

# A diffusion limit for generalized correlated random walks

*Urs Gruber*  
*Allianz AG*  
*Allianz Global Risks*  
*Königinstraße 28*  
*D – 80802 München*  
*Germany*  
*ursgruber@gmx.net*

*Martin Schweizer* \*  
*ETH Zürich*  
*Departement Mathematik*  
*ETH-Zentrum, HG G28.2*  
*CH – 8092 Zürich*  
*Switzerland*  
*martin.schweizer@math.ethz.ch*

**Abstract:** A generalized correlated random walk is a process of partial sums  $X_k = \sum_{j=1}^k Y_j$  such that  $(X, Y)$  forms a Markov chain. For a sequence  $(X^n)$  of such processes where each  $Y_j^n$  takes only two values, we prove weak convergence to a diffusion process whose generator is explicitly described in terms of the limiting behaviour of the transition probabilities for the  $Y^n$ . Applications include asymptotics for option replication under transaction costs and an approximation for a given diffusion by regular recombining binomial trees.

**Key words:** correlated random walk, diffusion limit, weak convergence, mathematical finance, large investor, transaction costs, binomial trees

**MSC 2000 Classification Numbers:** 60F17, 60J60, 91B28

**JEL Classification Numbers:** C60

This version: 27.10.2005

*(Journal of Applied Probability 43 (2006), 60–73)*

---

\* corresponding author

## 0. Introduction

Let  $(Y_j)_{j \in \mathbb{N}_0}$  be a sequence of random variables and  $X_k = \sum_{j=1}^k Y_j$ ,  $k \in \mathbb{N}_0$ , the corresponding process of partial sums. If the  $Y_j$  are independent,  $X$  is a random walk and Donsker's theorem shows under suitable assumptions that after rescaling, a sequence of such  $X^n$  converges in distribution to a Brownian motion. If the  $Y_j$  form a Markov chain,  $X$  is a *correlated random walk*, and there are some weak convergence results for such sequences as well; details are given later. In this paper, we study the more general case where the pair  $(X, Y)$  is a Markov chain and prove a functional central limit theorem for a sequence of such processes  $X^n$ .

More precisely, we consider the situation where each  $Y_j^n$  takes only two values. The limit process  $X^\infty$  is a diffusion and we explicitly describe its generator in terms of the limiting behaviour of the transition probabilities for the  $Y^n$ . The motivation for this problem comes from mathematical finance where it arose in the context of option pricing for a large investor. We briefly sketch this connection and give two other applications of the main convergence result. One is about asymptotics for option replication under transaction costs; the other shows how one can approximate a given diffusion by a regular recombining binomial tree.

The paper is structured as follows. Section 1 contains the precise setup and the main result as well as comments on the literature. In section 2, we prove the main result and discuss variations and extensions. Section 3 presents the applications.

## 1. Setup and main result

Our goal in this paper is to prove a weak convergence result for a class of generalized correlated random walks. This section contains the basic setup, the assumptions and the main result.

For each  $n \in \mathbb{N}$ , we start with a probability space  $(\Omega^n, \mathcal{F}^n, P^n)$  on which we have binary random variables  $(Z_k^n)_{k=0,1,\dots,n}$  taking the values  $+1$  and  $-1$ . We define a stochastic process  $X^n = (X_k^n)_{k=0,1,\dots,n}$  via its increments

$$(1.1) \quad \Delta X_k^n := X_k^n - X_{k-1}^n := \mu_n + \sigma_n Z_k^n \quad \text{for } k = 1, \dots, n$$

with constants  $\mu_n$  and  $\sigma_n$  to be specified later; hence

$$(1.2) \quad X_k^n = X_0^n + \sum_{j=1}^k (\mu_n + \sigma_n Z_j^n) \quad \text{for } k = 0, 1, \dots, n$$

is a sequence of partial sums. Set  $t_k^n := k/n$ . Piecewise constant interpolation on the intervals  $[t_{k-1}^n, t_k^n)$  yields a process  $X^{(n)} = (X_t^{(n)})_{0 \leq t \leq 1}$  with RCLL trajectories via

$$(1.3) \quad X_t^{(n)} := X_{\lfloor nt \rfloor}^n \quad \text{for } 0 \leq t \leq 1$$

so that

$$X_{t_k^n}^{(n)} = X_k^n \quad \text{for } k = 0, 1, \dots, n.$$

We denote by  $\mathbb{F}^{(n)} = (\mathcal{F}_t^{(n)})_{0 \leq t \leq 1}$  the filtration generated by  $X^{(n)}$ ; hence

$$\mathcal{F}_t^{(n)} = \sigma(X_0^n, X_1^n, \dots, X_{k-1}^n) =: \mathcal{F}_{k-1}^n \quad \text{for } t \in [t_{k-1}^n, t_k^n) \text{ and } k = 1, \dots, n.$$

The distribution of  $X^{(n)}$  under  $P^n$  is a probability measure  $\varrho_n$  on the Skorohod space  $D[0, 1]$  of RCLL functions. Our goal is to prove a weak convergence result for the sequence  $(\varrho_n)_{n \in \mathbb{N}}$  under suitable assumptions on  $\mu_n$ ,  $\sigma_n$  and the behaviour of  $(Z_k^n)_{k=0,1,\dots,n}$  under  $P^n$ .

**Remark.** It is purely for ease of notation that we work on the time interval  $[0, 1]$ . Analogous results can be obtained for  $[0, T]$  with  $T \in (0, \infty)$  or for  $[0, \infty)$ .  $\diamond$

Since we are interested in diffusion limits, we work with the usual Donsker type scaling. So let  $\delta_n := 1/\sqrt{n}$  and impose the condition

**(A1)** There are constants  $\sigma > 0$ ,  $\mu \in \mathbb{R}$  and  $\beta \in (0, 1)$  such that

$$(1.4) \quad \sigma_n = \sigma \delta_n + O(\delta_n^{1+\beta}),$$

$$(1.5) \quad \mu_n = \mu \delta_n^2 + O(\delta_n^{2+\beta}).$$

Thus each increment  $\Delta X_k^n$  has mean and variance of the order  $\delta_n^2 = 1/n$  like in Donsker's theorem. But our main assumption is that each pair  $(X^n, Z^n)$  is under  $P^n$  a Markov chain whose transition probabilities have a suitable form. More precisely, we assume that

$$(A2) \quad P^n[Z_k^n = +1 \mid \mathcal{F}_{k-1}^n] = p_n(k, X_{k-1}^n, Z_{k-1}^n) \quad \text{for } k = 1, \dots, n$$

with

$$(1.6) \quad p_n(k, x, z) = \frac{1}{2} \left( 1 + za(t_k^n, x) + \delta_n b(t_k^n, x) \right) + O(\delta_n^{1+\beta})$$

for  $k \in \{1, \dots, n\}$ ,  $x \in \mathbb{R}$  and  $z \in \{-1, +1\}$ . The assumptions on the functions  $a, b : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$  will be specified presently.

