

# A DISCRETE VALUATION OF SWING OPTIONS

ALI LARI-LAVASSANI, MOHAMADREZA SIMCHI, AND ANTONY WARE

ABSTRACT. A discrete forest methodology is developed for swing options as a dynamically coupled system of European options. Numerical implementation is fully developed for one- and two-factor, mean-reverting, underlying processes, with application to energy markets. Convergence is established via finite difference methods. Qualitative properties and sensitivity analysis are considered.

## 1. INTRODUCTION

Swing contracts have traditionally been used in the natural gas industry to provide hedging via the introduction of limited dynamic flexibility in the quantity of gas acquired. More generally this type of option is of value in any market where the physical transfer of the underlying asset must take place through interconnected networks, and is thus subject to volume constraints. This is the case for natural gas and pipelines, electricity and cable based telecommunications and their transmission lines, wireless telecommunications and their bandwidths. Going beyond issues of trading, dynamic risk management of real assets such as these complex interconnected networks could benefit from the methodology presented here.

Our aim in writing this paper is two-fold. First, in spite of their important applications, little fundamental work seems to have been done in modeling and valuing swing options; the few available published papers are qualitative in nature and do not provide a mathematical model or a solid ground for numerical implementation. Secondly, mean-reverting processes are increasingly being used to model the price processes of commodities and energy; we provide here a unified treatment of the numerical implementation of discrete mean-reverting processes.

The swing option can be viewed as a generalization of a Bermudan option, due to its multiple exercise feature. We model it as a dynamically coupled system of European options, and opt for a discrete forest model, which results in a realistic and practical numerical algorithm. We limit ourselves to modeling a swing option on one given underlying asset. Generalization to the case of multiple assets

---

*Date:* July 11, 2000.

*1991 Mathematics Subject Classification.* Primary 91B28, 91-08; Secondary 65C.

*Key words and phrases.* Option Pricing, Swing Option, Energy Commodity Modeling and Derivatives, Mean-reverting Models, Two-Factor Model, Forest, Numerical Methods, Finite Difference.

A. Lari-Lavassani is partially supported by a grant from the University of Calgary and NSERC. M. Simchi is supported by grants from the University of Calgary and TransAlta Energy Marketing Corp (TEM). A. Ware is supported by MITACS and a grant from TEM. All three authors are grateful to Dan Calistrate for many constructive discussions, to Shaun Hatch of TEM for introducing them to this problem and providing valuable insight and support, and to Tom Coyle from TEM for supporting the entire project.

This paper is in final form and no version of it will be submitted for publication elsewhere.

is straightforward but leads to challenging computational problems due to high dimensionality. One promising avenue would be to investigate numerical implementation on parallel computer. We do however allow the single underlying asset to follow a multi-factor stochastic process. A distinctive characterization of the methodology developed here is its abstract and general nature, which makes it applicable to any discrete modeling of the underlying asset, in particular binomial or trinomial trees, or indeed trees with any number of branching jumps. This abstract setting lends itself to a European style of exercise of the swing rights. We fully develop numerical implementations of swing options on mean-reverting underlying processes via binomial trees.

The paper is written in a self contained manner and full arguments are presented to the extent possible. It is organized as follows: Section 2 discusses some classical examples of one- and two-factor mean-reverting linear stochastic differential equations. These model energy price or interest rate processes, and throughout this work, they will be used to illustrate the methods and results. Section 3 presents a general abstract look at discrete modeling of stochastic processes. Section 4 investigates numerical implementations of mean-reverting processes on binomial trees. Section 5 describes a general swing option contract. Section 6 develops a mathematical model for the swing option via a forest methodology. Section 7 is devoted to numerical investigation of convergence, hedge parameters and comparison with a basket of American options. Section 8 proves the convergence of the binomial tree swing option valuation, showing that the upper bound on the error is inversely proportional to the number of time steps in the tree.

## 2. UNDERLYING PROCESSES AND APPLICATIONS TO ENERGY MODELS

The general setting for swing options developed in Section 6 is quite abstract and applies to all discrete processes. To make ideas more transparent we first introduce in this section one- and two-factor models historically used for short term interest rates, and which are now finding applications in the energy commodity markets as in [S, 1998], [P, 1997], [HR, 1998], [JRT, 1998]. When using these models to price options, it is implicitly assumed that they reflect the *risk neutralized* price processes, perhaps via the incorporation of a model of the cost-of-carry or the market-price-of-risk functions, see [H, 1999]. This will be the working assumption for the remainder of this paper.

A general diffusion one-factor model is:

$$(2.1) \quad dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dZ_t,$$

where  $\mu$  and  $\sigma$  are  $C^2$  differentiable and satisfy the usual linear growth conditions to guarantee the existence of solutions of the above stochastic differential equation (see [KP, 1999] Theorem 4.5.3, p. 131), and  $dZ_t$  is a standard Brownian motion,  $\mathbb{E}(dZ_t) = 0$ , and  $\mathbb{E}((dZ_t)^2) = dt$ . We will consider two classical particular cases:

$$(2.2) \quad dS_t = \mu S_t dt + \sigma S_t dZ_t$$

$$(2.3) \quad dS_t = \alpha(L - S_t) dt + \sigma S_t dZ_t,$$

where  $\mu$ ,  $\alpha$ ,  $\sigma$  and  $L$  are constants. The following system of stochastic linear differential equations is a possible two-factor model for energy and commodity spot

price processes:

$$(2.4) \quad \begin{aligned} dS_t &= \alpha(L_t - S_t) dt + \sigma S_t dZ_t \\ dL_t &= \mu L_t dt + \xi L_t dW_t, \end{aligned}$$

where  $\alpha$ ,  $\mu$ ,  $\sigma$  and  $\xi$  are constants and  $dZ_t$  and  $dW_t$  are uncorrelated standard Brownian motions.

Ignoring the stochastic parts in the mean-reverting equations (2.3) and (2.4) yields the deterministic differential equations governing the time evolution of the mean of each process:

$$(2.5) \quad \frac{d\mathbb{E}(S_t)}{dt} = \alpha(L - \mathbb{E}(S_t))$$

$$(2.6) \quad \frac{d}{dt} \begin{pmatrix} \mathbb{E}(S_t) \\ \mathbb{E}(L_t) \end{pmatrix} = \begin{pmatrix} -\alpha & \alpha \\ 0 & \mu \end{pmatrix} \begin{pmatrix} \mathbb{E}(S_t) \\ \mathbb{E}(L_t) \end{pmatrix}.$$

We shall establish that the orbits associated with the  $\mathbb{E}(L_t)$  equations are attracting invariant spaces for the flows of these deterministic stochastic differential equations. This justifies the use of the terms *long-run mean* and *mean-reverting*. For more on the notions of dynamical systems introduced in the next Proposition and its proof we refer to [R, 1995] Chapter 5.

**Proposition 1.** *For the process  $S_t$ , given by (2.3),  $L$  is the stable equilibrium of the hyperbolic differential equation governing (2.5) if  $\alpha > 0$ . For the process  $S_t$ , given by (2.4), the orbit of  $\mathbb{E}(L_t)$  is the unstable or stable manifold of the hyperbolic differential equation (2.6). This is the unstable manifold when  $\alpha > 0$ , and  $\mu > 0$ , which is the case for the financial applications we are considering.*

*Proof.* Let  $S_0$  be the value of  $S_t$  at  $t = t_0$ .  $\mathbb{E}(S_t) = L$  is the equilibrium of (2.5). Since the eigenvalue  $-\alpha < 0$ , stability follows. Therefore as  $t \rightarrow \infty$ , the solution curves of  $\mathbb{E}(S_t)$  decay exponentially to  $L$ . Denote the two-by-two matrix of (2.6) by  $A$ , and the value of  $L_t$  at  $t = t_0$  by  $L_0$ . This system is hyperbolic since the real part of the eigenvalues  $-\alpha$  and  $\mu$  of  $A$  are nonzero. Furthermore, the equilibrium at the origin is a saddle since  $-\alpha < 0$ , and  $\mu > 0$ . The stable and unstable manifolds of this hyperbolic equilibrium are the corresponding eigenspaces and, since the system is linear, they are global [R, 1995] Theorem 6.1, p. 111. They are respectively the lines,  $\{(1, 0) e^{-\alpha(t-t_0)}, t \in \mathbb{R}\}$  and  $\{(\frac{\alpha}{\mu+\alpha}, 1) L_0 e^{\mu(t-t_0)}, t \in \mathbb{R}\}$  in the phase space  $(\mathbb{E}(S_t), \mathbb{E}(L_t))$ . The solution to the above linear system is given by (see [R, 1995] Proposition 3.1, p. 97)

$$\begin{pmatrix} \mathbb{E}(S_t) \\ \mathbb{E}(L_t) \end{pmatrix} = e^{(t-t_0)A} \begin{pmatrix} S_0 \\ L_0 \end{pmatrix} = \begin{pmatrix} S_0 e^{-\alpha(t-t_0)} + \frac{\alpha}{\alpha+\mu} L_0 (e^{\mu(t-t_0)} - e^{-\alpha(t-t_0)}) \\ L_0 e^{\mu(t-t_0)} \end{pmatrix}$$

which can be decomposed in term of the direct sum of the stable and unstable manifolds as

$$\left( S_0 - \frac{\alpha}{\alpha + \mu} L_0 \right) e^{-\alpha(t-t_0)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + L_0 e^{\mu(t-t_0)} \begin{pmatrix} \frac{\alpha}{\alpha + \mu} \\ 1 \end{pmatrix}.$$

Dynamically speaking, in the phase space, any trajectory  $t \rightarrow (\mathbb{E}(S_t), \mathbb{E}(L_t))$ , off the stable or unstable manifolds, tends in forward time as  $t \rightarrow \infty$ , to  $L_0 e^{\mu(t-t_0)} \begin{pmatrix} \frac{\alpha}{\mu+\alpha} \\ 1 \end{pmatrix}$  which is the unstable manifold.  $\square$

Note that we used the notation  $\mathbb{E}(S_t)$  to denote the conditional expectation  $\mathbb{E}(S_t|S_{t_0})|_t$ . We next gather and present results on the first moments of the above processes:

**Proposition 2.** *a) Consider the process,  $dS_t = (\mu S_t + b)dt + \sigma S_t dZ_t$  over the time interval  $[t_i, t_k]$ , where  $\mu$ ,  $b$  and  $\sigma$  are constants and let  $S_{t_i} = S_i$ . We make the generic non-degeneracy assumptions:*

$$(H1) \quad \sigma^2 + \mu \neq 0, 2\mu + \sigma^2 \neq 0 \text{ and } \mu \neq 0.$$

*Then the first and second moments of  $S_t$  are given by:*

$$\begin{aligned} \mathbb{E}(S_t|S_i)|_{t_k} &= -\frac{b}{\mu} + (S_i + \frac{b}{\mu})e^{\mu(t_k-t_i)}, \\ \mathbb{E}(S_t^2|S_i)|_{t_k} &= c e^{(2\mu+\sigma^2)(t_k-t_i)} - \frac{2b(S_i + \frac{b}{\mu})}{\sigma^2 + \mu} e^{\mu(t_k-t_i)} + \frac{2b^2}{\mu(2\mu + \sigma^2)}, \end{aligned}$$

where

$$c = S_i^2 - \frac{2b(S_i + \frac{b}{\mu})}{\sigma^2 + \mu} + \frac{2b^2}{\mu(2\mu + \sigma^2)}.$$

*b) Consider the process given by (2.4) over the time interval  $[t_i, t_k]$ , and let  $S_{t_i} = S_i$ ,  $L_{t_i} = L_i$ . We make the generic non-degeneracy assumptions:*

$$(H2) \quad \mu + \alpha - \sigma^2 \neq 0 \neq (2\mu + \xi^2 - \sigma^2 + 2\alpha) \neq 0 \neq (\mu + \xi^2 + \alpha) \neq 0$$

*Then the first and second moments are given by:*

$$\begin{aligned} \mathbb{E}(S_t|S_i)|_{t_k} &= S_i e^{-\alpha(t_k-t_i)} + \frac{\alpha}{\alpha + \mu} L_i (e^{\mu(t_k-t_i)} - e^{-\alpha(t_k-t_i)}), \\ \mathbb{E}(L_t|L_i)|_{t_k} &= L_i e^{\mu(t_k-t_i)}, \\ \mathbb{E}((S_t, S_t)|S_i)|_{t_k} &= S_i^2 \varphi_1(t_k - t_i) + \frac{2\alpha}{\Psi} S_i L_i (\varphi_1(t_k - t_i) - \varphi_2(t_k - t_i)) \\ &\quad + 2\alpha^2 L_i^2 \left( \frac{-\varphi_1(t_k - t_i)}{\Psi\Gamma} + \frac{\varphi_2(t_k - t_i)}{\Psi\Theta} + \frac{\varphi_3(t_k - t_i)}{\Theta\Gamma} \right), \\ \mathbb{E}((S_t, L_t)|(S_i, L_i))|_{t_k} &= \varphi_2(t_k - t_i) S_i L_i + \frac{\alpha}{\Theta} L_i^2 (\varphi_3(t_k - t_i) - \varphi_2(t_k - t_i)) \\ \mathbb{E}((L_t, L_t)|L_i)|_{t_k} &= L_i^2 \varphi_3(t_k - t_i), \end{aligned}$$

where  $\varphi_1(t) = e^{(\sigma^2-2\alpha)(t_k-t_i)}$ ,  $\varphi_2(t) = e^{(\mu-\alpha)(t_k-t_i)}$ ,  $\varphi_3(t) = e^{(2\mu+\xi^2)(t_k-t_i)}$ ,  $\Psi = -\mu - \alpha + \sigma^2$ ,  $\Theta = \mu + \xi^2 + \alpha$ , and  $\Gamma = 2\mu + \xi^2 - \sigma^2 + 2\alpha$ .

