

# CONVERGENCE OF AMERICAN OPTION VALUES FROM DISCRETE- TO CONTINUOUS-TIME FINANCIAL MODELS<sup>1</sup>

KAUSHIK AMIN

*School of Business Administration  
University of Michigan, Ann Arbor, MI 48109-1234*

AJAY KHANNA

*Stern School of Business Administration  
New York University, New York, New York 01003*

Given a sequence of discrete-time option valuation models in which the sequence of processes defining the state variables converges weakly to a diffusion, we prove that the sequence of American option values obtained from these discrete-time models also converges to the corresponding value obtained from the continuous-time model for the standard models in the finance/economics literature. The convergence proof carries over to the case when the limiting risky asset price process follows a diffusion, except it pays discrete dividends on some fixed dates.

KEY WORDS: weak convergence, American options, martingales

## 1. INTRODUCTION

Since the seminal work of Cox, Ross, and Rubinstein (1979), discrete-time models have become very popular in option valuation as suitable approximations to continuous-time diffusion models. Due to their simplicity, these models provide us with a valuable pedagogic tool. They are also useful computational tools when option prices cannot be represented by simple closed-form solutions. In fact, many interesting results can be obtained when one takes limits of discrete-time models, and much work has been done in this area. Besides Cox, Ross, and Rubinstein (1979), other examples of such discrete-time approximations include Amin (1991), Boyle, Evnine, and Gibbs (1989), He (1990), Madan, Milne, and Shefrin (1989), and Nelson and Ramaswamy (1990).

In many cases cited earlier, the authors show that the prices of European-type options computed from their discrete-time models converge to their corresponding continuous-time model values. Since the sequence of stock price processes converges weakly to the appropriate diffusion process, if the option payoff is a continuous function of the underlying state variables and satisfies a uniform integrability condition, the convergence result is a direct consequence of the continuous mapping theorem. However, since the American option valuation problem involves a control, the corresponding results are not available in the American option case. Further, this problem has so far not been tackled in the finance/economics literature. For example, Nelson and Ramaswamy (1990) state: "Unfortunately, we have not been able to extend this argument (proof of convergence of European option values) to the American case rigorously."

In this paper, our objective is to prove that the convergence result is true even for Ameri-

<sup>1</sup>Helpful conversations with Professor S. R. S. Varadhan are gratefully acknowledged.  
*Manuscript received May 1992; final revision received August 1993.*

can options under a fairly general set of conditions. Our results are applicable in both theoretical work and in practical applications in which discrete-time models are used to approximate continuous-time diffusion models for the purpose of option valuation. The conditions that we require can be easily verified to guarantee that both American- and European-type option values obtained from the discrete-time models converge to their respective continuous-time values.

We provide a convergence proof for the following cases. First, we prove the convergence result when the limiting process is a diffusion process satisfying certain regularity conditions. This case includes most of the standard discrete-time models in the financial economics literature. Second, we show that the convergence result also holds when the underlying continuous-time process is a diffusion, but with the state variables jumping by discrete amounts on certain fixed dates. These fixed dates can be interpreted as dividend payment dates on which the stock price declines by the amount of the dividend. This situation corresponds to Roll's (1977) celebrated American call formula with discrete dividends. Other interpretations corresponding to lumpy information arrival are also possible.

The formal steps necessary to show convergence of the value function in the context of optimal stopping problems have been outlined by Kushner (1977). We follow these steps, trying to use available results and proofs as much as possible. We cannot directly use many of Kushner's results since his results are developed within the context of a specific discretization scheme, and he requires certain assumptions such as boundedness of the drift and dispersion coefficients of the diffusion and boundedness of the payoff functions. These assumptions are typically not satisfied by problems of interest in finance/economics.

## 2. THE FORMAL SETUP

Let  $D_{R^d}[0, T]$  be the space of  $\mathbb{R}^d$ -valued functions which are right continuous with left limits (RCLL) on some interval<sup>2</sup>  $[0, T]$ . Let  $C_{R^d}[0, T]$  be the subspace of  $D_{R^d}[0, T]$  of all  $\mathbb{R}^d$ -valued continuous functions on  $[0, T]$ . Throughout this paper, we assume that the space  $D_{R^d}[0, T]$  is endowed with the Skorokhod topology (see Billingsley 1968). The Skorokhod topology restricted to the space  $C_{R^d}[0, T]$  is the topology of uniform convergence.

A sequence of probability measures  $P_n$  on  $D_{R^d}[0, T]$  is said to converge weakly (converge in distribution) to a probability measure  $P$  on  $D_{R^d}[0, T]$  if  $E_{P_n}f(\cdot)$  converges to  $E_P f(\cdot)$  for every continuous, bounded function on  $D_{R^d}[0, T]$ . Given a sequence of  $d$ -dimensional stochastic processes  $X_n$  with sample paths in  $D_{R^d}[0, T]$ , let  $\mathcal{L}(X_n)$  denote the sequence of probability measures induced on  $D_{R^d}[0, T]$  by  $(X_n)$ . Then, we say that  $X_n$  converges weakly to a  $(D_{R^d}[0, T]$ -valued) process  $X$  if  $\mathcal{L}(X_n)$  converges weakly to  $\mathcal{L}(X)$ , the probability measure induced by  $X$  on  $D_{R^d}[0, T]$ , and we write  $X_n \Rightarrow X$ . Each of these processes  $(X_n)$  need not be defined on the same probability space.

The general setup for this paper is as follows. We assume that we are given a sequence of Markov chains (interpolated to continuous-time processes by defining the process to be constant over the interpolation interval) which satisfy certain conditions guaranteeing weak convergence to a given diffusion limit. Given an American option, we can compute the price of this option for each element of this sequence and for the continuous diffusion. Our task is to show that the sequence of American option prices computed with respect to

<sup>2</sup>Even though we will concentrate on only finite maturity options in this paper, all our proofs are valid as long as the sequence of optimal stopping times for the discrete-time problems are uniformly integrable in the interpolation interval.

the discrete processes converges to the continuous-time American option value determined from the continuous-time diffusion process.

### 2.1. Continuous Setup

Our economic model consists of a continuous trading interval  $[0, T]$  and a probability space  $(\Omega, \mathcal{M}, \mathcal{M}_t, P)$ . On  $(\Omega, \mathcal{M}, \mathcal{M}_t, P)$  is defined a  $d$ -dimensional Brownian motion  $W(t): 0 \leq t \leq T$ . We assume that  $\mathcal{M}_t$  satisfies the usual conditions.

**ASSUMPTION 2.1.** Let  $\mu(t, x): [0, T] \times \mathfrak{R}^d \rightarrow \mathfrak{R}^d$  be a continuous vector-valued function and  $\sigma(t, x): [0, T] \times \mathfrak{R}^d \rightarrow \mathfrak{R}^d \times \mathfrak{R}^d$  be a continuous, matrix-valued function such that  $a(t, x) = \sigma(t, x)\sigma'(t, x)$  is symmetric, nonnegative definite for all  $(t, x) \in [0, T] \times \mathfrak{R}^d$ . We also assume that  $\mu$  and  $\sigma$  satisfy the global Lipschitz and linear growth conditions:

$$(2.1) \quad \|\mu(t, x) - \mu(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq K\|x - y\|,$$

$$(2.2) \quad \|\mu(t, x)\|^2 + \|\sigma(t, x)\|^2 \leq K^2(1 + \|x\|^2)$$

for every  $0 \leq t \leq T, x, y \in \mathfrak{R}^d$ , and some positive constant  $K$ .

Assumption 2.1 guarantees that there exists a strong solution to the  $d$ -dimensional stochastic differential equation

$$(2.3) \quad dX(t) = \mu(t, X(t)) dt + \sigma(t, X(t)) dW(t), \quad 0 \leq t \leq T,$$

where  $X(t): 0 \leq t \leq T$  is a  $d$ -dimensional process with a fixed starting point  $X(0) \in \mathfrak{R}^d$ .

