Mixed Lognormal Distributions for Financial Applications

Dietmar P.J. Leisen*

This version: September 20, 2003.

Abstract

This paper studies the pricing of European-style options using mixed lognormal distributions. We advocate such distributions as a computationally efficient way to calculate prices of such options: we derive higher truncated moments in analytic form, explain how to use them to derive closed-form approximations of European-style options and construct sequence of mixed lognormal distributions that approximate the terminal distribution of a stock that follows either a Black-Scholes model with jumps or one with stochastic volatility.

Keywords
(Mixed) lognormal distribution, Moments

JEL Classification
G13

* McGill University, Department of Mathematics and Statistics, 805 Sherbrooke St. W., Montreal, PQ H3A 2K6, `email: Dietmar.Leisen@McGill.CA`
The author gratefully acknowledges research assistance provided by FCAR, NSERC and IFM2.

1 Introduction

One of the intrinsic advantages of the Black-Scholes setup is that pricing formulas for many derivatives can be expressed in closed-form; thereby prices of large portfolios can be calculated and intrinisic parameters can be extracted in efficient ways. More recent models of securities' dynamics have much better statistical porperties [1] but lack the above-mentioned computational properties. This paper argues that mixed lognormal distributions (henceforth *MLD*) provide approximations that exhibit both better statistical properties [2] and retain the analytic tractability of the Black-Scholes setup.

Our argument rests on two contributions: in a first contribution this paper calculates truncated moments of MLD in closed-form and explains how this is used to approximate derivatives' payoffs and resulting prices. Truncated moments play an important role in the calculation of option prices, e.g. the European call and put option pricing formula can be decomposed into a sum of truncated moments of order zero and 1. Most option payoff functions of the literature only have a finite number of discontinuties; we discuss how a Taylor-series approximation over sub-intervals can yield approximations up to any oder of accuracy. This requires calculation of trunated moments up to third order, which the paper derives in closed-form. This provides the basis to calculate option prices efficiently for mixed lognormal distributions.

The second contribution is that we construct sequences of MLD that approximate the Black-Scholes setup and models of stochastic volatility [3]. Stochastic volatility and Black-Scholes with jumps are the two extensions of the Black-Scholes theory that are used most often in the literature. They provide a rich basis to describe the dynamics of the underlying securities. To provide MLD approximations of the Black-Scholes model with jumps we assume that over each period at most one jump occurs; this describes a sequence of distribution of stock

[1] For a recent discussion, see, e.g. Bates (2000) and Eraker, Johannes, and Polson (2003).
[2] It is a well documented fact that asset price distributions are skewed and exhibit kurtosis in excess compared to the lognormal. Mixed lognormal distributions can be used to construct distributions with skewness and excess kurtosis relative to the lognormal. For example, a mixture of two lognormals has five degress of freedom which allow matching the empirical mean, variance, skewness and kurtosis.
[3] This paper ignores the important question what determines in incomplete markets the pricing measure, we refer to the literature for details on that.

price at maturity and we prove that this is a sequence of MLD and that it converges to the continuous-time distribution at maturity; we use this to prove the convergence of prices to the continuous-time solution. Tree based approximations for stochastic volatility models have been constructed by Hilliard and Schwartz (1996) and Leisen (2000)[4]; we construct mixing lognormal distributions that are driven by a binomial path approximation for the volatility process[5].

To our knowledge, at present, MLD have not been used to provide an efficient way to calculate derivatives' prices. Previously, the literature studied them only marginaly and only to extract parameters of the underlying security price process or marginal distribution from traded options. Gemmill and Saflekos (2000) and Melick and Thomas (1997) used mixed lognormal distributions to calibrate the underlying stock price distribution such that prices of the stock and traded options on it are best matched. Brigo and Mercurio (2000) calibrate a mixture of geometric Brownian motions to traded option prices. Our technique could be adapted to these purposes and used to provide efficient model based approximations of marginal distributions. In this paper we refrain from doing so since our focus is on efficient calculation of prices.

The remainder of the paper is organized as follows: the following section introduces MLD, discuss option pricing using such distributions and introduces the concept of weak convergence of MLD and its implications for price convergence. The third section looks at the Black-Scholes model with jumps; it discuss convergence in distribution of random variables as the concept to ensure that prices calculated through MLD approximations converge to the continuous-time counterpart, construct sequences of MLD with that property and discuss efficiency of the numerical schemes. The fourth section parallels that of the third one looking at the Black-Scholes model with stochastic volatility. The fifth section concludes the paper.

[4] Ritchken and Trevor (1999), and Duan and Simonato (2000) provides approximations for GARCH models.

[5] The binomial approximation for the volatility process replicates the construction by Nelson and Ramaswamy (1990)

2 Mixed Lognormal Distributions

2.1 Notation

On a probability space $(\Omega, \mathcal{F}, Q)$ we adopt:

Definition 1 *A* mixed lognormal distribution (henceforth MLD) *with* $((\mu_1, \sigma_1, \gamma_1), \ldots, (\mu_M, \sigma_M, \gamma_M))$ *is a random variable* A *with*

$$A \stackrel{d}{=} A_0 \sum_{i=0}^{M} 1_{C=i} X_i, X_i \stackrel{d}{=} \exp\left(\mu_i + \sigma_i Y_i\right),$$

where Y_i *is a standard normal distributed random variable under* Q*, i.e. a normal distributed random variable with mean 0 and variance 1,* $Y_i \stackrel{d}{=} \mathcal{N}(0,1)$*,* $A_0, \mu_i \in \mathbb{R}, \sigma_i > 0, \gamma_i \geq 0$ *(*$i = 0, \ldots, M$*),* $\sum_{i=0}^{M} \gamma_i = 1$ *and* C *is a random variable on* $\{0, 1, \ldots, M\}$ *with* $Q[C = i] = \gamma_i$*.*

We denote by $\stackrel{d}{=}$ equality in distribution of random variables and by $E_Q[\cdot]$ the expectation operator with respect to Q. In the remainder of this section we look at the MLD of definition 1. For future reference we denote for $i = 1, \ldots, M$, $j = 0, 1, 2, 3$ and $K \geq 0$

$$\nu_{ij}(K) = E_Q[X_i^j \cdot 1_{X_i \geq K}], \nu_{Aj}(K) = E_Q[A^j \cdot 1_{A \geq K}].$$

Here $\nu_{ij}(K)$ is the function that describes the j-th moment of X_i truncated at K and $\nu_{Aj}(K)$ is the function that describes j-th moment of the MLD A. As usual we assume that any number taken to the power 0 is equal to 1, i.e. $x^0 = 1$ for all $x \in \mathbb{R}$ so that $\nu_{i0}(K)$ and $\nu_{A0}(K)$ describe the upper probability distribution. Note that the random variable A can not become negative and so we use the value at $K = 0$ in these functions to describe the non-truncated moments. We have

$$\nu_{Aj}(K) = E_Q[E[A^j \cdot 1_{A \geq K} | C]] = \sum_{i=1}^{M} \gamma_i E_Q[A^j \cdot 1_{A \geq K} | C = i] = A_0^j \sum_{i=1}^{M} \gamma_i \cdot \nu_{ij}(K). \quad (1)$$

According to the appendix we have

$$E_Q[X_i] = A_0 \exp\left(\mu_i + \frac{1}{2}\sigma_i^2\right), E_Q[X_i^2] = A_0^2 \exp\left(2\mu_i + 2\sigma_i^2\right), \quad (2)$$

$$E_Q[X_i^2] = A_0^3 \exp\left(3\mu_i + \frac{9}{2}\sigma_i^2\right), E_Q[X_i^4] = A_0^4 \exp(4\mu_i + 8\sigma_i^2), \quad (3)$$

and

$$\nu_{i0}(K) = \Phi_n(d_{i2}), \nu_{i1}(K) = \exp\left(\mu_i + \frac{1}{2}\sigma_i^2\right) \cdot \Phi_n(d_{i1}), \tag{4}$$

$$\nu_{i2}(K) = \exp(2\mu_i + 2\sigma_i^2) \cdot \Phi_n(d_{i3}), \nu_{i3}(K) = \exp\left(3\mu_i + \frac{9}{2}\sigma_i^2\right) \cdot \Phi_n(d_{i4}) \tag{5}$$

where for $i = 1, \ldots, N$ the parameters [6]

$$d_{i1} = \frac{\ln(A_0/K) + \mu_i + \sigma_i^2}{\sigma_i}, d_{i2} = d_{i1} - \sigma_i, d_{i3} = d_{i1} + \sigma_i, d_{i4} = d_{i1} + 2\sigma_i,$$

and $\Phi_n(\cdot)$ denotes the standard normal cumulative distribution function.

2.2 Using MLD to approximate derivatives prices

Throughout this paper we are interested in pricing European-style options with maturity date $T > 0$. We therefore need to evaluate the distributions from the stock at that date and look at those MLD which can be used to price derivatives as discounted expected payoffs, i.e. for a derivative with payoff $f(A)$ at time T we calculate its price as

$$\frac{1}{B_T} E_Q[f(A)], \tag{6}$$

where $1/B_T$ denotes the discount factor [7] between 0 and T. Here the probability measure Q is the so-called the risk-neutral pricing measure; it will be discussed in the following sections in more detail.

Equation (6) can be calculated easily for many derivatives in closed-form. Our main example of interest here is the standard call option, e.g. the payoff $f(A) = (A - K)^+$ of a call option with strike K; it has the price

[6] The parameters $d_{i1}(K), d_{i2}(K)$ correspond to those used in the Black-Scholes pricing formula. The connection will be further explored below.

[7] In the following sections the short rate will be assumed to be constant over time and then $B_T = B_0 \exp(rT)$. Our setup could be extended to stochastic interest rate models using the change of measure technique derives similar representations, see, e.g. Bjork (1999).

$$\begin{aligned}\frac{1}{B_T}E_Q[(A-K)^+] &= \frac{1}{B_T}\left(E_Q[A\cdot 1_{A\geq K}] - KE_Q[1_{A\geq K}]\right) &(7)\\ &= \frac{1}{B_T}\left(\nu_{A1}(K) - K\cdot\nu_{A0}(K)\right) &(8)\\ &= \frac{1}{B_T}\sum_{i=1}^{N}\gamma_i\cdot\left(A_0\nu_{i1}(K) - K\cdot\nu_{i0}(K)\right) &(9)\\ &= \frac{1}{B_T}\sum_{i=1}^{N}\gamma_i\cdot\exp\left(\mu_i+\frac{\sigma_i^2}{2}\right)\cdot BS(A_0,K,\mu_i,\sigma_i), &(10)\end{aligned}$$

where we define

$$BS(A_0,K,\mu_i,\sigma_i) = A_0\Phi_n(d_{i1}(K)) - K\cdot\exp\left(-\mu_i-\frac{\sigma_i^2}{2}\right)\cdot\Phi_n(d_{i2}(K)), \qquad (11)$$

and $d_{i1}(K), d_{i2}(K)$ are as above and $\Phi_n(\cdot)$ denotes the standard normal distribution function. We refer to the formula in equation (11) later as the *Black-Scholes (call option) formula*. This is a slight abuse of notation since we have defined μ_i, σ_i to contain the maturity T, whereas usually the Black-Scholes keeps these parameters separate: note that in the Black-Scholes setup the stock price is $S_T = S_0\exp((r-\sigma^2/2)T+\sigma W_T^Q)$, where W^Q is a standard Brownian motion under Q; the bond is $B_T = B_0\exp(rT)$ and then the price of the option with maturity T is given by

$$\begin{aligned}&BS\left(S_0,K,\left(r-\frac{\sigma^2}{2}\right)T,\sigma\sqrt{T}\right)\\ &= S_0\Phi_n\left(\frac{\ln(S_0/K)+\left(r-\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right) - K\cdot\exp\left(-rT\right)\cdot\Phi_n\left(\frac{\ln(S_0/K)+\left(r+\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right).\end{aligned}$$

Many derivatives' payoff functions are piecewise linear; we can then proceed similarly, e.g for a put option and others that are subject to truncation. When the function f is not piecewise linear but it is differentiable we can perform a Taylor-series expansion $f(a) = f(a_0) + f'(a_0)(a-a_0) + f''(a_0)(a-a_0)^2 + f'''(a_0)(a-a_0)^3 + \ldots$. Provided we can exchange integration and infinite summation the price of derivative f is approximatively

$$\frac{1}{B_T}\cdot\left(f(a_0) + f'(a_0)\left(\nu_{A1}(0)-a_0\right) + f''(a_0)(\nu_{A2}(0)-a_0)^2 + f'''(a_0)(\nu_{A3}(0)-a_0)^3 + \ldots\right)$$

Here we stopped our approximation somewhat artifically at the cubic term, but this could of course be extended easily to achieve higher accuracy.

