Mixed Lognormal Distributions for Derivatives Pricing and Risk-Management

Dietmar P.J. Leisen

McGill University[1], CIREQ[2] and CRM[3]

This version: February 8, 2004.

Abstract

Many derivatives prices and their Greeks are closed-form expressions in the Black-Scholes model; when the terminal distribution is a mixed lognormal, prices and Greeks for these derivatives are then a weighted average of these (closed-form) expressions. They can therefore be calculated easily and efficiently for mixed lognormal distributions. This paper constructs mixed lognormal distributions that approximate the terminal distribution in the Merton model (Black-Scholes model with jumps) and in stochastic volatility models. Main applications are the pricing of large portfolio positions and their risk-management.

Keywords

mixed lognormal distribution, jump-diffusion, stochastic volatility, Greeks, risk-management

JEL Classification

C63, G13

---

[1] McGill University, Faculty of Management, 1001 Sherbrooke St. W., Montreal, PQ H3A 1G5, email: Dietmar.Leisen@McGill.CA.

[2] Center for Inter-university Research in Quantitative Economics (CIREQ), Montreal, Quebec.

[3] Center for Research in Mathematics (CRM), Montreal, Quebec.

The author gratefully acknowledges financial assistance provided by FCAR, NSERC and IFM2.

1 Introduction

Financial institutions holding large positions in derivatives need to calculate prices efficiently for trading purposes; for risk-management they are interested in Greeks and for risk-measurement in tail probabilities (VaR,"Value-at-Risk"), expected tail losses or prices of market value insurance against losses [1]. One of several advantages of the Black-Scholes setup is that pricing formulas for many derivatives are known in closed-form so that the before-mentioned calculations can be performed efficiently. However, the lognormal distribution does not fit well the distribution of the underlying security; common extensions of the Black-Scholes setup that incorporate jumps or stochastic volatility have better statistical properties [2] but typically closed-form expressions are not available and calculations are cumbersome. This paper constructs sequences of mixed lognormal distributions that approximate their marginal distributions and retain the computational tractability of the Black-Scholes setup.

We construct approximations for the Merton model (Black-Scholes with jumps) and models of stochastic volatility. These common extensions of the Black-Scholes setup provide a rich framework for pricing and risk-management purposes. Our approximations are based on the observation that price changes are approximately lognormal over each period; in the Merton model we assume that over each period at most one jump occurs; for stochastic volatility models we first construct a Markov chain for the volatility process in the spirit of Nelson and Ramaswamy (1990) and then extend it to the securities' process. For both setups we prove that the resulting sequence of mixed lognormal distributions converges to the terminal distribution in the corresponding continuous-time model.

Many techniques have been developed to price derivatives in jump-diffusion and stochastic volatility models. The most versatile among these are Monte-Carlo methods, see Glasserman (2004) for a recent overview, and generalizations of the Cox-Ross-Rubinstein binomial models. Among others, Amin (1993) and Hilliard and Schwartz (2003) provide approxima-

[1] We refer to Hull (2000) for a discussion of the Greeks ("The Greek Letters," Chapter 13) and to Jorion (2001) for a discussion of risk-measures and VaR in particular.

[2] For a discussion, see, e.g. Campbell, Lo, and MacKinlay (1997), Bates (2000) and Eraker, Johannes, and Polson (2003).

tions for jump-diffusions, and Hilliard and Schwartz (1996), Ritchken and Trevor (1999), Duan and Simonato (2001), and Leisen (2000) construct approximations for GARCH and stochastic volatility models. However, these techniques do not leverage on our knowledge of closed-form expressions in the Black-Scholes model to provide efficient approximations for prices and Greeks.

Recently, the computational advantages[3] of mixed lognormal distributions have been pointed out by Brigo and Mercurio (2002) and Alexander, Brintalos, and Nogueira (2004). Brigo and Mercurio (2002) study conditions under which a terminal distribution based on a generalized Black-Scholes setup (time and state-dependent volatility, constant mean) is given by a mixture of lognormals. Alexander, Brintalos, and Nogueira (2004) provide an extension of this approach where the volatility path is driven by an ad-hoc binomial tree. Our contribution to the literature is to *construct* sequences of mixed lognormal distributions that converge to the jump-diffusion models currently used in the literature.

The remainder of the paper is organized as follows: the following section introduces mixed lognormal distributions and the concept of weak convergence. The third section constructs sequences of mixed lognormal distributions for the Merton model (Black-Scholes model with jumps) and discusses efficiency of the numerical schemes. The fourth section parallels that of the third one but looks at the Black-Scholes model with stochastic volatility. The fifth section concludes the paper.

2 Mixed Lognormal Distributions

Definition 1 *A* mixed lognormal distribution *is a random variable A with*

$$A \stackrel{d}{=} A_0 \sum_{i=0}^{M} 1_{C=i} X_i,$$

where $A_0 \in \mathbb{R}^+$, and each X_i is lognormally distributed, i.e. $X_i \stackrel{d}{=} \exp(\mu_i + \sigma_i Y_i)$, $\mu_i \in \mathbb{R}, \sigma_i > 0$, Y_i independent standard normal random variables under the probability measure

[3] Mixed lognormal distributions have been used mainly in the literature to find a marginal distribution that consistently prices traded derivatives, see, e.g. Gemmill and Saflekos (2000) and Melick and Thomas (1997).

Q; the random variable C is assumed independent of the X_i and takes values on $\{0, 1, \ldots, M\}$ with $Q[C = i] = \gamma_i \geq 0$ $(i = 0, \ldots, M)$, $\sum_{i=0}^{M} \gamma_i = 1$.

Throughout we denote by "$\stackrel{d}{=}$" equality in distribution of random variables on the probability space $(\Omega, \mathcal{F}, Q)$ and by 1_F the indicator random variable for $F \in \mathcal{F}$, i.e. the random variable $1_F(\omega)$ is equal to 1 if $\omega \in F$, and zero otherwise. The interest rate r is always constant over time.

We are interested in pricing European-style derivatives with maturity $T > 0$ written on a single underlying security; according to Harrison and Kreps (1979), and Harrison and Pliska (1981) we price derivatives as discounted expected payoffs under a so-called *risk-neutral probability measure*. Here we assume Q describes that probability measure, and calculate the price of the derivative with payoff $f(A)$ at time T as $e^{-rT}E[f(A)]$. For lognormal distributions, closed-form expressions are known for many derivatives prices. These can be generalized to

$$\text{price under } A = e^{-rT}E[f(A)] = \sum_{i=1}^{n} Q[C = i] e^{-rT} E[f(A_0 X_i)] \tag{1}$$

$$= \text{weighted sum of prices under } A_0 X_i. \tag{2}$$

Since differential operators are linear, we can write the Greeks for $A^{(n)}$ also as a weighted average of the Greeks in $A_0 X_i$, e.g.,

$$\begin{aligned}\Delta \text{ under } A &= \frac{\partial e^{-rT}E[f(A)]}{\partial A_0} = \sum_{i=1}^{n} Q[C = i] \cdot \frac{\partial e^{-rT}E[f(A_0 X_i)]}{\partial A_0} \\ &= \text{weighted sum of } \Delta\text{'s under } A_0 X_i.\end{aligned}$$

For example, in the Black-Scholes setup the price of the underlying security is (under Q) $S_T = S_0 \exp\{(r - \sigma^2/2)T + \sigma W_T\}$, the standard call option with strike K is the payoff $f(A) = (A - K)^+$ and has price given by the *Black-Scholes (call option) formula* $BS\left(S_0, K, (r - \sigma^2/2)\, T, \sigma\sqrt{T}\right)$, where [4]

[4] This is an extension of the formula of Black and Scholes (1973): here we defined μ, σ to contain the maturity T, whereas the original keeps these parameters separate; furthermore the original takes $\mu = (r - \sigma^2/2)T$.

$$BS(A_0, K, \mu, \sigma) = A_0 \Phi_n\Big(d(K, \mu, \sigma)\Big) - K \cdot \exp\left(-\mu - \frac{\sigma^2}{2}\right) \cdot \Phi_n\Big(d(K, \mu, \sigma) - \sigma\Big), \tag{3}$$

$$d(K, \mu, \sigma) = \frac{\ln(S_0/K) + \mu + \sigma^2}{\sigma}, \tag{4}$$

and $\Phi_n(\cdot)$ denotes the standard normal distribution function. The derivatives price for a call option is then

$$e^{-rT} E[(A-K)^+] = e^{-rT} \sum_{i=1}^{n} Q[C=i] \cdot E[(X_i - K)^+] \tag{5}$$

$$= e^{-rT} \sum_{i=1}^{N} \gamma_i \cdot \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right) \cdot BS(A_0, K, \mu_i, \sigma_i). \tag{6}$$

Also, e.g., the Δ in the Black-Scholes model is $\Phi_n\Big(d(K, (r - \frac{\sigma^2}{2})T, \sigma\sqrt{T})\Big)$ and so the Δ for the mixed lognormal is

$$\sum_{i=1}^{n} Q[C=i]\ \Phi_n\Big((K, \mu_i, \sigma_i)\Big).$$

Other derivatives for which prices and their Greeks can be calculated easily are, e.g., chooser, exchange, compound and binary options.

In the following sections we construct approximations for the distribution S_T of prices of the underlying security at time T using sequences of mixed lognormal distributions $A^{(n)}$. We assume each $A^{(n)}$ is characterized using $C^{(n)}, X_i^{(n)}$ $(i = 1, \ldots, M^{(n)}; n = 1, 2, \ldots)$ and calculate $e^{-rT} E[f(A^{(n)}]$. We are here interested in the convergence of "prices" $e^{-rT} E[f(A^{(n)})] \xrightarrow{n} e^{-rT} E[f(S_T)]$ or equivalently $E[f(A^{(n)})] \xrightarrow{n} E[f(S_T)]$ for *any* European-style derivative payoff function f. When this holds for all *bounded* payoff functions f this is equivalent to the mathematical concept of convergence in distribution $A^{(n)} \overset{d}{\Longrightarrow} S_T$. In the following sections this will be the convergence concept we strive for [5].