$X(t)$  represents the state variables of interest in the economy. Examples of possible state variables represented by  $X$  include the prices of risky assets, forward prices, futures prices, foreign currency exchange rates, and the values of domestic and foreign spot and forward interest rates.

We permit risky assets to pay a continuous dividend yield. This yield may depend on the path of the state variables and need not be deterministic. Without dividends, American call options on risky assets will not be exercised early; hence, these options are equivalent to European options for which the convergence results are well known. However, this non-exercise condition is not true for options on nontraded state variables such as futures, forwards, yields, certain kinds of averages, etc. Another interesting case involving dividends is when a risky asset pays a discrete dividend(s) on some known date(s). However, we postpone discussion of this case.

**DEFINITION 2.1.** Let the *spot interest rate* be a function  $r(t, x): [0, T] \times D_{\mathfrak{R}^d}[0, T] \rightarrow \mathfrak{R}^+$  such that  $r(\cdot)$  is continuous in the product topology and given any two elements  $x, y \in D_{\mathfrak{R}^d}[0, T]$  such that  $x(s) = y(s): s \leq t \leq T$ , then  $r(t, x) = r(t, y)$ .

Notice that the interest rate is not assumed to be deterministic or even bounded. Corresponding to  $r$ , we can define a discount function  $B(t, x): [0, T] \times D_{\mathfrak{R}^d}[0, T] \rightarrow \mathfrak{R}^+$  such that  $B(t, x) = \exp[\int_0^t r(u, x) du]$ .

**DEFINITION 2.2.** An American option is a security with an expiration date  $T$ , a continuous cash flow rate  $c(t, x): [0, T] \times D_{\mathfrak{R}^d}[0, T] \rightarrow \mathfrak{R}$ , and a terminal payoff  $f(\rho, x): [0, T] \times D_{\mathfrak{R}^d}[0, T] \rightarrow \mathfrak{R}$  on the exercise date  $\rho \in [0, T]$ . Further,  $c(\cdot)$  and  $f(\cdot)$  satisfy the same properties as  $r(\cdot)$  in Definition 2.1.

To simplify matters, we will directly assume that the probability measure  $P$  is a martingale probability measure in the sense of Harrison and Pliska (1981). This implies that the value of an American option in our model, given that  $X(0) = y \in \mathfrak{N}^d$ , is given by

$$(2.4) \quad V(y) = \sup_{\rho \in \mathcal{T}[0, T]} E_P \left[ \int_0^\rho \frac{c(u, X)}{B(u, X)} du + \frac{f(\rho, X)}{B(\rho, X)} \right],$$

where  $\mathcal{T}[0, T]$  is the class of all stopping times in  $[0, T]$  with respect to  $\mathcal{F}_t$ , the filtration generated by  $X(t)$ . The price of the option would not increase even if we allowed (randomized) stopping times w.r.t.  $\mathcal{M}_t$ . The corresponding values in our discrete models to be defined in subsequent sections will be termed  $V_n(\cdot)$ .

Further, we assume that for some  $\delta > 0$ ,

$$(2.5) \quad E_P \left[ \sup_{t \in [0, T]} \left[ \int_0^t \frac{c(s, X)}{B(s, X)} ds + \frac{f(t, X)}{B(t, X)} \right]^{1+\delta} \right] < \infty.$$

This uniform integrability condition ensures that option values are finite and in the case of complete markets, one can actually replicate these options by trading in the underlying securities. For a justification of these statements, see Amin and Jarrow (1992).

### 2.2. Discrete Setup

For every integer  $n$ , divide the trading interval  $[0, T]$  into  $n$  intervals of width  $h_n = T/n$ . Assume that for each  $n$ , we are given a Markov chain  $Y_n(ih_n)$ ,  $i = 1, 2, \dots, n$ , on  $(\Omega_n, \mathcal{M}_n, P_n)$  with values in  $\mathfrak{N}^d$  and transition functions  $\Pi_n(ih_n)$ . Note that the Markov chain is not assumed to be time homogeneous. Now we can construct continuous-time processes  $X_n(t)$  by interpolating the process  $Y_n(ih_n)$  such that  $X_n(t) = Y_n(\lfloor t/h_n \rfloor h_n)$ . Let  $\mathcal{F}_{n,t} = \sigma(X_n(s) : s \leq t)$ . The processes  $X_n$  have sample paths in  $D_{R^d}[0, T]$  and are meant to approximate the continuous-time process  $X$  in the sense that  $X_n \Rightarrow X$ .

The value of an American option under  $X_n$ , given that  $X_n(0) = y \in \mathfrak{N}^d$ , is given by

$$(2.6) \quad V_n(y) = \sup_{\rho \in \mathcal{T}_n[0, T]} E_{P_n} \left[ \int_0^\rho \frac{c(u, X_n)}{B(u, X_n)} du + \frac{f(\rho, X_n)}{B(\rho, X_n)} \right],$$

where  $\mathcal{T}_n[0, T]$  is the class of all stopping times in  $[0, T]$  with respect to  $\mathcal{F}_{n,t}$ , the filtration generated by  $X_n(t)$ . We note that for each  $n$ , an optimal stopping time  $\rho_n$  (w.r.t  $\mathcal{F}_{n,t}$ ) always exists.

To ensure that  $X_n \Rightarrow X$ , we require some conditions on the moments of  $X_n$  which will also be used in a subsequent proof (Lemma 3.3). These conditions can also serve as a ready reference to verify that a particular sequence of discrete processes converges to the required diffusion. Define

$$(2.7) \quad \mu_n(t, x) = \frac{1}{h_n} \int (y - x) \Pi_n(t, x, dy),$$

$$(2.8) \quad a_n(t, x) = \frac{1}{h_n} \int (y - x - \mu_n(t, x)h)(y - x - \mu_n(t, x)h)' \Pi_n(t, x, dy),$$

$$(2.9) \quad \tilde{a}_n(t, x) = \frac{1}{h_n} \int (y - x)(y - x)' \Pi_n(t, x, dy).$$

ASSUMPTION 2.2. For every  $K > 0$ ,

$$(2.10) \quad \lim_{n \rightarrow \infty} \sup_{\|x\| < K, t \leq T} \|\mu_n(t, x) - \mu(t, x)\| = 0,$$

$$(2.11) \quad \lim_{n \rightarrow \infty} \sup_{\|x\| < K, t \leq T} \|a_n(t, x) - a(t, x)\| = 0,$$

and there exists  $\delta > 0$  such that

$$(2.12) \quad \lim_{n \rightarrow \infty} \sup_{\|x\| < K, t \leq T} \frac{1}{h_n} \int \|y - x\|^{2+\delta} \Pi_n(t, x, dy) = 0.$$

The first two conditions in Assumption 2.2 ensure that the instantaneous “mean and covariance” of the processes  $X_n$  converge to the corresponding values for  $X$  over compact intervals. Finally, the third condition ensures convergence to a continuous process.