If the payoff function f is not differentiable then in practice it will be piecewise differentiable.

We can then use the truncation method described above together with the Taylor series expansion method to derive closed-form approximations up to the cubic term using equations (1, 4, 5).

2.3 Convergence of MLD

In the following sections we construct approximations of the marginal distribution S_T (of the stock price at time T) using sequences of MLD $A^{(n)}$. We assume each $A^{(n)}$ is characterized using $C^{(n)}, X_i^{(n)}$ $(n = 1, 2, \ldots; i = 1, \ldots, i_{max} < \infty)$. We are then interested in using these to calculate derivatives' prices.

To apply general mathematical theorems we need to translate this into the right convergence concept. If Q represents the pricing measure and A the distribution of stock prices we wish to approximate, then we calculate the "true" price as $B_0 \cdot E_Q[f(A)]$; for each refinement n we calculate similarly $B_0 \cdot E_Q[f(A^{(n)}]$. We are here interested in the convergence $E_Q[f(A^{(n)})] \xrightarrow{n} E_Q[f(A)]$ for *any* European-style derivative payoff function f. The definition of the mathematical concept of convergence in distribution $A^{(n)} \stackrel{d}{\Longrightarrow} A$ is to require this to hold for all *bounded* payoff functionals F. In the following sections we adopt this as a consistency requirement[8]. Note that this requires $A^{(n)}$ to be (approximations of) risk-neutral distributions but they do not correspond to risk-neutral distributions nor do the numbers $B_0 \cdot E_Q[f(A^{(n)}]$ correspond to prices themselves. They are to be interpreted merely as approximations of the prices of driven by A under Q.

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

We assume a market in which a stock and a bond are traded. The interest rate r is assumed to be constant over time; without loss of generality we normalize the price of today's (date 0) bond to 1 so that $B_t = \exp\{rt\}$. We fix a probability space $(\Omega, \mathcal{F}, P)$; P is called the *objective*

[8] The "boundedness" condition excludes some payoffs, e.g. call options. It excludes payoffs like call options. However, there are tricks to get around this, e.g. using put–call parity and the fact that the put option is bounded.

probability measure and describes the actual distribution the econometrician would observe. It is to be distinguished from the so-called risk-neutral probability measure Q we introduce later under which pricing will occur.

3.1 The Continuous-Time Dynamics

We introduce here the model of Merton (1976) which is an extension of the Black-Scholes setup to jumps. Although there are many other ways to introduce jumps this has become the most frequently used one and for that reason it will be studied here. We assume that the stock price dynamics is under P

$$S_t = S_0 \cdot \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right\} \cdot \prod_{i=1}^{N_t} U_i. \tag{12}$$

where W is a standard Wiener process and $\mu \in R$, $\sigma > 0$. $(N_t)_t$ is a Poisson process with constant parameter $\lambda > 0$, $(U_i)_i$ a sequence of lognormal random variables, $U_i \stackrel{d}{=} \exp(\alpha + \beta Y_i)$, where $Y_i \stackrel{d}{=} \mathcal{N}(0,1)$. 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 other.

To illustrate the dynamics of stock prices let us define the process $(G_t)_t$ with $G_t = S_0 \cdot \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right\}$. It describes the first part in equation (12), has stationary, Gaussian returns and is the continuous process known as geometric Brownian motion. Originally suggested by Samuelson (1965) and later used by Black and Scholes (1973) in their seminal contribution it has become one of the standard financial models for derivatives pricing. S evolves according to G until the next jump time τ of the Poisson process at which N changes from, say, i to $i+1$. We then observe a per-cent change $U_i - 1$, i.e., the stock changes value from $S_{\tau-}$ before the jump to $S_{\tau-} \cdot U_i$. Therefore, the two parts in equation (12) have the following two properties: $(G_t)_t$ models the "typical" evolution of the stock under the "normal" arrival of information, whereas $\prod_{i=1}^{N_t} U_i$ models jumps in the stock prices, due to some rare strong information shock. Since the Poisson process is "memoryless," the expected time until the next shock occurs is equal to $1/\lambda$, independent of current time.

The Fundamental Theorem of Asset Pricing (Harrison and Kreps (1979), Harrison and Pliska (1981)) implies that there are no arbitrage opportunities if and only if a so-called *equivalent* *martingale measure* (henceforth *EMM*) Q exists, i.e. a probability measure Q that is equivalent to Q and under which discounted price processes of all traded securities are martingales. If the derivative with payoff $f(S_T)$ is traded in addition to the stock then this means that price processes for the stock S_t and for the derivative D_t are under Q

$$S_t = B_t E_Q \left[\frac{S_T}{B_T} \middle| \mathcal{F}_t \right], D_t = B_t E_Q \left[\frac{f(S_T)}{B_T} \middle| \mathcal{F}_t \right], \tag{13}$$

where $\mathcal{F}_t = \sigma(S_u | 0 \leq u | leqt)$ describes the σ-algebra generated by S. We refer to this as the *risk-neutral technique* and to Q as the *pricing measure*. This setup represents an incomplete market, i.e., not every contingent claim on the stock can be hedged (see Duffie (1992)). In the language of EMM this means that under the assumption of absence of arbitrage opportunities there is a multiplicity of EMM, i.e. there is multiplicity of probability measure Q, equivalent to P, under which $(S_t/B_t)_t$ is a martingale. From a purely theoretical perspective the probability measures P can be changed separately on three components:

(1) the Wiener process part: Girsanov's theorem tells us how to change the mean (drift) and that changes of variance (volatility) are impossible.
(2) Poisson process part: We refer to Brémaud (1981) how to change the intensity of the Poisson process.
(3) The sequence of random variables U_i: The distribution of each of these can be changed to any other distribution whose density function w.r.t. the Lebesgues measure is absolutely continuous.

Throughout we adopt the following:

Restriction 2 *We only study probability measures Q that are equivalent to P with the properties that*

(1) U_i remains a sequence of independent, identically, lognormally distributed random variables, and
(2) the intensity of the Poisson process remains constant over time.

This restricts us to measure changes that leave us in the same model class we started with in equation (12). We do not impose any other restrictions on the Poisson process and no restrictions on the Wiener process part. For any equivalent probability measure Q under which the drift of the Wiener process is constant over time[9] we denote by λ_Q the intensity of N, ν the drift of the Wiener process, ν_U the mean of U_i, and by θ, ζ the mean and variance of Y_i, all under the probability measure Q, i.e.[10]

$$\lambda_Q = \frac{E_Q[N_t]}{t}, \nu = \frac{E_Q[W_t]}{t}, \nu_U = E_Q[U_i] = \exp\left(\alpha + \beta\theta + \frac{(\beta\zeta)^2}{2}\right), \theta = E_Q[Z_i], \zeta = var_Q(Z_i).$$

Note also that the process W^Q, defined by setting $W_t^Q = W_t - \nu t$, is a standard Wiener process under Q. According to equation (13) a pricing measure Q is a probability measure Q equivalent to P with the property that for all $0 \leq t \leq T$, $E[S_t/S_0] = \exp(rt)$. Since

$$E_Q\left[\prod_{i=1}^{N_t} U_i\right] = \exp\left(\lambda_Q \cdot E_Q[U_i - 1]t\right) = \exp\left(\lambda_Q \cdot (\nu_U - 1) \cdot t\right),$$

see, e.g., Protter (1990) or Jacod and Shiryaev (1987), and $E_Q[G_t/G_0] = \exp((\mu + \nu\sigma)t)$ any pricing measure Q is characterized by

$$r - \mu = \nu\sigma + \lambda_Q(\nu_U - 1) = \nu\sigma + \lambda_Q\left(\exp\left(\alpha + \beta\theta + \frac{(\zeta\beta)^2}{2}\right) - 1\right). \tag{14}$$

We therefore have four variables for measure changes: changes in $\nu, \lambda_Q > 0, \theta, \zeta$. Throughout we will not discuss how to derive the probability measure under which pricing should occur in the market; there are many excellent papers on that topic in the mathematical finance literature to which we refer the reader. Here we assume suitable parameters λ_Q, θ, ζ are chosen correspondingly; we then know that the dynamics of S is under Q

$$S_t = S_0 \cdot \exp\left\{\left(r - \frac{\sigma^2}{2} - \lambda_Q(\nu_U - 1)\right)t + \sigma \cdot W_t^Q\right\} \cdot \prod_{i=1}^{N_t} U_i \tag{15}$$

$$= S_0 \cdot \exp\left\{\mu_{rn} t + \sigma \cdot W_t^Q\right\} \cdot \prod_{i=1}^{N_t} U_i, \tag{16}$$

$$\text{where } \mu_{rn} = r - \frac{\sigma^2}{2} - \lambda_Q(\nu_U - 1). \tag{17}$$

This representation is what we will use in the sequel when discussing approximations and derive option prices as discounted expectations $\frac{1}{B_T} E_Q[f(S_T)]$.

---

[9] A consequence of restriction 2 is that we will only be interested in changes of the Wiener process part where the drift is constant over time, see equation (14).

[10] Note that by the first condition in restriction 2 we are permitted to change volatility and variance of Z_i.

For given integer n we discretize the interval $[0, T]$ into n equidistant intervals $\mathcal{I}_i = [t_{ni}, t_{n,i+1})$ $(i = 0, \ldots, n-1)$ with $t_{ni} = i\Delta t_n$, $\Delta t_n = \frac{T}{n}$. Discretization and convergence are linked: the appropriate discretization is one where taking the limits we recover the continuous processes. As discussed before our goal is to construct sequences that converge in disctribution to their continuous-time counterpart.

