} \quad \alpha_n = \alpha(t_n), \qquad n \geq 0.$$

It then can be verified that $\{\alpha_n\}$ so defined is a discrete-time Markov chain with transition probability matrix

$$P^{n,n+1} = \left(p_{ij}^{n,n+1}\right)_{m\times m} = \exp((t_{n+1} - t_n)Q) = \exp(\varepsilon_n Q). \tag{6.1}$$

Using the ideas in stochastic approximation, define

$$m(t) = \max\{n : t_n \leq t\}. \tag{6.2}$$

It is clear that now the Markov chain α_n is not time homogeneous. The approximate solution for the SDE with jumps and regime switching equation (1.2) is given by

$$\begin{cases} x_{n+1} = x_0 + \sum_{k=0}^{n} f(x_k, \alpha_k)\varepsilon_k + \sum_{k=0}^{n} g(x_k, \alpha_k)\Delta w_k + \sum\limits_{v_j \leq n} g(\rho_j, \alpha_j, x_{j-1}), \\ x_0 = x, \quad \alpha_0 = \alpha, \end{cases} \tag{6.3}$$

where $\Delta w_n = w(t_{n+1}) - w(t_n)$. With the algorithm proposed, we can then proceed to study its performance. The proof is along the same line as that of Section 4.

References

[1] Allen, E.J., Novosel, S.J. and Zhang, Z., 1998, Finite element and difference approximation of some linear stochastic partial differential equations, *Stochastics and Stochastics Reports*, **64**, 117–142.

[2] Billingsley, P., 1968, *Convergence of Probability Measures* (New York: J. Wiley).

[3] Björk, T., 1980, Finite dimensional optimal filters for a class of Ito processes with jumping parameters, *Stochastics*, **4**, 167–183.

[4] Buffington, J. and Elliott, R.J., 2002, American options with regime switching, *International Journal of Theoretical and Applied Finance*, **5**, 497–514.

[5] Chung, K.L., 1967, *Markov Chains with Stationary Transition Probabilities*, 2nd ed. (New York, NY: Springer-Verlag).

[6] Di Masi, G.B., Kabanov, Y.M. and Runggaldier, W.J., 1994, Mean variance hedging of options on stocks with Markov volatility, *Theory of Probability and Applications*, **39**, 173–181.

[7] Dufresne, F. and Gerber, H.U., 1991, Risk theory for the compound Poisson process that is perturbed by diffusion, *Insurance; Mathematics and Economics*, **10**, 51–59.

[8] Ethier, S.N. and Kurtz, T.G., 1986, *Markov Processes: Characterization and Convergence* (New York: J. Wiley).

[9] Kloeden, P.E. and Platen, E., 1992, *Numerical Solution of Stochastic Differential Equations* (New York: Springer-Verlag).

[10] Kloeden, P.E., Platen, E. and Schurz, H., 1994, *Numerical Solution of SDE through Computer Experiments* (Berlin: Springer-Verlag).

[11] Kushner, H.J., 1990, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems* (Boston: Birkhäuser).

[12] Kushner, H.J. and Dupuis, P., 1992, *Numerical Methods for Stochastic Control Problems in Continuous Time* (New York: Springer).

[13] Kushner, H.J. and Yin, G., 2003, *Stochastic Approximation and Recursive Algorithms and Applications* (New York: Springer-Verlag).

[14] Milstein, G.N., 1995, *Numerical Integration of Stochastic Differential Equations* (New York: Kluwer).

[15] Moller, C.M., 1995, Stochastic differential equations for ruin probability, *Journal of Applied Probability*, **32**, 74–89.

[16] Platen, E., 1999, An introduction to numerical methods for stochastic differential equations, *Acta Numerica*, 197–246.

[17] Protter, P. and Talay, D., 1997, The Euler scheme for Levy driven stochastic differential equations, *Annals of Probability*, **25**, 397–423.

[18] Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J., 1999, *Stochastic Processes for Insurance and Finance* (New York: Wiley and Sons).

[19] Yao, D.D., Zhang, Q. and Zhou, X.Y. A regime-switching model for European option pricing, preprint.

[20] Yang, H. and Yin, G., 2003, Ruin probability for a model under Markovian switching regime. *Proc. Workshop on Probability with Applications in Finance and Insurance*.

[21] Yin, G. and Zhang, Q., 1998, *Continuous-Time Markov Chains and Applications: A Singular Perturbation Approach* (New York: Springer-Verlag).

[22] Yin, G. and Zhang, Q., 2005, *Discrete-time Markov Chains: Two-time-scale Methods and Applications* (New York: Springer-Verlag).

[23] Yin, G., Liu, R.H. and Zhang, Q., 2002, Recursive algorithms for stock liquidation: A stochastic optimization approach, *SIAM Journal on Optimization*, **13**, 240–263.

[24] Yin, G., Zhang, Q. and Badowski, G., 2003, Discrete-time singularly perturbed Markov chains: Aggregation, occupation measures, and switching diffusion limit, *Advances in Applied Probability*, **35**, 449–476.

[25] Zhang, Q., 2001, Stock trading: An optimal selling rule, *SIAM Journal on Control and Optimization*, **40**, 64–87.