Our conditions imply the conditions used by Stroock and Varadhan (1979, (2.4)–(2.6), p. 268), hereafter SV, to guarantee the convergence of Markov chains to diffusions. Equation (2.7) and the Lyapunov condition (2.12) imply that (2.5) and (2.6) in SV are satisfied. Equation (2.4) in SV requires that (2.11) holds with  $a(\cdot \cdot \cdot)$  and  $a_n(\cdot \cdot \cdot)$  replaced by  $\tilde{a}(\cdot \cdot \cdot)$  and  $\tilde{a}_n(\cdot \cdot \cdot)$  respectively. To see that this is true, use (2.9) and note that (2.11) holds if and only if

$$\lim_{n \rightarrow \infty} \sup_{\|x\| < K, t \leq T} \|\tilde{a}_n(t, x) - a(t, x)\| = 0.$$

This follows since

$$\|\tilde{a}_n(t, x) - a_n(t, x)\| \leq \|\tilde{a}_n(t, x) - a_n(t, x)\| + \|\tilde{a}_n(t, x) - a(t, x)\|.$$

By hypothesis, the second term on the right converges to zero over compact intervals. For the first term, on  $\|x\| \leq K, t \leq T$ ,

$$(2.13) \quad \begin{aligned} |a_n^{ij}(t, x) - \tilde{a}_n^{ij}(t, x)| &\leq |\mu_n^i(t, x) \mu_n^j(t, x)| \times h_n + |\mu_n^i(t, x)| \\ &\times \int |y^j - x^j| \Pi_n(t, x^i, dy^j) + |\mu_n^j(t, x)| \int |y^i - x^i| \Pi_n(t, x^i, dy^i). \end{aligned}$$

Given that  $\mu(\cdot)$  is continuous, we can bound  $|\mu_n^i(t, x)|$  over compact intervals. Then, the first term goes to zero trivially and the second and third terms converge to zero by hypothesis (2.12). The reverse implication follows in an identical fashion.

The following lemma is now a direct consequence of Lemma 11.2.1 in SV (p. 268).

LEMMA 2.1. If  $PX_n(0)^{-1} \Rightarrow X(0)$ , then  $X_n$  converges in distribution to  $X$ , the solution to (2.3).

This completes our initial setup.

### 3. CONVERGENCE OF STOCHASTIC PROCESSES AND OPTIMAL STOPPING TIMES

Let  $\rho_n$  be the sequence of optimal stopping times for the discrete problems. As  $\rho_n \in [0, T]$  and  $X_n$  is tight (Lemma 2.1),  $(X_n, \rho_n)$  is tight in  $D_{R^d}[0, T] \times [0, T]$ , and consequently some subsequence converges to a weak limit. Let  $(X, \rho)$  be the limit of one such convergent subsequence. The distribution of  $\rho$  may depend on the particular subsequence chosen, but the distribution of  $X$  will not. Henceforth, we fix this convergent subsequence. To keep subsequent notation manageable, we continue to use the subscript  $n$  for this subsequence. We need to show that  $\rho$  is “in some appropriate sense” a legitimate stopping time with respect to  $X$ . This will be shown in Theorem 4.1. As a prelude to this theorem, we show that for every  $\varepsilon > 0$ , we can construct a sequence of processes  $(X_{n,\varepsilon}, W_{n,\varepsilon})$  from  $X_n$  and an independent  $d$ -Brownian motion (B.M.)  $Z$  such that  $\rho_n$  is nonanticipative with respect to  $W_{n,\varepsilon}$ . Furthermore, there exists a subsequence such that  $(X_{n_k,\varepsilon_k}, W_{n_k,\varepsilon_k})$  converges jointly to  $(X, W)$  which satisfies the stochastic differential equation (2.3). We prove this in two steps. First, we show that there exists a subsequence  $n_k$  such that  $W_{n_k,\varepsilon_k} \Rightarrow W$ . It follows from the continuity of  $X$  and  $W$  that  $(X_{n_k,\varepsilon_k}, W_{n_k,\varepsilon_k})$  is tight on  $D_{R^{2d}}[0, T]$ .

Let  $(\zeta, B)$  be the limit of a particular subsequence of  $(X_{n_k,\varepsilon_k}, W_{n_k,\varepsilon_k})$ . In the second step, we show that  $(\zeta, B)$  is indeed a solution to

$$(3.1) \quad d\zeta(t) = \mu(t, \zeta(t)) dt + \sigma(t, \zeta(t)) dB(t), \quad 0 \leq t \leq T,$$

with  $\zeta(0) = X(0)$  and  $B$  a  $d$ -Brownian motion. This implies that  $(\zeta, B)$  has the same joint distribution as  $(X, W)$  and hence  $(X_{n_k,\varepsilon_k}, W_{n_k,\varepsilon_k}) \Rightarrow (X, W)$ .

Construction of the processes  $W_{n_k,\varepsilon_k}$  will now follow along the lines of the proof of a martingale representation theorem given in Karatzas and Shreve (1988, p. 170). For ease of notation, define the increments in the process  $X_n$  over intervals of width  $h_n = T/n$  as

$$\Delta X_n(t) = X_n(t + h_n) - X_n(t).$$

Define the conditional mean  $E[\Delta X_n(t)|X_n(t) = x]$  to be  $\mu_n(t, x)h_n$  and the conditional covariance  $E[(\Delta X_n(t) - \mu_n(t, x)h_n)^2|X_n(t) = x]$  to be  $a_n(t, x)h_n$ . For every  $\varepsilon > 0$ , define a sequence of processes  $X_{n,\varepsilon}$  with  $X_{n,\varepsilon}(0) = X_n(0)$  and

$$(3.2) \quad \Delta X_{n,\varepsilon}(t) = \Delta X_n(t) + \varepsilon \times (Z(t + h_n) - Z(t)),$$

where  $Z(t)$  is a  $d$ -B.M. on  $(\Omega_n, \mathcal{M}_n, P_n)$  which is independent of the processes  $X_n$  and  $X$  (may need to augment the space). Corresponding to  $a_n$ , define  $a_{n,\varepsilon}$  as

$$(3.3) \quad a_{n,\varepsilon}(t, x) = a_n(t, x) + \varepsilon I,$$

where  $I$  is a  $d \times d$  identity matrix. The term  $\varepsilon I$  is added as  $a_n(t, x)$  is not necessarily invertible.

It now follows that  $E[\Delta X_{n,\varepsilon}(t)|\mathcal{F}_{n,t}] \equiv \mu_{n,\varepsilon}(t, X_{n,\varepsilon}(t))h_n = \mu_n(t, X_{n,\varepsilon}(t))h_n$  and

$$(3.4) \quad E[(\Delta X_{n,\varepsilon}(t) - \mu_n(t, X_{n,\varepsilon}(t))h_n)(\Delta X_{n,\varepsilon}(t) - \mu_n(t, X_{n,\varepsilon}(t))h_n)' | \mathcal{F}_{n,t}] = a_{n,\varepsilon}(t, X_{n,\varepsilon}(t)),$$

where  $\mathcal{F}_{n,t}$  is the augmented filtration generated by  $X_n$  and  $Z$ . Furthermore,  $\int(\Delta X_{n,\varepsilon}(t) - \mu_n(t, X_{n,\varepsilon}(t))h_n)$  is a martingale with respect to  $\mathcal{F}_{n,t}$ .