[5] The "boundedness" condition excludes some payoffs, e.g. call options. It excludes payoffs like call options. However, for example for the call options put-call parity and the fact that the put option is bounded implies this property.

3 Mixed Lognormal Distributions as an Approximation to Black-Scholes with Jumps

3.1 Continuous-time Dynamics

The model of Merton (1976) is an extension of the Black-Scholes setup that incorporates jumps; we assume that on a finite interval $[0, T]$ the dynamics of security S under the risk-neutral pricing measure Q is

$$S_t = S_0 \cdot \exp\{\mu t + \sigma W_t\} \prod_{i=1}^{N_t} U_i, \text{ where } \mu = r - \frac{\sigma^2}{2} - \lambda \cdot (\nu - 1), \nu = \exp\left(\alpha + \frac{\beta^2}{2}\right), \quad (7)$$

and $(W_t)_{0 \leq t \leq T}$ is a standard Wiener process, $\mu \in R$, $\sigma > 0$, $(N_t)_{0 \leq t \leq T}$ is a Poisson process with constant parameter $\lambda > 0$, $(U_i)_i$ a sequence of serially independent lognormal random variables, i.e. each $U_i \stackrel{d}{=} \exp(\alpha + \beta Y_i)$ with Y_i a standard normal random variable and $\alpha \in \mathbb{R}, \beta > 0$. The processes N, W and the random variables U_i (respectively Y_i), $i = 1, 2, \ldots$ are assumed to be mutually independent of the each others.

Securities prices in this model follow a geometric Brownian motion from one jump time until the next jump time τ of the Poisson process. If N then changes from, say, i to $i+1$ we observe a per-cent change $U_i - 1$, i.e., the security changes value from $S_{\tau-}$ before the jump to $S_{\tau-} \cdot U_i$. Therefore, the first part in equation (7) models the evolution of the security in "normal" times, and the second part $\prod_{i=1}^{N_t} U_i$ models the additional dynamics in "extraordinary" times. Note that the Poisson process is "memoryless;" therefore the expected waiting time for the next shock is equal to $1/\lambda$, independent of current time.

We write the continuous-time price as $e^{-rT} E[(S_T - K)^+] = E\left[e^{-rT} E[(S_T - K)^+ | N_T]\right]$ and calculate then, based on equations (5, 6, 7), $E[(S_T - K)^+ | N_T = j] = E[E[(S_T - K)^+ | U_1, \ldots, U_j] | N_T = j] = E[BS(S_0 \exp(-\lambda \cdot (\nu - 1)T) \prod_{k=1}^{j} U_k, K, r - \frac{\sigma^2}{2}, \sigma^2) | N_T = j]$. This gives the continuous-time price $e^{-rT} E[(S_T - K)^+]$ as, see Merton (1976),

$$e^{-\lambda T} \sum_{j=0}^{\infty} \frac{(\lambda T)^j}{j!} E\left[BS\left(S_0 \exp\Big(-\lambda \cdot (\nu - 1)T\Big) \prod_{k=1}^{j} U_k, K, r - \frac{\sigma^2}{2}, \sigma^2\right)\right].$$

We restrict ourselves to lognormal U_i, while Merton (1976) studies also more general cases.

With our assumptions we can then write the call option price as

$$e^{-\lambda T}\sum_{j=0}^{\infty}\frac{(\lambda T)^j}{j!}E\left[BS\left(S_0\exp\Big(-\lambda\cdot(\nu-1)T+j\cdot(\alpha+\beta Y_j)\Big),K,r-\frac{\sigma^2}{2},\ \sigma^2\right)\right]. \tag{8}$$

3.2 Constructing a Sequence of Mixed Lognormal Distributions Based on an Approximation of the Process N

In this subsection we construct sequences of mixed lognormal distributions based on an approximation of the process $(N_t)_t$ over time. We start with a sequence Y_k $(k=0,1,2,\ldots)$ of serially independent standard normal random variables and for given integer n we discretize the interval $[0,T]$ into n equidistant time spots $t_k^{(n)}=k\Delta t^{(n)}$, $\Delta t^{(n)}=\frac{T}{n}$ $(k=0,\ldots,n)$.

The processes $\left(N_k^{(n)},S_k^{(n)}\right)_{k=0,\ldots,n}$ will be constructed by forward induction: First, we set $N_0^{(n)}=N_0$ and $S_0^{(n)}=S_0$. Then we assume that the processes have been defined for all dates 0 to k. The distribution of $N_{t_{k+1}^{(n)}}$ conditional on $N_{t_k^{(n)}}=N_k^{(n)}$ is

$$Q\left[N_{t_{k+1}^{(n)}}=N_k^{(n)}+i\,\middle|\,N_{t_k^{(n)}}=N_k^{(n)}\right]=\exp\left\{-\lambda\Delta t^{(n)}\right\}\cdot\frac{\left(\lambda\Delta t^{(n)}\right)^i}{i!}.$$

Note that, for "small" $\Delta t^{(n)}$, $Q[N_{t_{k+1}^{(n)}}=N_k^{(n)}|N_{t_k^{(n)}}=N_k^{(n)}]=\exp\left\{-\lambda\Delta t^{(n)}\right\}\approx 1-\lambda\Delta t^{(n)}$, and $Q[N_{t_{nk+1}}>N_k^{(n)}|N_{t_k^{(n)}}=N_k^{(n)}]\approx\lambda\Delta t^{(n)}$. Our idea is to approximate $N_{t_{k+1}^{(n)}}$ by a bivariate random variable taking value $N_k^{(n)}$ or $N_{k+1}^{(n)}$ with probabilities $1-\lambda\Delta t^{(n)}$, $\lambda\Delta t^{(n)}$, respectively.

The distribution of the security at time $t_{k+1}^{(n)}$ is, conditional on $S_{t_k^{(n)}}=S_k^{(n)}$ and $N_{t_k^{(n)}}=N_k^{(n)}$,

$$S_k^{(n)}\cdot\exp\left\{\mu\Delta t^{(n)}+\sigma\cdot\left(W_{t_{k+1}^{(n)}}-W_{t_k^{(n)}}\right)\right\}\cdot\prod_{i=N_k^{(n)}}^{N_{t_{nk+1}}}U_i.$$

Here the random variable $W_{t_{k+1}^{(n)}}-W_{t_k^{(n)}}$ is normal distributed with mean 0 and variance $\Delta t^{(n)}$ and is independent of $S_k^{(n)}$. Therefore we define random variables

$$Z_{k,0}^{(n)}=\exp\left\{\mu\Delta t^{(n)}+\sqrt{\sigma^2\Delta t^{(n)}}\ Y_k\right\},Z_{k,1}^{(n)}=\exp\left\{\mu\Delta t^{(n)}+\alpha+\sqrt{\sigma^2\Delta t^{(n)}+\beta^2}\ Y_k\right\},$$

and set, conditional on $\left(N_k^{(n)},S_k^{(n)}\right)$,

$$\left(N_{k+1}^{(n)}, S_{k+1}^{(n)}\right) = \begin{cases} \left(N_k^{(n)}+1, S_k^{(n)} \cdot Z_{k,1}^{(n)}\right) & \text{with probability } \lambda\Delta t^{(n)} \\ \left(N_k^{(n)}, S_k^{(n)} \cdot Z_{k,0}^{(n)}\right) & \text{with probability } 1-\lambda\Delta t^{(n)} \end{cases}.$$

Doing this for $k = 0, \ldots, n-1$ defines random variables $(N_n^{(n)}, S_n^{(n)})$ at date n. For our purposes we are only interested in the properties of $S_n^{(n)}$:

Theorem 2 [6] *The sequence of random variables $S_n^{(n)}$ at time T converges in distribution to the distribution of securities price S_T in the continuous-time model of equation (7), i.e. $S_n^{(n)} \stackrel{d}{\Longrightarrow} S_T$.*

To prove theorem 2 note that the Central Limit Theorem implies $\sqrt{\Delta t^{(n)}}\sum_{k=0}^{n-1} Y_k \stackrel{d}{\Longrightarrow} W_T$; therefore

$$S_0 \prod_{k=0}^{n-1} Z_{n,k,0} = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)T + \sigma\sqrt{\Delta t^{(n)}} \sum_{k=0}^{n-1} Y_k\right) \stackrel{d}{\Longrightarrow} S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right)T + \sigma W_T\right).$$

Since $Z_{n,k,1} \stackrel{d}{=} Z_{n,k,0} \cdot U_k$ and since $N_n^{(n)} \stackrel{d}{\Longrightarrow} N_T$ we conclude $S_n^{(n)} \stackrel{d}{\Longrightarrow} S_T$.

[Fig. 1 about here.]

Over two periods, figure 1 provides a snapshot of our approximation for the Poisson process and the resulting random variables that describe the terminal distribution of securities prices conditional on the total number of jumps between dates 0 and 2. It is important to note that the resulting random variables that describe the terminal securities price distribution are equal (in distribution), independent of the actual path.