We will now construct processes $(S_k^{(n)}, N_k^{(n)})_{k=0,\ldots,n}$ for given n. Under Q we start with sequences of standard normal random variables $\tilde{Y}_{ni0}, \tilde{Y}_{ni1}, \tilde{Y}_{ni2}, \tilde{Y}_{ni3}$ where $n = 1, 2, \ldots$ and $i = 0, 1, 2, \ldots, n$; elements within and between the sequences are supposed to be independent of the others. We calculate under Q that $\sigma W_T + \sum_{i=1}^{k} \beta Y_i \overset{d}{=} \mathcal{N}(0, \sigma^2 T + i\beta^2)$, since the sequence $U_i \overset{d}{=} \exp(\alpha + \beta Y_i)$ is iid with $Y_i \overset{d}{=} \mathcal{N}(0,1)$, $W_t^Q \overset{d}{=} \mathcal{N}(0,t)$. A straightforward approximation can then be obtained based on equation (16) by setting for $i = 0, 1, \ldots, n$

$$X_{1i}^{(n)} = S_0 \exp\left\{\mu_{rn} T + \alpha i + \sqrt{\sigma^2 T + i\beta^2} \cdot \tilde{Y}_{ni3}\right\}, \tag{18}$$

denoting by $C_1^{(n)}$ the random variable N_T truncated to the set $\{0, \ldots, n\}$ and by

$$A_1^{(n)} = \sum_{i=0}^{n} 1_{C_1^{(n)}=i} X_{1i}^{(n)} \tag{19}$$

the resulting sequence of MLD. Note that this is converges in distribution to S_T, i.e. $A_1^{(n)} \overset{d}{\Longrightarrow} S_T$, since $C_1^{(n)} \overset{d}{\Longrightarrow} N_T$; therefore this is a first approximation in line with our goals. (We recall that by equation (17), $\mu_{rn} = r - \frac{\sigma^2}{2} - \lambda_Q(\nu_U - 1)$.)

In the remainder of this section we construct a second MLD approximation based directly on an approximation of processes. There are two reasons for us: First, a process approximation might be of interest to those readers who look for process approximations, e.g. to calculate prices of path-dependent securities. The second is that our construction for the stochastic volatility process will be based on a process approximation and so this allows us to introduce in a simplified setup the main concepts that guide us there.

We construct the processes $(S_k^{(n)}, N_k^{(n)})_{k=0,\ldots,n}$ by forward induction as follows: At date $0 \leq k \leq n$ the distribution of $S_{k+1}^{(n)}$ at the next date conditional on $S_{t_{nk}} = S_k^{(n)}$, $N_{t_{nk}} = N_k^{(n)}$ at date k would be equal in distribution to

$$S_k^{(n)} \cdot \exp\left\{\mu_{rn} t + \sigma\sqrt{\Delta t_n}\tilde{Y}_{nk0}\right\} \cdot \prod_{i=N_k^{(n)}}^{N_{t_{nk+1}}} U_i,$$

if we assume that the dynamics over that interval follows the one described in equation (16). We will now approximate the Poisson process part: the Poisson process jumps in integers and is memoryless, i.e. the number of jumps over the interval of length Δt_n ahead of t_{nk} is independent of the number of jumps that occured up to that date and depends only on the length of the time period; it is known that the distribution of $N_{t_{nk+1}}$ conditional on $N_{t_{nk}} = N_k^{(n)}$ is

$$P[N_{t_{nk+1}} = N_k^{(n)} + n | N_{t_{nk}} = N_k^{(n)}] = e^{-\lambda_Q \Delta t_n} \frac{(\lambda_Q \Delta t_n)^n}{n!}.$$

For sufficiently small Δt_n we approximate that distribution by $P[N_{t_{nk+1}} = N_k^{(n)} + n | N_{t_{nk}} = N_k^{(n)}] = e^{-\lambda_Q \Delta t_n} \approx 1 - \lambda_Q \Delta t_n$, and $P[N_{t_{nk+1}} > N_k^{(n)} + n | N_{t_{nk}} = N_k^{(n)}] \approx \lambda_Q \Delta t_n$. We define sequences of mutually independent random variables $Z_{n,k,0}, Z_{n,k,1}, \epsilon_{nk}$ $(n = 1, 2, \ldots; k = 0, 1, 2, \ldots, n)$ with

$$\begin{aligned}
Z_{n,k,0} &= \exp\left\{\mu_{rn}\Delta t_n + \sigma\sqrt{\Delta t_n}\ \tilde{Y}_{nk0}\right\}, \\
Z_{n,k,1} &= \exp\left\{\mu_{rn}\Delta t_n + \alpha + \sqrt{\sigma^2 \Delta t_n + \beta^2} \cdot \tilde{Y}_{nk1}\right\} \stackrel{d}{=} Z_{n,k,0} \cdot U_{nk},
\end{aligned}$$

and

$$\epsilon_{nk} = \begin{cases} 1; & \text{with probability } \lambda_Q \Delta t_n \\ 0; & \text{with probability } 1 - \lambda_Q \Delta t_n \end{cases}.$$

We then set

$$N_{k+1}^{(n)} = N_k^{(n)} + \epsilon_{nk}, S_{k+1}^{(n)} = S_k^{(n)} \cdot Z_{n,k,\epsilon_{nk}} = S_k^{(n)} \cdot Z_{n,k,0} \cdot 1_{\epsilon_{nk}=0} + S_k^{(n)} \cdot Z_{n,k,1} \cdot 1_{\epsilon_{nk}=1}.$$

Note that this implies that conditional on $S_k^{(n)}$ we have

$$S_{k+1}^{(n)} = \begin{cases} S_k^{(n)} \cdot Z_{n,k,1}; & \text{with probability } \lambda_Q \Delta t_n \\ S_k^{(n)} \cdot Z_{n,k,0}; & \text{with probability } 1 - \lambda_Q \Delta t_n \end{cases}.$$

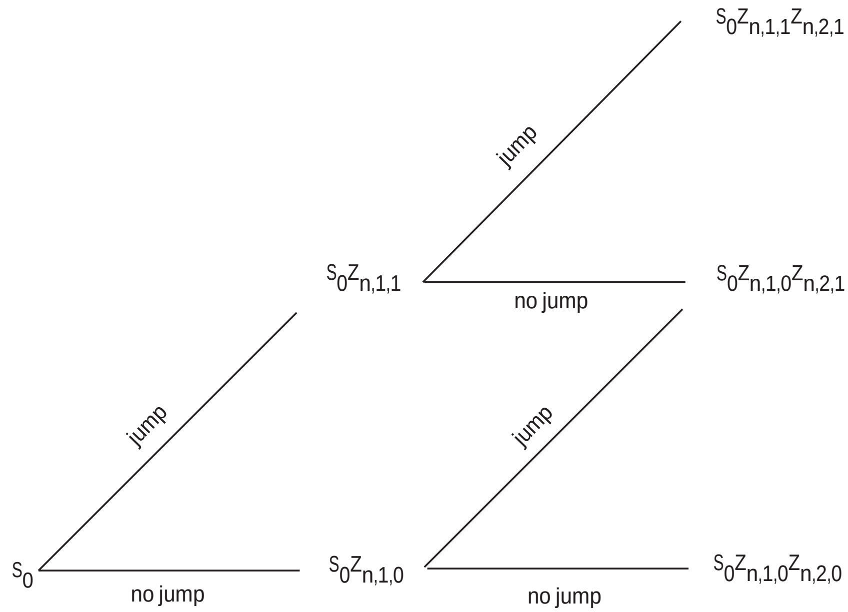

Fig. 1. Conditional dynamics of the Poisson process over two periods and the resulting random variables that describe the stock price.

Figure 1 provides a snapshot of our approximation for the Poisson process and the resulting random variables to describe the dynamics of stock prices conditional on a jump event or the absence of a jump ("no jump") between dates 0,1, and 2. Note that the resulting random variables to describe the stock prices are equal in distribution on whether we first go up and then down or the other way round. Therefore this behaves like a a so-called recombining "tree;" this will reduce significantly the computational requirements.

We conjecture that $(S^{(n)}, N^{(n)}) \overset{d}{\Longrightarrow} (S, N)$ but will not prove it since our main interest is in European-style options and for that purpose convergence for functionals based on time T, only, are sufficient. Instead we note that this gives us a sequence of distributions at date n:

Theorem 3 *The sequence of random variables $S_n^{(n)}$ at time T converges in distribution to the distribution of stock price S_T in the continuous-time model of equation (12), i.e. $S_n^{(n)} \overset{d}{\Longrightarrow} S_T$.*

To prove theorem 3 note first that the Central Limit Theorem implies that $\sqrt{\Delta t_n} \sum_{i=1}^n \tilde{Y}_{ni0} \overset{d}{\Longrightarrow} W_T^Q$; therefore

$$S_0 \prod_{i=1}^n Z_{n,i,0} = S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right) T + \sigma\sqrt{\Delta t_n} \sum_{i=1}^n \tilde{Y}_{ni0}\right) \overset{d}{\Longrightarrow} S_0 \exp\left(\left(r - \frac{\sigma^2}{2}\right) T + \sigma W_T^Q\right).$$

Furthermore we have $N_n^{(n)} \overset{d}{\Longrightarrow} N_T$ and so we derive $S_n^{(n)} \overset{d}{\Longrightarrow} S_T$, which ends the proof of theorem 3.

It therefore remains to write $S_n^{(n)}$ as an MLD to finish our construction of the second MLD that approximates S_T. For $i = 0, 1, \ldots, n$ we set

$$X_{2i}^{(n)} = S_0 \exp\left\{\mu_{rn}T + \alpha i + \sqrt{\sigma^2 T + i\beta^2} \cdot \tilde{Y}_{ni2}\right\}, \tag{20}$$

and denote by $C_2^{(n)}$ the n-step binomial distributed random variable with probability $\lambda_Q \Delta t_n$ for an up-move and $1 - \lambda_Q \Delta t_n$ for a down move and

$$A_2^{(n)} = \sum_{i=0}^{n} 1_{C_2^{(n)}=i} X_{2i}^{(n)}. \tag{21}$$

(We recall that by equation (17), $\mu_{rn} = r - \frac{\sigma^2}{2} - \lambda_Q(\nu_U - 1)$.) Note that $C_2^{(n)} \stackrel{d}{=} N_n^{(n)}$ and that

$$S_n^{(n)} \stackrel{d}{=} S_0 \cdot \prod_{k=0}^{N_n^{(n)}} Z_{n,k,1} \cdot \prod_{k=0}^{n-N_n^{(n)}} X_{n,k,0} \stackrel{d}{=} \sum_{i=0}^{n} 1_{N_n^{(n)}=i} X_{2i}^{(n)} \stackrel{d}{=} A_2^{(n)}. \tag{22}$$

Therefore $A_2^{(n)}$ is a sequence of MLD that approximate $S_n^{(n)}$. This ends the second construction of a sequence of MLD.

In this section we have presented two sequences of MLD approximations mixing $n+1$ lognormal distributed random variables given by equations (18, 19) and (20, 21). Note that they coincide in the lognormal random variable to be mixed, i.e. $X_{1i}^{(n)} \stackrel{d}{=} X_{2i}^{(n)}$; the difference is that one takes the exponential random variable N_T (with paramter $\lambda_Q T$) truncated to the interval $\{0, \ldots, n\}$ whereas the other one takes a binomial random variable on that set. Note that both converge to the continuous-time limit in distribution.