By construction, the matrix-valued process  $a_{n,\varepsilon}(t, X_{n,\varepsilon}(t))$  is symmetric and positive definite for every  $t$ . Every such matrix can be diagonalized by an orthogonal matrix  $Q_{n,\varepsilon}(t, X_{n,\varepsilon}(t))$  such that (see Karatzas and Shreve 1988, p. 170)

$$(3.5) \quad Q_{n,\varepsilon}^{-1} a_{n,\varepsilon} Q_{n,\varepsilon} = \Lambda_{n,\varepsilon},$$

and the process  $\Lambda_{n,\varepsilon}(t, X_{n,\varepsilon}(t))$  is diagonal with the nonzero eigenvalues of  $a_{n,\varepsilon}(t, X_{n,\varepsilon}(t))$  as its diagonal elements. In our case, given that  $a_{n,\varepsilon}$  is invertible, all the diagonal elements are nonzero and positive. Further, define  $\sigma_{n,\varepsilon} = Q_{n,\varepsilon} \sqrt{\Lambda_{n,\varepsilon}}$ , where  $\sqrt{\Lambda_{n,\varepsilon}}$  is a diagonal matrix with diagonal elements as square roots of the corresponding elements of  $\Lambda_{n,\varepsilon}$ . Then,  $\sigma_{n,\varepsilon} \sigma'_{n,\varepsilon} = a_{n,\varepsilon}$  and for every  $K > 0$ ,

$$(3.6) \quad \lim_{\varepsilon \downarrow 0} \lim_{n \rightarrow \infty} \sup_{\|x\| \leq K, t \leq T} \|\sigma_{n,\varepsilon}(t, x) - \sigma(t, x)\| = 0.$$

Let  $(Q_{n,\varepsilon}(t, \omega))_{ij} = q_{n,\varepsilon}^{ij}(t, \omega)$  and  $(\Lambda_{n,\varepsilon}(t, \omega))_{ij} = \delta_{ij} \lambda_{n,\varepsilon}^i(t, \omega)$ , where  $\delta_{ij}$  is an indicator function which equals 1 if  $i = j$  and zero otherwise. Then

$$(3.7) \quad \sum_{k=1}^d q_{n,\varepsilon}^{k,i}(t, \omega) q_{n,\varepsilon}^{k,j}(t, \omega) = \sum_{k=1}^d q_{n,\varepsilon}^{i,k}(t, \omega) q_{n,\varepsilon}^{j,k}(t, \omega) = \delta_{ij},$$

$$(3.8) \quad \sum_{k=1}^d \sum_{l=1}^d q_{n,\varepsilon}^{k,i}(t, \omega) a_{n,\varepsilon}^{k,l}(t, \omega) q_{n,\varepsilon}^{l,j}(t, \omega) = \delta_{ij} \lambda_{n,\varepsilon}^i(t, \omega),$$

and  $|q_{n,\varepsilon}^{k,i}| \leq 1$ . Define the process  $W_{n,\varepsilon}^i(t)$ ,  $i = 1, \dots, d$ , with  $W_{n,\varepsilon}^i(0) = 0$ , by

$$(3.9) \quad W_{n,\varepsilon}^i(t + h_n) - W_{n,\varepsilon}^i(t) = \sum_{k=1}^d q_{n,\varepsilon}^{k,i}(t, X_{n,\varepsilon}(t)) [\Delta X_{n,\varepsilon}^k(t) - \mu_{n,\varepsilon}^k(t, X_{n,\varepsilon}(t), X_{n,\varepsilon}(t))h_n] \frac{1}{\sqrt{\lambda_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t))}}.$$

LEMMA 3.1. For each  $\varepsilon > 0$ , the sequence of processes  $W_{n,\varepsilon}$  is tight,

$$(3.10) \quad \langle W_{n,\varepsilon}^i, W_{n,\varepsilon}^j \rangle_t = \delta_{ij} \lfloor t/h_n \rfloor h_n,$$

and  $W_{n,\varepsilon}$  is a square-integrable martingale.

*Proof.* First, we show that  $W_{n,\varepsilon}$  is an  $\mathcal{F}_{n,t}$ -martingale. From (3.9)

$$W_{n,\varepsilon}^i(t) = W_{n,\varepsilon}^i(s) + \sum_{k=1}^d \sum_{b=\lfloor s/h_n \rfloor}^{\lfloor t/h_n \rfloor} q_{n,\varepsilon}^{k,i}(bh_n, X_{n,\varepsilon}(bh_n)) \times [\Delta X_{n,\varepsilon}^k(bh_n) - \mu_{n,\varepsilon}^k(bh_n, X_{n,\varepsilon}(bh_n))h_n] \frac{1}{\sqrt{\lambda_{n,\varepsilon}^i(bh_n, X_{n,\varepsilon}(bh_n))}}.$$

Using the law of iterated expectations and the definition of  $\mu_{n,\varepsilon}$  now yields the martingale condition.

Now, from the definition of  $q_{n,\varepsilon}^i$ ,

$$\begin{aligned}
 E([W_{n,\varepsilon}^i(t+h_n) - W_{n,\varepsilon}^i(t)] [W_{n,\varepsilon}^j(t+h_n) - W_{n,\varepsilon}^j(t)] \mid \mathcal{F}_{n,t}) \\
 = \sum_{k=1}^d \sum_{l=1}^d q_{n,\varepsilon}^{k,i} a_{n,\varepsilon}^{k,l} h_n q_{n,\varepsilon}^{l,j} \frac{1}{\sqrt{\lambda_{n,\varepsilon}^i(t)}} \frac{1}{\sqrt{\lambda_{n,\varepsilon}^j(t)}} = \delta_{ij} h_n.
 \end{aligned}
 \tag{3.11}$$

An explicit computation using (3.11) yields that  $W_{n,\varepsilon}$  is a square integrable martingale (alternately, see (4.1)). Further, summing over time in (3.11) and using the definition of the squared covariation, we have

$$(t - h_n) \leq \langle W_{n,\varepsilon}^j, W_{n,\varepsilon}^j \rangle_t = \lfloor t/h_n \rfloor h_n \leq t.
 \tag{3.12}$$

Hence,  $\sum_{j=1}^d \langle W_{n,\varepsilon}^j, W_{n,\varepsilon}^j \rangle$  converges in law to a deterministic process. This property is sufficient to guarantee tightness (Jacod and Shiryaev, 1987, Theorem 4.13, p. 322).  $\square$

The above lemma guarantees that  $\lim_{n \rightarrow \infty} \langle W_{n,\varepsilon}^i, W_{n,\varepsilon}^j \rangle_t = \delta_{ij} t$ . We now show that  $W_{n,\varepsilon}$  is C-tight.<sup>3</sup> Hence, by Levy’s theorem we can conclude that for every  $\varepsilon > 0$ , the limit points of  $\{W_{n,\varepsilon}\}_n$  are in fact  $d$ -dimensional Brownian motions. First, we need a lemma which is a modified version of one in Duffie and Protter (1991).

LEMMA 3.2. *Let  $U_n \Rightarrow U$  in  $D_{R^d}[0, T]$  and  $U$  is continuous. Further,  $f_n: [0, T] \times \mathfrak{R}^d \rightarrow \mathfrak{R}^d$  is a sequence of functions that are bounded over compact intervals such that for a given continuous function  $f: [0, T] \times \mathfrak{R}^d \rightarrow \mathfrak{R}^d, \forall K > 0$ ,*

$$\lim_{n \rightarrow \infty} \sup_{|x| \leq K, t \leq T} \|f_n(t, x) - f(t, x)\| = 0.
 \tag{3.13}$$

Let  $H_n(t) = f_n(\lfloor t/h_n \rfloor h_n)$ ,  $U_n(\lfloor t/h_n \rfloor h_n)$ , and  $H(t) = f(t, U(t))$ . Then  $(H_n, U_n) \Rightarrow (H, U)$  in  $D_{R^{2d}}[0, T]$ .

*Proof.* This is a modified version of Lemma 5.1 in Duffie and Protter (1992). Equation (3.13) is sufficient to ensure that their proof holds.  $\square$

Now, we can prove that  $W_{n,\varepsilon}$  is C-tight.

LEMMA 3.3. *For a fixed  $\varepsilon > 0$ , the sequence of processes  $W_{n,\varepsilon}$  is C-tight. Furthermore, there exists a subsequence  $(n_k, \varepsilon_k)$  such that*

$$W_{n_k, \varepsilon_k} \Rightarrow W \text{ and } W_{n_k, \varepsilon_k} \Rightarrow X \text{ in } D_{R^d}[0, T].
 \tag{3.14}$$

*Proof.* To show C-tightness, we will use Proposition 3.26 of Jacod and Shiryaev (1987, p. 315). A sequence of processes  $U_n$  is C-tight if and only if, for every  $\gamma > 0$ ,

$$\lim_{n \rightarrow \infty} P_n[\sup_{t \leq T} |\Delta U_n(t)| > \gamma] \rightarrow 0.
 \tag{3.15}$$

<sup>3</sup>A sequence of processes  $(X^n)$  is called C-tight (definition is from Jacod and Shiryaev 1987) if it is tight and if all limit points of the sequence  $\mathcal{L}(X^n)$  are laws of continuous processes.