It remains to write $S_n^{(n)}$ as a mixed lognormal distribution. For $i = 0, 1, \ldots, n$ we set

$$X_i^{(n)} = \exp\left\{\mu T + \alpha i + \sqrt{\sigma^2 T + i\beta^2} \cdot Y_i\right\} \text{ and } A^{(n)} = S_0 \sum_{i=0}^{n} 1_{N^{(n)}=i} X_i^{(n)}. \tag{9}$$

For $S_n^{(n)}$ only the total number of jumps between 0 and T is relevant to determine its

[6] The processes $\left(N^{(n)}, S^{(n)}\right)$ we constructed are discrete-time processes; we could use them to define continuous-time process $N_t^{(n)} = N_{\lfloor t/\Delta t^{(n)} \rfloor}^{(n)}$, $S_t^{(n)} = S_{\lfloor t/\Delta t^{(n)} \rfloor}^{(n)}$ and then conjecture that $\left(N_t^{(n)}, S_t^{(n)}\right)_{0 \le t \le T} \stackrel{d}{\Longrightarrow} (N, S)$. We refrain from doing so, since our focus in this paper is on terminal distributions to price European-style derivatives. We refer the interested reader to Jacod and Shiryaev (1987) for background material.

distribution at the terminal date; therefore

$$S_n^{(n)} \stackrel{d}{=} S_0 \cdot \prod_{k=0}^{N_n^{(n)}} Z_{k,1}^{(n)} \cdot \prod_{k=0}^{n-N_n^{(n)}} Z_{k,0}^{(n)} \stackrel{d}{=} S_0 \sum_{i=0}^{n} 1_{N_n^{(n)}=i} X_i^{(n)} \stackrel{d}{=} A^{(n)}, \tag{10}$$

i.e. $A^{(n)}$ is mixed lognormal distribution. Note that by theorem 2 the sequence $A^{(n)}$ of mixed lognormal distribution converges in distribution to $S_n^{(n)}$.

3.3 Accuracy and Efficiency

We discuss accuracy and efficiency to price call options for the mixed lognormal distribution of equation (9). The distribution of $N^{(n)}$ is that of the n-step binomial distribution on $\{0, \ldots, n\}$ where over each step the probability is $\lambda\Delta t^{(n)}$ for an increase by 1 and $1-\lambda\Delta t^{(n)}$ to remain unchanged, i.e. $Q[A_n^{(n)} = i] = \binom{n}{i}(\lambda\Delta t^{(n)})^i(1-\lambda\Delta t^{(n)})^{n-i}$. Therefore, according to equations (5, 6) the n-th approximation $e^{-rT}E[(A^{(n)}-K)^+]$ of the call price is equal to[7]

$$\sum_{j=0}^{n} \binom{n}{j} \left(\lambda\Delta t^{(n)}\right)^j \cdot \left(1-\lambda\Delta t^{(n)}\right)^{n-j} \exp\left((\mu-r)T + \alpha j + \frac{\sigma^2 T + j\beta^2}{2}\right)$$
$$\cdot BS\left(S_0, K, \mu T + \alpha j, \sqrt{\sigma^2 T + j\beta^2}\right). \tag{11}$$

[Table 1 about here.]

Table 1 calculates call price approximations when $S_0 = 100, r = 0.05, T = 1, \sigma = 0.1$ varying the strikes $K = 90; 100; 110$, varying parameters $\alpha = -0.2; -0.5, \beta = 0.1; 0.3$ that describe "mean" and "variance" of the jump sizes and varying the frequency of jumps λ. We chose only cases with negative α because we connect market "crashes" and other stress periods with jumps and so downward jumps seemed more natural for us than upward jumps. *MLD* presents price approximations for $n = 20$ according to equation (11) and *Merton* presents price approximations calculated using Merton's integration formula (8) for $n = 20$ using an integral approximation of the expectation on the interval $[0, 6]$ with step size 0.00001. For us,

[7] Cox and Rubinstein (1985), p. 370, simplify the call pricing formula in the Merton model with lognormal random variable U_i and derive an equation similar to this; but they do not link it to the general properties of mixed lognormal distributions, efficiency and approximations for other processes.

Merton's integration formula serves as a benchmark and we expect prices calculated using it to be accurate to the penny presented.

When λ tends to 0, we tend to the Black-Scholes setup. Prices in that setup are 14.6288; 6.8050; 2.1739 respectively for the three options. When λ is 0.01, i.e. small, prices should be, and they are in fact, close to Black-Scholes prices. The larger α, β are in absolute terms the stronger are jumps; we see that prices become then larger. This is an effect known since Merton (1976): the risk-neutral probability is set such that the expectation is always fixed but these parameters increase the overall variance in prices and call prices are larger the larger the price variation in the underlying security.

All prices calculated using our approach differ at most one penny from our benchmark prices; the only exceptions occur where $\lambda = 0.2$ and $\beta = 0.3$, i.e. when jumps are frequent and exhibit a large variance. The maximal error in these cases is 17 cents. Our approximation was based on the product $\lambda\Delta t^{(n)}$ being small; this discrepancy therefore shows us that for λ of this size we should perform calculations using a larger n than $n = 20$ in the calculations of table 1. We do not present these results here because we believe the results presented are convincing that with sufficiently small n we can achieve sufficient accuracy for derivatives pricing purposes.

To assess the efficiency of our method we point out that we needed to evaluate the Black-Scholes formula only 20 times, whereas the numerical integration scheme we used evaluated them 600,000 times. Therefore our approach based on mixed lognormals provided a results of similar accuracy but using much less of computational intense calculations. Therefore our method we believe our method is computationally more efficient. These gains carry over to calculations of and prices and their Greeks for other derivatives, as well.

4 Mixed Lognormal Distributions as an Approximation to Stochastic Volatility Models

4.1 The Continuous-Time Dynamics

An important extension of the Black-Scholes setup is the bivariate diffusion, where the dynamics under the risk-neutral probability measure Q is given (jointly) by

$$dV_t = \mu_v(V_t)dt + \varphi(V_t)dW_{1t}, \quad dS_t = rS_tdt + \psi(V_t)S_tdW_{2t}. \tag{12}$$

Here (W_1, W_2) is a bivariate independent Wiener process with instantaneous correlation ρ. The process S describes the securities dynamics and V plays the role of the process that drives volatility. For simplicity we refer throughout to the term V as *volatility.*

[Table 2 about here.]

The functions φ, ψ are mappings from the positive real line into the positive real line and the function μ_v is a mapping from the positive real line into the real line. Using functions φ and ψ instead of concrete parametrizations, the models that are common in the literature can be treated in a unified way (see table 2) [8] . We will not impose specific functional forms for these functions but adopt throughout the following two assumptions:

Assumption 3 *With probability 1, a solution of the stochastic differential equation (12) exists, is distributionally unique and* $Q\left[\min_{0\leq t\leq T} V_t > 0\right] = 1$.

We refer the reader to the theory of stochastic differential equations for conditions that ensure assumption 3; see, e.g. Karatzas and Shreve (1991), Protter (1990) or Oksendal (1995). We denote by f a function on the positive real line with $f'(x) = \frac{1}{\varphi(x)}$ and by g the inverse of f.

Assumption 4 *The functions φ, ψ are twice continuously differentiable, g is three times differentiable and the volatility drift μ_v is strictly positive on the interval $[0, \epsilon)$ for suitable ϵ.*

[Table 3 about here.]

[8] The literature typically takes a mean-reverting dynamics for V under the objective measure; we allow for a general function μ_v to capture, among others, that the risk-neutral probability measure contains a drift-adjustment.

Table 3 provides the functions f, g for those stochastic volatility models that are common in the literature. They fulfill assumptions 3 and 4.

4.2 Constructing a Sequence of Mixed Lognormal Distributions

For given integer n we discretize the interval $[0, T]$ into n equidistant intervals $[t_k^{(n)}, t_{k+1}^{(n)})$ $(k = 0, \ldots, n-1)$ with $t_0^{(n)} = 0$ and $t_{k+1}^{(n)} = (k+1)\Delta t^{(n)}$, $\Delta t^{(n)} = \frac{T}{n}$. We will now construct a bivariate Markov chain $(V_k^{(n)}, S_k^{(n)})_{k=0,\ldots,n}$ that takes values at times $t_k^{(n)} = k\Delta t^{(n)}$; from this we then derive a bivariate Markov process setting

$$\left(V_t^{(n)}, S_t^{(n)}\right) = \left(V_{\lfloor t/\Delta t^{(n)} \rfloor}^{(n)}, S_{\lfloor t/\Delta t^{(n)} \rfloor}^{(n)}\right). \tag{13}$$

Note that $\left\lfloor t_k^{(n)}/\Delta t^{(n)} \right\rfloor = k$ and that at times t with $t_k^{(n)} \leq t < t_{k+1}^{(n)}$ this parameter gives us the "last" date k before t. Therefore the processes in (13) are right-continuous with left-hand limits; they move in steps, and are constant on intervals $(t_k^{(n)}, t_{k+1}^{(n)})$. The continuous-time versions of these processes form a sequence and our goal is to construct them such that $(V^{(n)}, S^{(n)}) \stackrel{d}{\Longrightarrow} (V, S)$ under Q. The actual convergence of processes will be stated as theorem 5 below and proven in the appendix after additional technical conditions have been checked; as a direct corollary we then have $S_n^{(n)} \stackrel{d}{\Longrightarrow} S_T$. Throughout, our goal is to construct the discrete processes such that local mean and co-variances of the discrete-time Markov chain converge to their continuous-time counterpart.

We proceed in two steps: in the first we construct an approximation of the volatility process $(V_k^{(n)})_{k=0,\ldots,n}$ and in the second step we then extend this to the securities price process. For the construction of the volatility process we follow the idea of Nelson and Ramaswamy (1990): Since $f'(V) = \frac{1}{\varphi(V)}$, Itô's formula implies that

$$df(V)_t = \left\{\frac{\mu_v(V_t)}{\varphi(V_t)} + \frac{1}{2}f''(V_t)\varphi^2(V_t)\right\} dt + dW_{1t}, \tag{14}$$

i.e. the dynamics of the transformed process $f(V)$ is homoscedastic. Further following Nelson and Ramaswamy (1990) we then define the points

$$C_i^{(n)} = f(V_0) + i\sqrt{\Delta t^{(n)}}\ ,\ D_i^{(n)} = g\left(C_i^{(n)}\right),\ i \text{ an integer}, \tag{15}$$

and define the process $(V_k^{(n)})_k$ by forward induction: we set $V_0^{(n)} = V_0$. Then we assume that the volatility process has been defined at all dates from 0 to k. Conditional on $V_k^{(n)} = D_i^{(n)}$, we define the random variable $V_{k+1}^{(n)}$ to take the value $D_{i+1}^{(n)}$ with probability $q_i^{(n)}$ and $D_{i-1}^{(n)}$ with probability $1 - q_i^{(n)}$, where we set

$$q_i^{(n)} = Q\left[V_{k+1}^{(n)} = D_{i+1}^{(n)} | V_k^{(n)} = D_i^{(n)}\right] = \min\left\{\frac{\mu_v\left(D_i^{(n)}\right)\Delta t^{(n)} - \left(D_{i-1}^{(n)} - D_i^{(n)}\right)}{D_{i+1}^{(n)} - D_{i-1}^{(n)}}, 1\right\}. \quad (16)$$

The truncation here is necessary to ensure that $q_i^{(n)}$ is a probability[9], i.e. that $0 \le q_i^{(n)} \le 1$. This defines a process $(V_k^{(n)})_{k=0,\dots,n}$; note that it corresponds to a recombining binomial tree with transition probability conditional on the position in the tree.