3.3 Accuracy and Efficiency

We now discuss accuracy and efficiency in pricing call options. We first recall that by equation (17), $\mu_{rn} = r - \frac{\sigma^2}{2} - \lambda_Q(\nu_U - 1)$. For both MLD approximations ($i = 1, 2$) we calculate approximations of the price of the call option with strike K as

$$
\begin{aligned}
\frac{1}{B_T} E_Q[(A_i^{(n)} - K)^+] &= E_Q\left[\frac{1}{B_T} E_Q[(A_i^{(n)} - K)^+ | C_i^{(n)}]\right] \\
&= \sum_{j=0}^{n} P[C_i^{(n)} = j] \frac{1}{B_T} E_Q[(A_i^{(n)} - K)^+ | C_i^{(n)} = j].
\end{aligned}
$$

The difference between the two MLD approximations is in the probability $Q[C_i^{(n)} = j]$. For the first it is $Q[C_1^{(n)} = j] = e^{-\lambda_Q T} \frac{(\lambda_Q T)^j}{j!}$ and so n-th approximation of the derivative's price is, see also the discussion leading to equations (7, 10),

$$
\begin{aligned}
\frac{1}{B_T} E_Q[(A_1^{(n)} - K)^+] &= e^{-(r+\lambda_Q)T} \sum_{j=0}^{n} \frac{(\lambda_Q T)^j}{j!} \exp\left(\mu_{rn} T + \alpha j + \frac{\sigma^2 T + j\beta^2}{2}\right) \\
&\qquad \cdot BS\left(S_0, K, \mu_{rn} T + \alpha j, \sqrt{\sigma^2 T + j\beta^2}\right).
\end{aligned} \tag{23}
$$

For the second MLD we have $Q[C_2^{(n)} = j] = \binom{n}{j} (\lambda_Q \Delta t_n)^j (1 - \lambda_Q \Delta t_n)^{n-j}$ and so the n-th approximation of the derivative's price is

$$
\begin{aligned}
&\frac{1}{B_T} E_Q[(A_2^{(n)} - K)^+] \\
&= e^{-rT} \sum_{j=0}^{n} \binom{n}{j} (\lambda_Q \Delta t_n)^j (1 - \lambda_Q \Delta t_n)^{n-j} \exp\left(\mu_{rn} T + \alpha j + \frac{\sigma^2 T + j\beta^2}{2}\right) \\
&\qquad \cdot BS\left(S_0, K, \mu_{rn} T + \alpha j, \sqrt{\sigma^2 T + j\beta^2}\right).
\end{aligned} \tag{24}
$$

Both approximations (23, 24) converge to the true price[11].

To discuss accuracy and efficiency of these approximations we now derive the continuous-time price: We write the continuous-time price as $\frac{1}{B_T} E_Q[(S_T - K)^+] = E_Q\left[\frac{1}{B_T} E_Q[(S_T - K)^+ | N_T]\right]$ and then based on equation (16) we calculate by analogy with the Black-Scholes derivation ($S_0 \exp(-\lambda_Q(\nu_U - 1)T) \prod_{k=1}^{j} U_k$ becomes the equivalent of the current stock price) that

$$
\begin{aligned}
&\frac{1}{B_T} E_Q[(S_T - K)^+ | N_T = j] \\
&= E_Q[E_Q[(S_T - K)^+ | U_1, \ldots, U_j] | N_T = j] \\
&= E_Q\left[BS\left(S_0 \exp(-\lambda_Q \cdot (\nu_U - 1)T) \prod_{k=1}^{j} U_k, K, r - \frac{\sigma^2}{2}, \sigma^2\right) \middle| N_T = j\right].
\end{aligned}
$$

[11] Although the call payoff options is not a bounded function the put payoff is and so put-call parity leads to that conclusion.

This gives the continuous-time price as

$$
\begin{aligned}
&\frac{1}{B_T} E_Q[(S_T - K)^+] \qquad (25)\\
&= e^{-\lambda_Q T} \sum_{j=0}^{\infty} \frac{(\lambda_Q T)^j}{j!} E_Q \left[BS \left(S_0 \exp(-\lambda_Q \cdot (\nu_U - 1)T) \prod_{k=1}^{j} U_k, K, r, \sigma^2 \right) \right] \\
&= e^{-\lambda_Q T} \sum_{j=0}^{\infty} \frac{(\lambda_Q T)^j}{j!} E_Q \left[BS \left(S_0 \exp\left(-\lambda_Q \cdot (\nu_U - 1)T + j(\alpha + \beta Y_j)\right), K, r - \frac{\sigma^2}{2}, \sigma^2 \right) \right].
\end{aligned}
$$

This coincides with the equation in Merton (1976) for the price of a call option [12].

For the following approximations we take $\lambda_Q = \lambda$, $\theta = 0$, $\zeta = 1$, i.e. we leave the jump part independent as in Merton (1976). We then have $\nu_U = \exp(\alpha + \frac{\beta^2}{2})$. Table 1 presents prices using our approach according to equation (24) (*MLD*) and those of the Merton model according to equation (25) (*Merton*). In table 1 Merton sums up $j = 0, \ldots, 20$ while MLD only takes $j = 0$ and $j = 1$. All cases take $S_0 = 100, r = 0.05, T = 1, \sigma = 0.1$. We compare prices for three call options with strikes $K = 90; 100; 110$; each line differs in their risk-neutral jump-components, i.e. in the variables describing mean and variance α, β in the events of jumps and in the frequency of jumps λ_Q. Note that λ_Q equal to 0 corresponds to the Black-Scholes setup. Prices in that setup are $14.6288, 6.8050, 2.1739$ respectively for the three options. All approximations are off at most one decimal in the third digit to the Merton solution despite the simplicity of our approximation. Higher order mixing ($j = 0, 1, 2$) would produce even more accurate results but we leave this to the reader since we believe our implementation is both simple and fairly accurate.

4 Mixed Lognormal Distributions as an Approximation to Stochastic Volatility Models

As in the previous section we assume a market in which a stock and a bond with price $B_t = \exp\{rt\}$ are traded. We fix a probability space $(\Omega, \mathcal{F}, P)$; P is called the *objective* probability measure and describes the actual distribution the econometrician would observe.

[12] Merton (1976) advocates a risk-neutral stock price dynamics with specific parameters for λ_Q, θ, ζ in equation (16). We do not impose this here and extend his formula as discussed here to the general case.

	$\lambda_Q = 0.01$					
	$K = 90$		$K = 100$		$K = 110$	
	MLD	Merton	MLD	Merton	MLD	Merton
$\alpha = -0.2, \beta = 0.1$	14.6823	14.6823	6.8723	6.8723	2.2156	2.2156
$\alpha = -0.2, \beta = 0.3$	14.7260	14.7253	6.9111	6.9105	2.2494	2.2489
$\alpha = -0.5, \beta = 0.1$	14.8506	14.8506	7.0144	7.0144	2.2881	2.2881
$\alpha = -0.5, \beta = 0.3$	14.8485	14.8484	7.0083	7.0083	2.2857	2.2856

	$\lambda_Q = 0.05$					
	$K = 90$		$K = 100$		$K = 110$	
	MLD	Merton	MLD	Merton	MLD	Merton
$\alpha = -0.2, \beta = 0.1$	14.8912	14.8893	7.1398	7.1390	2.3862	2.3859
$\alpha = -0.2, \beta = 0.3$	15.1142	15.0951	7.3394	7.3238	2.5604	2.5477
$\alpha = -0.5, \beta = 0.1$	15.7088	15.7088	7.8631	7.8631	2.7811	2.7811
$\alpha = -0.5, \beta = 0.3$	15.7074	15.7042	7.8349	7.8326	2.7667	2.7650

	$\lambda_Q = 0.20$					
	$K = 90$		$K = 100$		$K = 110$	
	MLD	Merton	MLD	Merton	MLD	Merton
$\alpha = -0.2, \beta = 0.1$	15.6100	15.5770	8.1078	8.0939	3.0724	3.0669
$\alpha = -0.2, \beta = 0.3$	16.5606	16.2665	8.9816	8.7406	3.8408	3.6435
$\alpha = -0.5, \beta = 0.1$	18.4783	18.4783	11.0384	11.0384	5.0853	5.0853
$\alpha = -0.5, \beta = 0.3$	18.6237	18.5672	10.9943	10.9520	5.0138	4.9818

Table 1
Comparing option prices calculated using Merton's formula with those calculated based on our approach (MLD) mixing two lognormals

4.1 The Continuous-Time Dynamics

Extending the Black-Scholes model by incorporating jumps captures the presence of rare events affecting securities. Another important extension of the Black-Scholes setup is the bivariate diffusion, where the dynamics under the objective probability measure P is given jointly by

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

Table 2
Parameter specifications for different models (σ a constant)

$$dV_t = \kappa(\nu - V_t)dt + \varphi(V_t)dW_{1t}, dS_t = \mu(V_t)S_t dt + \psi(V_t)S_t d\left(\rho W_{1t} + \sqrt{1-\rho^2}W_{2t}\right). \quad (26)$$

Here (W_1, W_2) is a bivariate independent Wiener process and $-1 \leq \rho \leq 1$. Note that $\rho W_{1t} + \sqrt{1-\rho^2}W_{2t}$ is also a standard Wiener process and that the instantaneous correlation of that process with W_1 is ρ. We define by $\mathcal{F}_t = \sigma(V_u, S_u | 0 \leq u \leq t)$ the filtration generated by V, S. We assume that the size of volatility is part of the information structure; this is a common assumption in the mathematical finance literature.

The process S describes the stock dynamics and V plays the role of the process that drives volatility; $\psi(V_t)$ enters as the current volatility. The function μ describes the stock drift and will not be discussed further. If the function $\psi(V) = V$, then V models volatility; for $\psi(V) = \sqrt{V}$, V models the variance. For simplicity we refer to V as *volatility* throughout. The volatility process is mean-reverting to ν at a rate κ; its dispersion coefficient ("volatility of volatility") is $\varphi(V_t)$. By specifying φ and ψ, the models in the literature can be treated in a unified way (see table 2). We will not impose specific functional forms for φ and ψ in (26), (26); we require φ, ψ only to be twice continuously differentiable as well as to fulfill growth conditions that ensure the existence of a solution to the system (26, 26). (We refer the reader to the literature on stochastic differential equations for a detailed treatment of this topic.)

For two functions η_1, η_2 we study the processes $(R_{1t})_t$ and $(R_{2t})_t$ that are solutions to the SDE's $dR_{1t} = \eta_1(t, R_{1t})R_{1t}dW_{1t}$, $dR_{2t} = \eta_2(t, R_{2t})R_{2t}dW_{2t}$ and assume that both are martingales; we can then define a probability measure Q_{η_1,η_2} that is equivalent to P. According to Girsanov's theorem under Q_{η_1,η_2} the processes $(W_{1t}^Q, W_{2t}^Q)_t$ with $dW_{1t}^Q = d(W_{1t} - \eta_{1t}t)$ and $dW_{2t}^Q = d(W_{2t} - \eta_{2t}t)$ are two standard (independent) Wiener processes under Q. We can write

$$dV_t = \{\kappa(\nu - V_t) + \varphi(V_t)\eta_{1t}\}\, dt + \varphi(V_t) dW_{1t}^Q,$$
$$dS_t = \left\{\mu(V_t) + \psi(V_t) \cdot \left(\rho\eta_{1t} + \sqrt{1-\rho^2}\eta_{2t}\right)\right\} S_t dt + \psi(V_t) S_t d\left\{\rho W_{1t}^Q + \sqrt{1-\rho^2} W_{2t}^Q\right\}.$$

The stock is traded while volatility is not. Therefore, according to the EMM technique we need to find the probability measure(s) such that the discounted stock price process is a martingale. (No such restriction is imposed on the volatility process.) Changes of the volatility and instantaneous correlation structure are not possible and therefore the condition to be a martingale is that

$$\psi(V_t) \cdot \left(\rho\eta_{1t} + \sqrt{1-\rho^2}\eta_{2t}\right) = r - \mu.$$