Fix some  $\gamma > 0$ . Noting the definition of  $W_{n,\varepsilon}$ , the inequality  $|q_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t))| \leq 1$  and the fact that for every set of random variables  $(A_i; i = 1, \dots, l)$ ,  $\text{Prob}[\sum_{i=1}^l A_i > \gamma] \leq \sum_{i=1}^l \text{Prob}[|A_i| > \gamma]$ ,

$$\begin{aligned}
 \lim_{n \rightarrow \infty} P[\sup_{t \leq T} |\Delta W_{n,\varepsilon}^i(t)| > \gamma] &= \lim_{n \rightarrow \infty} P \left[ \sup_{t \leq T} \left| \sum_{k=1}^d q_{n,\varepsilon}^{k,i}(t, X_{n,\varepsilon}(t)) (\Delta X_{n,\varepsilon}^{k,i}(t) \right. \right. \\
 (3.16) \qquad \qquad \qquad &\quad \left. \left. - \mu_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t)) h_n \right) \frac{1}{\sqrt{\lambda_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t))}} \right| > \gamma \right] \\
 &\leq \sum_{k=1}^d \lim_{n \rightarrow \infty} P \left[ \sup_{t \leq T} |\Delta X_{n,\varepsilon}^i(t) - \mu_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t)) h_n| \frac{1}{\sqrt{\lambda_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t))}} > \frac{\gamma}{d} \right].
 \end{aligned}$$

Noting that  $\sqrt{\lambda_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t))} \geq \sqrt{\varepsilon}$ , (3.16) is

$$\leq \sum_{k=1}^d \lim_{n \rightarrow \infty} P \left[ \sup_{t \leq T} |\Delta X_{n,\varepsilon}^i(t) - \mu_{n,\varepsilon}^i(t, X_{n,\varepsilon}(t)) h_n| > \frac{\gamma \sqrt{\varepsilon}}{d} \right].$$

Now, by Lemma 2.1,  $X_{n,\varepsilon}$  is C-tight since it converges to some process  $X_\varepsilon$  where  $X_\varepsilon$  is a diffusion with drift and dispersion coefficients  $\mu_\varepsilon(t, x) = \mu(t, x)$  and  $a_\varepsilon(t, x) = a(t, x) + \varepsilon I$  respectively. Hence,

$$\lim_{n \rightarrow \infty} P \left[ \sup_{t \leq T} |\Delta X_{n,\varepsilon}^i| > \frac{\gamma \sqrt{\varepsilon}}{d} \right] = 0.$$

Furthermore, by Lemmas 2.1 and 3.2 and Assumption 2.2,  $\mu_{n,\varepsilon}(t, X_{n,\varepsilon}) \Rightarrow \mu(t, X_\varepsilon)$ . Therefore,

$$\lim_{n \rightarrow \infty} P \left[ \sup_{t \leq T} |\mu_{n,\varepsilon}^i(t, X_{n,\varepsilon}) h_n| > \frac{\gamma \sqrt{\varepsilon}}{d} \right] = 0 \quad \text{as } n \rightarrow \infty.$$

This implies that  $W_{n,\varepsilon}$  is C-tight. As a direct consequence of Levy’s theorem,  $W_{n,\varepsilon} \Rightarrow W_\varepsilon$  where  $W_\varepsilon$  is a  $d$ -B.M. As  $W_\varepsilon$  is a  $d$ -B.M. for each  $\varepsilon > 0$ ,  $\lim_{\varepsilon \rightarrow 0} W_\varepsilon \Rightarrow W$ , which is also a  $d$ -B.M. Furthermore, by observing the drift and diffusion coefficients of  $X_\varepsilon$ , it is clear that  $\lim_{\varepsilon \rightarrow 0} X_\varepsilon \Rightarrow X$ . Hence, by a diagonalization argument, there exists a subsequence  $(n_k, \varepsilon_k)$  such that

$$(3.17) \qquad W_{n_k, \varepsilon_k} \Rightarrow W \quad \text{and} \quad X_{n_k, \varepsilon_k} \Rightarrow X \quad \text{in } D_{R^d}[0, T]. \qquad \square$$

We have accomplished the first step in our convergence proof, i.e., shown the existence of  $(X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k})$  such that  $\rho_{n_k}$  is nonanticipative with respect to  $W_{n_k, \varepsilon_k}$  (by construction) and  $X_{n_k, \varepsilon_k} \Rightarrow X$ , a solution to (2.3) and  $W_{n_k, \varepsilon_k} \Rightarrow W$ , a  $d$ -Brownian motion. We now turn to the second step, i.e., to show that if  $(\zeta, B)$  is the limit of a convergent subsequence of  $(X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k})$ , then  $(\zeta, B)$  is a solution to (2.3). To keep the notation simple, we will continue to denote the convergent subsequence by  $(X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k})$ .

At this stage, we rewrite  $(X_{n_k, \varepsilon_k})$  as a discrete stochastic differential equation similar to

the continuous-time case. Define the sequence of processes  $\alpha_{n_k, \varepsilon_k}(\cdot, \cdot)$  implicitly according to the equation

$$(3.18) \quad \begin{aligned} X_{n_k, \varepsilon_k}(t) &= X_{n_k, \varepsilon_k}(0) + \int_0^t \mu_{n_k, \varepsilon_k}(s, X_{n_k, \varepsilon_k}) \, ds \\ &+ \int_0^t \sigma_{n_k, \varepsilon_k}(s, X_{n_k, \varepsilon_k}) \, dW_{n_k, \varepsilon_k}(s) + \alpha_{n_k, \varepsilon_k}(t, X_{n_k, \varepsilon_k}) \quad \text{a.s.} \end{aligned}$$

From the definition of  $\sigma_{n, \varepsilon}$  and (3.7) and (3.9) note that  $\alpha_{n_k, \varepsilon_k}(t, X_{n_k, \varepsilon_k})$  is zero at dates given by integer multiples of  $h_{n_k}$ . Further,

$$|\alpha_{n_k, \varepsilon_k}(t, X_{n_k, \varepsilon_k})| \leq |\mu_{n_k, \varepsilon_k}(\lfloor t/h_{n_k} \rfloor h_{n_k}, X_{n_k, \varepsilon_k})| h_{n_k}.$$

Since  $\mu(\cdot)$  satisfies a linear growth condition (Assumption 2.1), and  $X_{n_k}$  is tight,  $E[\alpha_{n_k, \varepsilon_k}(t)]^2 \rightarrow 0$ . Hence, it can be shown that  $\alpha_{n_k, \varepsilon_k}$  converges in law to the zero process.

Our next task is to show that the right-hand side of (3.18) converges weakly to a solution to the stochastic differential equation (2.3). First, observe that by Lemmas 3.2 and 3.3, Assumption 2.2, and the continuity of the limit processes,

$$(3.19) \quad (\sigma_{n_k, \varepsilon_k}(\cdot, X_{n_k, \varepsilon_k}), \mu_{n_k, \varepsilon_k}(\cdot, X_{n_k, \varepsilon_k}), X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k}) \Rightarrow (\sigma(\cdot, \zeta), \mu(\cdot, \zeta), \zeta, B)$$

in  $D_{M^{dd} \times R^{3d}}[0, T]$  where  $M^{dd}$  is the space of real-valued  $d \times d$  matrices. To complete the proof, we need another lemma from Duffie and Protter (1991).