If (A2) holds, the process  $X^n$  is under  $P^n$  a *generalized correlated random walk*. Suppose there is no  $O$ -term in (1.6). In the simplest case where  $a \equiv 0$  and  $b \equiv 0$ ,  $X^n$  is just a binary random walk, and (A1) yields via Donsker's theorem that  $(X^{(n)})_{n \in \mathbb{N}}$  converges in distribution to a Brownian motion with drift  $\mu$  and volatility  $\sigma$ . If  $a \equiv 0$ , then  $p_n$  does not depend on  $z$ , and  $X^n$  alone is a Markov chain under  $P^n$ . If  $b \equiv 0$  and  $a$  does not depend on  $x$ , then  $Z^n$  is an inhomogeneous Markov chain and  $X^n$  is a cumulative sum of Markovian increments. This is called (in the time-homogeneous case) a correlated random walk (CRW);

see Chen/Renshaw (1994). The novel feature in the present paper is that the transition probabilities for  $Z^n$  (or  $X^n$ ) are allowed to depend on both the current state  $X_{k-1}^n$  and the current increment  $Z_{k-1}^n$  (or, equivalently,  $\Delta X_{k-1}^n$ ). This makes the analysis more delicate and produces more interesting limiting behaviour.

A detailed overview of much of the existing literature on CRWs is given in section 5.1 of Gruber (2004). Hence we focus here only on papers concerned with weak convergence. Renshaw/Henderson (1981) show such results for (classical) symmetric CRWs, i.e., for  $b \equiv 0$  and constant  $a$ . These CRWs constitute a special case of directionally reinforced random walks for which weak convergence has been analyzed by Mauldin/Monticino/von Weizsäcker (1996) and Horváth/Shao (1998). Szász/Tóth (1984) study symmetric and some more general CRWs in a random environment, and weak convergence to Brownian motion for a family of homogeneous CRWs has also been obtained by Opitz (1999). However, all these results are for CRWs which are homogeneous in time and space; no dependence on the current state is allowed.

Let us now return to our generalized correlated random walks. For the functions  $a, b$  that determine the transition probabilities via (1.6), we assume

**(A3)** The functions  $a, b : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$  satisfy

$$(1.7) \quad \|a\|_\infty := \sup \{|a(t, x)| \mid (t, x) \in [0, 1] \times \mathbb{R}\} < 1 \text{ and } \|b\|_\infty < \infty;$$

$$(1.8) \quad a'(t, x) := \frac{\partial}{\partial x} a(t, x) \text{ exists and is bounded on } [0, 1] \times \mathbb{R};$$

$$(1.9) \quad a'(t, x) \text{ is globally Hölder}(\beta)\text{-continuous in } x, \text{ uniformly in } t, \text{ i.e.,}$$

$$|a'(t, x) - a'(t, y)| \leq K|x - y|^\beta \quad \text{for all } x, y \in \mathbb{R} \text{ and } t \in [0, 1];$$

$$(1.10) \quad b(t, x) \text{ is globally Hölder}(\beta)\text{-continuous in } x, \text{ uniformly in } t;$$

$$(1.11) \quad a(t, x), a'(t, x) \text{ and } b(t, x) \text{ are all continuous in } t.$$

Without loss of generality, we may and do take the same  $\beta \in (0, 1)$  for (1.4), (1.5), (1.6), (1.9) and (1.10). Under (A3), we define the operator  $L$  on  $C^2$  functions  $f(x)$  by

$$(Lf)(t, x) := \frac{1}{2}\sigma^2 \frac{1+a(t, x)}{1-a(t, x)} f''(x) + \left( \mu + \frac{\sigma b(t, x)}{1-a(t, x)} + \frac{\sigma^2 a'(t, x)}{(1-a(t, x))^2} \right) f'(x).$$

Our main result is then

**Theorem 1.** *Assume (A1) – (A3) and that  $(X_0^n)_{n \in \mathbb{N}}$  converges in distribution to some  $X_0$  with distribution  $\nu$  on  $\mathbb{R}$ . If the martingale problem for  $L$  is well-posed in  $C[0, 1]$ , then  $(X^{(n)})_{n \in \mathbb{N}}$  converges in distribution to the solution  $X$  of the stochastic differential equation*

$$(1.12) \quad dX_t = \left( \mu + \frac{\sigma b(t, X_t)}{1-a(t, X_t)} + \frac{\sigma^2 a'(t, X_t)}{(1-a(t, X_t))^2} \right) dt + \sigma \sqrt{\frac{1+a(t, X_t)}{1-a(t, X_t)}} dW_t$$

with initial value  $X_0$ .

A proof of Theorem 1 is given in section 2. We provide here some comments instead.

First of all, we always have existence of a solution to the martingale problem for  $(L, \nu)$ . This follows from Theorem 6.1.7 in Stroock/Varadhan (1979) because (A3) implies that both coefficients (of  $f''(x)$  and of  $f'(x)$ ) in  $L$  are bounded in  $(t, x)$  and continuous in  $x$  for each  $t$ .

If we only suppose that the martingale problem for  $(L, \nu)$  has a unique solution, the conclusion of Theorem 1 is still true. This is not surprising and can be seen from the original proof in Gruber (2004). Even if the martingale problem has no unique solution, one can show that any subsequence of  $(X^{(n)})$  contains a further subsequence which converges in distribution to *some* solution of (1.12). (Of course, the latter need then not be unique.) For details, we refer to section 5.3.7 of Gruber (2004).

If  $a \equiv 0$ , the proof of Theorem 1 is a straightforward application of Theorem 2.1 in Nelson (1990). However, that result does not extend to  $a \neq 0$  because it requires the convergence of the conditional moments  $\delta_n^{-2} E[\Delta X_k^n | \mathcal{F}_{k-1}^n]$ .

For some applications, it is desirable to have the functions  $a$  and  $b$  depend on  $n$  as well. This is possible within certain limits; see the remark at the end of section 2.

## 2. Proof of the main result

In this section, we show how to prove Theorem 1. We do this in detail when  $a, b$  do not depend on  $t$  and then explain how to deal with the time-dependent case. We also comment on the original argument in Gruber (2004) and sketch how to extend Theorem 1 to functions  $a_n, b_n$  depending on  $n$ . Although this section is the most important contribution of the paper, readers interested only in applications can skip it and continue directly with section 3.

We do not prove Theorem 1 from first principles. Our main tool is the following result which is — up to notational changes — Theorem 7.4.1 from Ethier/Kurtz (1986).