**Remark 2.1.** By generic is meant that in the parameter space  $\mathbb{R}\{\sigma, \mu\}$  or  $\mathbb{R}\{\sigma, \mu, \alpha\}$ , the set of values satisfying (H1) or (H2) is open and dense. See [HS, 1974] p 154. From a mathematical modeling point of view, it suffices to only study generic models since any model can be made generic via a small perturbation.

*Proof.* The differential equations governing time evolution of  $\mathbb{E}(S_t)$  and  $\mathbb{E}(S_t^2)$  are give in [KP, 1999] p. 113, see also, [G, 1997] p. 113. For the case at hand, they become,

$$\begin{aligned} \frac{d}{dt}(\mathbb{E}(S_t)) &= \mu\mathbb{E}(S_t) + b \\ \frac{d}{dt}(\mathbb{E}(S_t^2)) &= (2\mu + \sigma^2) \mathbb{E}(S_t^2) + 2b\mathbb{E}(S_t). \end{aligned}$$

Then a simple integration yields the result. The system of ordinary differential equations governing the time evolution of the first and second moments of  $S_t$  and

$L_t$  are quoted in general in [KP, 1999] p. 152. For the first moments  $\mathbb{E}(S_t)$  and  $\mathbb{E}(L_t)$ , these equations are (2.6), and they were solved in the proof of Proposition 1. The equations for the second moments become:

$$\frac{d}{dt} \begin{pmatrix} \mathbb{E}(S_t, S_t) \\ \mathbb{E}(S_t, L_t) \\ \mathbb{E}(L_t, L_t) \end{pmatrix} = \begin{pmatrix} (\sigma^2 - 2\alpha) & 2\alpha & 0 \\ 0 & (\mu - \alpha) & \alpha \\ 0 & 0 & (\xi^2 + 2\mu) \end{pmatrix} \begin{pmatrix} \mathbb{E}(S_t, S_t) \\ \mathbb{E}(S_t, L_t) \\ \mathbb{E}(L_t, L_t) \end{pmatrix}.$$

Denoting the above three-by-three matrix by  $B$ , the solution to this system is given by

$$\begin{pmatrix} \mathbb{E}(S_t, S_t) \\ \mathbb{E}(S_t, L_t) \\ \mathbb{E}(L_t, L_t) \end{pmatrix} = e^{(t_k - t_i)B} \begin{pmatrix} S_i^2 \\ S_i L_i \\ L_i^2 \end{pmatrix}.$$

The computation of  $e^{(t_k - t_i)B}$  is routine in an eigenbasis, see [HS, 1974] chapter 7, or can be carried out in one operation in the symbolic language Maple.  $\square$

### 3. AN ABSTRACT APPROACH TO DISCRETE PROCESSES

Traditionally in mathematical finance literature various discrete modeling of stochastic processes under the guise of binomial or trinomial trees (and beyond) have been introduced. A single abstract unifying framework can be developed for all these trees. Consider a single or multi-factor stochastic process  $S_t$  modeled discretely over some time interval. At each time step  $i$ , the possible states of the world are given by finitely many vectors  $S^{(i)} = (S_j^{(i)})$ , where  $j$  is in a finite index set  $J^{(i)}$ . Note that we are not restricting ourselves to one-dimensional models by this notation. The state vector  $S^{(i)}$  can incorporate the values of the underlying asset as well as secondary factors in the multi-factor case. As time flows from the time step  $i$  to  $i + 1$ , the discrete node  $S_j^{(i)}$  can go to any  $S_{j'}^{(i+1)}$ ,  $j' \in J^{(i+1)}$ . What determines the discrete process is a transition probability matrix  $P^{(i)} = (P_{j,j'}^{(i)})$ , so that the element  $P_{j,j'}^{(i)}$  represents the probability of  $S_j^{(i)}$  going to  $S_{j'}^{(i+1)}$ . The number of rows in  $P^{(i)}$  is the cardinal of  $J^{(i)}$ , and the number of columns is the cardinal of  $J^{(i+1)}$ ; and for all  $j \in J^{(i)}$ ,  $\sum_{j' \in J^{(i+1)}} P_{j,j'}^{(i)} = 1$ .

**Example 3.1.** A one-factor non-recombining binomial discrete stochastic process is  $S^{(i)} = (S_j^{(i)})$  with  $j \in J^{(i)} = \{1, \dots, 2^i\}$ . The transition probabilities are given for every  $i$  by a vector  $(p_j^i), j \in J^{(i)}$ , such that  $P_{j,j'}^{(i)}$  is defined to be  $p_j^i$  if  $j' = 2j$ ,  $1 - p_j^i$  if  $j' = 2j - 1$ , or 0 otherwise, as depicted in Figure 1. The dashed lines in that figure represent prohibited connections which are ruled out by setting their corresponding probabilities equal to zero. In a multiplicative process one defines the up and down jump sizes  $u_j^{(i)}$  and  $d_j^{(i)}$  such that,  $S_{2j}^{(i+1)} = S_j^{(i)} u_j^{(i)}$  and  $S_{2j-1}^{(i+1)} = S_j^{(i)} d_j^{(i)}$ . We note that this tree is not numerically efficient since the number of nodes grows exponentially with time.

**Example 3.2.** A one-factor recombining binomial discrete stochastic process is  $S^{(i)} = (S_j^{(i)})$  with  $j \in J^{(i)} = \{1, \dots, i + 1\}$ . The transition probabilities, are given for every  $i$  by a vector  $(p_j^i), j \in J^{(i)}$ , such that  $P_{j,j'}^{(i)}$  is defined to be  $p_j^i$  if  $j' = j + 1$ ,

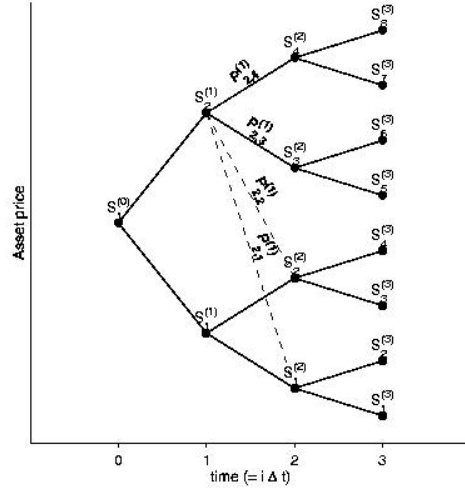


FIGURE 1. A non-recombining binomial tree, showing the labeling of the asset prices at the nodes, and of selected elements of the transition probability matrix  $P^{(1)}$ . Dashed lines indicate zero values for the associated probabilities. The non-zero transition probabilities are indicated by solid lines joining the nodes.

$1 - p_j^i$  if  $j' = j$ , or 0 otherwise, see Figure 2. In a multiplicative process one defines the up and down jump sizes  $u_j^{(i)}$  and  $d_j^{(i)}$  so that,  $S_{j+1}^{(i+1)} = S_j^{(i)} u_j^{(i)}$  and  $S_j^{(i+1)} = S_j^{(i)} d_j^{(i)}$ . By definition of recombination  $u_j^{(i)} d_{j+1}^{(i+1)} = d_j^{(i)} u_j^{(i+1)}$ . A standard choice is to set the up and down jumps constant across the tree:  $u_j^{(i)} = u$  and  $d_j^{(i)} = d$ , and to have a single probability  $p$  such that  $p_{jj'}^{(i)} = p \delta_{jj'-1} + (1 - p) \delta_{jj'}$ , where  $\delta_{jj'}$  is equal to 1 if  $j = j'$  and 0 otherwise. Then  $S_j^{(i)} = S_0 u^j d^{i-j}$ , with  $S_0 = S_1^{(0)}$ . A corresponding additive process might be defined via  $S_j^{(i)} = S_0 + ju + (i - j)d$ .

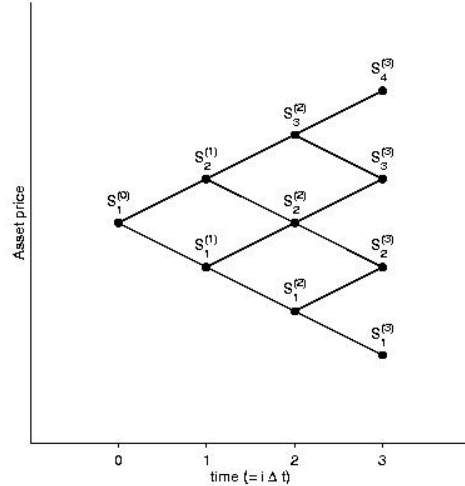


FIGURE 2. A recombining binomial tree, illustrating the labeling of the asset-prices at the nodes, as well as the links associated with non-zero probabilities.

**Example 3.3.** A one-factor recombining trinomial discrete stochastic process is  $S^{(i)} = (S_j^{(i)})$  with the index set  $J^{(i)} = \{1, \dots, 2i + 1\}$  and the transition probabilities, given for every  $i$  by two vectors  $(p_j^i), (q_j^i), j \in J^{(i)}$ , such that  $P_{j,j'}^{(i)}$  is defined to be  $p_j^i$  if  $j' = j + 2, q_j^i$  if  $j' = j + 1, 1 - p_j^i - q_j^i$  if  $j' = j$ , or 0 otherwise. See Figure 3.