We next extend this to the process $(S_k^{(n)})_{k=0,\dots,n}$; our construction will use a sequence Y_k $(k = 0, 1, \dots)$ of serially independent, standard normal random variables and proceeds by forward induction: First we set $S_0^{(n)} = S_0$; then we assume that the discrete security process has been defined for all dates from 0 to k. When the volatility that enters into the security process (equation (12)) would be constant at $V_k^{(n)}$ over the time period $[t_k^{(n)}, t_{k+1}^{(n)})$ then $S_{k+1}^{(n)}$ would be equal to

$$S_k^{(n)} \exp\left\{\left(r - \frac{\psi^2\left(V_k^{(n)}\right)}{2}\right) t_k^{(n)} + \psi\left(V_k^{(n)}\right) \cdot \left(W_{2,t_{k+1}^{(n)}} - W_{2,t_k^{(n)}}\right)\right\}.$$

We define a process $\epsilon^{(n)}$ by setting for $k = 0, \dots, n-1$

$$\epsilon_k^{(n)} = f\left(V_{k+1}^{(n)}\right) - f\left(V_k^{(n)}\right) - \left\{\frac{\mu_v\left(V_k^{(n)}\right)}{\varphi\left(V_k^{(n)}\right)} + \frac{1}{2}\varphi^2\left(V_k^{(n)}\right) f''\left(V_k^{(n)}\right)\right\}\Delta t^{(n)}. \quad (17)$$

Our goal is $V^{(n)} \stackrel{d}{\Longrightarrow} V$, and *if* this convergence holds then it *suggests* that $\sum_{k=1}^{\lfloor t/\Delta t^{(n)} \rfloor} \epsilon_k^{(n)} \stackrel{d}{\Longrightarrow} W_{1t}$ and

$$\left(\sum_{k=1}^{\lfloor t/\Delta t^{(n)} \rfloor} \epsilon_k^{(n)}, \sum_{k=1}^{\lfloor t/\Delta t^{(n)} \rfloor} \left(\rho\epsilon_k^{(n)} + \sqrt{(1-\rho^2)\Delta t^{(n)}}\ Y_k\right)\right) \stackrel{d}{\Longrightarrow} (W_{1t}, W_{2t}).$$

[9] Nelson and Ramaswamy (1990) introduce multiple jumps to ensure the transition probability that "matches" its continuous-time counterpart is between 0 and 1. Our construction permits us to choose an "arbitrary" transition probability in those events. Negativity is not an issue and so we only truncate $q_i^{(n)}$ above.

This *remains to be proved*, but it motivates $\rho\epsilon_k^{(n)} + \sqrt{(1-\rho^2)\Delta t^{(n)}}\ Y_k$ as an approximation for $W_{2,t_{k+1}^{(n)}} - W_{2,t_k^{(n)}}$. We set

$$S_{k+1}^{(n)} = S_k^{(n)} \exp\left\{\left(r - \frac{\psi^2\left(V_k^{(n)}\right)}{2}\right)\Delta t^{(n)} + \psi\left(V_k^{(n)}\right)\cdot\left(\rho\epsilon_k^{(n)} + \sqrt{(1-\rho^2)\Delta t^{(n)}}Y_k\right)\right\}. \tag{18}$$

This ends our construction of the bivariate Markov chain $\left(V_k^{(n)}, S_k^{(n)}\right)_{k=0,\dots,n}$. The appendix proves:

Theorem 5 *The sequence of processes* $(V_t^{(n)}, S_t^{(n)})_{0\le t\le T}$ *(based on equation (13)) converges in distribution to the process* $(V_t, S_t)_{t\in[0,T]}$ *of equation (12).*

In particular this implies $S_n^{(n)} \overset{d}{\Longrightarrow} S_T$. It remains to write the distribution at time T (date n) as a mixed lognormal distribution: Let us denote by $\Gamma^{(n)} = \{\gamma = (\gamma_0, \dots, \gamma_{n-1}) | \gamma_i \in \{+1, -1\}\}$ the set of *volatility paths* between 0 and T, and for $\gamma \in \Gamma$ by $\delta_0(\gamma) = 0$ and $\delta_k(\gamma) = \sum_{i=0}^{k-1} \gamma_i$ (for $k > 0$) the current tree node at date k when the volatility path is γ. We define for $\alpha = -1, +1$

$$\eta_{i,\alpha}^{(n)} = \alpha \cdot \sqrt{\Delta t^{(n)}} - \left\{\frac{\mu_v\left(D_i^{(n)}\right)}{\varphi\left(D_i^{(n)}\right)} + \frac{1}{2}\varphi^2\left(D_i^{(n)}\right) f''\left(D_i^{(n)}\right)\right\}\Delta t^{(n)}, \tag{19}$$

Note that $f\left(D_{i+1}^{(n)}\right) - f\left(D_i^{(n)}\right) = C_{i+1}^{(n)} - C_i^{(n)} = \sqrt{\Delta t^{(n)}}$ and similarly $f\left(D_{i-1}^{(n)}\right) - f\left(D_i^{(n)}\right) = -\sqrt{\Delta t^{(n)}}$. Therefore, according to equation (17) and conditional on $V_k^{(n)} = D_i^{(n)}$, $\epsilon_k^{(n)}$ adopts $\eta_{i,+1}^{(n)}$ with probability $q_i^{(n)}$ and $\eta_{i,-1}^{(n)}$ with probability $1 - q_i^{(n)}$. Denote for $\alpha = -1, +1$

$$Z_{k,i,\alpha}^{(n)} = \exp\left\{\left(r - \frac{\psi^2(D_i^{(n)})}{2}\right)\Delta t^{(n)} + \rho\psi(D_i^{(n)})\eta_{i,\alpha}^{(n)} + \psi(D_i^{(n)})\sqrt{(1-\rho^2)\Delta t^{(n)}}\ Y_k\right\}. \tag{20}$$

When the volatility moves up, the product of the current securities price with $Z_{k,i,1}^{(n)}$ determines next periods securities prices; similarly $Z_{k,i,-1}^{(n)}$ for a down move in volatility.

[Fig. 2 about here.]

Figure 2 illustrates this; it describes the volatility tree over two periods. Over the first period volatility can increase from $D_0^{(n)}$ to $D_1^{(n)}$ or decrease to $D_{-1}^{(n)}$. Over the second period volatility can further increase or decrease; there are three potential values for the volatility at date 2: $D_{-2}^{(n)}$, $D_0^{(n)}$ and $D_2^{(n)}$. The figure also depicts at each node the two lognormal variables that will be mixed over each period.

By construction, a volatility decrease followed by an increase leads to the same volatility $D_0^{(n)}$ as an increase followed by a decrease. But the individual lognormal random variables in the final product depend on the volatility path leading there; here a volatility decrease followed by an increase will lead to a securities price determined by $S_0 \cdot Z_{0,0,-1}^{(n)} \cdot Z_{1,-1,1}^{(n)}$ and this is *not* equal in distribution to the securities price determined by $S_0 \cdot Z_{0,0,1}^{(n)} \cdot Z_{1,1,-1}^{(n)}$ when we see a volatility increase followed by a decrease.

We also define a random variable $A^{(n)}$ on $\Gamma^{(n)}$, setting for $\gamma \in \Gamma^{(n)}$

$$Q[A^{(n)} = \gamma] = \prod_{k=0}^{n-1} \left(q_{\delta_k(\gamma)}^{(n)} \cdot 1_{\gamma_k=+1} + \left(1 - q_{\delta_k(\gamma)}^{(n)}\right) \cdot 1_{\gamma_k=-1} \right)$$

the probability of that volatility path. Note that each of the product terms is either $q_{\delta_k(\gamma)}^{(n)}$ or $1 - q_{\delta_k(\gamma)}^{(n)}$. Furthermore we denote $\Theta_1^{(n)}(\gamma)$ the average of the squared ("effective") volatility for the security S and by $\Theta_2^{(n)}(\gamma)$ the total contribution to the volatility of security S due to its correlation with the volatility process, i.e.

$$\Theta_1^{(n)}(\gamma) = \frac{1}{n} \sum_{k=0}^{n-1} \psi^2 \left(D_{\delta_k(\gamma)}^{(n)} \right), \Theta_2^{(n)}(\gamma) = \sum_{k=0}^{n-1} \left(\eta_{\delta_k(\gamma),\gamma_k}^{(n)} \cdot \psi \left(D_{\delta_k(\gamma)}^{(n)} \right) \right),$$

For a set of independent standard normal random variables Y_γ ($\gamma \in \Gamma$) we then define random variables

$$Z_\gamma^{(n)} = \prod_{\gamma \in \Gamma} \exp \left\{ \left(r - \frac{\Theta_1^{(n)}(\gamma)}{2} \right) T + \rho \Theta_2^{(n)}(\gamma) + \sqrt{(1-\rho^2)\Theta_1^{(n)}(\gamma)}\; Y_\gamma \right\}. \tag{21}$$

Then $\prod_{k=0}^{n-1} Z_{k,\delta_k(\gamma),\gamma_k}^{(n)} \stackrel{d}{=} Z_\gamma^{(n)}$ which implies $S_n^{(n)} \stackrel{d}{=} \sum_{\gamma \in \Gamma} 1_{A^{(n)}=\gamma} Z_\gamma^{(n)}$. Hence the distribution of $S_n^{(n)}$ at date n is a mixed lognormal distribution[10] and according to theorem 5, $S_n^{(n)} \stackrel{d}{\Longrightarrow} S_T$.