**Proposition 2.** *Let  $c : \mathbb{R} \rightarrow [0, \infty)$  and  $\gamma : \mathbb{R} \rightarrow \mathbb{R}$  be continuous, define the operator  $G$  on  $C^\infty$  functions  $f$  with compact support by  $Gf := \frac{1}{2}cf'' + \gamma f'$  and suppose that the martingale problem for  $G$  is well-posed in  $C[0, 1]$ . In the setting of section 1, let  $\Gamma^{(n)}$  and  $C^{(n)}$  be  $\mathbb{F}^{(n)}$ -adapted processes such that  $C^{(n)}$  is increasing and both*

$$N^{(n)} := X^{(n)} - X_0^{(n)} - \Gamma^{(n)}$$

and  $(N^{(n)})^2 - C^{(n)}$  are local  $(P^n, \mathbb{F}^{(n)})$ -martingales for each  $n$ . Set

$$\tau_n^r := \inf \left\{ t \geq 0 \mid |X_t^{(n)}| \geq r \text{ or } |X_{t-}^{(n)}| \geq r \right\} \wedge 1$$

and suppose that for each  $r > 0$ ,

$$(2.1) \quad \lim_{n \rightarrow \infty} E^n \left[ \sup_{0 \leq t \leq \tau_n^r} |X_t^{(n)} - X_{t-}^{(n)}|^2 \right] = 0,$$

$$(2.2) \quad \lim_{n \rightarrow \infty} E^n \left[ \sup_{0 \leq t \leq \tau_n^r} |\Gamma_t^{(n)} - \Gamma_{t-}^{(n)}|^2 \right] = 0,$$

$$(2.3) \quad \lim_{n \rightarrow \infty} E^n \left[ \sup_{0 \leq t \leq \tau_n^r} |C_t^{(n)} - C_{t-}^{(n)}| \right] = 0$$

as well as

$$(2.4) \quad \sup_{0 \leq t \leq \tau_n^r} \left| \Gamma_t^{(n)} - \int_0^t \gamma(X_s^{(n)}) ds \right| \longrightarrow 0 \quad \text{in probability as } n \rightarrow \infty,$$

$$(2.5) \quad \sup_{0 \leq t \leq \tau_n^r} \left| C_t^{(n)} - \int_0^t c(X_s^{(n)}) ds \right| \longrightarrow 0 \quad \text{in probability as } n \rightarrow \infty.$$

If the distributions of  $X_0^{(n)}$  under  $P^n$  converge weakly to a probability measure  $\nu$  on  $\mathbb{R}$ , then  $(X^{(n)})_{n \in \mathbb{N}}$  converges in distribution to the solution of the martingale problem for  $(G, \nu)$ .

The tricky bit in the application of Proposition 2 is to find the decomposition of a given  $X^{(n)}$  into  $N^{(n)}$  and  $\Gamma^{(n)}$ , and this is linked in turn to the knowledge of the functions  $c$  and  $\gamma$  in the generator  $G$ . One of the main difficulties in Gruber (2004) was to find these functions in the first place, and so a completely different (and much longer) proof was given there. We comment on this below in some more detail.

Because each  $X^{(n)}$  is piecewise constant and so is  $\mathbb{F}^{(n)}$ , it is enough to do everything in discrete time. More precisely, we can start with  $X^n$  and look for processes  $Y^n = (Y_k^n)_{k=0,1,\dots,n}$  with  $Y \in \{N, \Gamma, C\}$  such that  $C^n$  is increasing and both  $N^n := X^n - X_0^n - \Gamma^n$  and  $(N^n)^2 - C^n$  are  $P^n$ -martingales for the filtration  $\mathbb{F}^n := (\mathcal{F}_k^n)_{k=0,1,\dots,n}$ . The corresponding processes  $Y^{(n)}$  obtained by piecewise constant interpolation like in (1.3) can then be used for Proposition 2. Moreover, the obvious choice for  $C^n$  is clearly the increasing  $\mathbb{F}^n$ -predictable process from the Doob decomposition of  $(N^n)^2$  so that we take

$$\Delta C_k^n = C_k^n - C_{k-1}^n = E^n [(N_k^n)^2 - (N_{k-1}^n)^2 | \mathcal{F}_{k-1}^n] = E^n [(\Delta N_k^n)^2 | \mathcal{F}_{k-1}^n] = \text{Var}^n[\Delta N_k^n | \mathcal{F}_{k-1}^n].$$

Here and in the sequel, we use the notation  $\Delta Y_k := Y_k - Y_{k-1}$  for the increments of a discrete-time process  $Y$ . So the first (and most laborious) step is to find the process  $\Gamma^n$ .

In order to ease the notation, we drop in the subsequent computations all sub- and superscripts  $n$  and think of a fixed  $n$ . The only exceptions are  $\mu_n$  and  $\sigma_n$  since we need to distinguish these from the constants  $\mu$  and  $\sigma$ . We also omit all time arguments since we first consider the case where  $a$  and  $b$  do not depend on  $t$ .

To find a decomposition of  $X$  like in Proposition 2, an obvious first idea is to try and use the Doob decomposition  $X = X_0 + M + A$ . So we attempt with

$$\Delta A_k := E[\Delta X_k | \mathcal{F}_{k-1}] = \mu_n + \sigma_n E[Z_k | \mathcal{F}_{k-1}] = \mu_n + \sigma_n (2p(X_{k-1}, Z_{k-1}) - 1)$$

due to (1.1) and (A2). By using (1.6) and (1.4), (1.5), we obtain

$$(2.6) \quad \Delta A_k = \mu \delta^2 + \sigma b(X_{k-1}) \delta^2 + \sigma_n Z_{k-1} a(X_{k-1}) + O(\delta^{2+\beta}).$$

To proceed with the computation of  $\Delta A_k$ , the simplest (but too naive) way is to use a Taylor expansion for  $a(X_{k-1})$  around  $X_{k-2}$ , multiply the result with  $\sigma_n Z_{k-1}$  and simplify. If we do this, we obtain on the right-hand side of (2.6) a term  $a(X_{k-2}) \Delta A_{k-1}$ , while there is  $\Delta A_k$  on the left-hand side. Asymptotically, (2.6) thus produces an expression for  $(1-a)dA$  whereas we should like to have  $dA$  itself. Hence it seems useful to divide by  $1-a$  before doing more computations.

We have deliberately not given any details in the reasoning just above since its only purpose is to provide the motivation for our next step. The upshot is that we now apply a Taylor expansion to the ratio  $\frac{a(X_{k-1})}{1-a(X_{k-1})}$ . Using also (1.9), (1.7), (1.1) and (1.4), (1.5) yields

$$(2.7) \quad \frac{a(X_{k-1})}{1-a(X_{k-1})} = \frac{a(X_{k-2})}{1-a(X_{k-2})} + \frac{a'(X_{k-2})}{(1-a(X_{k-2}))^2} \Delta X_{k-1} + O(\delta^{1+\beta});$$

the error term comes from evaluating the derivative at  $X_{k-2}$  instead of at an intermediate point between  $X_{k-2}$  and  $X_{k-1}$ , and we also use that  $|\Delta X_{k-1}| = O(\delta)$ . Now multiply (2.7) by  $\sigma_n Z_{k-1} = \Delta X_{k-1} - \mu_n$  and use (1.4), (1.5) and (1.1) to get

$$\sigma_n Z_{k-1} \frac{a(X_{k-1})}{1-a(X_{k-1})} = \frac{a(X_{k-2})}{1-a(X_{k-2})} (\Delta X_{k-1} - \mu \delta^2) + \frac{a'(X_{k-2})}{(1-a(X_{k-2}))^2} \sigma^2 \delta^2 + O(\delta^{2+\beta}).$$