**Example 3.4.** In some cases, two-factor stochastic processes can be modeled as a direct product of two trees, see Theorem 2 below. Consider two recombining binomial trees  $S^{(i)} = (S_j^{(i)}), j \in J^{(i)}$  and  $L^{(i)} = (L_k^{(i)}), k \in K^{(i)}$  as in Example 3.2, possessing the same number of nodes along their time axes, that is  $J^{(i)} = K^{(i)}$  for all  $i$ . Denote the transition probabilities of  $S^{(i)}$  by  $p_j^i$  with up and down jumps  $u_j^i, d_j^i$  and let those of  $L^{(i)}$  be  $q_k^i, w_k^i$  and  $h_k^i$  respectively. The direct product  $S^{(i)} \times L^{(i)}$

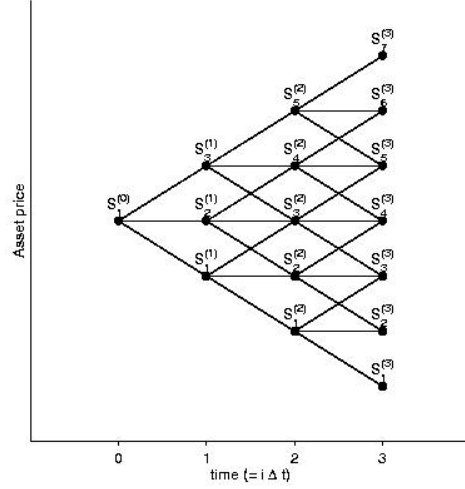


FIGURE 3. A recombining trinomial tree, illustrating the labeling of the asset-prices at the nodes, as well as the links associated with non-zero probabilities.

is naturally defined to be a tree  $T^{(i)}$  with a typical node  $T_{j,k}^{(i)} = (S_j^{(i)}, L_k^{(i)})$  at time  $i$ . In the next time step,  $T_{j,k}^{(i)}$  branches into four nodes  $T_{j+1,k+1}^{(i+1)}$ ,  $T_{j,k+1}^{(i+1)}$ ,  $T_{j+1,k}^{(i+1)}$  and  $T_{j,k}^{(i+1)}$ , defined along with their probabilities by

	Nodes	Probabilities
$T_{j+1,k+1}^{(i+1)}$	$= (S_j^{(i)} u_j^i, L_k^{(i)} w_k^i)$	$p_j^i q_k^i$
$T_{j,k+1}^{(i+1)}$	$= (S_j^{(i)} d_j^i, L_k^{(i)} w_k^i)$	$(1 - p_j^i) q_k^i$
$T_{j+1,k}^{(i+1)}$	$= (S_j^{(i)} u_j^i, L_k^{(i)} h_k^i)$	$p_j^i (1 - q_k^i)$
$T_{j,k}^{(i+1)}$	$= (S_j^{(i)} d_j^i, L_k^{(i)} h_k^i)$	$(1 - p_j^i)(1 - q_k^i)$ .



## 4. NUMERICAL IMPLEMENTATIONS

In this section we derive formulae for numerical implementations of various underlying processes discussed in Section 2 in terms of the multiplicative recombining binomial trees of Section 3. We first discuss the general, equation (2.1) and deduce the cases of log-normal and one factor mean-reverting processes as an application. We next show that up to the first order approximation, the two-factor model (2.4) can be discretized as a direct product of the two one-factor trees above.

Consider the discrete binomial process  $S^{(i)}$  wherein the value of the process at the time step  $i$ , given by  $S_j^{(i)}$ , can undergo an up jump to  $S_{j'+1}^{(i+1)} = S_j^{(i)} u_j^i$  with the probability  $p_j^i$ , or a down jump to  $S_{j'}^{(i+1)} = S_j^{(i)} d_j^i$ . Note that this situation is general enough to encompass both Example 1 and 2 of Section 3, and to also include the possibilities of multiple jumps. To approximate a continuous process, we match the first and the second moments of the discontinuous process on the tree with those of the continuous process over time intervals of length  $\Delta t$ .

**Proposition 3.** *Consider the continuous process  $dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dZ_t$  as in (2.1), and the recombining binomial tree  $(S_j^{(i)}) = S^{(i)}$  described above. Assume that at  $t = t_i$ ,  $S_{t_i} = S_j^{(i)}$ . Over the time interval  $[t_i, t_i + \Delta t]$ , we denote the conditional expectations of the continuous process by  $\mathbb{E}(S_t | S_{t_i})|_{t_i + \Delta t}$ , and that of the discrete process by  $\mathbb{E}_d(S^{(i)} | S_j^{(i)})|_{t_i + \Delta t}$ .*

*Matching the first and second moments of these two processes results in the binomial tree specified by:*

$$(4.1) \quad p_j^i = \frac{A(S_{t_i}, \Delta t) - d_j^i}{u_j^i - d_j^i}, \quad u_j^i = e^{\cosh^{-1}(\theta)}, \quad d_j^i = 1/u_j^i$$

where,

$$\begin{aligned} A(S_{t_i}, \Delta t) &:= \mathbb{E}(S_t | S_{t_i})|_{t_i + \Delta t} / S_j^{(i)}, \\ B(S_{t_i}, \Delta t) &:= \mathbb{E}(S_t^2 | S_{t_i})|_{t_i + \Delta t} / (S_j^{(i)})^2, \\ \theta(S_{t_i}, \Delta t) &:= \frac{1 + B(S_{t_i}, \Delta t)}{2A(S_{t_i}, \Delta t)} \end{aligned}$$

**Remark 2.2.** This general result applies to all binomial approximation schemes. Some classical particular cases are discussed in [K,1998] Chapter 5.1.

*Proof.* Using the definition of  $\mathbb{E}_d(\cdot)$ , the matching equations become

$$\begin{aligned} \mathbb{E}_d(S^{(i)} | S_j^{(i)})|_{t_i + \Delta t} &: = p_j^i S_{j'+1}^{(i+1)} + (1 - p_j^i) S_{j'}^{(i+1)} = \mathbb{E}(S_t | S_{t_i})|_{t_i + \Delta t} \\ \mathbb{E}_d((S^{(i)})^2 | S_j^{(i)})|_{t_i + \Delta t} &: = p_j^i (S_{j'+1}^{(i+1)})^2 + (1 - p_j^i) (S_{j'}^{(i+1)})^2 = \mathbb{E}(S_t^2 | S_{t_i})|_{t_i + \Delta t} \end{aligned}$$

Since  $S_{j'+1}^{(i+1)} = S_j^{(i)} u_j^i$  and  $S_{j'}^{(i+1)} = S_j^{(i)} d_j^i$ , the above equations reduce to

$$(4.2) \quad \begin{aligned} p_j^i u_j^i + (1 - p_j^i) d_j^i &= A(S_{t_i}, \Delta t) \\ p_j^i u_j^{i2} + (1 - p_j^i) d_j^{i2} &= B(S_{t_i}, \Delta t) \end{aligned}$$

where  $A(S_{t_i}, \Delta t) := \mathbb{E}(S_t | S_{t_i})|_{t_i+\Delta t} / S_j^i$  and  $B(S_{t_i}, \Delta t) := \mathbb{E}(S_t^2 | S_{t_i})|_{t_i+\Delta t} / S_j^{i2}$ . Solving for  $p_j^i$  results in  $p_j^i = \frac{A(S_{t_i}, \Delta t) - d_j^i}{u_j^i - d_j^i}$ . To remove an extra degree of freedom in these equations we assume that  $d_j^i = 1/u_j^i$ , then substituting  $p_j^i$  into the second equation and solving gives  $u_j^i + d_j^i = \frac{1+B(S_{t_i}, \Delta t)}{A(S_{t_i}, \Delta t)}$ , which yields the result.  $\square$

**Remark 2.3.** One should note<sup>1</sup> that  $0 \leq p_j^i \leq 1$ . Moreover, although the above values for  $u_j^i$  and  $d_j^i$  match (4.2) precisely, in general they will not result in a recombining tree. In practice, for every particular choice of  $\mu(S_t, t)$  and  $\sigma(S_t, t)$  in Equation (2.1) appropriate approximations must be chosen to ensure numerical feasibility as in the following theorem.

**Theorem 1.** a) Consider the linear stochastic differential equation

$$dS_t = (\mu S_t + b)dt + \sigma S_t dZ_t$$

with constant coefficients over the time interval  $[t_0, T]$ . Let  $S^{(i)} = (S_j^{(i)})$  be a recombining binomial tree as described in Example 3.2, with time step  $\Delta t$ . Suppose there exists a lower bound  $S_{\min} \neq 0$  such that  $S_j^{(i)} > S_{\min}$  for all  $i, j$ . Matching the first and second moments of  $S^{(i)} = (S_j^{(i)})$  with those of  $S_t$ , up to order  $\Delta t$  and for  $S_j^i$  above  $S_{\min}$ , results in

$$p_j^i = \frac{1}{2} + \frac{\mu + b/S_j^{(i)} - \sigma^2/2}{2\sigma} \sqrt{\Delta t}, \quad u_j^i = e^{\sigma\sqrt{\Delta t}}, \quad d_j^i = 1/u_j^i.$$

b) Without assuming a lower bound  $S_{\min}$ , the log-normal process  $dS_t = \mu S_t dt + \sigma S_t dZ_t$  can be discretized by a recombining binomial tree with

$$p_j^i = \frac{1}{2} + \frac{\mu - \sigma^2/2}{2\sigma} \sqrt{\Delta t}, \quad u_j^i = e^{\sigma\sqrt{\Delta t}}, \quad d_j^i = 1/u_j^i.$$

<sup>1</sup>To see why this is so, note first that the definitions of  $A(S_{t_i}, \Delta t)$  and  $B(S_{t_i}, \Delta t)$  as normalized expectations of  $S(t)$  and  $(S(t))^2$  respectively imply (dropping extraneous notation) that

$$B \geq A^2.$$

One immediate consequence of this is that

$$\theta = \frac{1+B}{2A} \geq \frac{1+A^2}{2A} = 1 + \frac{(A-1)^2}{2A} \geq 1.$$

A second consequence is that

$$2A\theta = 1+B \geq 1+A^2,$$

so that

$$1+A^2-2A\theta \leq 0.$$

By adding  $\theta^2 - 1$  to both sides we obtain

$$(A-\theta)^2 \leq \theta^2 - 1.$$

It follows now that

$$\theta - \sqrt{\theta^2 - 1} \leq A \leq \theta + \sqrt{\theta^2 - 1}.$$

Noting that  $\cosh^{-1} \theta = \ln(\theta + \sqrt{\theta^2 - 1})$  we see that we have

$$d \leq A \leq u.$$

The required result now follows directly from the definition of  $p$ .

c) The mean-reverting process  $dS_t = \alpha(L - S_t) dt + \sigma S_t dZ_t$ , can be discretized by a recombining binomial tree with

$$p_j^i = \frac{1}{2} + \left( \alpha \frac{L/S_j^{(i)} - 1}{2\sigma} - \frac{\sigma}{4} \right) \sqrt{\Delta t}, \quad u_j^i = e^{\sigma\sqrt{\Delta t}} \quad \text{and} \quad d_j^i = 1/u_j^i.$$

**Remark 2.4.** The above approximations are valid for a  $\Delta t$  which is independent of the node. Consequently, the up and down jumps  $u_j^i$  and  $d_j^i$  are constant throughout the tree and do not depend on the position of the node, forcing this tree to recombine. Furthermore, for  $\Delta t$  small enough,  $0 \leq p_j^i \leq 1$ . Therefore the above trees are numerically efficient.