Again, this is an incomplete market setup and here we have one degree of freedom. In the following we assume that the "correct" process η_1 that describes the pricing measure in the market has been chosen; for a detailed discussion of that topic we refer to the vast mathematical finance literature. Since η_1 is adapted to the filtration $\sigma(W_u^1 | 0 \le u \le t) = \sigma(V_u | 0 \le u \le t)$ generated by V, there exists a function of V_t, which we call η_{1t} by simplicity, such that $\eta_{1t} = \eta_{1t}(V_t)$. Our construction will be for the equivalent martingale measure Q and for standard independent Brownian motion W_1^Q, W_2^Q under Q with

$$dV_t = \mu_{rn}^V(t, V) dt + \varphi(V_t) dW_{1t}^Q, \text{ where } \mu_{rn}^V(t, V) = \kappa(\nu - V_t) + \varphi(V_t)\eta_{1t}(t, V_t), \tag{27}$$

and

$$dS_t = rS_t dt + \psi(V_t) S_t d\left\{\rho W_{1t}^Q + \sqrt{1-\rho^2} W_{2t}^Q\right\}. \tag{28}$$

4.2 Constructing a Sequence of Mixed Lognormal Distributions

For given integer n we discretize the interval $[0, T]$ into n equidistant intervals $[t_{ni}, t_{n,i+1})$ $(i = 0, \ldots, n-1)$ with $t_{ni} = i\Delta t_n$, $\Delta t_n = \frac{T}{n}$. We will now construct joint Markov processes $(V_k^{(n)}, S_k^{(n)})_{k=0,\ldots,n}$ for this n; this will give us a sequence of processes and our goal is to construct sequences of processes such that $(V^{(n)}, S^{(n)}) \overset{d}{\Longrightarrow} (V, S)$ under Q. Heuristically the standardized mean and (co-)variances to should converge to their continuous-time counterpart. The following theorem formalizes mathematically that essentially imposing these conditions is sufficient for weak convergence; it is an immediate consequence of the martingale central limit theorem (see, Ethier and Kurtz (1986), p. 354 and He (1990)).

Theorem 4 *Define, for $m > 0$, $\mathcal{M}_m = (-m, m) \times (-m, m)$. If, for all $m > 0$,*

(1) drifts converge on $\mathcal{M}_m$, i.e. for $(V_k^{(n)}, S_k^{(n)}) \in \mathcal{M}_m$ with $t_{nk} \xrightarrow{n} t$, $V_k^{(n)} \xrightarrow{n} V$, $S_k^{(n)} \xrightarrow{n} S$ we have

$$\frac{1}{\Delta t_n} E_Q \left[\left(\begin{matrix} V_{k+1}^{(n)} - V_k^{(n)} \\ S_{k+1}^{(n)} - S_k^{(n)} \end{matrix} \right) \middle| V_k^{(n)}, S_k^{(n)} \right] \xrightarrow{n} \left(\begin{matrix} \mu_{rn}^V(t, V) \\ rS \end{matrix} \right),$$

(2) variance/covariances converge on $\mathcal{M}_m$, i.e. for $(V_k^{(n)}, S_k^{(n)}) \in \mathcal{M}_m$ with $V_k^{(n)} \xrightarrow{n} V$, $S_k^{(n)} \xrightarrow{n} S$ we have

$$\frac{1}{\Delta t_n} Var_Q \left[\left(\begin{matrix} V_{k+1}^{(n)} - V_k^{(n)} \\ S_{k+1}^{(n)} - S_k^{(n)} \end{matrix} \right) \middle| V_k^{(n)}, S_k^{(n)} \right] \xrightarrow{n} \left(\begin{matrix} \varphi^2(V) & \rho\varphi(V)\psi(V)S \\ \rho\varphi(V)\psi(V)S & \psi^2(V)S^2 \end{matrix} \right),$$

(3) and jump-sizes vanish on $\mathcal{M}_m$ in the limit Q-a.s., i.e.

$$\max_{(V_k^{(n)}, S_k^{(n)}) \in \mathcal{M}_m} \max \left(\left| V_{k+1}^{(n)} - V_k^{(n)} \right|, \left| S_{k+1}^{(n)} - S_k^{(n)} \right| \right) \xrightarrow{n} 0,$$

then, $(V^{(n)}, S^{(n)}) \stackrel{d}{\Longrightarrow} (V, S)$.

The local conditions of this theorem will guide us in our construction. Our main goal in the following construction is mainly to ensure conditions 1 and 2, i.e. convergence of the first two moments to their corresponding continuous-time counterparts; we will see that condition 3 on vanishing jump sizes is easily fulfilled on $\mathcal{M}_m$.

We now construct two processes $(S_k^{(n)})_{k=0,\ldots,n}$, $(V_k^{(n)})_{k=0,\ldots,n}$ for which we will prove at the end of this subsection $(V^{(n)}, S^{(n)}) \stackrel{d}{\Longrightarrow} (V, S)$ (as processes). As a direct corollary we then also know that $S_n^{(n)} \stackrel{d}{\Longrightarrow} S_T$. Our construction is based on sequences of standard normal random variables Y_{ni1}, Y_{ni2} ($n = 0, \ldots, n; i = 0, \ldots, 2^n$). Elements of each series are assumed to be independent of all others.

We first construct an approximation of the volatility process $(V_k^{(n)})_{k=0,\ldots,n}$ using the technique of Nelson and Ramaswamy (1990). For that purpose we define the function

$$f(x) = \int^x \frac{1}{\varphi(z)} dz,$$

Model	$f(x)$	$g(z)$
Hull and White (1987)	$\ln\left(\frac{x}{\sigma}\right)$	$\exp(\sigma z)$
Heston (1993)	$\frac{2\sqrt{x}}{\sigma}$	$\frac{(\sigma z)^2}{4}$; if $z > 0$ 0 ; otherwise
Stein and Stein (1991)	σx	$\frac{1}{\sigma z}$
Chesney and Scott (1989)	σx	$\frac{1}{\sigma z}$

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

and denote g its inverse (see table 3 for those of the models commonly used in the literature). Since $f'(V) = \frac{1}{\varphi(V)}$, Itô's formula implies that

$$df(V)_t = \left\{ \frac{\kappa(\nu - V_t)}{\varphi(V_t)} + \eta_1(t, V_t) + \frac{1}{2} f''(V_t)\varphi(V_t) \right\} dt + dW_{1t}^Q, \tag{29}$$

i.e. the dynamics of the transformed process $f(V)$ is homoscedastic. We define the points

$$D_i^{(n)} = g\left(f(V_0) + i\sqrt{\Delta t_n} \right), \; i \text{ an integer},$$

and the Markov-chain[13] $(M_k^{(n)})_k$ on the time-homogeneous grid $(D_i^{(n)})_i$ with the following transition probabilities: at date k conditional on $M_k^{(n)} = D_i^{(n)}$ we assume that the transition to grid point $D_{m_{k,i}+1}^{(n)}$ occurs with probability $q_{k,i}^{(n)}$ and to $D_{m_{k,i}-1}^{(n)}$ occurs with complementary probability $1 - q_{k,i}^{(n)}$, i.e.

$$Q\left[M_{k+1}^{(n)} = D_{m_{k,i}+1}^{(n)} | M_k^{(n)} = M_i^{(n)} \right] = q_{k,i}^{(n)}, \tag{30}$$

$$Q\left[M_{k+1}^{(n)} = D_{m_{k,i}-1}^{(n)} | M_k^{(n)} = D_i^{(n)} \right] = 1 - q_{k,i}^{(n)}, \text{ where} \tag{31}$$

$$q_i^{(n)} = \frac{\mu_{rn}^V(t_{nk}, D_i^{(n)})\Delta t_n - (D_{m_{k,i}-1}^{(n)} - D_i^{(n)})}{D_{m_{k,i}+1}^{(n)} - D_{m_{k,i}-1}^{(n)}}, \text{ and} \tag{32}$$

$$m_{k,i} = \min\left\{ j \mid \{\mu_{rn}^V(t_{nk}, D_i^{(n)})\Delta t_n < D_j^{(n)} - D_i^{(n)} \right\} \tag{33}$$

Note that typically the drift $\mu_{rn}^V(t_{nk}, D_i^{(n)})\Delta t_n$ will be of smaller order than $D_{i+1}^{(n)} - D_i^{(n)} = g'\left(f(V_0) + i\sqrt{\Delta t_n} \right) \sqrt{\Delta t_n} + \mathcal{O}(\Delta t_n)$; therefore typically $m_i = i$. The general case is necessary, however, at grid points for which g' itself is small in order to ensure[14] $p_i^{(n)} \in (0, 1)$.

[13] The Markov chain will be time-homogeneous if and only if η_1 is independent of t.
[14] For a detailed discussion we refer to Nelson and Ramaswamy (1990); for example they construct an example of a recombining binomial model that approximates the square root process which is used in the Heston (1993) model and explain why in that case sometimes $m_i \neq i$ is necessary.

Equation (33) defines in these exceptional cases the transition to be to two grid points that lie around $D_i^{(n)} + \mu_{rn}^V(t_{nk}, D_i^{(n)})\Delta t_n$. The Markov chain defines a process which we adopt as $(V_k^{(n)})_{k=0,\dots,n}$. Doing this for each n defines a sequence of processes $(V^{(n)})_n$. Note that each process corresponds to a recombining binomial tree with transition probability depending on where you are in the tree. We denote $\mathcal{M}_{1m} = (-m, m)$. We have

$$\begin{aligned} E\left[M_{k+1}^{(n)} - M_k^{(n)} \middle| M_k^{(n)} = D_i^{(n)} \right] &= D_{m_{k,i}+1}^{(n)} q_i^{(n)} + (D_{m_{k,i}-1}^{(n)} - D_i^{(n)})(1 - q_i^{(n)}) \\ &= D_{m_{k,i}-1}^{(n)} + q_i^{(n)}(D_{m_{k,i}+1}^{(n)} - D_{m_{k,i}-1}^{(n)}) - D_i^{(n)} \\ &= \mu_{rn}^V(t_{nk}, D_i^{(n)})\Delta t_n. \end{aligned}$$

Therefore the first condition in theorem 4 is always fulfilled for the volatility process, i.e. the volatility drift converges on $\mathcal{M}_{1m}$, i.e. for $(V_k^{(n)} \in \mathcal{M}_m$ with $t_{nk} \xrightarrow{n} t$, $V_k^{(n)} \xrightarrow{n} V$ we have

$$\frac{1}{\Delta t_n} E_Q\left[V_{k+1}^{(n)} - V_k^{(n)} \middle| V_k^{(n)} \right] \xrightarrow{n} \mu_{rn}^V(t_{nk}, D_i^{(n)}). \tag{34}$$

We also have

$$\begin{aligned} & E\left[\left(M_{k+1}^{(n)} - D_{m_{k,i}}^{(n)} \right)^2 \middle| M_k^{(n)} = D_i^{(n)} \right] \\ &= (D_{m_{k,i}+1}^{(n)} - D_{m_{k,i}}^{(n)})^2 q_i^{(n)} + (D_{m_{k,i}-1}^{(n)} - D_{m_{k,i}}^{(n)})^2 (1 - q_i^{(n)}) \\ &= \left(g'\left(f(V_0) + i\sqrt{\Delta t_n} \right) \right)^2 \Delta t_n q_i^{(n)} + \left(g'\left(f(V_0) + i\sqrt{\Delta t_n} \right) \right)^2 \Delta t_n (1 - q_i^{(n)}) + \mathcal{O}(\Delta t_n^{3/2}) \\ &= \left(g'\left(f(V_0) + i\sqrt{\Delta t_n} \right) \right)^2 \Delta t_n + \mathcal{O}(\Delta t_n^{3/2}), \end{aligned}$$

which implies that

$$\begin{aligned} & Var\left(M_{k+1}^{(n)} - M_k^{(n)} \middle| M_k^{(n)} = D_i^{(n)} \right) \\ &= E\left[\left(M_{k+1}^{(n)} - D_{m_{k,i}}^{(n)} \right)^2 \middle| M_k^{(n)} = D_i^{(n)} \right] \\ &\quad - \left(E\left[M_{k+1}^{(n)} - D_i^{(n)} \middle| M_k^{(n)} = D_i^{(n)} \right] - \left(D_{m_{k,i}}^{(n)} - D_i^{(n)} \right) \right)^2. \end{aligned}$$