Plugging this into (2.6) and using  $\Delta X = \Delta M + \Delta A$  and the identity  $\frac{a}{1-a} = \frac{1}{1-a} - 1$  gives

$$(2.8) \quad \begin{aligned} \frac{\Delta X_k}{1-a(X_{k-1})} &= \frac{\Delta M_k}{1-a(X_{k-1})} + \frac{1}{1-a(X_{k-2})} \Delta X_{k-1} - \Delta X_{k-1} \\ &\quad + \frac{\mu + \sigma b(X_{k-1})}{1-a(X_{k-1})} \delta^2 + \mu \delta^2 \left( 1 - \frac{1}{1-a(X_{k-2})} \right) + \frac{a'(X_{k-2})}{(1-a(X_{k-2}))^2} \sigma^2 \delta^2 \\ &\quad + O(\delta^{2+\beta}). \end{aligned}$$

**Lemma 3.** *Define the martingale  $N$  by*

$$(2.9) \quad N_m := \sum_{k=1}^m \frac{\Delta M_k}{1-a(X_{k-1})} \quad \text{for } m = 0, 1, \dots, n$$

and the process  $\Gamma$  by  $\Gamma := X - X_0 - N$ . Then we have

$$(2.10) \quad \Gamma_m = \sum_{k=1}^m \left( \mu + \frac{\sigma b(X_{k-1})}{1 - a(X_{k-1})} + \frac{\sigma^2 a'(X_{k-1})}{(1 - a(X_{k-1}))^2} \right) \delta^2 + O(\delta^\beta) \quad \text{for } m = 0, 1, \dots, n.$$

**Proof.** We start from (2.8) which we rewrite as

$$(2.11) \quad \begin{aligned} \Delta X_{k-1} &= \frac{\Delta M_k}{1 - a(X_{k-1})} + \left( \mu + \frac{\sigma b(X_{k-1})}{1 - a(X_{k-1})} + \frac{\sigma^2 a'(X_{k-1})}{(1 - a(X_{k-1}))^2} \right) \delta^2 \\ &+ \left( \frac{\Delta X_{k-1}}{1 - a(X_{k-2})} - \frac{\Delta X_k}{1 - a(X_{k-1})} \right) + \mu \delta^2 \left( \frac{1}{1 - a(X_{k-1})} - \frac{1}{1 - a(X_{k-2})} \right) \\ &+ \sigma^2 \delta^2 \left( \frac{a'(X_{k-2})}{(1 - a(X_{k-2}))^2} - \frac{a'(X_{k-1})}{(1 - a(X_{k-1}))^2} \right) + O(\delta^{2+\beta}). \end{aligned}$$

Because of (1.7) and

$$|\Delta X_k| = O(\delta),$$

the errors we make by summing over  $k$  from 1 or 2 and to  $m$  or  $m + 1$  are of the order  $O(\delta)$ . The third, fourth and fifth terms in (2.11) all yield telescoping series whose sums are of the order  $O(\delta)$  due to (1.7) and (1.8). Finally, since we sum at most  $n$  terms and  $\delta = \delta_n = 1/\sqrt{n}$ , the sum of the terms in  $O(\delta^{2+\beta})$  is of the order  $O(\delta^\beta)$ . Hence the assertion follows. **q.e.d.**

**Lemma 4.** Define the martingale  $N$  by (2.9) and the increasing process  $C$  by

$$(2.12) \quad C_m := \sum_{k=1}^m \frac{\text{Var}[\Delta M_k | \mathcal{F}_{k-1}]}{(1 - a(X_{k-1}))^2} \quad \text{for } m = 0, 1, \dots, n.$$

Then  $N^2 - C$  is a martingale and

$$(2.13) \quad C_m = \sum_{k=1}^m \sigma^2 \frac{1 + a(X_{k-1})}{1 - a(X_{k-1})} \delta^2 + O(\delta^\beta) \quad \text{for } m = 0, 1, \dots, n.$$

**Proof.** Since  $N$  is a martingale, it is clear from (2.9) that  $N^2 - C$  with  $C$  from (2.12) is also a martingale. To prove (2.13), it is enough to show that

$$(2.14) \quad \text{Var}[\Delta M_k | \mathcal{F}_{k-1}] = \text{Var}[\Delta X_k | \mathcal{F}_{k-1}] = (1 - a^2(X_{k-1})) \sigma^2 \delta^2 + O(\delta^{2+\beta}).$$

But if we note that (1.1), (A2) and (1.6) yield

$$\begin{aligned} \text{Var}[\Delta X_k | \mathcal{F}_{k-1}] &= \sigma_n^2 \text{Var}[Z_k | \mathcal{F}_{k-1}] \\ &= 4\sigma_n^2 p(X_{k-1}, Z_{k-1}) (1 - p(X_{k-1}, Z_{k-1})) \\ &= \sigma_n^2 (1 - Z_{k-1}^2 a^2(X_{k-1}) + O(\delta)), \end{aligned}$$



we see that (2.14) follows from (1.4).

**q.e.d.**

Having found the processes  $N$ ,  $\Gamma$  and  $C$ , we want to verify that they satisfy the assumptions of Proposition 2. So we reinstate the indices  $n$  from now on and define the functions

$$(2.15) \quad \gamma(x) := \mu + \frac{\sigma b(x)}{1 - a(x)} + \frac{\sigma^2 a'(x)}{(1 - a(x))^2}, \quad c(x) := \sigma^2 \frac{1 + a(x)}{1 - a(x)}.$$

These appear in  $L$  and as coefficients of  $\delta_n^2$  in (2.10) and (2.13), and are bounded due to (1.7) and (1.8). The conditions (2.1) – (2.3) are easy to check. In fact, (1.1) and (1.4), (1.5) give

$$\sup_{0 \leq t \leq 1} |X_t^{(n)} - X_{t-}^{(n)}| = \max_{k=1, \dots, n} |\Delta X_k^n| = O(\delta_n);$$

recall that  $X^{(n)}$  is the piecewise constant interpolation of  $X^n$ . In the same way, we obtain

$$\sup_{0 \leq t \leq 1} |\Gamma_t^{(n)} - \Gamma_{t-}^{(n)}| + \sup_{0 \leq t \leq 1} |C_t^{(n)} - C_{t-}^{(n)}| = O(\delta_n^\beta)$$

since  $\gamma$  and  $c$  are bounded.