*Proof.* a) Using the expressions for  $\mathbb{E}(S_t|S_{t_i})|_{t_i+\Delta t}$  and  $\mathbb{E}(S_t^2|S_{t_i})|_{t_i+\Delta t}$  from Proposition 2 in Proposition 3, leads to:

$$\begin{aligned} A(S_{t_i}, \Delta t) &= -\frac{b}{\mu S_j^{(i)}} + \left(1 + \frac{b}{\mu S_j^{(i)}}\right) e^{\mu\Delta t} \\ B(S_{t_i}, \Delta t) &= \frac{c}{(S_j^{(i)})^2} e^{(2\mu+\sigma^2)\Delta t} - \frac{2b(S_j^{(i)} + \frac{b}{\mu})}{(\sigma^2 + \mu)(S_j^{(i)})^2} e^{\mu\Delta t} + \frac{2b^2}{\mu(2\mu + \sigma^2)(S_j^{(i)})^2}. \end{aligned}$$

In the interval  $(S_{\min}, \infty)$ ,  $A(S_{t_i}, \Delta t)$  and  $B(S_{t_i}, \Delta t)$  can be expanded in Taylor series in  $\Delta t$  uniformly with respect to  $S_i$ . Indeed,  $A(S_{t_i}, \Delta t) = 1 + (\mu + \frac{b}{S_j^{(i)}})\Delta t +$

$R(\Delta t, S_j^{(i)})$  where  $R(\Delta t, S_j^{(i)}) = (1 + \frac{b}{\mu S_j^{(i)}}) r(\Delta t)$ , with  $\lim_{\Delta t \rightarrow 0} \frac{r(\Delta t)}{\Delta t} = 0$ . Since

$\frac{b}{\mu S_j^{(i)}} \leq \frac{b}{\mu S_{\min}}$  for all  $i, j$ ,  $\lim_{\Delta t \rightarrow 0} \frac{R(\Delta t, S_j^{(i)})}{\Delta t} = 0$ . A similar argument holds for  $B(S_{t_i}, \Delta t)$ . Therefore,

$$(4.3) \quad \begin{aligned} A(S_{t_i}, \Delta t) &= 1 + \left(\mu + \frac{b}{S_j^{(i)}}\right)\Delta t + O(\Delta t^2) \\ B(S_{t_i}, \Delta t) &= 1 + \left(\frac{2b}{S_j^{(i)}} + 2\mu + \sigma^2\right)\Delta t + O(\Delta t^2) \end{aligned}$$

and hence using (4.1) one has  $\theta(S_{t_i}, \Delta t) = 1 + \frac{\sigma^2}{2}\Delta t + O(\Delta t^2)$ . On the other hand,  $\cosh(\sigma\sqrt{\Delta t}) = 1 + \frac{\sigma^2}{2}\Delta t + O(\Delta t)^2$ , therefore

$$\theta(S_{t_i}, \Delta t) = \cosh(\sigma\sqrt{\Delta t}) \text{ mod}O(\Delta t^2).$$

Substituting this value for  $\theta$  in (4.1) yields:

$$u_j^i = e^{\sigma\sqrt{\Delta t}} \text{ mod}(\Delta t)^2 \quad \text{and} \quad p_j^i = \frac{1}{2} + \frac{\mu + b/S_j^{(i)} - \sigma^2/2}{2\sigma} \sqrt{\Delta t} + O(\Delta t^{3/2})$$

The notation  $u_j^i = e^{\sigma\sqrt{\Delta t}} \text{ mod}(\Delta t)^2$  signifies that  $u_j^i$  and  $e^{\sigma\sqrt{\Delta t}}$  agree up to and including terms of degree  $(\Delta t)$ . Cases b) and c) are particular cases of a). If  $b = 0$ , the uniform expansion of  $A$  and  $B$  will be valid on  $(0, \infty)$ . Part b) is a special case of a) with  $b = 0$ , and part c) is a special case of a) with  $\mu = -\alpha$  and  $b = \alpha L$ .  $\square$

We now turn to the two-factor model and give

**Theorem 2.** Consider the system of linear stochastic differential equations (2.4) over the time interval  $[t_0, T]$ ,

$$\begin{aligned} dS_t &= \alpha(L_t - S_t) dt + \sigma S_t dZ_t \\ dL_t &= \mu L_t dt + \xi L_t dW_t, \end{aligned}$$

a) Over any interval  $[t_i, t_i + \Delta t] \subset [t_0, T]$ , the processes  $S_t$  and  $L_t$  will remain uncorrelated up to the order  $O(\Delta t, S_{t_i}, L_{t_i})$ .

b) Let  $S^{(i)} = (S_j^{(i)})$  and  $L^{(i)} = (L_k^{(i)})$  be recombining binomial trees as described in Example 3.2, with time step  $\Delta t$ . Suppose there exist lower and upper bounds  $S_{\min} \neq 0$ , respectively  $L_{\max} < \infty$  such that  $S_j^{(i)} > S_{\min}$  and  $L_k^{(i)} < L_{\max}$  for all  $i, j, k$ . Then, away from  $S_{\min}$  and  $L_{\max}$ , matching the first and second moments of  $S^{(i)}$ , respectively  $L^{(i)}$  with those of  $S_t$ , respectively  $L_t$ , and up to order  $\Delta t$  results in

$$\begin{aligned} L^{(i)} &: \quad q_k^i = \frac{1}{2} + \frac{\mu - \xi^2}{2\xi} \sqrt{\Delta t}, \quad w_k^i = e^{\xi \sqrt{\Delta t}}, \quad h_k^i = 1/w_k^i. \\ S^{(i)} &: \quad p_j^i = \frac{1}{2} + \left( \alpha \frac{L_k^{(i)}/S_j^{(i)} - 1}{2\sigma} - \frac{\sigma}{4} \right) \sqrt{\Delta t}, \quad u_j^i = e^{\sigma \sqrt{\Delta t}}, \quad d_j^i = 1/u_j^i. \end{aligned}$$

The above system can be discretized as a direct product of the two trees  $S^{(i)} \times L^{(i)}$  described in Example 3.4.

*Proof.* a) Assume that at  $t = t_i$  the value of  $S_t$  is  $S_j^{(i)}$  and that the value of  $L_t$  is  $L_k^{(i)}$ . A first order approximation of the results obtained in Proposition 2 part b) yields

$$\begin{aligned} \mathbb{E}(S_t | S_j^{(i)})|_{t_i + \Delta t} &= S_j^{(i)} + \alpha(L_k^{(i)} - S_j^{(i)})\Delta t + O(\Delta t^2, S_{t_i}, L_{t_i}) \\ \mathbb{E}(L_t | L_k^{(i)})|_{t_i + \Delta t} &= L_k^{(i)} + \mu L_k^{(i)} \Delta t + O(\Delta t^2, L_{t_i}) \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}((S_t, S_t) | S_j^{(i)})|_{t_i} &= (S_j^{(i)})^2 + ((\sigma^2 - 2\alpha)(S_j^{(i)})^2 + 2\alpha S_j^{(i)} L_k^{(i)})\Delta t + O(\Delta t^2, S_{t_i}, L_{t_i}) \\ \mathbb{E}((S_t, L_t) | (S_j^{(i)}, L_k^{(i)}))|_{t_i} &= S_j^{(i)} L_k^{(i)} + ((\mu - \alpha)S_j^{(i)} L_k^{(i)} + \alpha(L_k^{(i)})^2)\Delta t + O(\Delta t^2, S_{t_i}, L_{t_i}) \\ \mathbb{E}((L_t, L_t) | L_k^{(i)})|_{t_i} &= (L_k^{(i)})^2 + (\xi^2 + 2\mu)\Delta t + O(\Delta t^2, L_{t_i}). \end{aligned}$$

Therefore, up to  $O(\Delta t, S_{t_i}, L_{t_i})$ ,

$$\text{Cov}(S_t, L_t)|_{t_i} = \mathbb{E}((S_t, L_t) | (S_j^{(i)}, L_k^{(i)}))|_{t_i} - \mathbb{E}(S_t | S_j^{(i)})|_{t_i + \Delta t} \mathbb{E}(L_t | L_k^{(i)})|_{t_i + \Delta t} = 0$$

and hence at  $(S_j^{(i)}, L_k^{(i)})$  and up to  $O(\Delta t, S_{t_i}, L_{t_i})$  the processes  $S_t$  and  $L_t$  will remain uncorrelated over  $[t_i, t_i + \Delta t]$ .

b) By part a) at each  $(S_j^{(i)}, L_k^{(i)})$  the two processes  $S_t$  and  $L_t$  remain uncorrelated over  $[t_i, t_i + \Delta t]$ . Therefore at such a point and over  $[t_i, t_i + \Delta t]$ ,  $S_t$  and  $L_t$  can be approximated by independent binomial trees, and hence (2.4) can in turn be approximated by the direct product of these two trees over the same time interval. Since  $L$  is a log-normal process, part b) of Theorem 1 yields the tree  $L^{(i)}$ . Note that, at  $(S_j^{(i)}, L_k^{(i)})$ ,  $S_t$  follows a one-factor mean-reverting process (2.3) with  $L = L_k^{(i)}$ .

Therefore from Theorem 1 we have

$$A(S_{t_i}, \Delta t) = \frac{L_k^{(i)}}{S_j^{(i)}} + \left(1 - \frac{L_k^{(i)}}{S_j^{(i)}}\right) e^{-\alpha \Delta t}$$

$$B(S_{t_i}, \Delta t) = \frac{c}{(S_j^{(i)})^2} e^{(\sigma^2 - 2\alpha)\Delta t} - \frac{2\alpha L_k^{(i)}(S_j^{(i)} - L_k^{(i)})}{(\sigma^2 - \alpha)(S_j^{(i)})^2} e^{-\alpha \Delta t} - \frac{2\alpha L_k^{(i)2}}{(\sigma^2 - 2\alpha)(S_j^{(i)})^2}.$$

Note that since  $S_{min}$  and  $L_{max}$  exist  $\frac{L_k^{(i)}}{S_j^{(i)}}$  is bounded. Therefore  $A$  and  $B$  can be uniformly expanded in Taylor series, as in Theorem 1. Hence the tree of  $S$  follows from part c) of Theorem 1.  $\square$

## 5. A GENERAL SWING CONTRACT

We begin by introducing the parameters necessary for modeling a fairly general swing option contract which accomodates a wide range of applications.

The economy considered has two assets, a risk free interest rate, and a risky asset  $S$  exchanged in units, and two players, a buyer and a seller. Consider three time values  $T_0 \leq T_1 < T_2$  where  $T_0$  is the time when the swing option is priced, and the interval  $[T_1, T_2]$  is the swing contract period during which the buyer is assumed to be purchasing from the seller a determined amount of  $S$ , called the base load. Base loads can be easily priced, we will therefore only focus on modeling and pricing the swing. Beyond the base load, the swing option contract provides the possibility of exchanging the asset  $S$  in fixed quantities and at determined strike prices. By definition, an up swing consists in the buyer acquiring  $V_u$  units of  $S$  immediately upon request at a strike price of  $K_u$  per unit, and a down swing consists in the buyer delivering  $V_d$  units of  $S$  to the seller at a strike price of  $K_d$  per unit, immediately after notification of the latter. The cost of a swing is to be settled immediately. The swing option entitles the buyer to exercise, during the time interval  $[T_1, T_2]$ , up to  $N_u$  up swings and  $N_d$  down swings. An exercise can only occur at a discrete set of times  $\{\tau_1, \dots, \tau_e\} \subset [T_1, T_2]$ , and consists of at most one up or down swing at each time. A penalty is to be imposed if the net amount of  $S$  acquired by the buyer via swing exercises is not bounded by  $V_{min}$  and  $V_{max}$  at expiry. More precisely let  $n_u$  and  $n_d$  denote respectively the actual number of up and down swing exercises that occurred during  $[T_1, T_2]$ , then to avoid a penalty one must have

$$(5.1) \quad V_{min} \leq n_u V_u - n_d V_d \leq V_{max}.$$

If not, there are cash penalties of  $A_1$  (respectively  $A_2$ ) per unit of the net amount of units acquired,  $n_u V_u - n_d V_d$ , short of  $V_{min}$ , (respectively in excess of  $V_{max}$ ). Then the resulting penalties are  $A_1(V_{min} - n_u V_u + n_d V_d)$  (respectively,  $A_2(n_u V_u - n_d V_d - V_{max})$ ).

## 6. DISCRETE FOREST METHODOLOGY

We now turn to modeling the swing option contract described in the above section via a *forest*. It is assumed that the underlying asset  $S$ , follows a one- or multi-factor risk-neutralized stochastic process, expressed in terms of a single or a system of stochastic differential equations. We use the notation of the previous section and recall that a swing exercise can only occur at discrete times  $\{\tau_1, \dots, \tau_e\} \subset$

$[T_1, T_2]$ . We may without loss of generality assume  $\tau_1 = T_1$  and  $\tau_e = T_2$ . To simplify the presentation, we assume that the rights to exercise occur at equally spaced time intervals, or *ticks*, of length  $\Delta t_e$ . We also assume that there are integers  $N_1$  and  $N_2$  such that  $T_1 = N_1\Delta t_e$  and  $T_2 = N_2\Delta t_e$ . To obtain an accurate valuation of the stochastic process for the underlying asset  $S$  we model it on a discrete tree with a finer time-scale obtained, by dividing each tick into  $n$  *computational timesteps*  $\Delta t = \Delta t_e/n$ . Therefore  $S$  follows a process discretized by a tree  $S^{(i)}, P^{(i)}$ , with  $nN_2$  time grid nodes, that is  $i \in I = \{0, \dots, nN_2\}$ , whereas option exercise is permitted at nodes whose indices belong to the set  $I_e = \{N_1n, (N_1+1)n, \dots, N_2n\}$ . The local value of the swing option, denoted by  $V_{n_u, n_d}^i(j)$ , depends on the generalized index set  $(i, j, n_u, n_d)$ , where  $n_u$  and  $n_d$  denote the number of remaining up and down swings respectively,  $i \in I$  and  $j \in J^{(i)}$ . The forest terminology finds its origin in this setting: for each  $(n_u, n_d)$ , the tree inhabited by the values  $V_{n_u, n_d} = \{V_{n_u, n_d}^i(j), j \in J^{(i)}, i \in I\}$  may be viewed as a tree isomorphic to the tree  $\{S^{(i)}, i \in I\}$ ; thus we have not one tree but  $(N_u + 1) \times (N_d + 1)$  trees.