LEMMA 3.4 (Duffie and Protter 1992, Theorem 4.3). *Let  $U^n$  be a sequence of semimartingales with decompositions  $U^n = M^n + A^n$ , and let  $K^n$  be adapted RCLL processes on some stochastic basis for each  $n \in \mathcal{N}$ . Further, let  $|A^n|_t$  be the total variation of  $A^n$  at time  $t$ . Suppose that  $E_n|A^n|_T$  and  $E_n(\sup_{s \leq T} |\Delta M^n(s)|)$  are bounded. If the joint distribution of  $(K^n, U^n)$  converges to the joint distribution of  $(K, U)$ , then  $U$  is a semimartingale and  $\int K^n(s^-) dU^n(s) \Rightarrow \int K(s^-) dU(s)$ .*

To use Lemma 3.4, we need to show that for  $1 \leq i \leq d$ ,  $\sup_n E_n(\sup_{s \leq t} |\Delta W_n^i(s)|)$  is bounded. This follows since

$$(3.20) \quad \sup_n E_n(\sup_{s \leq t} |\Delta W_n^i(s)|) \leq 1 + \sup_n E_n \sum_{k=0}^{\lfloor t/h_n \rfloor} (\Delta W_n^i)^2 \leq 1 + t.$$

Using Lemma 3.4 and (3.19) the sequence of processes

$$(3.21) \quad \begin{aligned} X_{n_k, \varepsilon_k}(0) + \int_0^t \mu_{n_k, \varepsilon_k}(s, X_{n_k, \varepsilon_k}) \, ds + \int_0^t \sigma_{n_k, \varepsilon_k}(s, X_{n_k, \varepsilon_k}) \, dW_{n_k, \varepsilon_k}(s) \\ \Rightarrow \zeta(0) + \int_0^t \mu(s, \zeta) \, ds + \int_0^t \sigma(s, \zeta) \, dB(s) \quad \text{in } D_{R^d}[0, T]. \end{aligned}$$

Now, since  $X_{n_k, \varepsilon_k} \Rightarrow \zeta$  and  $\alpha_{n_k, \varepsilon_k} \Rightarrow$  the zero process, (3.21) implies that

$$(3.22) \quad d\zeta(t) = \mu(t, \zeta(t)) dt + \sigma(t, \zeta(t)) dB(t),$$

i.e.,  $(\zeta, B)$  has the same joint distribution as  $(X, W)$ . Hence,  $(X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k})$  jointly converges to  $(X, W)$ .

It follows from our previous discussion that  $(X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k}, \rho_{n_k})$  is tight in  $D_{R^{2d}}[0, T] \times [0, T]$ . Let  $(X, W, \rho)$  be the limit of any convergent subsequence. Henceforth, fix this convergent subsequence. We now have the following lemma.

LEMMA 3.5. *By Skorokhod embedding (Ethier and Kurtz 1986, Theorem 2.1.8), we can assume that  $\rho_{n_k}, X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k}, X, W,$  and  $\rho$  are all defined on the same probability space*

$$\begin{aligned} X_{n_k, \varepsilon_k}(t) &= X_{n_k, \varepsilon_k}(0) + \int_0^t \mu_{n_k, \varepsilon_k}(s, X_{n_k, \varepsilon_k}) ds + \int_0^t \sigma_{n_k, \varepsilon_k}(s, X_{n_k, \varepsilon_k}) dW_{n_k, \varepsilon_k}(s) + \alpha_{n_k, \varepsilon_k}(t, X_{n_k, \varepsilon_k}) \quad a.s., \\ X(t) &= X(0) + \int_0^t \mu(s, X) ds + \int_0^t \sigma(s, X) dW(s) \quad a.s. \end{aligned}$$

where  $(X_{n_k, \varepsilon_k}, W_{n_k, \varepsilon_k}, \rho_{n_k})$  converges to  $(X, W, \rho)$  a.s.

#### 4. CONVERGENCE OF AMERICAN OPTION VALUES

We now show that  $V_n(y)$  converges to  $V(y, \rho)$  where  $V(y, \rho)$  is defined as follows. Let  $F: D_{R^d}[0, T] \times [0, T] \rightarrow \mathfrak{R}$ ,

$$F(t, x) = \int_0^t \frac{c(u, x)}{B(u, x)} du + \frac{f(t, x)}{B(t, x)},$$

and for any random time  $\tau$  (defined on the same probability space as  $X$  above) and  $X(0) = y$  define

$$V(\tau, y) = E[F(\tau, X)].$$

Now, by the assumptions on  $c, r, B$  (see Definition 2.1) we see that  $F$  is a continuous function on  $C_{R^d}[0, T] \times [0, T]$ . Therefore, by the continuous mapping theorem (Billingsley 1968, Theorem 5.5.4),  $F(\rho_n, X_n) \Rightarrow F(\rho, X)$ . Furthermore, if we assume that  $F(\rho_n, X_n)$  is uniformly integrable (as will be shown to be true for several interesting cases in the next section), then

$$V_n(y) = E(F(\rho_n, X_n)) \rightarrow E(F(\rho, X)) = V(\rho, y).$$

We now need to show that  $V(\rho, y)$  actually equals  $V(y)$ . This cannot be shown directly because  $\rho$  may not be a stopping time with respect to the filtration generated by  $X$ . In fact, very little can be said in general about  $\rho$ . However,  $\rho$  is nonanticipative with respect to  $W$ , the Brownian motion driving  $X$  (by Lemma 3.5 and Kushner 1977, Theorem 8.2.2). Furthermore, Theorem 4.1 shows that it is possible to construct a probability space with  $\bar{X}$ ,

$\bar{\rho}$  defined on it such that  $(\bar{X}, \bar{\rho})$  and  $(X, \rho)$  have the same distribution and  $\bar{\rho}$  is a stopping time with respect to a filtration under which  $\bar{X}$  is a strong Markov process. This implies that  $V(\bar{\rho}, y) = V(\rho, y)$  as  $V(\rho, y)$  depends only on the joint distribution of  $(X, \rho)$ . But this then implies that  $V(\rho, y) \leq V(y)$  and hence<sup>4</sup>  $\limsup_{n \rightarrow \infty} V_n(y) \leq V(y)$ .

**THEOREM 4.1.** *There exists a stochastic basis  $(\Omega, \mathcal{B}, \mathcal{B}_t, Q_\nu)$  on which one can define a Brownian motion  $(\bar{W})$  and a strong Markov process  $\bar{X}$  which is a solution to (2.3) with  $Q_\nu^{-1} \bar{X}(0) = X(0)$ . Further, on this space,  $\exists$  a random variable  $\bar{\rho}$  s.t.  $(\bar{X}, \bar{\rho})$  have the same distribution as  $(X, \rho)$  and  $\bar{\rho}$  is a (randomized) stopping time with respect to  $\mathcal{B}_t$ .*

*Proof.* The proof is identical to the proof of Theorem 8.2.3 in Kushner (1977).  $\square$

So far, we have shown that  $V(\rho, y) \leq V(y) \forall y \in \mathfrak{N}^d$ . Hence, to show the necessary convergence, all we need is to show the reverse implication, i.e.,  $V(\rho, y) \geq V(\tau, y)$  for every stopping time  $\tau$  with respect to the filtration generated by  $X$ . The identical proof of Theorem 8.2.4 in Kushner (1977) gives us this reverse implication once we show that  $W_{n_k, \varepsilon_k}$  is uniformly integrable in  $k$  and observe that if  $\tau$  is any stopping time with respect to the filtration generated by  $X_n$  and  $Z$ , then  $V_n(y, \tau) \leq V_n(y)$ . Stopping times that take advantage of information from  $Z$  cannot yield a larger value function than those with respect to  $X_n$ .