**Lemma 5.** *For  $\Gamma^{(n)}$ ,  $C^{(n)}$  and  $\gamma, c$  defined by (2.10), (2.13) and (2.15), the conditions (2.4) and (2.5) are satisfied.*

**Proof.** Since the argument is the same in both cases, we only prove (2.4). For  $t \in [t_m^n, t_{m+1}^n)$ , using that  $X^{(n)}$ ,  $\Gamma^{(n)}$  are piecewise constant and Lemma 3 yields

$$(2.16) \quad \begin{aligned} \left| \Gamma_t^{(n)} - \int_0^t \gamma(X_s^{(n)}) ds \right| &= \left| \Gamma_m^n - \sum_{k=1}^m \int_{t_{k-1}^n}^{t_k^n} \gamma(X_s^{(n)}) ds - \int_{t_m^n}^t \gamma(X_s^{(n)}) ds \right| \\ &\leq O(\delta_n^\beta) + \sum_{k=1}^m \left| \gamma(X_{k-1}^n) \delta_n^2 - \int_{t_{k-1}^n}^{t_k^n} \gamma(X_{k-1}^n) ds \right| + \|\gamma\|_\infty \delta_n^2 \\ &= O(\delta_n^\beta), \end{aligned}$$

and since this is uniform in  $t$ , (2.4) follows.

**q.e.d.**

Lemmas 3, 4 and 5 imply that in the setting of Theorem 1, all the assumptions of Proposition 2 are satisfied. Thus  $(X^{(n)})_{n \in \mathbb{N}}$  converges in distribution to the solution of the martingale problem for  $(G, \nu)$ , which in view of (2.15) is the same as the solution of (1.12). This completes the proof of Theorem 1 if  $a$  and  $b$  do not depend on  $t$ .

Before we start on the time-dependent case, we introduce a bit of notation.

**Definition.** Consider functions  $f : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ ,  $(t, x) \mapsto f(t, x)$ . For  $\beta, \varepsilon \in (0, 1)$ ,  $H^{\frac{1}{2}\beta, \varepsilon}$  is the space of all  $f$  that are continuous and bounded, Hölder( $\frac{1}{2}\beta$ )-continuous in  $t$ , uniformly in  $x$ , and Hölder( $\varepsilon$ )-continuous in  $x$ , uniformly in  $t$ .  $H^{\frac{1}{2}(1+\beta), 1+\varepsilon}$  is the space of all  $f$  that are continuous and bounded, Hölder( $\frac{1}{2}(1+\beta)$ )-continuous in  $t$ , uniformly in  $x$ , and differentiable in  $x$  such that  $f'(t, x) = \frac{\partial}{\partial x} f(t, x)$  is bounded, Hölder( $\frac{1}{2}\beta$ )-continuous in  $t$ , uniformly in  $x$ , and Hölder( $\varepsilon$ )-continuous in  $x$ , uniformly in  $t$ .  $H^{\frac{1}{2}\beta, 1+\varepsilon}$  is the space of all  $f$  that are continuous and bounded, Hölder( $\frac{1}{2}\beta$ )-continuous in  $t$ , uniformly in  $x$ , and differentiable in  $x$  such that  $f'(t, x) = \frac{\partial}{\partial x} f(t, x)$  is bounded, Hölder( $\frac{1}{2}\beta$ )-continuous in  $t$ , uniformly in  $x$ , and Hölder( $\varepsilon$ )-continuous in  $x$ , uniformly in  $t$ .

Suppose now that  $a$  and  $b$  are allowed to depend on  $t$ . Then we can prove Theorem 1 by almost the same arguments as above if we additionally assume that

$$(2.17) \quad a \in H^{\frac{1}{2}(1+\beta), 1+\beta} \quad \text{and} \quad b \in H^{\frac{1}{2}\beta, \beta}.$$

(See below for comments how this is related to the original assumptions in Theorem 1.) In fact, (2.17) guarantees that we can again do a Taylor expansion (now in both  $t$  and  $x$ ) for  $f(t, x) := \frac{a(t, x)}{1-a(t, x)}$  and obtain (2.7) with all arguments  $X_{k-j}$  replaced by  $(t_{k-j+1}^n, X_{k-j}^n)$ . The crucial point here is that under (2.17),  $f(t, x)$  is also Hölder( $\frac{1}{2}(1+\beta)$ )-continuous in  $t$ . Once we have (2.7) in this time-dependent form, the same telescoping argument still yields Lemma 3. Lemma 4 and (2.1) – (2.3) are proved like before, and the only other change occurs in the proof of Lemma 5. First note that due to (2.17), both  $c$  and  $\gamma$  are Hölder( $\frac{1}{2}\beta$ )-continuous in  $t$ , uniformly in  $x$ . Rewriting (2.16) in time-dependent form gives

$$(2.18) \quad \begin{aligned} \left| \Gamma_t^{(n)} - \int_0^t \gamma(s, X_s^{(n)}) ds \right| &\leq O(\delta_n^\beta) + \sum_{k=1}^m \left| \gamma(t_k^n, X_{k-1}^n) \delta_n^2 - \int_{t_{k-1}^n}^{t_k^n} \gamma(s, X_{k-1}^n) ds \right| + \|\gamma\|_\infty \delta_n^2 \\ &\leq O(\delta_n^\beta) + \sum_{k=1}^m \int_{t_{k-1}^n}^{t_k^n} |\gamma(t_k^n, X_{k-1}^n) - \gamma(s, X_{k-1}^n)| ds, \end{aligned}$$

and each integrand is  $O((\delta_n^2)^{\frac{1}{2}\beta}) = O(\delta_n^\beta)$ . Thus all of the at most  $n$  summands are  $O(\delta_n^{2+\beta})$  because  $t_k^n - t_{k-1}^n = \frac{1}{n} = \delta_n^2$ ; this holds uniformly in  $t$ , and so we get

$$\sup_{0 \leq t \leq 1} \left| \Gamma_t^{(n)} - \int_0^t \gamma(s, X_s^{(n)}) ds \right| \leq O(\delta_n^\beta)$$

which clearly implies (2.4). The argument for (2.5) is completely analogous. Finally, we note that Proposition 2 can easily be extended to the time-dependent form we need here. This shows how to prove Theorem 1 under the additional assumption (2.17).

**Remark.** Comparing (A3) with (2.17) reveals that the latter is only a quantitative strengthening of (1.11); we replace mere continuity of  $a'(t, x)$  and  $b(t, x)$  in  $t$  by Hölder( $\frac{1}{2}\beta$ )-continuity in  $t$ , uniformly in  $x$ . The above sketch of the proof also shows how this is exploited.  $\diamond$

If  $a$  and  $b$  satisfy the assumptions of (A3) but not (2.17), the preceding proof no longer seems to work. The reason is that the Taylor expansion in (2.7) can then only be done in  $x$  alone. Hence it involves arguments  $(t_k^n, X_{k-1}^n)$  on the left-hand side and  $(t_k^n, X_{k-2}^n)$  — instead of  $(t_{k-1}^n, X_{k-2}^n)$  — on the right-hand side, and so the telescoping sum argument no longer works. So far, we have not been able to overcome the resulting complications in (2.8).