Define the discounted expected value

$$(6.1) \quad W_{n_u, n_d}^i(j) = R_i \sum_{j' \in J^{(i+1)}} P_{j, j'}^{(i)} V_{n_u, n_d}^{i+1}(j'),$$

where  $R_i = \exp(-r\Delta t)$  is the present value of 1 unit at a time  $\Delta t$  in the future. When not at an exercise node ( $i \notin I_e$ ), each tree  $V_{n_u, n_d}$  in the forest is back-folded independently of the others, and  $V_{n_u, n_d}^i(j) = W_{n_u, n_d}^i(j)$ .

At a point on the tree when exercise is permitted ( $i \in I_e$ ), one can still use (6.1) to calculate a discounted expected value on each tree, but the exercise of a swing (up or down) will cause a shift from the tree  $V_{n_u, n_d}$  to one of the trees  $V_{n_u-1, n_d}$  or  $V_{n_u, n_d-1}$ . Exercising the swing will result in an immediate cash flow corresponding to the relevant payoff formula. However, the option now has one fewer exercise rights, and a jump to a neighboring tree has taken place. Then the option value  $V_{n_u, n_d}^i(j)$  is given as the maximum of three possibilities:

$$\begin{aligned} & W_{n_u, n_d}^i(j) && \text{(no swing)} \\ & W_{n_u-1, n_d}^i(j) + (S_j^i - K_u) && \text{(swing up)} \\ & W_{n_u, n_d-1}^i(j) + (K_d - S_j^i) && \text{(swing down)}. \end{aligned}$$

This forest methodology is depicted in Figure 4 for a swing option with just one up- and one down-swing exercise right using 4 computational timesteps. Each tree carries the values of an option with a particular combination of up- and down-swing rights used up, over the course of 4 time steps. The calculation proceeds by stepping backwards in time down through the trees. At the top time level, the option values are determined solely by the penalties imposed. Then, at each state of the world, a decision is made as to whether it is optimal to exercise an up- or down-swing right. The choices made at the final time are shown in (a), where the links indicate the transfer of information from tree-to-tree as a result of an exercise of a swing right. The resulting values are then transferred to the next time level by a discounted expectation, and a set of decisions made as to whether or not to swing once more. Those decisions are shown in (b), and also in (c) at the next time level. The final two time levels are not shown because in this calculation, at no point was it found to be optimal to exercise the swing rights. The total history of all swing choices in all states of the world is shown in (d). The final option value is to be

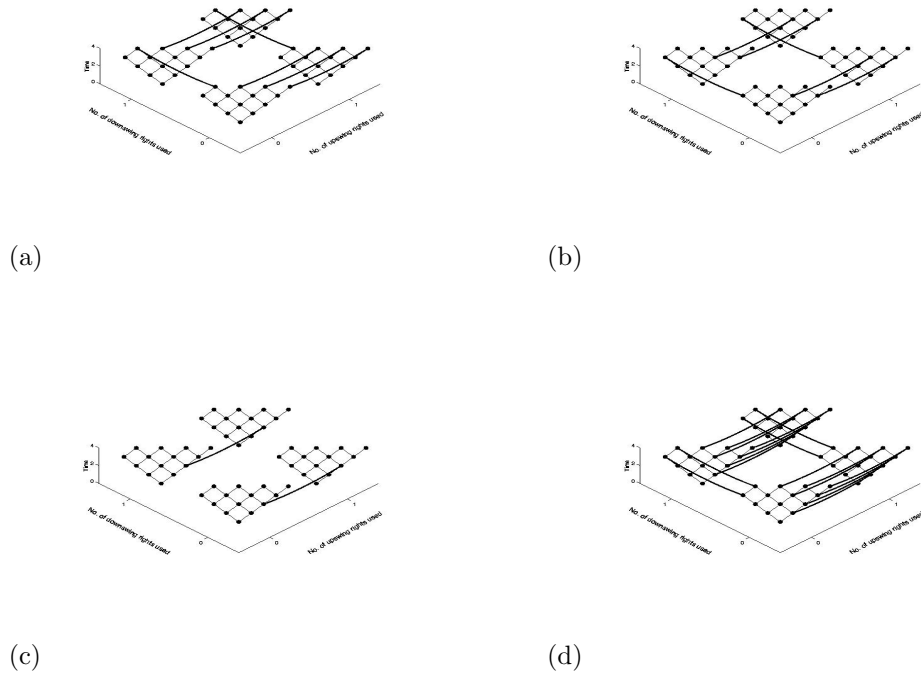


FIGURE 4. A  $2 \times 2$  forest for pricing a swing option with one up- and one down-swing exercise right using 4 computational timesteps.

found at the bottom node of the front tree. One feature to note is that the trees have all been truncated. This is because of the imposition of an  $S_{\min}$  as described in Theorem 1.

### 7. SENSITIVITY ANALYSIS

Pricing a swing contract involves a considerable number of input parameters and as such a detailed sensitivity analysis of such a contract would be the subject of a publication on its own. To exhibit a fair amount of information while keeping the presentation of different scenarios tractable, we consider two swing contracts with their parameters defined below. These parameters are fixed, unless otherwise stated. We follow the notation of Section 6.

For the one-factor mean-reverting model  $dS_t = \alpha(L - S_t)dt + \sigma S_t dW_t$  we label the following setting by (A) and implement it according to the binomial tree given in Theorem 1, (c).

$n$	$\alpha = 0.2, L = 30$			$\alpha = 1.2, L=30$		
	$S_{\min} = 5$	$S_{\min} = 15$	$S_{\min} = 25$	$S_{\min} = 5$	$S_{\min} = 15$	$S_{\min} = 25$
80	196.2270	196.2270	197.2613	183.7410	183.7411	184.7600
160	196.2245	196.2246	197.2006	183.7372	183.7373	184.6986
320	196.2208	196.2208	197.1931	183.7338	183.7339	184.6915

TABLE 1. Convergence behaviour and its dependence on  $\alpha$  and  $S_{\min}$ .

(A):  $S_0 = 45, S_{\min} = 5, r = 0.5, \alpha = 10, L = 40, \sigma = 1.5,$   
 $T_2 - T_0 = 30$  days,  
 $\Delta t_e = \tau_{i+1} - \tau_i = 1$  in units of day,  $N_1 = 1, N_2 = 30, n = 80,$   
 $N_u = 1, N_d = 0, V_u = 20, V_d = 20, K_u = 40, K_d = 40,$   
 $A_u = 3, A_d = 3, V_{\max} = 150, V_{\min} = -150.$

For the two-factor mean-reverting model  $dS_t = \alpha(L_t - S_t) dt + \sigma S_t dZ_t, dL_t = \mu L_t dt + L_t \xi dW_t$ , we label the following setting by (B) and implement it according to Theorem 2.

(B):  $S_0 = 1.5, S_{\min} = 0.1, r = 0.05, \alpha = 0.2, L_0 = 1, \sigma = 1.5, \mu = 1,$   
 $\xi = 1,$   
 $T_2 - T_0 = 30$  days,  
 $\Delta t_e = \tau_{i+1} - \tau_i = 1$  in units of day,  $N_1 = 1, N_2 = 30, n = 80,$   
 $N_u = 3, N_d = 3, V_u = 1, V_d = 1, K_u = 1, K_d = 1,$   
 $A_u = 0, A_d = 0, V_{\max} = 15, V_{\min} = -15.$

**7.1. Numerical Convergence.** We compute a sequence of the values of the swing option by doubling the number of computational time steps  $n$ . The convergence pattern of the value of the swing as  $n$  increases is illustrated by verifying that the error goes to zero. More specifically, the value of the swing option is calculated with  $n = 10^9$  for the one-factor scenario A, and with  $n = 2^7$  for the two-factor scenario B. These values are then subtracted from the approximate values of the swing option obtained with  $n = 10 \cdot 2^i, i = 0, \dots, 5$  for A, respectively  $n = 2^i, i = 1 \dots 6$  for B. These yield estimates for the error, which are plotted versus  $n$ , in Figure 5, respectively Figure 6 in a base 2 logarithmic scale. Note that in Figure 5 all of the values agree to 3 significant figures. A fourth digit of accuracy is added with more than 640 computational timesteps (see Section 6).

We finally note that the convergence pattern is a function of the various parameters of the swing contract as discussed in the next section.

**7.2. Parameter Sensitivity.** Table 1 is generated for scenario (A) wherein the values of  $\alpha, S_{\min}$  and  $n$  are varied as prescribed, and  $L = 30$ .

First note that the above setting (A) with long-run mean  $L = 30$  only involves one up-swing with an strike price,  $K_u = 40$ , which is above  $L$  and below the initial asset price. As a result, the swing option behaves as an in the money Bermudan call option. Next observe that the closer  $S_{\min}$  is to  $L$ , the higher is the value of the option. This is to be expected for, setting  $S_{\min}$  closer to  $L$  increases the probability of the asset prices being closer to the strike price, resulting in a higher value for



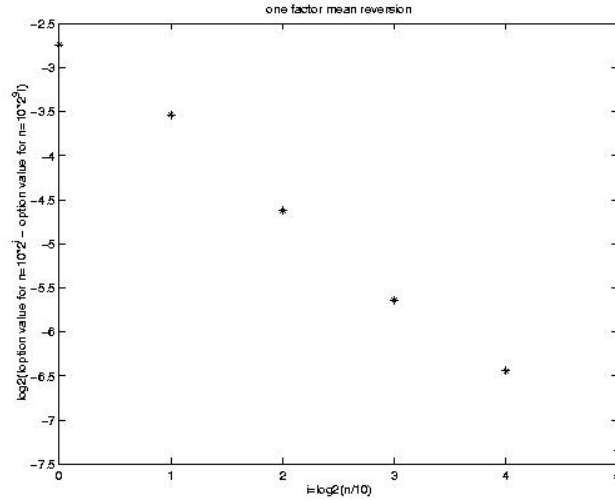


FIGURE 5. Option values for the one-factor scenario (A), calculated with increasing numbers of computational timesteps.

the option since it is a call option. This behaviour is presented in Figure 7. One may also note that the value of the swing option is not sensitive to changes in  $S_{\min}$  for  $S_{\min}$  less than 20, which is about half the long-run mean  $L$ . As the process is mean-reverting, asset values far from its long-run mean are less probable to occur and as such excluding them does not effect the price of the option. Finally, the stronger the strength of mean reversion is, the less valuable is the option. Indeed, as the strength of mean reversion increases, the process is pulled to its mean, which is below the strike price in this scenario, making the call option less valuable. Figure 8 captures this behaviour.

**7.3. Hedging Parameters.** Generally speaking, there are at least as many hedge parameters for a contract as input parameters. As such the hedge parameters of a swing contract constitute a long list. For the sake of exposition we provide two conventional hedge parameters,  $\Delta = \frac{\partial V}{\partial S}|_{S_0}$  and  $\Lambda = \frac{\partial V}{\partial \sigma}|_{S_0}$ , and a less conventional

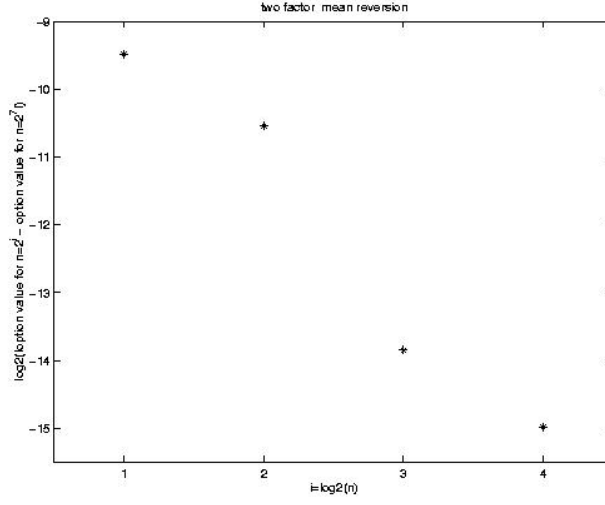


FIGURE 6. Option values for the two-factor scenario ( $B$ ), calculated with increasing numbers of computational timesteps.

