On Pricing Derivatives in the Presence of Auxiliary State Variables

J. Lin [*]   P. Ritchken[†]

September 25, 2006

---

[*]Bear Stearns & Co. Inc, 383 Madison Ave., New York, NY 10179, Tel: (212) 272-7225, E-mail: jlin@bear.com

[†]Department of Banking and Finance,Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Ave., Cleveland, OH 44106-7235, Tel: (216) 368–3849, E-mail: phr@po.cwru.edu

Abstract

This article investigates the pricing of options when a need arises to carry a path dependent auxiliary state variable. Examples of such problems include the pricing of interest rate claims in a Heath Jarrow Morton paradigm, where the underlying forward rates follow a Markovian process, and the pricing of equity options, when the underlying asset price follows a GARCH process. In the former case, the primary state variable is the spot interest rate, and the auxiliary state variable is the accrued variance of the current spot rate. In the latter case, the primary state variable is the asset price, and the auxiliary state variable is the path dependent statistic representing local volatility. An efficient algorithm is developed for pricing claims under these type of processes. Illustrative examples are presented that demonstrate the efficiency of the algorithm and conditions are developed that ensure the algorithm will produce accurate prices.

This article is concerned with the efficient pricing of contingent claims when a need arises to carry a path dependent auxiliary state variable. These problems are often encountered when pricing specialized derivative claims such as average rate options, where the auxiliary state variable is the average price to date, or lookback options, where the auxiliary state variable might be the maximum price to date. There are other problems, however, where the auxiliary variable arises more naturally in the underlying dynamics of the process rather than being defined by the nature of the financial claim. For example, consider pricing interest rate claims under the assumption that volatilities of forward rates belong to a one factor Heath Jarrow Morton [1992] (hereafter HJM) family, where bond prices can be represented by a low dimensional Markovian system. In such models, the primary underlying state variable is the short rate and the second, auxiliary state variable, is a path statistic that corresponds to the accrued variance of the spot rate. As a second example, consider the pricing of an equity derivative when the underlying dynamics are given by a GARCH process. Here the primary state variable is the stock price and the second auxiliary state variable is a path statistic, representing the local volatility.

Several authors have established algorithms for implementing option models with auxiliary state variables. Most methods rely on a lattice scheme for the primary state variable, and the auxiliary state variable is represented by an array of values at each node in the tree. Examples of this approach include Ritchken, Sankarasubramanian, and Vijh [1993], Hull and White [1993a], and Barraquand and Pudet [1996].[1] In most applications, the number of possible values of the auxiliary variable at a node equals the number of distinct paths to the node. In this case, it is not possible to track every possible value of the auxiliary variable, and instead a grid of representative values is established.[2] Option pricing then proceeds at each point in the grid, for each node in the lattice. Typically, it is necessary to interpolate option prices at successor points when solving backwards through the tree. Li, Ritchken, and Sankarasubramanian [1995](hereafter LRS) apply this approach to pricing Markovian interest rate models in the HJM family, where volatility varies with the level of rates, and Ritchken and Trevor [1999] apply similar ideas to price options under GARCH processes, where volatility is persistent and path dependent. For the most part, formal proofs of convergence of prices in these papers have been absent, but illustrations of convergence of prices have been provided, as the lattice (for the primary variable) and the grid (for the auxiliary variable) have been refined.

In this article we focus primarily on the efficient implementation of the low dimensional Markovian HJM models that were described by Ritchken and Sankarasubramanian [1995]. They identified the family of volatility structures for forward rates that permitted a HJM models to be represented by a two state Markovian system, thereby allowing American claims to be priced, without the need to use bushy trees. Under their restrictions, the dynamics of the entire term structure depends on the short rate and a second auxiliary state variable that can be viewed as an accrued variance statistic. Li, Ritchken, and Sankarasubramanian [1995] provide an algorithm

for pricing interest rate sensitive claims, under any of the volatility structures in this restricted HJM family. Our main contribution here will be to develop a new algorithm which is an order of magnitude more efficient than the LRS algorithm.

While our primary focus is to price interest sensitive claims, we also consider applications involving auxiliary state variable methods to pricing equity options under GARCH processes. We develop an algorithm that can be used to price claims under the NGARCH model of Duan [1995]. In light of the connections between GARCH and bivariate stochastic volatility, we can also use this algorithm to obtain American option prices for claims under an array of bivariate stochastic volatility processes.