Nevertheless, Theorem 1 is true in the generality given here. An alternative proof can be found in Gruber (2004), but it is rather long and technically involved. We have therefore decided to prove Theorem 1 here only under slightly less general assumptions. Very briefly, the proof in Gruber (2004) can be summarized as follows. In a first step, assuming (1.4), (1.5), (1.7) and (1.8), the sequence  $(X^{(n)})_{n \in \mathbb{N}}$  is shown to be tight in  $D[0, 1]$  by deriving precise bounds on product moments of the increments of  $X^n$  and then employing techniques from Billingsley (1968) for controlling the fluctuations of partial sums of not necessarily independent or identically distributed random variables. The second step then shows that the weak limit of any convergent subsequence of  $(X^{(n)})$  solves the martingale problem for  $(L, \nu)$ . Since the correlation between two successive increments of  $X^{(n)}$  does not vanish as  $n \rightarrow \infty$ , this requires a careful consideration of conditional moments on a time scale of the order  $O(\delta_n)$ . By further refining the arguments in Gruber (2004), one can probably even abandon the Hölder-continuity assumptions in the  $x$ -variable as well.

**Remark.** If we want to allow in (1.6)  $a$  and  $b$  that depend on  $n$ , this can be done as follows. For  $f \in \{a, b\}$ , replace  $f(t, x)$  by  $f_n(t, x) = f(t, x) + \xi_n^f$  with constants  $(\xi_n^f)_{n \in \mathbb{N}}$  satisfying

$$(2.19) \quad \xi_n^f = O(\delta_n^\beta) \quad \text{for some } \beta \in (0, 1).$$

All other conditions are unchanged. If we have (2.17), we can still prove Theorem 1 almost as above by replacing  $f$  with  $f_n$ ; the only change is that the integrand in (2.18) becomes

$$|\gamma_n(t_k^n, X_{k-1}^n) - \gamma(s, X_{k-1}^n)| \leq |(\gamma_n - \gamma)(t_k^n, X_{k-1}^n)| + |\gamma(t_k^n, X_{k-1}^n) - \gamma(s, X_{k-1}^n)|.$$

But combining (2.19) with the conditions in (A2) easily yields  $\|\gamma_n - \gamma\|_\infty \leq O(\delta_n^\beta)$  as well as  $\|c_n - c\|_\infty \leq O(\delta_n^\beta)$ , and this allows us to finish the proof as before.  $\diamond$

### 3. Examples and applications

In this section, we present three situations where Theorem 1 can be useful.

#### 3.1. Option pricing with a large investor

Our first example actually provided the motivation and setup for Theorem 1. However, its details are too involved for a full presentation here. Hence we only sketch the main ideas and refer to Gruber (2004) and forthcoming work for more information.

The basic question is easy to explain. A large investor in a financial market wants to price an option, and “large” means that the hedging strategy he plans to construct for replicating the option has an impact on the price process of the underlying asset (stock, say). What are then the value of the option and the resulting stock price evolution?

To formalize this situation, we start with a discrete-time model where uncertainty is generated by a binary random walk  $X^n$ ; this describes some fundamentals in the financial market. We construct a mechanism for the price formation of the stock and thus obtain its price process for any exogenously given strategy of the large investor. However, the hedging strategy we want must be determined endogenously since it must replicate a payoff on the given stock. So it will be given via a fixed point argument, and the resulting transition probabilities for the evolution of uncertainty can only be obtained from an implicit equation.

It turns out under smoothness conditions on the payoff that the hedging function for the option can be recursively described backward in time. This yields a difference equation whose continuous-time limit is a nonlinear PDE. Under suitable assumptions, one can show that this PDE has a unique solution  $\varphi$  and that the transition probabilities in discrete time can be described like in (1.6) in terms of  $\varphi$ . More precisely, this is true for the transition probabilities of  $X^n$  under a measure which turns into martingales both the stock price process resulting for the large investor and his valuation process for the option. Using Theorem 1 then produces continuous-time models for option valuation with a large investor. In particular, this provides new insights into the precise impact of the model chosen for the market mechanism.

In a little bit more detail, the price formation mechanism is described by two ingredients: a reaction function  $\psi(t, x, \vartheta)$  of time  $t$ , current fundamental value  $x$  and stock holding  $\vartheta$  of the large investor, and a measure  $\varrho$  which models the timing in forming the price at which the large investor can trade. The diffusion coefficient of the limit process  $X$  is then given by

$$\sigma_\varphi^2(t, x) = \frac{\psi_x(t, x, \varphi(t, x)) + 2d(\varrho)\psi_\vartheta(t, x, \varphi(t, x))\varphi_x(t, x)}{\psi_x(t, x, \varphi(t, x)) + \psi_\vartheta(t, x, \varphi(t, x))\varphi_x(t, x)}$$

with  $d(\varrho) = \int_{\mathbb{R}} z \varrho(dz) - \frac{1}{2}$ ; see Theorem 4.4 and (4.2.22) in Gruber (2004).

Even without going into any further detail, we can explain why generalized correlated random walks come up in this context. It is well known that the value of an option whose payoff at time 1 is of the form  $h(X_1)$  can usually be obtained as a function  $v(t, X_t)$  and that the corresponding hedging strategy is given in terms of the derivative  $\frac{\partial v}{\partial x}(t, X_t)$ . If we now look at a large investor in discrete time, the price formation mechanism involves not only the current price  $X_k^n$ , but via the strategy’s impact also the increment  $\Delta X_k^n$ . Hence the transition probabilities of  $X^n$  take the form (A2) and we no longer have a simple Markovian structure for  $X^n$  or  $Z^n$  alone.

### 3.2. Option pricing under transaction costs

Consider a financial market with a bank account  $B$  and a stock  $S$  traded with transaction costs as follows. If the (nominal) stock price at time  $t$  is  $S_t$ , one can buy a share for  $(1 + \kappa^b)S_t$ , but sell it only for  $(1 - \kappa^s)S_t$ , where  $\kappa^b$  and  $\kappa^s$  are both in  $[0, 1)$ . What is then a reasonable price for a European call option on  $S$ ?

This question has been studied and answered in the well-known Cox/Ross/Rubinstein model where the stock price process is given by a geometric binary random walk. More precisely, suppose that  $S^n = (S_k^n)_{k=0,1,\dots,n}$  is given by

$$(3.1) \quad S_k^n = S_0^n \exp(X_k^n) = S_0^n \exp\left(\sum_{j=1}^k \Delta X_j^n\right),$$

where the log-returns  $\Delta X_j^n = \log(S_j^n/S_{j-1}^n)$  take the values  $u_n$  and  $d_n$  with  $u_n > r_n > d_n$ , while the bank account evolves according to  $B_k^n = \exp(kr_n)$  for  $k = 0, 1, \dots, n$ . Boyle/Vorst (1992) have shown that there exists a unique strategy which is self-financing inclusive of transaction costs and has a final wealth  $(S_n^n - K)^+ = h(S_n^n)$  equal to the payoff of the call. The option's value is thus the initial cost for this strategy, and it turns out that it can be computed as the expectation of the discounted payoff  $h(S_n^n)/B_n^n$  under a probability measure  $P_n^*$  under which the  $\Delta X_k^n$  form a Markov chain. Using the central limit theorem and an appropriate scaling of parameters, Boyle/Vorst (1992) also show that this option pricing formula converges to the Black/Scholes formula, but with a higher variance than in the case of no transaction costs. For a general payoff function  $h$ , Kusuoka (1995) gives an expression for the limit of the superreplication price for  $h(S_n^n)$ ; his results also show that the superreplication price coincides with the above replication price if  $h$  is convex and monotonic.