$\frac{\partial V}{\partial L}|_{S_0}$ . Consider the following setting for a swing contract  $C$ , written on a one-factor mean-reverting process:

$$\begin{aligned}
 &S_0 = 1.5, \quad S_{\min} = .01, \quad r = 0.05, \quad \alpha = 2, \quad L = .8, \quad \sigma = 1.5, \\
 &T_2 - T_0 = 30 \text{ days}, \\
 &\Delta t_e = \tau_{i+1} - \tau_i = 1 \text{ in units of day}, \quad N_1 = 2, \quad N_2 = 30, \quad n = 30, \\
 &N_u = 4, \quad N_d = 3, \quad V_u = 1, \quad V_d = 1, \quad K_u = .4, \quad K_d = .4, \\
 &A_u = .1, \quad A_d = .1, \quad V_{\max} = 3, \quad V_{\min} = -2.
 \end{aligned}$$

We compute the above Greeks for values of  $S_0$  varying in a range which covers the cases in the money, at the money and out of the money. Table 2 introduces a few values from the set of data which is used to generate Figures 9, 10 and 11.

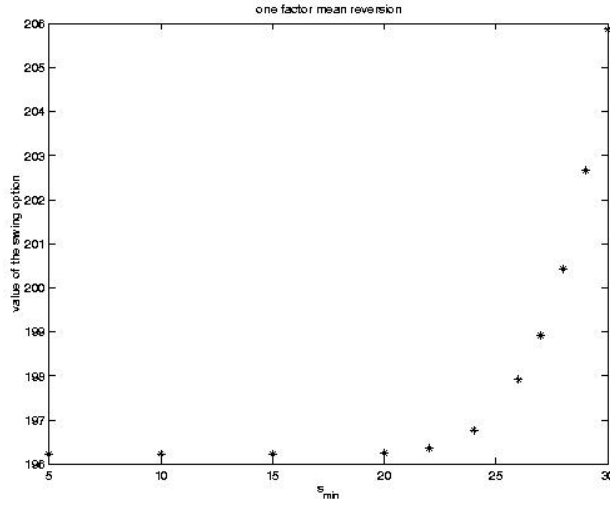


FIGURE 7. The effect of varying  $S_{\min}$  on the option price for the one-factor scenario  $A$ .

Hedge parameter	In the money $S_0 = .2$	At the money $S_0 = .4$	Out of the Money $S_0 = .6$
$\Delta$	-1.95	1.35	2.94
$\Lambda$	0.08	0.26	0.19
$\frac{\partial V}{\partial L}  _{S_0}$	0.11	0.29	0.41

TABLE 2. Sample values of some hedge parameters for various values of  $S_0$ .

Note that contract  $C$  is a combination of Bermudan put and call options. For asset values  $S$  below the strike price  $K_u = K_d = 0.4$ , the swing contract is out of the money and behaves as a put option. On the contrary, asset values above the strike price result in the contract being similar to a call option. This observation

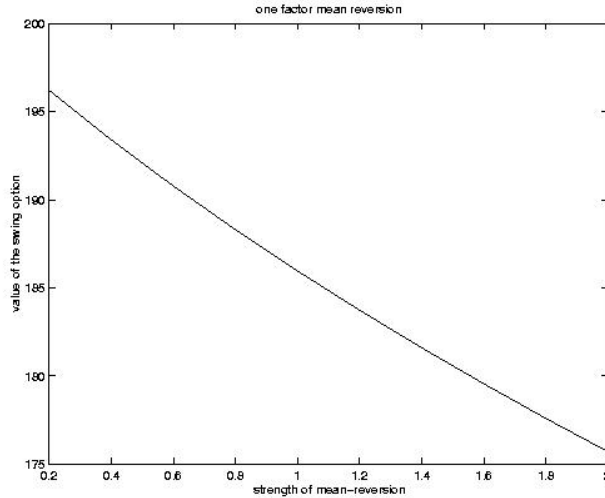


FIGURE 8. The effect of varying strengths of mean-reversion on the option value. As one might expect, higher values for  $\alpha$  mean lower option values.

explains the concavity and curvature in the  $\Delta$  and  $\Lambda$  curves above. As for the  $\frac{\partial V}{\partial L}$  curve the change occurs at  $S = L_0$ , which is predictable.

**7.4. Swings versus American Options.** We compare a swing contract with a basket of American options. Having a swing contract in hand, the buyer of the contract will have the opportunity of exercising his swing rights at particular dates at his leisure. Another way of replicating such a right is to purchase the same number of rights via American options having the same expiry date as that of the swing contract. Of course a buyer with a basket of American options in hand is able to exercise his rights anytime. He will also be able to exercise all or some of those rights at a single moment of time. However such extra freedom comes with an associated cost. Figure 12 captures this cost. Each point on the curve represents the value of a swing option with  $N_u$  upswing rights in a month,  $N_u = 1, 2, \dots, 30$ ,

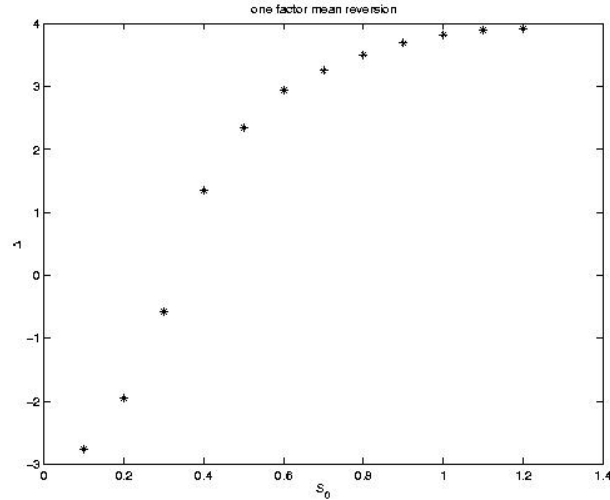


FIGURE 9.  $\Delta$  as a function of the initial asset price  $S_0$ .

whereas each point on the line indicates the value of  $N_u$  individual American call options expiring within the same period. As seen from the figure, due to the flattening out of the graph, the buyer of the swing option does not have to pay much for the extra rights purchased. He is losing some flexibility but at the same time he is saving money. The buyer of the American basket has more flexibility at a much higher cost. Therefore the choice between a swing contract and a basket of American options depends on the flexibility required and the price attached to it.

## 8. PROOF OF CONVERGENCE

In this section we demonstrate the convergence of the value of the swing option calculated on a forest of recombining binomial trees to the solution of a partial differential equation model of the swing option. The error is shown to decrease linearly with the number of computational time steps.

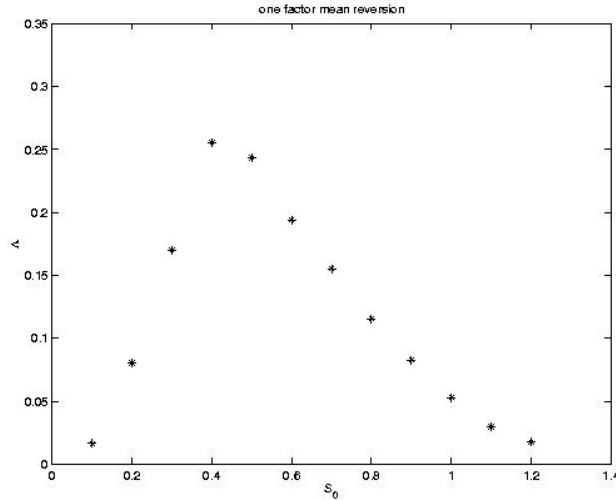


FIGURE 10.  $\Lambda$  as a function of the initial asset price  $S_0$ .

We limit the exposition to the proof of convergence for the one-factor mean-reverting binomial model described in Theorem 1, and note that the result could be readily adapted to cover other one-factor trees. The proof for the two-factor case would be similar in structure but rather more involved.

**8.1. The Swing Option as a Coupled System of European Options.** Consider the option described in Section 5, and write  $\underline{\nu} = (n_u, n_d)$ . Let  $V_{\underline{\nu}}(S, t)$  be the value at time  $t \in [0, T_2]$  of an option with  $n_u$  up-swing rights and  $n_d$  down-swing rights remaining, given an asset price of  $S$  at time  $t$ . Noting that the holder of such an option may choose to exercise at the exercise times  $t \in \{\tau_1, \dots, \tau_e\}$ , we see that  $V_{\underline{\nu}}(S, t)$  may have a jump discontinuity at those times. We write  $V_{\underline{\nu}}(S, \tau_i^+) = \lim_{t \rightarrow \tau_i^+} V_{\underline{\nu}}(S, t)$ .

As described in Section 6, the exercise of a swing right brings with it an associated cash flow, but also means that the option value is exchanged for the value of

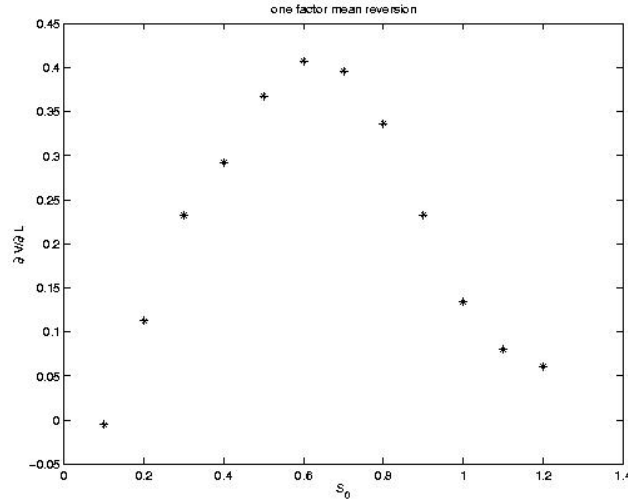


FIGURE 11.  $\frac{\partial V}{\partial L}$  as a function of the initial asset price  $S_0$ .

an option with one fewer right. At each exercise time, for each value of the underlying asset  $S$ , it is assumed that the optimal choice is made among the available possibilities. Thus, for  $\underline{\nu} \in (1, \dots, N_u) \times (1, \dots, N_d)$ , and for  $i = 1, \dots, e - 1$ , the (payoff) functions can be defined by

$$(8.1) \quad P_{\underline{\nu}}(S, \tau_i) = \max \begin{cases} V_{\underline{\nu}}(S, \tau_i^+) \\ V_{\underline{\nu}^u}(S, \tau_i^+) + \text{Pay}_u(S) \\ V_{\underline{\nu}^d}(S, \tau_i^+) + \text{Pay}_d(S), \end{cases}$$

where  $\underline{\nu}^u = (n_u - 1, n_d)$ ,  $\underline{\nu}^d = (n_u, n_d - 1)$ , and

$$\text{Pay}_u(S) = V_u(S - K_u), \quad \text{Pay}_d(S) = V_d(K_d - S),$$

with obvious changes to the definition for the cases  $n_u = 0$  or  $n_d = 0$ . The payoff function at the final exercise time is determined by the penalties being applied.

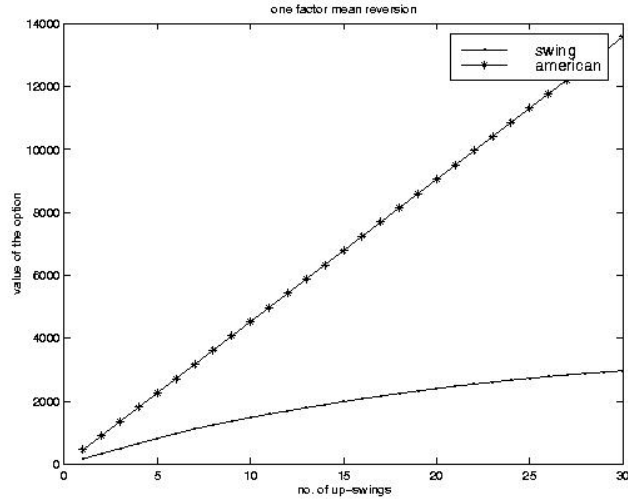


FIGURE 12. The difference between a swing option with only up-swing rights and an American option with an equivalent total volume. The option values are plotted as a function of the number of up-swing rights (each with unit volume) for the swing option, and as a function of the total volume for the American option.