[10] Formally, to fit into our definition of mixed lognormal distributions at the beginning of this page, we need to associate each path with an integer number and then define $A^{(n)}$ on those number and index the lognormal random variables based on them.

The Hull-White model is given by $\varphi(V) = \sigma V, \psi(V) = \sqrt{V}$, where σ is a constant. In the following discussion we assume that $\mu_v(V) = -\kappa V$; then V_T is a lognormal distributed random variable, i.e. $\ln(V_T/V_0)$ is normal distributed with mean $\left(-\kappa - \frac{\sigma^2}{2}\right)T$ and variance $\sigma^2 T$. We have $f(x) = \frac{\ln x}{\sigma}$ and $g(z) = \exp(\sigma z)$; therefore our grid points are

$$C_i^{(n)} = \frac{\ln V_0}{\sigma} + i\sqrt{\Delta t^{(n)}},\ D_i^{(n)} = \exp\left(\sigma C_i^{(n)}\right) = V_0 \cdot \exp\left(i\sigma\sqrt{\Delta t^{(n)}}\right).$$

A series expansion of the exponential function around $\sigma C_i^{(n)}$ reveals that $D_{i+1}^{(n)} - D_i^{(n)} = D_i^{(n)}\sigma\sqrt{\Delta t^{(n)}} + \frac{1}{2}D_i^{(n)}\sigma^2\Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right)$, $D_{i-1}^{(n)} - D_i^{(n)} = -D_i^{(n)}\sigma\sqrt{\Delta t^{(n)}} + \frac{1}{2}D_i^{(n)}\sigma^2\Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right)$ so that $D_{i+1}^{(n)} - D_{i-1}^{(n)} = 2D_i^{(n)}\sigma\sqrt{\Delta t^{(n)}} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right)$. Therefore we calculate for $q_i^{(n)}$ of equation (16)

$$\begin{aligned} q_i^{(n)} &= \frac{D_i^{(n)}\sigma\sqrt{\Delta t^{(n)}} - \frac{1}{2}D_i^{(n)}\sigma^2\Delta t^{(n)} - \kappa D_i^{(n)}\Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right)}{2D_i^{(n)}\sigma\sqrt{\Delta t^{(n)}} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right)} \\ &= \frac{1}{2} + \frac{-\kappa - \frac{\sigma^2}{2}}{2\sigma}\sqrt{\Delta t^{(n)}} + \mathcal{O}\left(\Delta t^{(n)}\right), \end{aligned}$$

uniformly on compact sets. (Note that, for sufficiently large n, this is always between 0 and 1 and so the truncation in equation (16) will never come into effect.) We calculate based on equation (19) that $\eta_{i,\alpha}^{(n)} = \alpha\sqrt{\Delta t^{(n)}} - \frac{-\kappa-\sigma^2/2}{\sigma}\Delta t^{(n)}$ since $f''(x) = -\frac{1}{\sigma V^2}$. Also, we have, for $\gamma \in \Gamma^{(n)}$, $\Theta_1^{(n)}(\gamma) = \frac{1}{n}\sum_{k=0}^{n-1} D_{\delta_k(\gamma)}^{(n)}$, and we set $\Theta_2^{(n)} = -\frac{-\kappa-\sigma^2/2}{\sigma}T + \sum_{k=0}^{n-1}\left(\alpha D_{\delta_k(\gamma)}^{(n)}\right)\sqrt{\Delta t^{(n)}}$. Using $Z_\gamma, A^{(n)}$ of the previous subsection this can be implemented in computer code using a recursive procedure.

[Fig. 3 about here.]

To assess the accuracy of our approximation we assume that the processes for the securities price and volatility are uncorrelated, i.e. $\rho = 0$. Under this assumption, call prices are given as $e^{-rT}E[(S_T - K)^+] = e^{-rT}E[E[(S_T - K)^+|V_T]] = E\left[BS\left(S_0, K, rT - V_T/2, \sqrt{V_T}\right)\right]$; therefore we calculate prices as a numerical integration over Black-Scholes prices weighted by a lognormal density function. (A parameter $\rho \neq 0$ would considerably complicate our calculation of continuous-time prices.) Here we take a step size of 0.00001 and 600,000 steps, i.e. we integrate over the interval $(0, 6)$; a further decrease of step sizes did not increase further the accuracy and the interval $(0, 6)$ should capture almost all the probability mass of the volatility lognormal distribution for the cases we study below.

Figure 3 presents for $n = 5, 6, \ldots, 20$ the price approximation for a call option calculated by our mixed lognormal; the underlying parameters are $\kappa = 0$; $\sigma = 0.4$; $S_0 = 100$; $r = 0$; $T = 1$; $K = 100$ and $v_0 = 0.3^2$. (This corresponds to an initial volatility of $\psi(v_0) = 0.3$.) The figure also presents the continous-time price as a flat line. We see that the convergence behavior is very smooth and the approximation approaches fast the continuous-time price. For comparison we also calculated Monte-Carlo price approximations by simulating 1,000,000 paths for the same time-refinements. The Monte-Carlo numbers fluctuate somewhat, reflecting the randomness of the approximation. However, overall these numbers seem to converge slower than those of our mixed lognormal to the continuous-time price.

[Fig. 4 about here.]

These figures are indicative of the efficiency of our mixed lognormal approximation. However to nail down the actual gains we carry out the following test: we calculate price approximations for three strikes $K = 90, 100, 110$, varying the dispersion of the volatility process $\sigma = 0.3, 0.4, 0.5$ and the interest rate $r = 0.05, 0.1, 0.15$. (The other parameters are as in the previous figure.) This gives 27 prices and errors to their continuous-time price. We determine the average computing time needed for a MATLAB implementation on a Pentium M 1.30 GHz machine, the average error and plot them in figure 4 on a logarithmic scale. Approximations for the mixed lognormal are based on varying $n = 5, 6, \ldots, 19$ and those for Monte-Carlo are based on $2^{15} = 32,768; 2^{16}; \ldots; 2^{20} = 1,048,576$ simulated paths and 10 time-steps.

Figure 4 exhibits that the Monte-Carlo is relatively inefficient in approximating call prices: for the same computing time its error is up to ten times as high than that for our mixed lognormal; put differently to achieve the same level of accuracy the figure suggests we need initially 10000 times as much computing time.

Further investigations into the convergence pattern of our mixed lognormal approach revealed that pricing errors seemed to be of the order $1/n^2$. Since figure 3 shows a very smooth convergence pictures, we implemented an extrapolation; the extrapolated price for given n defined as $\frac{(n+1)^2 p_{n+1} - n^2 p_n}{(n+1)^2 - n^2}$, where p_n, p_{n+1} denotes prices calculated directly based on our mixed lognormal approach[11]. The extrapolation further improves our method relative to the Monte-Carlo approach. We conclude that the mixed lognormal is an efficient method to calculate call prices. Therefore we believe that prices and their Greeks can be efficiently calculated using our technique.

5 Conclusion

This paper constructed sequences of mixed lognormal distributions that converge in distribution to the Black-Scholes setup with jumps (Merton model) and of the Black-Scholes with stochastic volatility model, respectively. We explained how to calculate derivatives prices and their Greeks and discussed efficiency and accuracy. The techniques should be of interest for derivatives pricing and risk-management in financial institutions with large derivative positions.

A Proof of Theorem 5

We define for all $t \in [0,T], k = 0, \ldots, n-1$, $N_k^{(n)} = f\left(V_k^{(n)}\right)$, $N_t = f(V_t)$, $R_k^{(n)} = \ln S_k^{(n)}$, $R_t = \ln S_t$, and

$$X_k^{(n)} = \left(N_t^{(n)}, R_k^{(n)}\right), X_t = (N_t, R_t).$$

Instead of proving $(V^{(n)}, S^{(n)}) \stackrel{d}{\Longrightarrow} (V, S)$, our goal is to prove that the continuous-time version[12] of $X^{(n)}$ converges in distribution to X. Since the function g and the exponential function are continuous this is sufficient to prove theorem 5.

Based on the definition of f we calculate directly that $f'(V) = \frac{1}{\varphi(V)}$ and $f''(V) = \frac{\varphi'(V)}{\varphi^2(V)}$. Therefore, by equation (14) and Itô's Lemma,

[11] Denote p_∞ the continuous-time price. *If* $p_n = p_\infty + \frac{c}{n^2}$ for some suitable constant c and $p_{n+1} = p_\infty + \frac{c}{(n+1)^2}$ for the same constant c, then a linear transformation gives our extrapolation rule. The assumptions here are motivated by the observed (quadratic) convergence rate.