As in section 1,  $S^{(n)}$  denotes the piecewise constant interpolation on  $[0, 1]$  of  $S^n$ . While convergence of option prices is a result on the one-dimensional marginal distributions of  $S^{(n)}$  at time 1, we are here interested in weak convergence for the entire processes  $S^{(n)}$ . In the case without transaction costs, the  $\Delta X_k^n$  are independent under  $P_n^*$  and it follows from Donsker's theorem that the Cox/Ross/Rubinstein models converge weakly to the Black/Scholes model of geometric Brownian motion; see Chapter 22 of Duffie (1988) for a detailed account. The case of transaction costs has been treated in the (unpublished German) diploma thesis of A. Opitz (1999), and we now show how to deduce this from Theorem 1.

Let us first specify the parameters. As in section 1, we work on  $[0, 1]$  and write  $\delta_n = 1/\sqrt{n}$ . The random variables  $\Delta X_k^n = \mu_n + \sigma_n Z_k^n$  take the values

$$(3.2) \quad u_n = \mu\delta_n^2 + \sigma\delta_n \quad \text{and} \quad d_n = \mu\delta_n^2 - \sigma\delta_n \quad \text{with} \quad \mu \in \mathbb{R}, \sigma > 0$$

so that we have

$$(3.3) \quad \sigma_n = \sigma\delta_n \quad , \quad \mu_n = \mu\delta_n^2.$$

The bank account is determined by

$$(3.4) \quad r_n = \varrho \delta_n^2 \quad \text{with } \varrho \in \mathbb{R},$$

and transaction costs are specified by

$$(3.5) \quad \kappa_n^x = \kappa^x \delta_n \quad \text{with } 0 \leq \kappa^x < 1 \quad \text{for } x \in \{b, s\}.$$

The probability measure  $P_n^*$  is defined by its transition probabilities

$$P_n^*[Z_k^n = +1 | \mathcal{F}_{k-1}^n] = P_n^*[Z_k^n = +1 | Z_{k-1}^n]$$

with

$$(3.6) \quad \begin{aligned} P_n^*[Z_k^n = +1 | Z_{k-1}^n = +1] &= p_{n,U}^* := \frac{e^{r_n}(1 + \kappa_n^b) - e^{d_n}(1 - \kappa_n^s)}{e^{u_n}(1 + \kappa_n^b) - e^{d_n}(1 - \kappa_n^s)}, \\ P_n^*[Z_k^n = +1 | Z_{k-1}^n = -1] &= p_{n,D}^* := \frac{e^{r_n}(1 - \kappa_n^s) - e^{d_n}(1 - \kappa_n^s)}{e^{u_n}(1 + \kappa_n^b) - e^{d_n}(1 - \kappa_n^s)}; \end{aligned}$$

the initial distribution of  $Z_0^n$  plays asymptotically no role. By writing (3.6) as

$$P_n^*[Z_k^n = +1 | Z_{k-1}^n] = \frac{1}{2} \left( p_{n,U}^* + p_{n,D}^* + Z_{k-1}^n (p_{n,U}^* - p_{n,D}^*) \right),$$

we can obtain

**Lemma 6.** *In the above setting, we have*

$$P_n^*[Z_k^n = +1 | \mathcal{F}_{k-1}^n] = p_n^*(k, Z_{k-1}^n)$$

with

$$(3.7) \quad p_n^*(k, z) = \frac{1}{2} (1 + z\lambda_n + \delta_n \varphi_n) + O(\delta_n^2)$$

and

$$(3.8) \quad \lambda_n := \lambda + O(\delta_n) := \frac{\kappa^b + \kappa^s}{2\sigma + \kappa^b + \kappa^s} + O(\delta_n) \quad \text{as } n \rightarrow \infty,$$

$$(3.9) \quad \varphi_n := \varphi + O(\delta_n) := \frac{2(\varrho - \mu) - (\kappa^b + \kappa^s + \sigma)\sigma}{2\sigma + \kappa^b + \kappa^s} + O(\delta_n) \quad \text{as } n \rightarrow \infty.$$

**Proof.** We omit the proof since it consists of straightforward computations.

If we compare (3.7) to (1.6), we see that the present situation does not exactly fit into the setting of Theorem 1. However, the extension explained in the last remark in section 2 allows us to derive the following result first obtained by Opitz (1999).

**Theorem 7.** Suppose  $S^{(n)}$  is the piecewise constant interpolation of  $S^n$  defined by (3.1) with (3.2) – (3.5). Let  $P_n^*$  be given by (3.6) and suppose that  $S_0^n \rightarrow S_0$  as  $n \rightarrow \infty$ . Then  $(S^{(n)})_{n \in \mathbb{N}}$  converges in distribution under  $P_n^*$  to

$$S_t = S_0 \exp\left(\bar{\sigma}W_t + \left(\varrho - \frac{1}{2}\bar{\sigma}^2\right)t\right) \quad , \quad 0 \leq t \leq 1$$

for a Brownian motion  $W$ , where

$$\bar{\sigma}^2 := \sigma(\sigma + \kappa^b + \kappa^s) = \sigma^2 \left(1 + \frac{\kappa^b + \kappa^s}{\sigma}\right).$$

**Proof.** Choose  $a_n(t, x) := \lambda_n$  and  $b_n(t, x) := \varphi_n$  to obtain from (2.15) and (3.8), (3.9) that

$$\gamma(t, x) = \mu + \frac{\sigma\varphi}{1-\lambda} = \varrho - \frac{1}{2}\bar{\sigma}^2 \quad , \quad c(t, x) = \sigma^2 \frac{1+\lambda}{1-\lambda} = \bar{\sigma}^2.$$

Hence the result follows by applying the  $n$ -dependent extension of Theorem 1 to  $X^{(n)} = \log(S^{(n)}/S_0^{(n)})$  under  $P_n^*$  and using the continuous mapping theorem. **q.e.d.**

**Remark.** Of course, Theorem 7 implies the convergence result of Boyle/Vorst (1992).  $\diamond$

### 3.3. Approximating diffusions by regular recombining binomial trees

In many situations, one needs to approximate a diffusion process  $X$  given by an SDE by a discrete-time process  $X^n$  in order to compute approximations for some functionals of  $X$ . A typical example occurs in mathematical finance if one wants to compute values and hedging strategies for options written on  $X$ . The value function  $v$  is usually given by some expectation, and the hedging strategy involves the derivative of  $v$ . To approximate this efficiently by discrete differences, one would like to have some control over the values taken by  $X^n$ . In particular, computations are often more efficient if these values lie on a regular grid. Our next application of Theorem 1 shows how this can be achieved. For those readers who skipped the proof of Theorem 1, we recall that the Hölder spaces  $H^{\frac{1}{2}\beta, \varepsilon}$  are defined in section 2.