We note that increments to the processes  $W_{n,\varepsilon}^i$  over finite intervals are uniformly integrable in  $n$  and  $\varepsilon$ . This follows since, for every  $0 \leq s \leq t \leq T$ ,

$$\begin{aligned}
 \sup_{n,\varepsilon} E_n[|W_{n,\varepsilon}(t) - W_{n,\varepsilon}(s)|]^2 &\leq \sup_{n,\varepsilon} E_n \left[ \left| \sum_{k=\lfloor s/h_n \rfloor}^{\lfloor t/h_n \rfloor} \Delta W_{n,\varepsilon}(kh_n) \right|^2 \right] \\
 (4.1) \qquad \qquad \qquad &\leq \sup_{n,\varepsilon} E \left[ \sum_{k=\lfloor s/h_n \rfloor}^{\lfloor t/h_n \rfloor} |\Delta W_{n,\varepsilon}(kh_n)|^2 \right] \\
 &= \left( \lfloor \frac{t}{h_n} \rfloor - \lfloor \frac{s}{h_n} \rfloor \right) h_n \leq (t - s) < \infty.
 \end{aligned}$$

This completes our convergence proof.

### 5. UNIFORM INTEGRABILITY OF THE DISCOUNTED PAYOFF FUNCTION

According to Definition 2.2, we can view the American option as paying off

$$F(t, X_n) = \left[ \int_0^t \frac{c(u, X_n)}{B(u, X_n)} du + \frac{f(t, X_n)}{B(t, X_n)} \right]$$

if the option is exercised at date  $t$ . Given an optimal stopping time  $\rho_n$  for a fixed  $n$ , the American option value equals  $E_{\rho_n}[F(\rho_n, X_n)]$ . In this section, we will develop some conditions for the most common types of options which guarantee that  $F(\rho_n, X_n)$  is uniformly integrable in  $n$  as required in Section 4.

To begin, we note that the option payoff to any put option is uniformly integrable since it is bounded. However, for other types of options, the uniform integrability condition must

<sup>4</sup> $\rho$  may be sequence dependent, hence we use the limsup.

be proven on a case-by-case basis. In the subsequent discussion, the condition that we will use is the following.  $F(\rho_n, X_n)$  is uniformly integrable if for some  $\delta > 1$ ,

$$(5.1) \quad \sup_n E_{\rho_n} [\sup_{t \leq T} |F(t, X_n)|^\delta] < \infty.$$

Initially, consider the univariate CRR (1979) binomial model. Let  $\sigma$  and  $r$  be the instantaneous standard deviation of returns and the spot interest rate respectively. These are assumed to be constant. If  $S_n(t)$  is the stock price, then

$$(5.2) \quad S_n(t) = S_n(\lfloor t/h_n \rfloor h_n) = S(0) \exp[j_n(t)\sigma\sqrt{h_n}],$$

where  $n = T/h_n$  and  $j_n(t)$  is the sum of  $\lfloor t/h_n \rfloor$  i.i.d. Bernoulli random variables taking on values in  $\{+1, -1\}$  with probabilities  $\{p(h_n), 1 - p(h_n)\}$  such that

$$(5.3) \quad p(h_n) = \frac{\exp(rh_n) - \exp(-\sigma\sqrt{h_n})}{\exp(\sigma\sqrt{h_n}) - \exp(-\sigma\sqrt{h_n})} = \frac{1}{2} + \frac{r}{2\sigma}\sqrt{h_n} + o(\sqrt{h_n}).$$

Further, for each  $n$ ,  $S_n(t) \exp(-r\lfloor t/h_n \rfloor)$  is a martingale. We will first show that, for every  $\gamma > 0$ ,

$$(5.4) \quad \sup_n E_{\rho(h_n)} \left[ \sup_{t \leq T} \left( \exp\left(-r\lfloor \frac{t}{h_n} \rfloor\right) S_n(t) \right)^\gamma \right] < \infty.$$

This will imply the uniform integrability condition in (5.1) for the common types of options.

For a fixed  $n$  and  $T$ ,  $E_{\rho}[S_n(T)]^\gamma$  is an increasing function of  $p$  if  $\gamma > 1$  (notice that this permits  $p$  to be different from  $p(h_n)$  implied by (5.3)). Further,  $p(h_n)$  is an increasing function of  $h_n$  (5.3). Let  $j_n = j_n(T)$ . Then

$$\begin{aligned} \sup_n E_{\rho(h_n)} [\exp(-rT)S_n(T)]^\gamma &= \sup_n S(0)^\gamma \exp(-r\gamma T) E_{\rho(h_n)} [\exp[\gamma j_n \sigma \sqrt{h_n}]] \\ &\leq \sup_n S(0)^{\gamma-1} \exp(-r(\gamma - \gamma^2)T) E_{\rho(\gamma^2 h_n)} S(0) \exp(-r\gamma^2 T) [\exp[\gamma j_n \sigma \sqrt{\gamma^2 h_n}]]. \end{aligned}$$

The term inside the expectation equals  $S(0)$  since it corresponds to a rescaling of the time step size by  $\gamma^2$  (notice that the states of  $j_n$  are not affected by this rescaling). Therefore,

$$(5.5) \quad \sup_n E_{\rho(h_n)} [\exp(-rT)S_n(T)]^\gamma = S(0)^\gamma \exp(-r(\gamma - \gamma^2)T) < \infty.$$

Equation (5.5) and Doob's inequality imply that for  $\gamma > 1$ ,

$$(5.16) \quad \begin{aligned} \sup_n E_{\rho(h_n)} \left[ \sup_{t \leq T} \left( \exp\left(-r\lfloor \frac{t}{h_n} \rfloor\right) S_n(t) \right)^\gamma \right] \\ \leq \left( \frac{\gamma}{\gamma - 1} \right)^\gamma \sup_n E_{\rho(h_n)} [(\exp(-rT)S_n(T))^\gamma] < \infty. \end{aligned}$$

This implies that any option for which the option payoff is bounded above by a polynomial function of the stock price has a uniformly integrable payoff function. The uniform inte-

grability condition also holds if the stock pays dividends and the ex dividend stock price is bounded above by  $S_n$ .

The payoff from a futures call option is also uniformly integrable. This is true since the futures price is a martingale which is also square integrable and the futures option payoff is dominated by the futures price. The square integrability follows since the futures price equals the spot price on the maturity date and the spot price is square integrable. Similarly, it is easy to see that the payoff from a call option on the average (arithmetic or geometric) of the stock price evaluated on a certain fixed set of dates (or an interval) or the payoff from a lookback option also satisfies the necessary criteria.<sup>5</sup>

In the multivariate case, the options of interest are options on the maximum or minimum of several assets or those based on some average of  $l$  risky asset prices. Given a sequence of  $l < d$  asset price processes  $S_n^i: i \leq l$ , since  $\sup_i S_n^i(t) \leq \sum_{i=1}^l S_n^i(t)$ , if (5.4) holds for individual stocks, then it must also hold for the maximum of a finite number of stock prices. As the minimum as well as any average of  $l$  stock prices is dominated by the maximum, the condition also holds for call options based on this quantity and similarly call options on futures based on these quantities.

Further, using (5.6) the uniform integrability condition is also true in He (1990), as He (1990) shows that the prices of European options whose payoffs are polynomial functions of the stock price converge as  $n \rightarrow \infty$ . This condition is also true in the discrete models of Boyle, Evnine, and Gibbs (1989) and Madan, Milne, and Shefrin (1989).