Note that when $m_{ki} = i$, we also have that conditions 1, 2 in theorem 4 is fulfilled. Therefore in the (typical) case convergence holds; we will not prove this here for the general case; instead we note that our application here is covered by theorem 3 in Nelson and Ramaswamy (1990); in addition to equation (34) they prove also that

(1) variance/covariances converge on $\mathcal{M}_{1m}$, i.e. for $(V_k^{(n)}, S_k^{(n)}) \in \mathcal{M}_m$ with $(t_{nk} \xrightarrow{n} t)$ $V_k^{(n)} \xrightarrow{n} V$ we have $\frac{1}{\Delta t_n} Var_Q\left[V_{k+1}^{(n)} - V_k^{(n)} \middle| V_k^{(n)}\right] \xrightarrow{n} \varphi^2(V)$,

(2) and jump-sizes vanish on $\mathcal{M}_m$ in the limit Q-a.s., i.e. $\max_{V_k^{(n)} \in \mathcal{M}_m} \left|V_{k+1}^{(n)} - V_k^{(n)}\right| \xrightarrow{n} 0$.

Nelson and Ramaswamy (1990) conclude that $V^{(n)} \overset{d}{\Longrightarrow} V$. We next construct the process $(S_k^{(n)})_{k=0,\ldots,n}$ by forward induction: First, note that when the volatility that enters into the stock process would be constant at $V_k^{(n)}$ over the time period $[t_{nk}, t_{n,k+1})$ in equation (28) then $S_{k+1}^{(n)}$ would be equal to

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

We use the Markov-chain $(M_k^{(n)})_k$ to define the Markov-chain ϵ_n by setting

$$\begin{aligned} \epsilon_{n,k} &= \frac{M_{k+1}^{(n)} - M_k^{(n)} - \mu_{rn}^V\left(t_{nk}, M_k^{(n)}\right)\Delta t_n}{\varphi(M_k^{(n)})}, \text{ i.e.} \\ M_{k+1}^{(n)} - M_k^{(n)} &= \mu_{rn}^V\left(t_{nk}, M_k^{(n)}\right)\Delta t_n + \varphi(M_k^{(n)}) \cdot \epsilon_{n,k}. \end{aligned}$$

and since $V^{(n)} \overset{d}{\Longrightarrow} V$ we therefore conclude immediately that

$$\sum_{k=1}^{\lfloor \frac{t}{\Delta t_n} \rfloor} \epsilon_{nk} \overset{d}{\Longrightarrow} W_{1t}^Q. \tag{35}$$

Note also that by equation (33)

$$\begin{aligned} \max\left\{D_{m_{k,i}+1}^{(n)} - D_i^{(n)}, D_i^{(n)} - D_{m_{k,i}-1}^{(n)}\right\} &= g'\left(f(V_0) + i\sqrt{\Delta t_n}\,\right)\sqrt{\Delta t_n} + \mathcal{O}(\Delta t_n) \\ &= \varphi(D_i^{(n)})\sqrt{\Delta t_n} + \mathcal{O}(\Delta t_n) \end{aligned}$$

uniformly on $\mathcal{M}_m$. Since $f'(V) = \frac{1}{\varphi(V)}$ and g is the inverse of f we have Q-a.s.

$$|\epsilon_{nk}| = \sqrt{\Delta t_n} + \mathcal{O}(\Delta t_n). \tag{36}$$

We then approximate $W_{1,t_{n,k+1}}^Q - W_{1,t_{nk}}^Q$ by ϵ_{nk}, approximate $W_{2,t_{n,k+1}}^Q - W_{2,t_{nk}}^Q$ by $Y_{nk,\epsilon_{nk}}$, and set, conditional on $V_k^{(n)} = D_i$,

$$
\begin{aligned}
S_{k+1}^{(n)} &= S_k^{(n)} \cdot \exp\left\{ \left(r - \frac{\psi^2(V_k^{(n)})}{2} \right) \Delta t_n + \psi(V_k^{(n)}) \cdot \left(\rho\epsilon_{nk} + \sqrt{(1-\rho^2)\Delta t_n} Y_{nk\epsilon_{nk}} \right) \right\} && (37) \\
&= 1_{\theta_{ki}=1} S_k^{(n)} Z_{nki1} + 1_{\theta_{ki}=2} S_k^{(n)} Z_{nki2}. && (38)
\end{aligned}
$$

where we denoted

$$
\begin{aligned}
Z_{n,k,i,1} &= \exp\left\{ \left(r - \frac{\psi^2(D_i^{(n)})}{2} \right) \Delta t_n + \rho \frac{\psi(D_i^{(n)})}{\varphi(D_i^{(n)})} \left(D_{m_{k,i}+1}^{(n)} - D_i^{(n)} - \mu_{rn}^V \left(t_{nk}, D_i^{(n)} \right) \Delta t_n \right) \right. \\
&\qquad \left. + \psi(D_i^{(n)}) \sqrt{(1-\rho^2)\Delta t_n}\ Y_{nk1} \right\}, && (39) \\
Z_{n,k,i,2} &= \exp\left\{ \left(r - \frac{\psi^2(D_i^{(n)})}{2} \right) \Delta t_n + \rho \frac{\psi(D_i^{(n)})}{\varphi(D_i^{(n)})} \left(D_{m_{k,i}-1}^{(n)} - D_i^{(n)} - \mu_{rn}^V \left(t_{nk}, D_i^{(n)} \right) \Delta t_n \right) \right. \\
&\qquad \left. + \psi(D_i^{(n)}) \sqrt{(1-\rho^2)\Delta t_n}\ Y_{nk2} \right\}, && (40)
\end{aligned}
$$

and

$$
\theta_k \sim \begin{cases} 1\ ;\ \text{if } \epsilon_{nk} > 0 \\ 2\ ;\ \text{otherwise} \end{cases}.
$$

In particular we have $Q[\theta_k = 1 | M_k^{(n)} = D_i^{(n)}] = q_i^{(n)}$ and $Q[\theta_{ki} = 2 | M_k^{(n)} = D_i^{(n)}] = 1 - q_i^{(n)}$. Based on equation (37) we calculate, since $E[\epsilon_{nk}] = 0$ and using equation (36) that

$$
\begin{aligned}
& E\left[S_{k+1}^{(n)} \middle| S_k^{(n)} \right] = E\left[E\left[S_{k+1}^{(n)} \middle| \epsilon_{nk} \right] \middle| S_k^{(n)} \right] \\
&= S_k^{(n)} E\left[\exp\left\{ \left(r - \frac{\psi^2(D_i^{(n)})}{2} + \frac{\psi^2(D_i^{(n)})}{2}(1-\rho^2) \right) \Delta t_n + \rho\psi(D_i^{(n)})\epsilon_{nk} \right\} \right] \\
&= S_k^{(n)} E\left[1 + \rho\psi(D_i^{(n)})\epsilon_{nk} + \left(r - \frac{\psi^2(D_i^{(n)})}{2}\rho^2 \right) \Delta t_n + \frac{\rho^2\psi^2(D_i^{(n)})}{2}\epsilon_{nk}^2 + \mathcal{O}(\Delta t_n^{3/2}) \right] \\
&= S_k^{(n)} \left(1 + r\Delta t_n + \mathcal{O}(\Delta t_n^{3/2}) \right).
\end{aligned}
$$

Together with the above this ensures the first condition in theorem 4. Similarly we calculate

$$
\begin{aligned}
&E\left[\left(S_{k+1}^{(n)}\right)^2\middle| S_k^{(n)}\right] = E\left[E\left[\left(S_{k+1}^{(n)}\right)^2\middle|\epsilon_{nk}\right]\middle| S_k^{(n)}\right] \\
&= \left(S_k^{(n)}\right)^2 E\left[\exp\left\{\left(2r - \psi^2(D_i^{(n)}) + 2\psi^2(D_i^{(n)})(1-\rho^2)\right)\Delta t_n + 2\rho\psi(D_i^{(n)})\epsilon_{nk}\right\}\right] \\
&= \left(S_k^{(n)}\right)^2 E\left[1 + 2\rho\psi(D_i^{(n)})\epsilon_{nk} + \left(2r + \psi^2(D_i^{(n)})(1-2\rho^2)\right)\Delta t_n + 2\rho^2\psi^2(D_i^{(n)})\epsilon_{nk}^2 + \mathcal{O}(\Delta t_n^{3/2})\right] \\
&= \left(S_k^{(n)}\right)^2 \cdot \left(1 + (2r + \psi^2(D_i^{(n)}))\Delta t_n + \mathcal{O}(\Delta t_n^{3/2})\right).
\end{aligned}
$$

Note that $\left(1 + r\Delta t_n + \mathcal{O}(\Delta t_n^{3/2})\right)^2 = 1 + 2r\Delta t_n + \mathcal{O}(\Delta t_n^{3/2})$ so that

$$
\begin{aligned}
Var\left(S_{k+1}^{(n)} - S_k^{(n)}\middle| S_k^{(n)}\right) &= \left(E\left[\left(S_{k+1}^{(n)}\right)^2\middle| S_k^{(n)}\right] - E\left[S_{k+1}^{(n)}\middle| S_k^{(n)}\right]^2\right) \\
&= \left(S_k^{(n)}\right)^2 \cdot \psi^2(D_i^{(n)}))\Delta t_n + \mathcal{O}(\Delta t_n^{3/2}),
\end{aligned}
$$

i.e. this proves the variance part of S in the theorem 4. It remains to check the covariance part: we have based on the above analysis and based on equation (35) that conditional on $V_k^{(n)}$ and $S_k^{(n)}$ that

$$
\begin{aligned}
&V_{k+1}^{(n)} \cdot S_{k+1}^{(n)} \\
&= \left(V_k^{(n)} + \mu_{rn}^V(t_{nk}, V_k^{(n)}) + \varphi(V_k^{(n)})\epsilon_{nk}\right) \\
&\quad \cdot S_k^{(n)}\left(1 + \rho\psi(V_k^{(n)})\epsilon_{nk} + \left(r - \frac{\psi^2(V_k^{(n)})}{2}\rho^2\right)\Delta t_n + \frac{\rho^2\psi^2(V_k^{(n)})}{2}\epsilon_{nk}^2 + \mathcal{O}(\Delta t_n^{3/2})\right) \\
&= S_k^{(n)}\Bigg\{\left(V_k^{(n)} + \mu_{rn}^V(t_{nk}, V_k^{(n)})\right) \\
&\qquad \cdot \left(1 + \rho\psi(V_k^{(n)})\epsilon_{nk} + \left(r - \frac{\psi^2(V_k^{(n)})}{2}\rho^2\right)\Delta t_n + \frac{\rho^2\psi^2(V_k^{(n)})}{2}\epsilon_{nk}^2\right) \\
&\qquad + \varphi(V_k^{(n)})\epsilon_{nk} \cdot \left(1 + \left(r - \frac{\psi^2(V_k^{(n)})}{2}\rho^2\right)\Delta t_n\right) \\
&\qquad + \varphi(V_k^{(n)})\rho\psi(V_k^{(n)})\epsilon_{nk}^2 + \varphi(V_k^{(n)})\frac{\rho^2\psi^2(V_k^{(n)})}{2}\epsilon_{nk}^3 + \mathcal{O}(\Delta t_n^{3/2})\Bigg\}
\end{aligned}
$$

which implies since $E[\epsilon_{nk}] = 0$, $E[\epsilon_{nk}^2] = \mathcal{O}(\Delta t_n)$, $E[\epsilon_{nk}^3] = \mathcal{O}(\Delta t_n^{3/2})$ that

$$
\begin{aligned}
&E\left[S_{k+1}^{(n)}V_{k+1}^{(n)}\middle| V_k^{(n)}, S_k^{(n)}\right] \\
&= S_k^{(n)}\left\{\left(V_k^{(n)} + \mu_{rn}^V(t_{nk}, V_k^{(n)})\right)\cdot(1 + r\Delta t_n) + \rho\varphi(V_k^{(n)})\psi(V_k^{(n)})\Delta t_n + \mathcal{O}(\Delta t_n^{3/2})\right\}.
\end{aligned}
$$