The function  $\text{Pen}_{\underline{v}}(S)$  is defined by

$$(8.2) \quad \text{Pen}_{\underline{v}}(S) = A_1(n_u V_u - n_d V_d - V_{\min})_+ - A_2(n_u V_u - n_d V_d - V_{\max})_+.$$

The final exercise payoff function is then

$$(8.3) \quad P_{\underline{v}}(S, \tau_e) = \max \begin{cases} -\text{Pen}_{\underline{v}}(S) \\ -\text{Pen}_{\underline{v}^u}(S) + \text{Pay}_u(S) \\ -\text{Pen}_{\underline{v}^d}(S) + \text{Pay}_d(S), \end{cases}$$

again for  $n_u = 1, \dots, N_u$  and  $n_d = 1, \dots, N_d$ , and with obvious changes for the cases  $n_u = 0$  or  $n_d = 0$ .



For  $t$  in the interval  $(\tau_i, \tau_{i+1})$ , the option value  $V_{k,l}(S, t)$  is the value of a European contract at time  $t$  with expiry at time  $\tau_{i+1}$  and payoff function  $P_{\underline{v}}(S, \tau_{i+1})$ . The value of this swing option in any given time interval thus depends (via the payoff function) on the values of related swing options in subsequent time intervals. In particular, the value of the original swing contract is  $V_{\underline{v}^0}(S_0, 0)$ . This is the value at time 0 and asset price  $S_0$  of a European option with expiry time  $\tau_1$  and payoff  $P_{\underline{v}^0}(S, \tau_1)$ , and thus depends on the values of all of the swing options in the subsequent time intervals.

Obtaining a value for  $V_{\underline{v}^0}(S_0, 0)$  thus involves, for  $i$  running backwards from  $e-1$  to 1, first calculating each  $P_{\underline{v}}(S, \tau_{i+1})$  via (8.1) or (8.3), and then evaluating the European options described above in the interval  $(\tau_i, \tau_{i+1})$ . These values are then used to calculate the next set of payoff functions. Finally, the European option with expiry time  $\tau_1$  and payoff  $P_{\underline{v}^0}(S, \tau_1)$  is used to provide the required value at  $S = S_0$  and  $t = 0$ .

**8.2. A Partial Differential Equation Model.** The question remains of how the European options described above are to be valued. Here we describe a partial differential equation model based on a generalization of the Black-Scholes equation,

$$V_t + \frac{\sigma^2}{2} S^2 V_{SS} + rSV_S = rV.$$

This models, for example, the price  $V$  of a stock with volatility  $\sigma$  and risk-free interest rate  $r$  contingent on a stock price  $S$  which follows the stochastic process (2.2). The complete-market assumptions which underlies this model may not apply in the context we are interested in. Relaxing these assumptions somewhat, we have the equation

$$V_t + \frac{\sigma^2}{2} S^2 V_{SS} + (\mu - \lambda\sigma S)V_S = rV,$$

where  $S$  is now assumed to follow the stochastic process (2.3), and  $\lambda$  is the market price of risk function (See [H, 1999], [W, 1998]). We assume for ease of exposition that  $\lambda = 0$ . The results obtained in this section would hold equally well for other forms of  $\lambda$  satisfying quite mild continuity assumptions.

We make two concessions to our numerical schemes in the model we describe. The first is that we impose a lower bound  $S_{\min} > 0$  on the asset price  $S$ . The second is due to the fact that the payoff functions described above are continuous, but are only piecewise differentiable. This lack of smoothness introduces theoretical problems for the convergence of numerical schemes such as those described here. We define a ‘mollified’ maximum function,  $\max_{\epsilon}$ , by means of a formula such as

$$\max_{\epsilon}(a, b) = \frac{a + b + \sqrt{\epsilon + (a - b)^2}}{2},$$

with

$$\max_{\epsilon}(a, b, c) = \max_{\epsilon}(a, \max_{\epsilon}(b, c)).$$

Note that when  $\epsilon = 0$  this definition agrees with the standard definition of the maximum. However, defining the payoff functions using  $\max_{\epsilon}$  instead of  $\max$  means that the payoff functions  $P_{\underline{v}}(S, \tau_i)$  in (8.1) are just as smooth as the functions  $V_{\underline{v}}(S, \tau_i^+)$  used to define them. As long as  $\epsilon$  is chosen sensibly, it is easy to verify that the effect on the option price is negligible.

We are now ready to describe our partial differential equation model for the value of a swing option.

Let  $S_{\min} > 0$ ,  $\sigma > 0$  and  $r \geq 0$  be constants. We seek a set of functions  $V_{\underline{\nu}}(S, t)$  satisfying

$$\frac{\partial}{\partial t} V_{\underline{\nu}}(S, t) + \frac{\sigma^2 S^2}{2} \frac{\partial^2}{\partial S^2} V_{\underline{\nu}}(S, t) + (AS + B) \frac{\partial}{\partial S} V_{\underline{\nu}}(S, t) = rS V_{\underline{\nu}}(S, t),$$

for  $S > S_{\min}$ ,  $t \in [0, T] \setminus \{\tau_i\}_{i=1}^e$ , with

$$(8.4) \quad V_{\underline{\nu}}(S, T) = \max_{\epsilon} (-\text{Pen}_{\underline{\nu}}(S), -\text{Pen}_{\underline{\nu}^u}(S) + \text{Pay}_u(S), \\ -\text{Pen}_{\underline{\nu}^d}(S) + \text{Pay}_d(S)),$$

$$(8.5) \quad V_{\underline{\nu}}(S, \tau_i) = \max_{\epsilon} (V_{\underline{\nu}}(S, \tau_i^+), V_{\underline{\nu}^u}(S, \tau_i^+) + \text{Pay}_u(S), \\ V_{\underline{\nu}^d}(S, \tau_i^+) + \text{Pay}_d(S)), \quad i = 1, \dots, e-1$$

$$(8.6) \quad \frac{\partial}{\partial S} V_{\underline{\nu}}(S_{\min}, t) = 0, \quad t \in [0, T].$$

The comparison with the binomial scheme will in fact be more straightforward if we describe the equivalent log-transformed problem for  $w_{\underline{\nu}}(x, t) = e^{-r(T-t)} V_{\underline{\nu}}(S, t)$ :

$$(8.7) \quad \frac{\partial}{\partial t} w_{\underline{\nu}}(x, t) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} w_{\underline{\nu}}(x, t) + (a + be^{-x}) \frac{\partial}{\partial x} w_{\underline{\nu}}(x, t) = 0,$$

for  $x > \ln S_{\min}$ ,  $t \in [0, T] \setminus \{\tau_i\}_{i=1}^e$ , with

$$(8.8) \quad w_{\underline{\nu}}(x, T) = \max_{\epsilon} (-\text{Pen}_{\underline{\nu}}(e^x), -\text{Pen}_{\underline{\nu}^u}(e^x) + \text{Pay}_u(e^x), \\ -\text{Pen}_{\underline{\nu}^d}(e^x) + \text{Pay}_d(e^x)),$$

$$(8.9) \quad w_{\underline{\nu}}(x, \tau_i) = \max_{\epsilon} (w_{\underline{\nu}}(x, \tau_i^+), w_{\underline{\nu}^u}(x, \tau_i^+) + \text{Pay}_u(e^x), \\ w_{\underline{\nu}^d}(x, \tau_i^+) + \text{Pay}_d(e^x)), \quad i = 1, \dots, e-1$$

$$(8.10) \quad \frac{\partial}{\partial x} w_{\underline{\nu}}(\ln S_{\min}, t) = 0, \quad t \in [0, T],$$

where  $a = A - \sigma^2/2$  and  $b = B$ .

**8.3. The One-factor Binomial Scheme.** The binomial approximation to the above partial differential equation model consists in finding vectors  $W_{\underline{\nu}}^i(j)$  satisfying, for  $j \geq j_{\min}$ ,

$$W_{\underline{\nu}}^{i,*}(j) = p_j W_{\underline{\nu}}^{i+1}(j+1) + (1-p_j) W_{\underline{\nu}}^{i+1}(j-1), \quad i < T/\Delta t = N_2 n,$$

$$W_{\underline{\nu}}^{N_2 n,*}(j) = -\text{Pen}_{\underline{\nu}}(e^{x_j}),$$

$$W_{\underline{\nu}}^i(j) = W_{\underline{\nu}}^{i,*}(j), \quad i+1 \notin I_e$$

$$W_{\underline{\nu}}^i(j) = \max_{\epsilon} \left( W_{\underline{\nu}}^{i,*}(j), W_{\underline{\nu}^u}^{i,*}(j) + \text{Pay}_u(e^{x_j}), W_{\underline{\nu}^d}^{i,*}(j) + \text{Pay}_d(e^{x_j}) \right), \quad i+1 \in I_e.$$

Here  $x_j = \ln(S_0 e^{j\Delta x})$ , with  $\Delta x = \sigma\sqrt{\Delta t}$ , and  $j_{\min}$  is the value of  $j$  for which  $S_j = e^{x_j} = S_{\min}$ . The above system is completed by setting

$$W_{\underline{\nu}}^{i+1}(j_{\min} - 1) = W_{\underline{\nu}}^{i+1}(j_{\min} + 1),$$

thus approximating the zero-derivative boundary condition at  $S_{\min}$ . The required option value is obtained once  $W_{\underline{\nu}^0}^0(0)$  is obtained. Note that not all values of  $W_{\underline{\nu}}^i(j)$  are required in order to achieve this. For instance,  $W_{\underline{\nu}}^i(j)$  with  $i+j$  odd never needs to be calculated. Moreover, only values with  $|j| \leq i$  can ever affect  $W_{\underline{\nu}^0}^0(0)$ .

**8.4. Convergence of the Value of the Swing Option.** Here we write  $w_{\underline{\nu}}^i(j) = w_{\underline{\nu}}(S_j, t_i^+)$ , with  $t_i = i\Delta t$ , and denote the difference between the result of the binomial calculation and the true solution of the partial differential equation model by  $e_{\underline{\nu}}^i(j) = W_{\underline{\nu}}^i(j) - w_{\underline{\nu}}^i(j)$ . Then we have the following global convergence result for the binomial forest method.