Now consider the case of options on assets which pay a continuous dividend yield. First, consider the CRR (1979) model with continuous dividends (for a description, see Cox and Rubinstein 1985). Define  $S_n^{\text{cum}}(t)$  as the stock price plus all dividends reinvested into the stock. Clearly, the process  $S_n^{\text{cum}}(t)$  is identical to the process defined in (5.2). In fact, the dividend yield may even be stochastic without affecting this. Hence, (5.4) holds for the  $S_n^{\text{cum}}$  process. Since  $S_n(t) \leq S_n^{\text{cum}}(t) \forall t \in [0, T]$ , the same condition must also hold for the stock price process without the dividends. This implies that our entire discussion for the various types of options in the no dividend case is valid. A similar argument also holds with the possibility of discrete dividends.

In other models in the literature, for example He (1990) or Boyle, Evnine, and Gibbs (1989), the authors do not specify how to handle dividends. But, we can always modify their models in an identical fashion as in the CRR model. Hence, we can prove the required uniform integrability condition. Finally, consider interest rate options. Call and put options based on the prices of various types of bonds or the value of a futures contract based on these bond prices inherently satisfy our condition as long as the spot interest rate is always nonnegative (or bounded from below). This is true since bond prices are bounded by the sum of the total cash flows during the life of the bond. However, for options based on yields or on spot rates, one would separately have to verify that the uniform integrability condition holds.

## 6. CONVERGENCE OF AMERICAN OPTIONS PRICES WITH DISCRETE JUMPS IN THE STATE VARIABLES

In this section, we consider the possibility of discrete jumps in the value of the state variables.

<sup>5</sup> In fact, He (1990) excludes such non-Markovian payoffs because his constructive proofs do not work in that case. However, we do not need to exclude these payoffs.

Suppose we are given a sequence of discrete processes  $X_n$  such that  $X_n \Rightarrow X$  and a sequence of optimal stopping times  $\tau_n \Rightarrow \tau$ . To obtain the convergence of American option values, we need to apply the continuous mapping theorem and correspondingly require  $\tau$  to be a continuity point of  $X$  (a.s.). With a discrete jump in  $X$  at a fixed date, this cannot be guaranteed. In fact, for an American call option on a stock which pays a discrete dividend,  $\tau$  has a probability mass only on dividend payment dates and at the maturity date of the option. Further, for the limiting problem, an optimal stopping time does not even exist. Therefore, we prove our result in an indirect fashion.

Consider the probability space  $(\Omega, \mathcal{M}, \mathcal{M}_t, P)$  defined in Section 2.1 and a diffusion  $X(t)$  satisfying (2.3). We assume<sup>6</sup> that the value of the underlying state variables is  $U(t) = X(t) + DB(t)1_{t < T_R}$  where  $D$  is a fixed vector of length  $d$  and  $B(t)$  is the discount factor at date  $t$ . For each integer  $n$ , there is defined a Markov chain  $U_n$  on  $(\Omega_n, \mathcal{M}_n, \mathcal{M}_{n,t}, P_n)$  such that  $U_n \Rightarrow U$  as  $n \rightarrow \infty$ . The American option value in (2.4) is obtained with respect to the process  $U$  and not  $X$ .

Fix  $\delta > 0$ . Define the  $\delta$ -restricted stopping problem to be one in which the owner of the option is not permitted to exercise her option in the interval  $[T_R - \delta, T_R]$ . In all other respects, the problem is identical to the unrestricted problem studied in previous sections. Let  $\rho_{n,\delta}$  be the sequence of optimal stopping times and  $V_{n,\delta}(y) = V_{n,\delta}(y, \rho_{n,\delta})$  be the optimal option values, with initial condition  $U_n(0) = y$ , for the  $\delta$ -restricted problem. In the continuous-time case, let the option value for the  $\delta$ -restricted problem be  $V_\delta(y)$  and  $V(y)$  for the unrestricted problem.

We can use the continuous mapping theorem and the results in Section 4 to conclude that, for every fixed  $y \in \mathfrak{R}^d$  and  $\varepsilon > 0$ , there exists an integer  $m$  such that  $|V_{n,\delta}(y) - V_\delta(y)| < \varepsilon$ , if  $n > m$ . Furthermore, from Shiryaev (1977, Lemma 3, Section 3.2.3) we know that for every  $\varepsilon > 0$ , there exists  $\delta > 0$  such that  $|V_\delta(y) - V(y)| < \varepsilon$ . Therefore, by the triangle inequality,

$$|V_{n,\delta}(y) - V(y)| \leq |V_{n,\delta}(y) - V_\delta(y)| + |V_\delta(y) - V(y)|.$$

Since each of the terms on the right-hand side can be made arbitrarily small by choosing  $\delta$  and  $n$ ,  $V_{n,\delta}(y)$  converges to  $V(y)$  for every  $y \in \mathfrak{R}^d$ .

### 7. SUMMARY AND CONCLUSIONS

In this paper, we have shown that the prices of American options obtained from a sequence of discrete processes, which converges weakly to a continuous-time diffusion process, converge to their continuous-time values under certain conditions. These conditions require the payoff function to be continuous and uniformly integrable and the limit diffusion to satisfy the linear growth and Lipschitz conditions.

### REFERENCES

AMIN, K. I. (1991): "On the Computation of Continuous-Time Options Prices Using Discrete-Time Models," *J. Financial Quant. Anal.*, 26, 477-496.  
 AMIN, K. I., and R. A. JARROW (1992): "Pricing Options on Risky Assets in a Stochastic Interest Rate Economy," *Math. Finance*, 2, 217-238.

<sup>6</sup>The proof can easily be generalized to multiple jumps and even jumps of random size.

- BILLINGSLEY, P. (1968): *Convergence of Probability Measures*. New York: Wiley.
- BOYLE, P. P., J. EVNINE, and S. GIBBS (1989): "Numerical Evaluation of Multivariate Contingent Claims," *Rev. Financial Stud.*, 2, 241–250.
- COX, J. C., S. A. ROSS, and M. RUBINSTEIN (1979): "Option Pricing: A Simplified Approach," *J. Financial Econ.*, 7, 229–263.
- COX, J. C., and M. RUBINSTEIN (1985): *Options Markets*. New York: Prentice Hall.
- DUFFIE, D., and P. PROTTER (1992): "From Discrete to Continuous Time Finance: Weak Convergence of the Financial Gains Process," *Math. Finance*, 2, 1–15.
- ETHIER, S. N., and T. G. KURTZ (1986): *Markov Processes: Characterisation and Convergence*. New York: Wiley.
- HARRISON, J. M., and S. R. PLISKA (1981): "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stoch. Process. Appl.*, 11, 215–260.
- HE, H. (1990): "Convergence from Discrete to Continuous Time Contingent Claims Prices," *Rev. Financial Stud.*, 3, 523–546.
- JACOD, J., and A. N. SHIRYAEV (1987): *Limit Theorems for Stochastic Processes*. New York: Springer-Verlag.
- KARATZAS, I., and S. SHREVE (1988): *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.
- KUSHNER, H. J. (1974): "On the Weak Convergence of Interpolated Markov Chains to a Diffusion," *Ann. Probab.*, 2, 40–50.
- KUSHNER, H. J. (1977): *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. New York: Academic Press.
- MADAN, D. B., F. MILNE, and H. SHEFRIN (1989): "The Multinomial Option Pricing Model and Its Brownian and Poisson Limits," *Rev. Financial Stud.*, 2, 251–265.
- NELSON, D., and K. RAMASWAMY (1990): "Simple Binomial Processes as Diffusion Approximations in Financial Models," *Rev. Financial Stud.*, 3, 393–430.
- ROLL, R. (1977): "An Analytic Formula for Unprotected American Call Options on Stocks with Known Dividends," *J. Financial Econ.*, 17, 251–258.
- SHIRYAEV, A. N. (1977): *Optimal Stopping Rules*. New York: Springer-Verlag.
- STROOCK, D., and S. VARADHAN (1979): *Multidimensional Diffusion Processes*. New York: Springer-Verlag.