[12] The continuous-time version of the processes in X are defined analogously to those in equation (13), see theorem 6 below.

$$dN_t = df(V)_t = \left\{ \frac{\mu_v(V_t)}{\varphi(V_t)} - \frac{1}{2}\varphi'(V_t) \right\} dt + dW_{1t}, \quad dR_t = \left(r - \frac{\psi^2(V_t)}{2} \right) dt + \psi(V_t) dW_{2t}.$$

Also, by equations (17, 18),

$$\begin{aligned} N_{k+1}^{(n)} - N_k^{(n)} &= \left\{ \frac{\mu_v\left(V_k^{(n)}\right)}{\varphi\left(V_k^{(n)}\right)} - \frac{1}{2}\varphi'\left(V_k^{(n)}\right) \right\} + \epsilon_k^{(n)}, \\ R_{k+1}^{(n)} - R_k^{(n)} &= \left(r - \frac{\psi^2\left(V_k^{(n)}\right)}{2} \right) \Delta t^{(n)} + \psi(V_k^{(n)}) \cdot \left(\rho \epsilon_k^{(n)} + \sqrt{(1-\rho^2)\Delta t^{(n)}} Y_k \right). \end{aligned}$$

(For simplicity of exposition we write these depending on $V_k^{(n)} = g(N_k^{(n)})$). We then define a transition function $Q_n(x, M)$ on $\mathbb{R}^2$ by setting $Q_n\left((x_1, x_2)), (M_1, M_2)\right)$ equal to

$$\begin{cases} q_i^{(n)} \cdot Q\left[x_2 + \ln Z_{k,i,1}^{(n)} \in M_2\right] & \text{if } x_1 = D_i^{(n)} \text{ and } D_{i+1}^{(n)} \in M_1 \\ \left(1 - q_i^{(n)}\right) \cdot Q\left[x_2 + \ln Z_{k,i,-1}^{(n)} \in M_2\right] & \text{if } x_1 = D_i^{(n)} \text{ and } D_{i-1}^{(n)} \in M_1 \\ \left.\begin{array}{l} q_i^{(n)} \cdot Q\left[x_2 + \ln Z_{k,i,1}^{(n)} \in M_2\right] \\ \quad + \left(1 - q_i^{(n)}\right) \cdot Q\left[x_2 + \ln Z_{k,i,-1}^{(n)} \in M_2\right] \end{array}\right\} & \text{if } x_1 = D_i^{(n)} \text{ and } D_{i-1}^{(n)}, D_{i+1}^{(n)} \in M_1 \\ 1 & \text{otherwise} \end{cases}.$$

Note that $Q_n(x, M)$ is the transition function for the discrete-time Markov Chain $\left(X_k^{(n)}\right)_k$ and that $Q\left[x_2 + \ln Z_{k,i,1}^{(n)} \in M_2\right]$ does not depend on k. We then define for $x = (x_1, x_2)$,

$$\mu^{(n)}(x) = \frac{1}{\Delta t^{(n)}} \int_0^\infty (y - x) Q_n(x, dy), \quad \sigma^{(n)}(x) = \frac{1}{\Delta t^{(n)}} \int_0^\infty (y - x)^t (y - x) Q_n(x, dy),$$

where "t" denotes the vector transpose. Note that $\mu^{(n)}$ is a function $\mathbb{R}^2 \to \mathbb{R}^2$ and $\sigma^{(n)}$ is a function $\mathbb{R}^2 \to \mathbb{R}^{2\times 2}$. We also define

$$\mu(x) = \left(\frac{\mu_v(x_1)}{\varphi(x_1)} - \varphi'(x_1), r - \frac{\psi^2(x_1)}{2} \right), \sigma(x) = \begin{pmatrix} 1 & \rho\psi(x_1) \\ \rho\psi(x_1) & \psi^2(x_1) \end{pmatrix}$$

the local mean vector and local covariance matrix function of process X. To prove Theorem 5 we apply the martingale central limit theorem of Ethier and Kurtz (1986), p. 354 in the form of their corollary 4.2, p. 355; for our problem it reads:

Theorem 6 *Suppose for each $\theta > 0$ and $\vartheta > 0$, that $\sup_{|x|\le\theta} |\mu^{(n)}(x) - \mu(x)| \to 0$, that $\sup_{|x|\le\theta} |\sigma^{(n)}(x) - \sigma(x)| \to 0$ and $\sup_{|x|\le\theta} \frac{1}{\Delta t^{(n)}} Q_n(x, \{y; |y - x| \ge \vartheta\}) \to 0$. We define the* *continuous-time version of* $X^{(n)}$ *by setting for all* $0 \le t \le T$, $X_t^{(n)} = X^{(n)}_{\lfloor t/\Delta t^{(n)} \rfloor}$. *Then*

$$\left(X_t^{(n)}\right)_{0\le t\le T} \overset{d}{\Longrightarrow} (X_t)_{0\le t\le T}.$$

It is therefore sufficient to check the conditions of this theorem. The conditions are that uniformly on compact sets the local mean and covariances converge to their continuous-time counterpart and that for each ϑ the probability for moves greater than that converge to 0 uniformly on compact sets. A Taylor series expansion gives

$$\begin{aligned}
D_{i+1}^{(n)} - D_i^{(n)} &= g'\left(C_i^{(n)}\right)\sqrt{\Delta t^{(n)}} + \frac{1}{2} g''\left(C_i^{(n)}\right)\Delta t^{(n)} + \mathcal{O}\Big(\sqrt{\Delta t^{(n)}}^3\Big), \text{ and}\\
D_{i-1}^{(n)} - D_i^{(n)} &= -g'\left(C_i^{(n)}\right)\sqrt{\Delta t^{(n)}} + \frac{1}{2} g''\left(C_i^{(n)}\right)\Delta t^{(n)} + \mathcal{O}\Big(\sqrt{\Delta t^{(n)}}^3\Big), \text{ so that}\\
D_{i+1}^{(n)} - D_{i-1}^{(n)} &= 2g'\left(C_i^{(n)}\right)\sqrt{\Delta t^{(n)}} + \mathcal{O}\Big(\sqrt{\Delta t^{(n)}}^3\Big).
\end{aligned}$$

Since g is the inverse of f and $f'(V) = \frac{1}{\varphi(V)}$ we find $g'\left(C_i^{(n)}\right) = \left(f'\left(D_i^{(n)}\right)\right)^{-1} = \varphi\left(D_i^{(n)}\right) = \varphi\left(g\left(C_i^{(n)}\right)\right)$. Therefore $g''\left(C_i^{(n)}\right) = \varphi'\left(g\left(C_i^{(n)}\right)\right)\cdot g'\left(C_i^{(n)}\right) = \varphi'\left(D_i^{(n)}\right)\cdot\varphi\left(D_i^{(n)}\right)$. Based on this we calculate

$$\begin{aligned}
&\frac{\mu_v\left(D_i^{(n)}\right) - \left(D_{i-1}^{(n)} - D_i^{(n)}\right)}{D_{i+1}^{(n)} - D_{i-1}^{(n)}}\\
&= \frac{g'\left(C_i^{(n)}\right)\sqrt{\Delta t^{(n)}} + \left\{\mu_v\left(D_i^{(n)}\right) - \frac{1}{2} g''\left(C_i^{(n)}\right)\right\}\Delta t^{(n)} + \mathcal{O}\Big(\sqrt{\Delta t^{(n)}}^3\Big)}{2g'\left(C_i^{(n)}\right)\sqrt{\Delta t^{(n)}} + \mathcal{O}\Big(\sqrt{\Delta t^{(n)}}^3\Big)}
\end{aligned}$$

This implies that

$$\begin{aligned}
\frac{\mu_v\left(D_i^{(n)}\right) - \left(D_{i-1}^{(n)} - D_i^{(n)}\right)}{D_{i+1}^{(n)} - D_{i-1}^{(n)}} &= \frac{1}{2} + \frac{1}{2}\left\{\frac{\mu_v\left(D_i^{(n)}\right)}{g'\left(C_i^{(n)}\right)} - \frac{g''\left(D_i^{(n)}\right)}{2g'\left(C_i^{(n)}\right)}\right\}\sqrt{\Delta t^{(n)}} + \mathcal{O}\left(\Delta t^{(n)}\right)\\
&= \frac{1}{2} + \frac{1}{2}\left\{\frac{\mu_v\left(D_i^{(n)}\right)}{\varphi\left(D_i^{(n)}\right)} - \frac{1}{2}\varphi'\left(D_i^{(n)}\right)\right\}\sqrt{\Delta t^{(n)}} + \mathcal{O}\left(\Delta t^{(n)}\right)
\end{aligned}$$

uniformly on compacts. Therefore, on compact sets and for sufficiently large n, this number is always between 0 and 1 and in equation (16)

$$q_i^{(n)} = \frac{1}{2} + \frac{1}{2}\left\{\frac{\mu_v\left(D_i^{(n)}\right)}{\varphi\left(D_i^{(n)}\right)} - \frac{1}{2}\varphi'\left(D_i^{(n)}\right)\right\}\sqrt{\Delta t^{(n)}} + \mathcal{O}\left(\Delta t^{(n)}\right).$$

Since $f'(V) = \frac{1}{\varphi(V)}$ implies $f''(V) = \frac{\varphi'(V)}{\varphi^2(V)}$ we write, based on equation (19), for $\alpha = +1, -1$,

$$\eta_{i,\alpha}^{(n)} = \alpha \cdot \sqrt{\Delta t^{(n)}} - \left\{ \frac{\mu_v\left(D_i^{(n)}\right)}{\varphi\left(D_i^{(n)}\right)} - \frac{1}{2}\varphi'\left(D_i^{(n)}\right) \right\} \Delta t^{(n)},$$