**Theorem 8.** Suppose that  $\hat{\sigma} \in H^{\frac{1}{2}\beta, 1+\beta}$  and  $\hat{\mu} \in H^{\frac{1}{2}\beta, \beta}$  with  $\beta \in (0, 1)$ . If  $\hat{\sigma} : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$  is bounded away from 0, uniformly in  $(t, x)$ , there exists a sequence of generalized correlated random walks  $X^n$  like in section 1 whose corresponding piecewise constant interpolations  $X^{(n)}$  converge in distribution to the process  $X$  given by the SDE

$$(3.10) \quad dX_t = \hat{\mu}(t, X_t) dt + \hat{\sigma}(t, X_t) dW_t.$$

**Proof.** If we define the functions  $a$  and  $b$  by

$$a(t, x) := \frac{\hat{\sigma}^2(t, x) - 1}{\hat{\sigma}^2(t, x) + 1} \quad \text{and} \quad b(t, x) := 2 \frac{\hat{\mu}(t, x) - \hat{\sigma}(t, x)\hat{\sigma}'(t, x)}{\hat{\sigma}^2(t, x) + 1},$$

a straightforward calculation yields

$$\frac{1 + a(t, x)}{1 - a(t, x)} = \hat{\sigma}^2(t, x) \quad \text{and} \quad \frac{b(t, x)}{1 - a(t, x)} + \frac{a'(t, x)}{(1 - a(t, x))^2} = \hat{\mu}(t, x).$$

We now define  $X^n$  by (1.2) with  $X_0^n = X_0$ ,  $Z_0^n = 1$ , parameters  $\mu_n \equiv 0$ ,  $\sigma_n = \delta_n = 1/\sqrt{n}$  and the transition function  $p_n$  as in (1.6) with the above  $a$  and  $b$ . Then the assertion follows directly from Theorem 1 with  $\mu = 0$  and  $\sigma = 1$ . **q.e.d.**

To see why Theorem 8 is useful, let  $\hat{\mu}(t, x)$  or  $\hat{\sigma}(t, x)$  depend on the space variable  $x$ . The most straightforward way of approximating  $X$  by a binomial process  $\tilde{X}^n$  is to set

$$P^n[\tilde{X}_k^n = \tilde{X}_{k-1}^n + \hat{\mu}(t_k^n, \tilde{X}_{k-1}^n)\delta_n^2 \pm \hat{\sigma}(t_k^n, \tilde{X}_{k-1}^n)\delta_n \mid \mathcal{F}_{k-1}^n] = \frac{1}{2}.$$

However, this leads to a non-recombining tree for the paths of  $\tilde{X}^n$  since the value after an up move followed by a down move need not coincide with the value attained if the steps are taken in reverse order. A common method to construct a recombining binomial approximation of the diffusion (3.10) is due to Nelson/Ramaswamy (1990). They first construct a suitable transformation  $g(X)$  of  $X$  with constant volatility, develop an approximation for  $g(X)$  by a simple binomial process on a recombining tree, and then apply the inverse of  $g$  to obtain a binomial approximation  $\bar{X}^n$  of  $X$  itself. The paths of  $\bar{X}^n$  are then still recombining, but the corresponding tree is compressed and stretched in space in a patchy way.

The proof of Theorem 8 now constructs a generalized correlated random walk  $X^n$  which satisfies  $|X_k^n - X_{k-1}^n| = \delta_n$  for all  $k$ . This means that the corresponding binomial tree is both recombining and homogeneous in space. The striking simplicity of this structure is obtained by allowing the transition probabilities of  $X^n$  to depend not only on  $t_k^n$  and  $X_k^n$ , but on the increment  $\Delta X_k^n = Z_k^n$  as well. It will be interesting to see if this can be used to improve calculation efficiency in financial models.

**Acknowledgments.** Financial support by the Deutsche Forschungsgemeinschaft through the Graduiertenkolleg ‘‘Stochastische Prozesse und probabilistische Analysis’’ at the Technical University of Berlin is gratefully acknowledged. Part of this research has been carried out within the project ‘‘Mathematical Methods in Financial Risk Management’’ of the National Centre of Competence in Research ‘‘Financial Valuation and Risk Management’’ (NCCR FINRISK). The NCCR FINRISK is a research program supported by the Swiss National Science Foundation. Parts of this paper were written in Cambridge at the Isaac Newton Institute for Mathematical Sciences, and MS is grateful for financial support there in connection with the programme on *Developments in Quantitative Finance*.



## References

- P. Billingsley (1968), “Convergence of Probability Measures”, *Wiley, New York*
- P. P. Boyle and T. Vorst (1992), “Option replication in discrete time with transaction costs”, *Journal of Finance* 47, 271–293
- A. Chen and E. Renshaw (1994), “The general correlated random walk”, *Journal of Applied Probability* 31, 869–884
- D. Duffie (1988), “Security Markets. Stochastic Models”, *Academic Press, Boston*
- S. Ethier and T. Kurtz (1986), “Markov Processes. Characterization and Convergence”, *Wiley, New York*
- U. Gruber (2004), “Convergence of Binomial Large Investor Models and General Correlated Random Walks”, *PhD thesis, Technical University of Berlin*, [http://www.edocs.tu-berlin.de/diss/2004/gruber\\_urs.htm](http://www.edocs.tu-berlin.de/diss/2004/gruber_urs.htm)
- L. Horváth and Q.-M. Shao (1998), “Limit distributions of directionally reinforced random walks”, *Advances in Mathematics* 134, 367–383
- S. Kusuoka (1995), “Limit theorem on option replication cost with transaction costs”, *Annals of Applied Probability* 5, 198–221
- R. D. Mauldin, M. Monticino and H. von Weizsäcker (1996), “Directionally reinforced random walks”, *Advances in Mathematics* 117, 239–252
- D. B. Nelson (1990), “ARCH models as diffusion approximations”, *Journal of Econometrics* 45, 7–38
- D. B. Nelson and K. Ramaswamy (1990), “Simple binomial processes as diffusion approximations in financial models”, *Review of Financial Studies* 3, 393–430
- A. Opitz (1999), “Zur Asymptotik der Bewertung von Optionen unter Transaktionskosten im Binomialmodell”, *diploma thesis, Technical University of Berlin*
- E. Renshaw and R. Henderson (1981), “The correlated random walk”, *Journal of Applied Probability* 18, 403–414
- D. Szász and B. Tóth (1984), “Persistent random walks in a one-dimensional random environment”, *Journal of Statistical Physics* 37, 27–38
- D. W. Stroock and S. R. S. Varadhan (1979), “Multidimensional Diffusion Processes”, *Springer, New York*