**Theorem 3** (Convergence). *Suppose that for each  $x \in [x_{\min}, \infty)$ ,  $w_{\underline{\nu}}(x, t)$  is continuous on  $(\tau_i, \tau_{i+1}]$  for  $i = 1, \dots, e$  and on  $[0, \tau_1]$ , and that  $\frac{\partial^2}{\partial t^2} w_{\underline{\nu}}(x, t)$  exists and is bounded on each of these intervals. Suppose also that for each  $t \in [0, T_2]$ ,  $w_{\underline{\nu}}(x, t)$  is three times continuously differentiable, with bounded derivatives up to order four on  $(x_{\min}, \infty)$ . Then there exists a constant  $C$  such that for each  $\underline{\nu}$ ,*

$$\max_i \max_j \left| e_{\underline{\nu}}^i(j) \right| \leq C\Delta t.$$

The remainder of this section will be taken up with the proof of Theorem 3. Substituting the vectors  $e_{\underline{\nu}}^i(j)$  into the binomial scheme yields, for  $j > j_{\min}$ ,

$$(8.11) \quad \begin{aligned} e_{\underline{\nu}}^{i,*}(j) &= p_j e_{\underline{\nu}}^{i+1}(j+1) + (1-p_j) e_{\underline{\nu}}^{i+1}(j-1) + T_{\underline{\nu}}^i(j), \quad i < N_2 n \\ e_{\underline{\nu}}^{i,*}(j) &= 0, \quad i = N_2 n \end{aligned}$$

$$(8.12) \quad e_{\underline{\nu}}^i(j) = e_{\underline{\nu}}^{i,*}(j), \quad i+1 \notin I_e,$$

$$(8.13) \quad \begin{aligned} e_{\underline{\nu}}^i(j) &= \max_{\epsilon} \left( W_{\underline{\nu}}^{i,*}(j), W_{\underline{\nu}^u}^{i,*}(j) + \text{Pay}_u(e^{x_j}), W_{\underline{\nu}^d}^{i,*}(j) + \text{Pay}_d(e^{x_j}) \right) - \\ &\quad \max_{\epsilon} \left( w_{\underline{\nu}}^{i,*}(j) + T_{\underline{\nu}}^i(j), w_{\underline{\nu}^u}^{i,*}(j) + T_{\underline{\nu}^u}^i(j) + \text{Pay}_u(e^{x_j}), \right. \\ &\quad \left. w_{\underline{\nu}^d}^{i,*}(j) + T_{\underline{\nu}^d}^i(j) + \text{Pay}_d(e^{x_j}) \right) \quad i+1 \in I_e, \end{aligned}$$

where the *truncation error*  $T_{\underline{\nu}}^i(j)$ , which will be the amount by which  $w_{\underline{\nu}}^i(j)$  fails to satisfy the equations defining the binomial approximation, is given by

$$(8.14) \quad T_{\underline{\nu}}^i(j) = w_{\underline{\nu}}^{i,*}(j) - w_{\underline{\nu}}^i(j),$$

and

$$w_{\underline{\nu}}^{i,*}(j) = p_j w_{\underline{\nu}}^{i+1}(j+1) + (1-p_j) w_{\underline{\nu}}^{i+1}(j-1) = W_{\underline{\nu}}^{i,*}(j) - e_{\underline{\nu}}^{i,*}(j)$$

for  $j > j_{\min}$ . To completely specify the error, we need the boundary equation

$$(8.15) \quad e_{\underline{\nu}}^i(j_{\min}) = e_{\underline{\nu}}^{i+1}(j_{\min}+1) + T_{\underline{\nu}}^i(j_{\min}),$$

where

$$(8.16) \quad T_{\underline{\nu}}^i(j_{\min}) = w_{\underline{\nu}}^{i+1}(j_{\min}+1) - w_{\underline{\nu}}^i(j_{\min}).$$

We shall restrict our attention to relevant values of  $j$ : for each  $i$  we only consider  $j \in J_i$  with

$$J_i = \{j : j+i \text{ even}, \quad j \geq j_{\min}, \quad |j| \leq i\}.$$

It will be shown inductively that the maximum error  $e^i = \max_{\nu, j \in J_i} \left| e_{\underline{\nu}}^i(j) \right|$  is  $O(\Delta t)$ . In order to do this, it is necessary to get a handle on the truncation error terms.

From (8.16), using the zero-derivative boundary condition together with Taylor-series expansions of  $w_{\underline{\nu}}(x, t)$ , we have

$$\begin{aligned} T_{\underline{\nu}}^i(j_{\min}) &= w_{\underline{\nu}}^{i+1}(j_{\min} + 1) - w_{\underline{\nu}}^i(j_{\min}) \\ &= (w_{\underline{\nu}}^{i+1}(j_{\min} + 1) - w_{\underline{\nu}}^{i+1}(j_{\min})) + (w_{\underline{\nu}}^{i+1}(j_{\min}) - w_{\underline{\nu}}^i(j_{\min})) \\ &= \frac{\Delta x^2}{2} \frac{\partial^2}{\partial x^2} w_{\underline{\nu}}(\xi_1, t_{i+1}) + \Delta t \frac{\partial}{\partial t} w_{\underline{\nu}}(x_{\min}, \xi_2), \end{aligned}$$

for some  $\xi_1 \in (x_{j_{\min}}, x_{j_{\min}+2})$  and  $\xi_2 \in (t_i, t_{i+1})$ . Given the bounded derivatives of  $w_{\underline{\nu}}$ , and the fact that  $\Delta x^2 = \Delta t/\sigma^2$ , one can conclude that there is a constant  $C_1$  such that

$$(8.17) \quad \left| T_{\underline{\nu}}^i(j_{\min}) \right| \leq C_1 \Delta t$$

for  $i < N_2 n$ .

We next consider the generic case  $j > j_{\min}, i + 1 \notin I_e$ , for which the truncation error involves the probabilities

$$p_j = \frac{1}{2} + \frac{(a + be^{-x_j})\Delta t}{2\Delta x} = \frac{1}{2} + \frac{\mu_j \Delta t}{\Delta x}.$$

We have, making use of the fact that  $w_{\underline{\nu}}(x, t)$  satisfies (8.7),

$$\begin{aligned} T_{\underline{\nu}}^i(j) &= p_j w_{\underline{\nu}}^{i+1}(j+1) + (1-p_j) w_{\underline{\nu}}^{i+1}(j-1) - w_{\underline{\nu}}^i(j) \\ &= w_{\underline{\nu}}^{i+1}(j) - w_{\underline{\nu}}^i(j) + \frac{1}{2} \left( w_{\underline{\nu}}^{i+1}(j+1) - 2w_{\underline{\nu}}^{i+1}(j) + w_{\underline{\nu}}^{i+1}(j-1) \right) \\ &\quad + \frac{\mu_j \Delta t}{2\Delta x} \left( w_{\underline{\nu}}^{i+1}(j+1) - w_{\underline{\nu}}^{i+1}(j-1) \right) \\ &= \Delta t \frac{\partial}{\partial t} w_{\underline{\nu}}(x_j, t_i) + \frac{\Delta t^2}{2} \frac{\partial^2}{\partial t^2} w_{\underline{\nu}}(x_j, \xi_3) + \frac{\Delta x^2}{2} \frac{\partial^2}{\partial x^2} w_{\underline{\nu}}(x_j, t_i) \\ &\quad + \frac{\Delta x^4}{12} \frac{\partial^4}{\partial x^4} w_{\underline{\nu}}(\xi_4, t_i) + \mu_j \Delta t \frac{\partial}{\partial x} w_{\underline{\nu}}(x_j, t_i) + \mu_j \Delta t \Delta x^2 \frac{\partial^2}{\partial x^2} w_{\underline{\nu}}(\xi_5, t_i) \\ &= \frac{\Delta t^2}{2} \frac{\partial^2}{\partial t^2} w_{\underline{\nu}}(x_j, \xi_3) + \frac{\Delta x^4}{12} \frac{\partial^4}{\partial x^4} w_{\underline{\nu}}(\xi_4, t_i) + \mu_j \Delta t \Delta x^2 \frac{\partial^2}{\partial x^2} w_{\underline{\nu}}(\xi_5, t_i) \end{aligned}$$

for some  $\xi_3 \in (t_i, t_{i+1})$  and  $\xi_4, \xi_5 \in (x_j, x_{j+1})$ . It follows from the boundedness of the derivatives of  $w_{\underline{\nu}}^i$  that

$$(8.18) \quad \max_{\underline{\nu}} \max_{j > j_{\min}} \left| T_{\underline{\nu}}^i(j) \right| \leq C_2 \Delta t^2, i < N_2 n.$$

Let

$$e^i = \max_{\underline{\nu}} \max_{j \geq j_{\min}} \left| e_{\underline{\nu}}^i(j) \right|.$$

Then (8.18), (8.17) and (8.11) imply that for  $i + 1 \notin I_e$ ,

$$(8.19) \quad e^i \leq \max(e^{i+1} + C_2 \Delta t^2, C_1 \Delta t).$$

In order to obtain a similar result for  $i + 1 \in I_e$  we bound the error  $e_{\underline{\nu}}^i(j)$ . By the Lipschitz continuity of the maximum function, there is a constant  $C_3$  such that

$$|\max_{\epsilon}(a, b, c) - \max_{\epsilon}(a', b', c')| \leq C_3 \max(|a - a'|, |b - b'|, |c - c'|)$$

for all  $a, b, c, a', b', c' \in \mathbb{R}$ . Then it follows from (8.13) that, when  $i + 1 \in I_e$  and  $j > j_{\min}$ ,

$$\left| e_{\underline{L}}^i(j) \right| \leq C_3 \max \left( \left| e_{\underline{L}}^{i+1}(j) \right| + \left| T_{\underline{L}}^i(j) \right|, \left| e_{\underline{L}^u}^{i+1}(j) \right| + \left| T_{\underline{L}^u}^i(j) \right|, \left| e_{\underline{L}^d}^{i+1}(j) \right| + \left| T_{\underline{L}^d}^i(j) \right| \right),$$

with obvious changes for the cases  $n_u = 0$  and  $n_d = 0$ , so that

$$(8.20) \quad e^i \leq C_3 \max (e^{i+1} + C_2 \Delta t^2, C_1 \Delta t).$$

Comparing (8.19) and (8.20), we find that they differ only by the extra factor of  $C_3$  when  $i+1 \in I_e$ . For each  $i$ , define  $k(i)$  to be the size of the set  $\{i \geq k : i+1 \in I_e\}$ . Then we make the inductive claim

$$e^i \leq C_3^{k(i)} (C_1 \Delta t + C_2 \Delta t^2 (nN_2 - i)).$$

Since  $e^{N_2 n} = 0$ , the inductive hypothesis holds for  $i = N_2 n$ . Suppose that it holds for  $i = i' + 1$ . Then if  $i' + 1 \in I_e$ ,

$$\begin{aligned} e^{i'} &\leq C_3 \max (e^{i'+1} + C_2 \Delta t^2, C_1 \Delta t) \\ &\leq C_3 C_3^{k(i'+1)} (C_1 \Delta t + C_2 \Delta t^2 (nN_2 - i' - 1) + C_2 \Delta t^2) \\ &= C_3^{k(i')} (C_1 \Delta t + C_2 \Delta t^2 (nN_2 - i')) \end{aligned}$$

so that the inductive step is complete in this case. The proof in the case  $i' + 1 \notin I_e$  is exactly similar, and so the theorem is proved, with

$$C = C_3^{k(0)} (C_1 + C_2 T_2).$$

## REFERENCES

- [BG, 1996] Berbieri, Angelo and Garman, Mark B. October 1996. Putting a Price on Swings. *Energy & Power Risk Management*.
- [G, 1997] Gardiner, Crispin W. 1997. *Handbook of Stochastic Methods*. Springer.
- [JRT, 1998] Jaillet Patrick, Ronn Ehud I. and Tompaidis Stathis. 1998. *Modeling Energy Prices and Pricing and Hedging Energy Derivatives*. Preprint.
- [H, 1999] Hull, John. 1999. *Options, Futures and Other Derivative Securities*. Prentice Hall.
- [HR, 1998] Hilliard Jimmy E and Reis Jorge. 1998. *Valuation of Commodity Futures and Options under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot*. J. of Financial and Quantitative Analysis, Vol. 33, No. 1, pp. 61-86.
- [HS, 1974] Hirsch Morris and Smale Stephen. 1974. *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press.
- [KP, 1999] Kloeden Peter E. and Platen Eckhard. 1995. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag.

- [K, 1998] Kwok Yue-Kuen. 1998. *Mathematical Models of Financial Derivatives*. Springer-Verlag.
- [P, 1997] Pilipovic Dragana. 1997. *Valuing and Managing Energy Derivatives*. McGraw-Hill.
- [R, 1995] Robinson Clark. 1995. *Dynamical Systems*. CRC Press.
- [S, 1998] Schwartz Eduardo S. 1998. *Valuing Long-Term Commodity Assets*. Financial Management, Vol. 27, No. 1, pp. 57-66.
- [W, 1998] Wilmott, Paul. 1998. *Derivatives: The Theory and Practice of Financial Engineering*. John Wiley & Sons.

THE MATHEMATICAL AND COMPUTATIONAL FINANCE LABORATORY. DEPARTMENT OF MATHEMATICS AND STATISTICS. UNIVERSITY OF CALGARY. CALGARY, ALBERTA N1G 4Z4 CANADA

*E-mail address:* `lavassan@math.ucalgary.ca`

`simchi@math.ucalgary.ca`

`ware@math.ucalgary.ca`