Conditional on $V_k^{(n)} = D_i^{(n)}$ (or equivalently, conditional on $N_k^{(n)} = C_i^{(n)}$), $\epsilon_k^{(n)}$ takes value $\eta_{i,+1}^{(n)}$ with probability $q_i^{(n)}$ and value $\eta_{i,-1}^{(n)}$ with probability $1 - q_i^{(n)}$. Therefore

$$\begin{aligned} E\left[\epsilon_k^{(n)}\right] &= \eta_{i,+1}^{(n)} q_i^{(n)} + \eta_{i,-1}^{(n)}\left(1 - q_i^{(n)}\right) \\ &= \sqrt{\Delta t^{(n)}} q_i^{(n)} - \sqrt{\Delta t^{(n)}}\left(1 - q_i^{(n)}\right) - \left\{ \frac{\mu_v\left(D_i^{(n)}\right)}{\varphi\left(D_i^{(n)}\right)} - \frac{1}{2}\varphi'\left(D_i^{(n)}\right) \right\} \Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right) \\ &= \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right), \end{aligned}$$

and $E\left[\left(\epsilon_k^{(n)}\right)^2\right] = \Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right)$ so that $Var\left(\epsilon_k^{(n)}\right) = \Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right)$; in all three terms the bounds are uniformly on compact sets. Since $E[Y_k] = 0$ these terms imply directly

$$\begin{aligned} E\left[N_{k+1}^{(n)} - N_k^{(n)} \middle| N_k^{(n)} = D_i^{(n)}, R_k^{(n)} \right] &= \left\{ \frac{\mu_v\left(D_i^{(n)}\right)}{\varphi\left(D_i^{(n)}\right)} - \frac{1}{2}\varphi'\left(D_i^{(n)}\right) \right\} \Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right) \\ Var\left(N_{k+1}^{(n)} - N_k^{(n)} \middle| N_k^{(n)} = D_i^{(n)}, R_k^{(n)} \right) &= \Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right) \\ E\left[R_{k+1}^{(n)} - R_k^{(n)} \middle| N_k^{(n)} = D_i^{(n)}, R_k^{(n)} \right] &= \left(r - \frac{\psi^2(D_i^{(n)})}{2} \right) \Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right), \text{ and} \\ Var\left(R_{k+1}^{(n)} - R_k^{(n)} \middle| N_k^{(n)} = D_i^{(n)}, R_k^{(n)} \right) &= \psi^2(D_i^{(n)}))\Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right), \end{aligned}$$

uniformly on compact sets. Also, $E\left[\left(R_{k+1}^{(n)} - R_k^{(n)}\right) \cdot \left(N_{k+1}^{(n)} - N_k^{(n)}\right) \middle| N_k^{(n)} = D_i^{(n)}, R_k^{(n)}\right] = \rho\psi\left(D_i^{(n)}\right) E\left[\left(\epsilon_k^{(n)}\right)^2\right]$ so that

$$Cov\left(N_{k+1}^{(n)} - N_k^{(n)}, R_{k+1}^{(n)} - R_k^{(n)} \middle| N_k^{(n)} = D_i^{(n)}, R_k^{(n)} \right) = \rho\psi\left(D_i^{(n)}\right)\Delta t^{(n)} + \mathcal{O}\left(\sqrt{\Delta t^{(n)}}^3\right),$$

uniformly on compact sets. Therefore, for all $x_2 \in \mathbb{R}$ we have $|\mu_2^{(n)}(x) - \mu_2(x)| \to 0$ and $|\sigma_2^{(n)}(x) - \sigma_2(x)| \to 0$, uniformly on compact sets. Furthermore $\sup_{|x|\le\theta} \frac{1}{\Delta t^{(n)}} Q_n(x, \{y; |y-x| \geq \varepsilon\}) \to 0$, since jumps of $\epsilon^{(n)}$ are bounded on compacts and the standard deviation in the normal distribution shrinks to 0. Therefore all conditions of theorem 6 are fulfilled and this ends our proof of theorem 5.

References

Alexander, C., G. Brintalos, and L. Nogueira (2004): "Short and Long Term Smile Effects: The Binomial Normal Mixture Diffusion Model," *Forthcoming: Journal of Banking and Finance*.

Amin, K. I. (1993): "Jump Diffusion Option Valuation in Discrete Time," *The Journal of Finance*, 43, 1833–1863.

Bates, D. S. (2000): "Post–'87 Crash Fears in the S&P 500 Futures Market," *Journal of Econometrics*, 94(1–2), 181–238.

Black, F., and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–659.

Brigo, D., and F. Mercurio (2002): "Lognormal-mixture dynamics and calibration to market volatility smiles," *International Journal of Theoretical and Applied Finance*, 5(4), 427–446.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997): *The Econometrics of Financial Markets*. Princeton University Press.

Chesney, M., and L. Scott (1989): "Pricing European Currency Options: A Comparison of the Modified Black–Scholes Model and a Random Variance Model," *Journal of Financial and Quantitative Analysis*, 24, 267–284.

Cox, J., and M. Rubinstein (1985): *Options Markets*. Prentice–Hall.

Duan, J.-C., and J.-G. Simonato (2001): "American Option Pricing under GARCH by a Markov Chain Approximation," *Journal of Economic Dynamics and Control*, 25(11), 1689–1718.

Eraker, B., M. Johannes, and N. Polson (2003): "The Impact of Jumps in Equity Index Volatility and Returns," *Journal of Finance*, 53, 1269–1300.

Ethier, S. N., and T. Kurtz (1986): *Markov Processes: Characterization and Convergence*. John Wiley & Sons.

Gemmill, G., and A. Saflekos (2000): "How Useful are Implied Distributions? Evidence from Stock-Index Options," *Journal of Derivatives*, 8(Spring), 1–17.

Glasserman, P. (2004): *Monte Carlo Methods in Financial Engineering*. Springer.

Harrison, J. M., and D. M. Kreps (1979): "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20, 381–408.

Harrison, J. M., and S. R. Pliska (1981): "Martingales and Stochastic Integrals in the Theory of Continous Trading," *Stochastic Processes and their Applications*, 11, 215–260.

Heston, S. (1993): "A Closed–Form Solution for Options with Stochastic Volatilities with Applications to Bond and Currency Options," *The Review of Financial Studies*, 6, 281–300.

Hilliard, J. E., and A. Schwartz (1996): "Binomial Option Pricing Under Stochastic Volatility and Correlated State Variables," *The Journal of Derivatives*, 4(1), 23–39.

——— (2003): "Pricing European and American Derivatives Under a Jump-Diffusion Process: A Bivariate Tree Approach," *Working Paper*.

Hull, J. C. (2000): *Options, Futures, And Other Derivative Securities*. Prentice-Hall.

Hull, J. C., and A. White (1987): "The Pricing of Options on Assets with Stochastic Volatilities," *The Journal of Finance*, 42, 281–300.

Jacod, J., and A. N. Shiryaev (1987): *Limit Theorems for Stochastic Processes*. Springer–Verlag.

Jorion, P. (2001): *Value at risk: the new benchmark for managing financial risk*. McGraw-Hill.

Karatzas, I., and S. E. Shreve (1991): *Brownian Motion and Stochastic Calculus*. Springer–Verlag.

Leisen, D. P. (2000): "Stock Evolution Under Stochastic Volatility: A Discrete Approach," *Journal of Derivatives*, 8(2), 9–27.

Melick, W., and C. Thomas (1997): "Recovering an Asset's PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis," *Journal of Financial and Quantitative Analysis*, 32(1), 91–115.

Merton, R. C. (1976): "Option Pricing When Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3(1/2), 125–44.

Nelson, D. B., and K. Ramaswamy (1990): "Simple Binomial Processes as Diffusion Approximations in Financial Models," *The Review of Financial Studies*, 3, 393–430.

Oksendal, B. (1995): *Stochastic Differential Equations: An Introduction with Application*. Springer–Verlag.

Protter, P. (1990): *Stochastic Integration and Differential Equations: A new approach*. Springer–Verlag.

Ritchken, P., and R. Trevor (1999): "Pricing Options under Generalized GARCH and Stochastic Volatility Processes," *The Journal of Finance*, 54(1), 377–402.

Stein, E., and J. Stein (1991): "Stock price distributions with stochastic volatility: an analytic approach," *The Review of Financial Studies*, 4, 727–752.

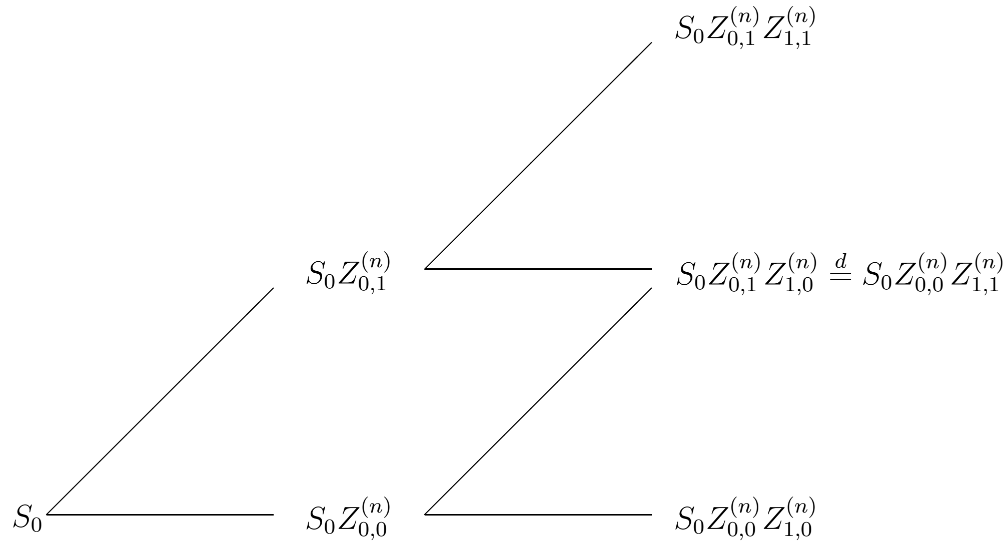

Fig. 1. Dynamics of the Poisson process over two periods and the resulting random variables that describe the securities price for our approximation.