Therefore, since $E[V_{k+1}^{(n)}|V_k^{(n)}] = V_k^{(n)} + \mu_{rn}^V(t_{nk}, V_k^{(n)})$,

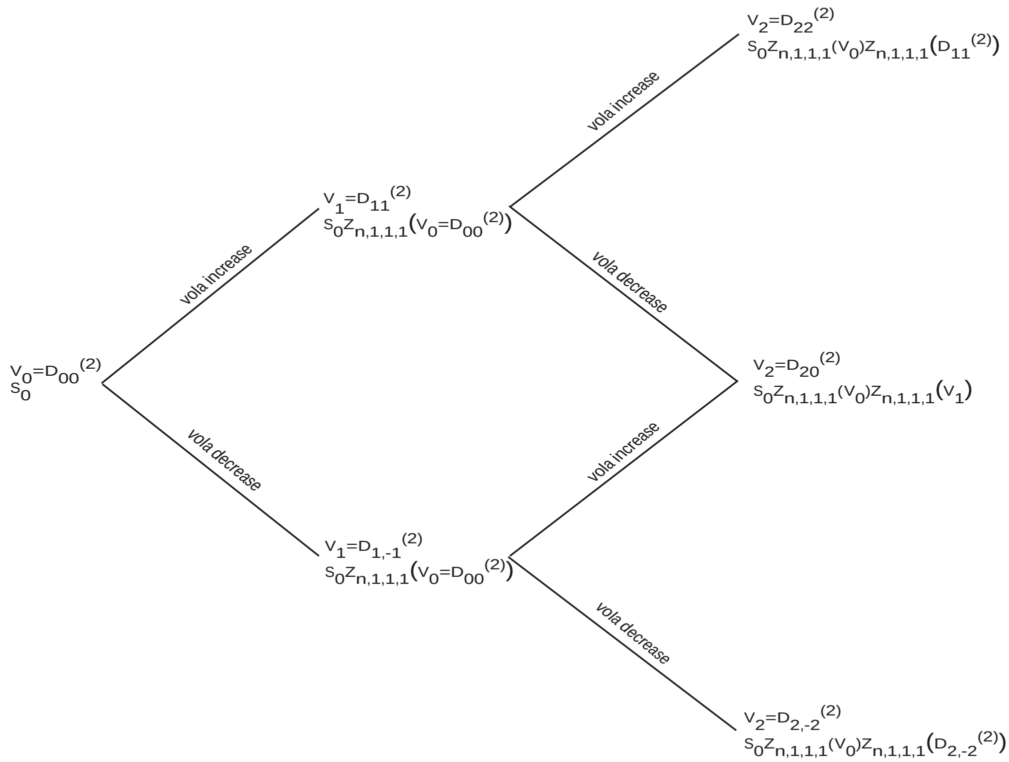

Fig. 2. The volatility dynamics and a description of the lognormal transition variables at each grid point.

$$Cov\left(S_{k+1}^{(n)}, V_{k+1}^{(n)} \middle| V_k^{(n)}, S_k^{(n)}\right) = \rho\varphi(V_k^{(n)})\psi(V_k^{(n)})\Delta t_n + \mathcal{O}(\Delta t_n^{3/2}),$$

which directly implies the covariance condition in theorem 4. Note that jump sizes in S are limited[15] on $\mathcal{M}_m$ and so we conclude based on theorem 4:

Theorem 5 *The sequence of processes* $(S_k^{(n)})_{k=0,\dots,n}$ *converges weakly to the process* $(S_t)_{t\in[0,T]}$ *of equations (27, 28).*

It remains to write this as an MLD. Before doing so let us take a look at the structure of our construction: Figure 2 describes the volatility tree for $n = 2$ over the two periods. Over the first period volatility can increase to $V_1 = D_{1,1}^{(n)}$ or decrease to $V_1 = D_{1,-1}^{(n)}$. Over the second period volatility can increase or decrease; note that by construction a decrease followed by an increase leads to the same volatility $D_{2,0}^{(n)}$ as an increase followed by a decrease, i.e. the volatility tree is recombining.

[15] Those of the volatility process are covered by Nelson and Ramaswamy (1990).

The figure also depicts at each node the two lognormal variables that will be mixed over each period; note that this depends on the volatility at the previous period. While the volatility tree is recombining the mixing lognormal random variables depend on the volatility path chosen, i.e. a volatility decrease followed by an increase will lead to a stock price determined by $S_0 \cdot Z_{2,1,1,1}(V_0) \cdot Z_{2,1,1,1}(D^{(n)}_{1,1})$ and this is *not* equal in distribution to the stock price determined by $S_0 \cdot Z_{2,1,1,1}(V_0) \cdot Z_{2,1,1,1}(D^{(n)}_{1,1})$ when we see a volatility increase followed by a decrease.

Let us denote $J^{(n)}_k$ the index j for which $V^{(n)}_k = D^{(n)}_j$. Note that this is a process and describes the evolution of the volatility process on the grid. We can then write using the definition of equation (38)

$$S^{(n)}_n = S_0 \prod_{k=1,\dots,n} Z_{n,k,J^{(n)}_k,\theta_k}.$$

Note that this is equal in distribution to all possible paths

$$S^{(n)}_n \stackrel{d}{=} S_0 \sum_{i_1,\dots,i_n \in \{1,2\}} \prod_{k=1,\dots,n} 1_{\theta_k = i_k} \prod_{k=1}^{n} X_{n,k,\tilde{J}^{(n)}_{k-1},i_k}$$

where $\tilde{J}^{(n)}_k = -k + \sum_{j=0}^{k} i_j$ denotes the current grid node. Here the first product term describes that we condition on the path of volatility and the second product term describes a lognormal random variable since it is a product of exponentials of normal random variables and can therefore be written as a single random variable which is the exponential of an appropriately defined normal random variable. Therefore the distribution of $S^{(n)}_n$ at date n is a mixture of 2^n lognormal appropriately chosen lognormal distributions that depend on the volatility path. We derive directly from Theorem 5:

Theorem 6 *The sequence of mixed lognormal distributions $S^{(n)}_n$ converges weakly to S_T.*

According to table 3 the Hull-White model is given by $\varphi(V) = \sigma V, \psi(V) = \sqrt{V}$. In the following discussion we assume that $\nu = 0$ and $\eta_1 = 0$. The risk-neutral process is then

$$dV_t = -\kappa V_t dt + \sigma V_t dW_{1t}^Q, dS_t = rS_t dt + \sqrt{V_t} S_t dW_{2t}^Q.$$

In particular V_T is a lognormal distributed random variable, i.e. we have $\ln(V_T/V_0) \sim \mathcal{N}\left(\left(-\kappa - \frac{\sigma^2}{2}\right) t; \sigma\sqrt{t}\ \right)$. We have $f(x) = \ln(x/\sigma)$ and $g(z) = \exp(\sigma z)$; therefore our grid points are

$$D_i^{(n)} = V_0 \cdot \exp\left(i\sigma\sqrt{\Delta t_n}\ \right).$$

We recall the definition in equation (27) and note that under our assumptions here $\mu_{rn}^V(t, V) = -\kappa V$. For sufficiently large n we have $\mathcal{M}_{2m} = (-m, m)$ that $\mu_{rn}^V(t, D_i^{(n)})\Delta t_n < D_i^{(n)}/\sigma\sqrt{\Delta t_n} + \mathcal{O}(\Delta t_n) = g'\left(i\sqrt{\Delta t_n}\ \right)\sqrt{\Delta t_n} + \mathcal{O}(\Delta t_n) = D_{i+1}^{(n)} - D_i^{(n)}$. In the sequel we will always take n large enough and therefore we can always assume that $m_{k,i} = i$. Also we calculate from equation (32)

$$q_i^{(n)} = \frac{D_i^{(n)} - \kappa D_i^{(n)}\Delta t_n - D_{i-1}^{(n)}}{D_{i+1}^{(n)} - D_{i-1}^{(n)}}.$$

Note that $D_i^{(n)} - \kappa D_i^{(n)}\Delta t_n \approx D_i^{(n)} \exp(-\kappa\Delta t_n)$ and so corresponds to the standard binomial approach. In particular this is time-homogeneous.

We approximate $\frac{D_{i+1}^{(n)} - D_i^{(n)}}{\varphi(D_i^{(n)})} = \exp(\sigma\sqrt{\Delta t_n}\ ) - 1$ by $\sqrt{\Delta t_n}$ and according to equations (39, 40) we take the mixing random variables

$$X_{n,k,i,1} = \exp\left\{\left(r - \frac{D_i^{(n)}}{2}\right)\Delta t_n + \rho\frac{D_{i+1}^{(n)} - D_i^{(n)} - \kappa D_i^{(n)}\Delta t_n}{\sigma\sqrt{D_i^{(n)}}} + \sqrt{(1-\rho^2)D_i^{(n)}\Delta t_n}\ Y_{nk1}\right\},$$
$$X_{n,k,i,2} = \exp\left\{\left(r - \frac{D_i^{(n)}}{2}\right)\Delta t_n + \rho\frac{D_{i-1}^{(n)} - D_i^{(n)} - \kappa D_i^{(n)}\Delta t_n}{\sigma\sqrt{D_i^{(n)}}} + \sqrt{(1-\rho^2)D_i^{(n)}\Delta t_n}\ Y_{nk2}\right\},$$

since $\frac{\psi(V)}{\varphi(V)} = \frac{1}{\sigma V}$. According to theorem 6 we have $S_n^{(n)} \stackrel{d}{\Longrightarrow} S_T$.