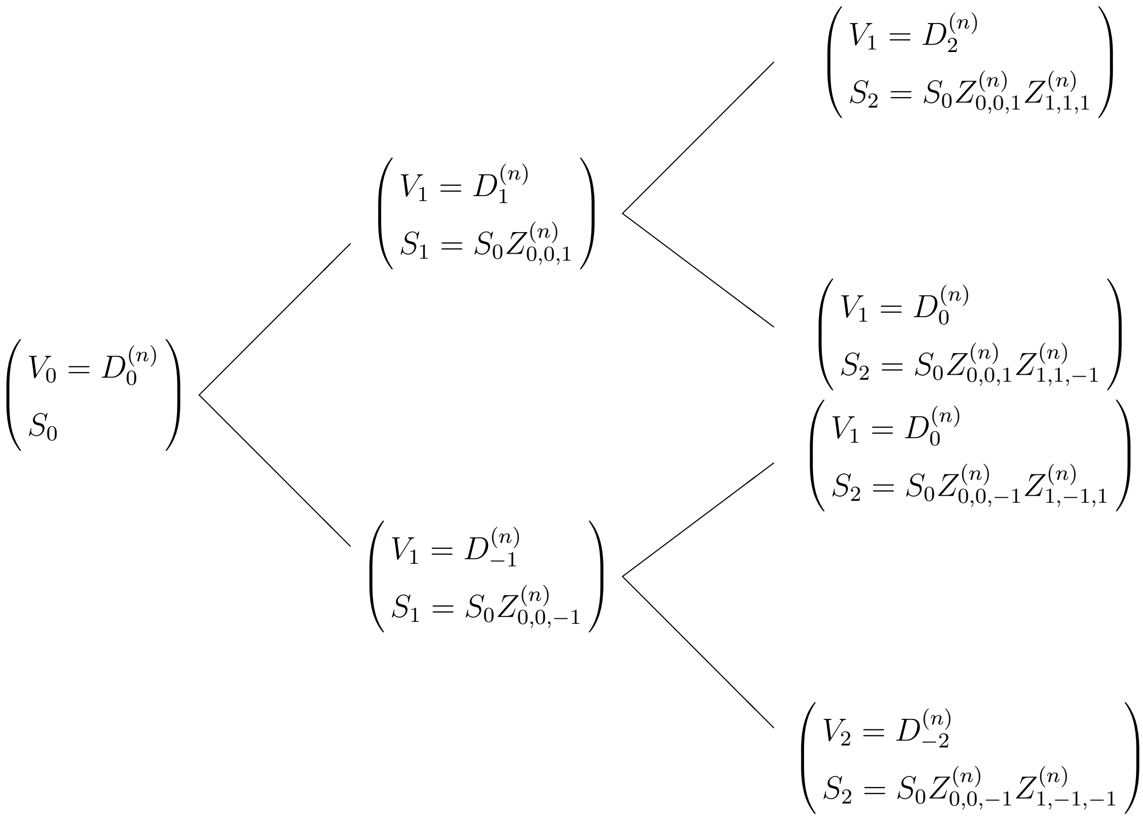

Fig. 2. A two-period description of the volatility dynamics and the lognormal random variables that describe the securities price.

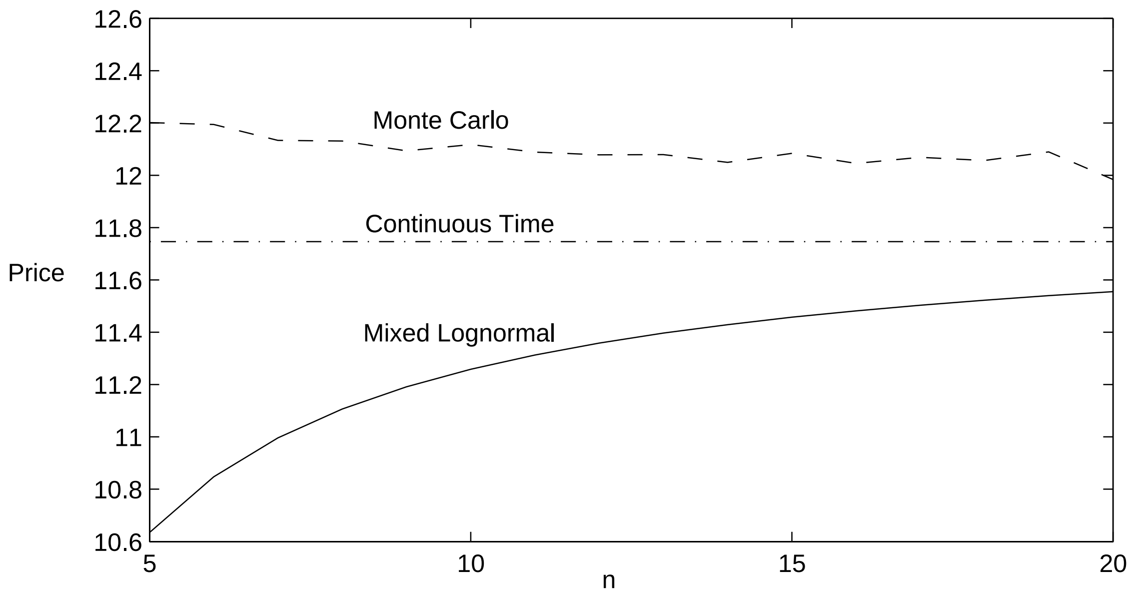

Fig. 3. Price approximations based on our mixed lognormals and Monte-Carlo depending on the refinement.

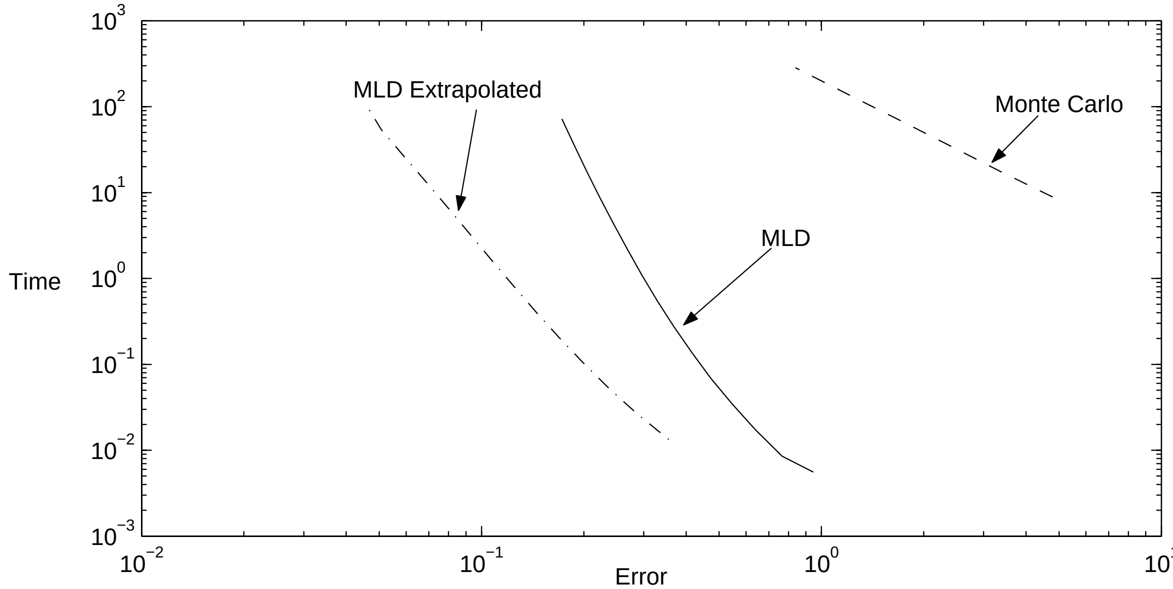

Fig. 4. Efficiency of our mixed lognormal (MLD), its extrapolation and Monte-Carlo.

	$\lambda = 0.01$					
	$K = 90$		$K = 100$		$K = 110$	
	Merton	MLD	Merton	MLD	Merton	MLD
$\alpha = -0.2, \beta = 0.1$	14.68	14.68	6.87	6.87	2.22	2.22
$\alpha = -0.2, \beta = 0.3$	14.73	14.73	6.91	6.91	2.25	2.25
$\alpha = -0.5, \beta = 0.1$	14.85	14.85	7.01	7.01	2.29	2.28
$\alpha = -0.5, \beta = 0.3$	14.85	14.85	7.01	7.01	2.29	2.29

	$\lambda = 0.05$					
	$K = 90$		$K = 100$		$K = 110$	
	Merton	MLD	Merton	MLD	Merton	MLD
$\alpha = -0.2, \beta = 0.1$	14.89	14.89	7.14	7.14	2.39	2.39
$\alpha = -0.2, \beta = 0.3$	15.12	15.11	7.34	7.33	2.56	2.55
$\alpha = -0.5, \beta = 0.1$	15.71	15.71	7.86	7.86	2.78	2.78
$\alpha = -0.5, \beta = 0.3$	15.71	15.70	7.84	7.83	2.77	2.76

	$\lambda = 0.20$					
	$K = 90$		$K = 100$		$K = 110$	
	Merton	MLD	Merton	MLD	Merton	MLD
$\alpha = -0.2, \beta = 0.1$	15.61	15.58	8.11	8.09	3.07	3.07
$\alpha = -0.2, \beta = 0.3$	16.60	16.43	9.02	8.86	3.88	3.73
$\alpha = -0.5, \beta = 0.1$	18.48	18.46	11.04	11.03	5.09	5.08
$\alpha = -0.5, \beta = 0.3$	18.63	18.56	11.00	10.95	5.02	4.98

Table 1
Approximations of call prices based on Merton's integration formula (Merton) and on our mixed lognormal distributions (MLD).

Model	$\mu_v(V)$	$\varphi(V)$	$\psi(V)$
Hull and White (1987)	$\nu - V$	σV	$\sqrt{V}$
Heston (1993)	$\nu - V$	$\sigma\sqrt{V}$	$\sqrt{V}$
Stein and Stein (1991)	$\nu - V$	σ	V
Chesney and Scott (1989)	$\nu - V$	σ	$\exp\{V\}$

Table 2
Parameter specifications for various models. (ν, σ are constants.)

Model	$f(x)$	$g(z)$
Hull and White (1987)	$\frac{\ln x}{\sigma}$	$\exp(\sigma z)$
Heston (1993)	$\frac{2\sqrt{x}}{\sigma}$	$\frac{(\sigma z)^2}{4}$
Stein and Stein (1991)	$\frac{x}{\sigma}$	σz
Chesney and Scott (1989)	$\frac{x}{\sigma}$	σz

Table 3
The transformation functions f, g for the models from the literature (σ as in table 2)