Note that we know from before

$$S_n^{(n)} \stackrel{d}{=} S_0 \sum_{i_1,\dots,i_n \in \{1,2\}} \prod_{k=1,\dots,n} 1_{\theta_k = i_k} \prod_{k=1}^{n} X_{n,k,\tilde{J}_{k-1}^{(n)},i_k}$$

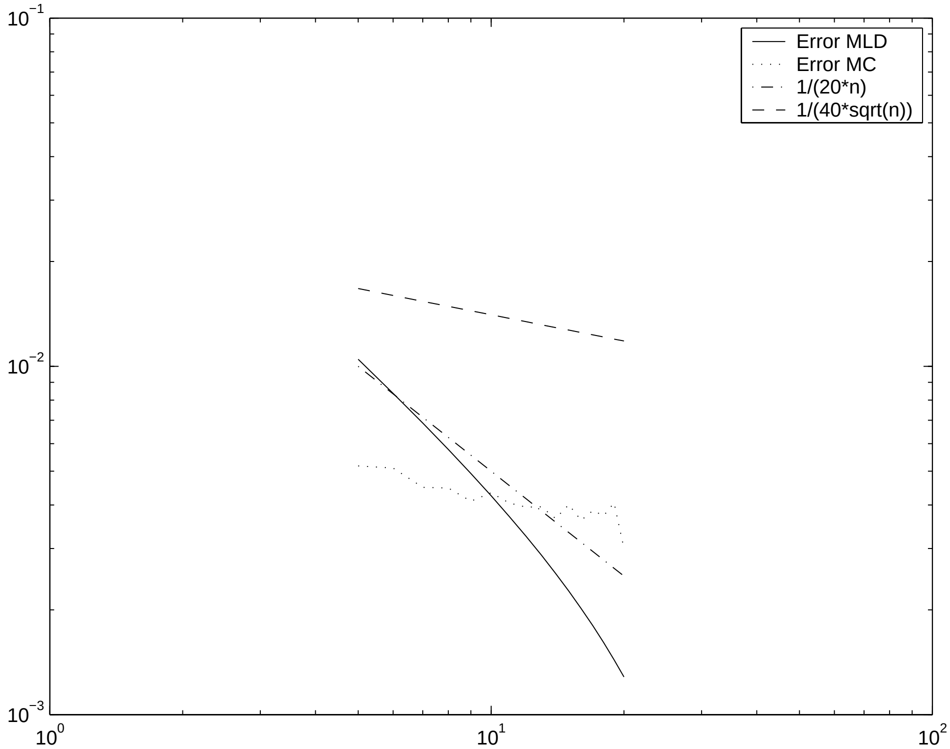

Fig. 3. Errors for MLC and Monte-Carlo depending on the refinement.

For any path $i_1, \ldots, i_n \in \{1, 2\}$ we can write the second product term as

$$\begin{aligned}
&\prod_{k=1}^{n} X_{n,k,\tilde{J}_{k-1}^{(n)},i_k} \\
&= \exp\left\{ rT - \frac{\sum_{k=1}^{n-1} D_{\tilde{J}_{k-1}^{(n)}}^{(n)}}{2} \Delta t_n + \rho \sum_{k=1}^{n-1} \left(\frac{D_{\tilde{J}_k^{(n)}}^{(n)} - D_{\tilde{J}_{k-1}^{(n)}}^{(n)} - \kappa D_{\tilde{J}_{k-1}^{(n)}}^{(n)} \Delta t_n}{\sigma \sqrt{D_{\tilde{J}_{k-1}^{(n)}}^{(n)}}} \right) \right. \\
&\qquad \left. + \sqrt{(1-\rho^2) \sum_{k=1}^{n-1} D_{\tilde{J}_{k-1}^{(n)}}^{(n)} \Delta t_n Y_{nk3}} \right\},
\end{aligned}$$

where Y_{nk3} is a sequence of iid standard normal random variables.

Figure 3 assesses the accuracy of our approximation when the stock and volatility process are uncorrelated, i.e. $\rho = 0$. Then we can write the price of the option as

$$\frac{1}{B_T} E_Q[f(S_T)] = \frac{1}{B_T} E_Q[E_Q[f(S_T)|V_T]] = E_Q\left[BS(S_0, K, r, \sqrt{V_T}\ )\right]. \tag{41}$$

This allows us to calculate prices as a numerical integration over Black-Scholes prices weighted by the normal density function. Figure 3 compares the errors that result from calculating

	$\rho = -0.5$		
	$K = 90$	$K = 100$	$K = 110$
$n = 5$	16.5068	11.1214	7.1498
$n = 10$	16.6812	11.3077	7.3254
$n = 15$	16.7671	11.3994	7.4120

	$\rho = 0$		
	$K = 90$	$K = 100$	$K = 110$
$n = 5$	16.4041	11.2264	7.4577
$n = 10$	16.5777	11.4213	7.6519
$n = 15$	16.6632	11.5172	7.7475

	$\rho = 0.5$		
	$K = 90$	$K = 100$	$K = 110$
$n = 5$	16.2725	11.3075	7.7382
$n = 10$	16.4378	11.5048	7.9447
$n = 15$	16.5193	11.6020	8.0465

Table 4
Option prices calculated our approach varying refinement n

that price using MLD and Monte-Carlo approach. We plot the errors depending on the refinement; for the Monte-Carlo approach we took that as the time-discretization and simulated 1000000 paths. We see that initially the MLD is less accurate but quickly catches up; it seems that MLD converges with order one in the refinement while Monte-Carlo only converges with the slower order 1/2. (The MLD line is not straight; the price calculation based on equation (41) used step-sizes of 0.00001 step sizes and was implemented in MATLAB; this did not improve by decreasing step-sizes; this indicates that we are beyond the accuracy of the price calculation based on equation (41).)

We also find that to achieve the same level of accuracy the MLD needs less computing time. We conclude that our MLD method is an accurate method to price derivatives. For comparative purposes we present in table 4 we take $v_0 = 0.3^2$; $\kappa = 0$; $\sigma = 0.5$; $S_0 = 100$; $r = 0$; $T = 1$ and vary $K = 90, 100, 110$ and $n = 5, 10, 15$.

5 Conclusion

This paper constructed approximations of the Black-Scholes setup with jumps (Merton model) and of the Black-Scholes with stochastic volatility model using MLD. We explained how to use them to calculate option prices within these models and discussed efficiency and accuracy.

A Calculating The First Three Moments

We assume here X_i as in defnition 1 and $d_{ij}(K)$ as defined in the paragraphs thereafter. Since only the ratio A_0 to matters we assume $A_0 = 1$ to simplify notation. Calculations of the necessary terms will make use of the following:

Lemma 7 *For a normal distributed random variable Z with $E[Z] = 0$, $Var(Z) = \sigma^2$ and $\mu \in \mathbb{R}$ we define the random variable $\frac{d\tilde{Q}}{dQ} = \exp\left(\mu Z - \frac{1}{2}\mu^2\sigma^2\right)$ and a new probability measure $\tilde{Q}$ on $(\Omega, \mathcal{F})$ by setting $\tilde{Q}[A] = E[1_A \frac{d\tilde{Q}}{dQ}]$ for $A \in \mathcal{F}$. (The expectation operator will be denoted $E_{\tilde{Q}}[\cdot]$.) Then the random variable $\bar{Z}$ has mean $\mu\sigma^2$ and variance σ^2 under $\tilde{Q}$, i.e. $Z - \mu\sigma^2$ has mean zero and variance σ^2 under $\tilde{Q}$.*

PROOF. The density of $\tilde{Q}$ (w.r.t. the Lebesgue measure λ on $\mathbb{R}$) is

$$\begin{aligned} \frac{d\tilde{Q}}{d\lambda} &= \frac{d\tilde{Q}}{dQ}\frac{dQ}{d\lambda} = \exp\left(\mu z - \frac{1}{2}\mu^2\sigma^2\right) \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{z^2}{2\sigma^2}\right) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{z^2 - 2\mu\sigma^2 z + \mu^2\sigma^4}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z-\mu\sigma^2)^2}{2\sigma^2}\right). \end{aligned}$$

To calculate expectations note first that $E[X_i] = \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right)$ and

$$\begin{aligned} E\left[X_i^2\right] &= \exp\left(2\mu_i + \frac{(-2\sigma_i)^2}{2}\right) \cdot E\left[\exp\left(-\frac{(-2\sigma_i)^2}{2} + 2Z\right)\right] = \exp\left(2\mu_i + 2\sigma_i^2\right), \\ E\left[X_i^3\right] &= \exp\left(3\mu_i + \frac{9}{2}\sigma_i^2\right) \cdot E\left[\exp\left(-\frac{(-3\sigma_i)^2}{2} + 3Z\right)\right] = \exp\left(3\mu_i + \frac{9}{2}\sigma_i^2\right). \end{aligned}$$

To calculate truncated moments note that standard arguments used in option pricing theory tell us immediately that $E\left[X_i 1_{X_i \geq K_i}\right] = \exp\left(\mu_i + \frac{\sigma_i^2}{2}\right)\Phi_n(d_{i1})$ and $E\left[1_{X_i \geq K_i}\right] = \Phi_n(d_{i2})$. Also,

$$\nu_{i2}(K) = E\left[X_i^2 1_{X_i \geq K_i}\right] = \exp(2\mu_i + 2\sigma_i^2) \cdot E\left[\exp\left(-\frac{(2\sigma_i)^2}{2} + 2Z\right) \cdot 1_{X_i \geq K_i}\right].$$

Using lemma 7 we interpret the term $\exp\left(-\frac{(2\sigma_i)^2}{2} + 2Z\right)$ as the density of a new probability measure Q_{i3} and find that under this measure $Z - 2\sigma_i^2$ is a normal distributed random variable with mean 0 and variance σ_i^2. Therefore $E_{Q_{i3}}\left[1_{X_i \geq K_i}\right] = \Phi_n(d_{i3})$ and so $\nu_{i2}(K) = \exp(2\mu_i + 2\sigma_i^2) \cdot \Phi_n(d_{i3})$. It remains to calculate

$$\nu_{i3}(K) = E\left[X_i^3 \cdot 1_{X_i \geq K_i}\right] = \exp\left(3\mu_i + \frac{9}{2}\sigma_i^2\right) \cdot E\left[\exp\left(-\frac{(3\sigma_i)^2}{2} + 3Z\right) \cdot 1_{X_i \geq K_i}\right]$$

Using lemma 7 we interpret the term $\exp\left(-\frac{(3\sigma_i)^2}{2} + 3Z\right)$ as the density of a new probability measure Q_{i4} under which $Z - 3\sigma_i^2$ is a normal distributed random variable with mean zero and variance σ_i^2. Therefore $E_{Q_{i1}}\left[1_{X_i \geq K_i}\right] = \Phi_n(d_{i4})$ and so $E\left[X_i^3 1_{X_i \geq K_i}\right] = \exp\left(3\mu_i + \frac{9}{2}\sigma_i^2\right) \cdot E_{Q_{i1}}\left[1_{X_i \geq K_i}\right] = \exp\left(3\mu_i + \frac{9}{2}\sigma_i^2\right) \cdot \Phi_n(d_{i4})$.

References

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

Bjork, T. (1999): *Arbitrage Theory in Continuous Time*. Oxford University Press.

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

Brémaud, P. (1981): *Point Processes and Queues: Martingale Dynamics*. Springer–Verlag.

Brigo, D., and F. Mercurio (2000): "Lognormal-mixture dynamics and calibration to market volatility smiles," *Working paper.*

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.

Duan, J.-C., and J.-G. Simonato (2000): "American Option Pricing under GARCH by a Markov Chain Approximation," Discussion paper, to appear: Journal of Economic Dynamics and Control.

Duffie, D. (1992): *Dynamic Asset Pricing Theory*. Princeton University Press, first edn.

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

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.

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.

He, H. (1990): "Convergence from Discrete– to Continuous–Time Contingent Claims Prices," *The Review of Financial Studies*, 3(4).

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.

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.

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

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.

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.

Samuelson, P. A. (1965): "Rational Theory of Warrant Pricing," *Industrial Management Review*, 6, 13–31.

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