The paper proceeds as follows. In section 1 we review HJM models where we focus on their Markovian family and the current procedure for pricing such claims. In section 2 we present our algorithm, and prove that it is capable of producing reliable prices. Then, in section 3, we illustrate the algorithm by considering the convergence rate of prices. In section 4 we consider a second application involving pricing equity option contracts under GARCH and stochastic volatility processes. Section 5 concludes.

1   Markovian HJM Models

Heath, Jarrow and Morton [1992] have developed a paradigm for pricing interest rate claims relative to a given term structure, that requires a specification of the volatility structure of forward rates. Given a volatility function, a unique martingale measure can be identified under which all interest rate claims can be priced as if the local expectations hypothesis holds. Let $f(t, T)$ be the forward rate at date t for instantaneous riskless borrowing and lending at date T. The evolution of forward rate under this martingale measure is given by

$$df(t, T) = \mu_f(t, T)dt + \sigma_f(t, T)dw(t)$$

with $f(0, t)$ given for all t, and, as shown by HJM, the drift term is uniquely determined by the no arbitrage condition :

$$\mu_f(t, T) = \sigma_f(t, T) \int_t^T \sigma_f(t, s)ds.$$

Unfortunately, the computation of prices is typically complex because the evolution of the term structure under the martingale measure is usually not Markovian with respect to a finite dimensioned state space. As a result, lattice based numerical procedures for pricing usually results in trees that grow exponentially with the number of time periods. Simulation models are often employed to solve typical pricing problems such as pricing claims on a 10 year quarterly deal, where there might be 40 cash flow dates, and hence at least 40 model dates. While simulation models can be reasonably effective, when pricing American claims, there still remain

some difficulties.[3] A major hurdle associated with implementing models in the HJM paradigm involves the estimation of parameters. The parameters are typically implied out from actively traded volatility products such as caps and swaptions. Since the highly non linear estimation procedure requires frequent calls to numerical routines that price these products, these routines need to be extremely efficient. The presence of path dependency here can create a severe hinderence to a successful implementation. Recently, a family of models have been developed that avoids some of the calibration issues without resolving the path dependency problems. These models, collectively referred to as *market* models, preclude arbitrage among bonds, keep rates positive, and non exploding, and price caps or swaptions according to the Black model. These models belong to the HJM family, but utilize discrete compounding, rather than continuous compounding, and limit themselves to a set of rates (maturities) that are commonly quoted in financial markets. The initial market models incorporate lognormal volatilities, but extensions to permit fatter tailed distributions have been made.[4] Market models permit easy calibration to caps or swaptions, but they do not resolve the path dependence problem. As a result, valuation of general interest rate claims remains complex because of the total path dependence.

It turns out that in the HJM framework there is an important family of volatility structures for which the dynamics of forward rates can be represented by a finite dimensioned Markovian system. Ritchken and Sankarasubramanian [1995] have identified necessary and sufficient conditions on volatility structures which do not completely remove the path dependence, but rather capture it by a single sufficient statistic. Specifically, they focus on classes of the form:

$$\sigma_f(t,T) = \sigma_f(t,t)e^{-\int_t^T \kappa(x)dx}.$$

Here $\kappa(x)$ is some exogenously provided deterministic function and $\sigma_f(t,t)$ is the volatility of spot rate at date t, which could depend on the entire set of term structure information at that date. The class of volatilities that permit this representation is quite large. When $\sigma_f(t,t) = \sigma$, and when $\kappa(x) = \kappa$, the model reduces to the Generalized Gaussian model studied by Vasicek [1977], HJM and by Hull and White [1993]. When $\sigma_f(t,t) = \sigma\sqrt{r(t)}$, then the short rate will be of a form similar to the Cox, Ingersoll, and Ross [1985] square root model, and volatilities of all forward rates will fluctuate according to the level of the short rate. The deterministic function, $\kappa(x)$ can be used to calibrate prices of caps or swaptions so that their theoretical prices equal those observed in the marketplace, or can be chosen to be a constant. The exact functional form for $\sigma_f(t,t)$ is important and is restricted to being some predictable function. Interestingly, the Markovian class imposes no particular restrictions on the structure of the spot rate volatility, and hence permits volatilities to fluctuate according to the level of the spot rate. In this paper, for simplicity, we will assume the volatility of the spot rate has constant elasticity, or:

$$\sigma_f(t,t) = \sigma[r(t)]^\gamma. \qquad \gamma \geq 0.$$

Of course, when $\gamma = 0$, the path statistic becomes a deterministic function, and the model

reduces to a one state variable model that permits simple analytical solutions for many European claims. With $\gamma > 0$, this type of volatility function allows the entire future term structure to be represented by a Markovian model with two state variables, namely the short rate, $r(t)$, and the auxiliary path dependent statistic, $\Phi(t) = \int_0^t \sigma_f^2(s,t)ds$. The dynamics of the state variables are:

$$dr(t) = \mu(r, \Phi, t)dt + \sigma_f(t, t)dw(t) \tag{1}$$

$$d\Phi(t) = (\sigma_f^2(t, t) - 2\kappa(t)\Phi(t))dt, \tag{2}$$

where

$$\mu(r, \Phi, t) = \kappa(t)[f(0, t) - r(t)] + \Phi(t) + \frac{d}{dt}f(0, t),$$

and analytical solutions exist that permit the entire term structure to be recovered at any date, given the state variables.[5] Given a Markovian representation for the dynamics of the term structure, American claims can be priced without the need to build bushy trees. An algorithm that accomplishes this is provided by Li, Ritchken and Sankarasubramanian [1995]. They first use the following transformation on the spot rate process (1) to get a form with constant volatility.

$$Y(t) = \int \frac{1}{\sigma_f(t, t)}dr(t). \tag{3}$$

Let $r(t) = h(Y(t))$ be the inverse function. Then the new state variables are $Y(t)$ and $\Phi(t)$ with dynamic

$$dY(t) = m(Y, \Phi, t)dt + dw(t) \tag{4}$$

$$d\Phi(t) = n(Y, \Phi, t)dt \tag{5}$$

where

$$m(Y, \Phi, t) = \frac{\partial Y(t)}{\partial t} + \mu(r, \Phi, t)\frac{\partial Y(t)}{\partial r(t)} + \frac{1}{2}\sigma_f^2(t, t)\frac{\partial^2 Y(t)}{\partial r(t)^2}$$

$$n(Y, \Phi, t) = \sigma_f^2(t, t) - 2\kappa(t)\Phi(t).$$

Notice that both $m(Y, \Phi, t)$ and $n(Y, \Phi, t)$ are linear in Φ. We will use this property to construct a recombining lattice in next section.

Since there is only one source of uncertainty and Φ is a path statistic, if we follow the traditional tree building approach, we will get a one dimensional bushy tree. However, notice that the instantaneous volatility of $Y(t)$ is constant, 1. At each slice of time t, there would be several nodes having the same value of $Y(t)$. If we aggregate those nodes with the same $Y(t)$ together, the topology of the tree would look like a reconnecting trinomial tree. But keep in mind that at each node there are many $\Phi(t)$ values associated with the same $Y(t)$ (as many as the number of distinct paths to reach $Y(t)$) and the transition probability for $(Y(t), \Phi(t))$ varies not only with $Y(t)$ but also with $\Phi(t)$.

4

As we roll forward through the tree, it is not possible to keep track of all the Φ values at each node. LRS use a forward scan through the tree to produce the minimum and maximum Φ values at each node. Li, Ritchken and Sankarasubramanian [1995] use a grid of K equally spaced values between the maximum and minimum to represent the state variables at each node. Given, the state variables at any node, an analytical solution for the term structure exists, and at the expiration date, boundary prices for the option contract can be established. The tree is then rolled back, with appropriate interpolations used, to recover the date 0 price. LRS show that in order to get convergence in prices it might be necessary to use a rather fine partition for the Φ variable. That is K may have to be fairly large. Our algorithm reduces the bushyness of their approach.

2 A New Model for Markovian HJM Models

In this section, we first build a reconnecting lattice where, at each node, the grid of values for the possible values of the auxiliary state variable, is replaced by a single value, representing the conditional expectation of the auxiliary state variable. In the second part of this section, we show how this "reduced lattice", which tracks less information than the original lattice, can be used to generate approximate prices for option contracts. Conditions are identified that ensure the resulting algorithm produces approximations that converge to the true prices as the time partition of the lattice is refined.

2.1 Construction of a Reduced Lattice

We start by building a trinomial lattice for the processes (4) and (5). At each time slice the Y values will be separated by a predetermined distance equal to $\eta\sqrt{\Delta t}$, where Δt is the time increment between time slices and η is a constant. If $\eta = 1$ then the trinomial lattice will reduce to the well known Binomial lattice.[6] Let y_t and ϕ_t denote specific values of the variables Y_t and Φ_t at time slice t, and let y_{t+1} and ϕ_{t+1} denote values of the two state variables at time slice $t + \Delta t$. The local drift and volatility at a node (y_t, ϕ_t) will not only depend on the value of y_t but also on ϕ_t. According to the lattice, from node (y_t, ϕ_t), we will move to nodes $(y_{t+1}^+, \phi_{t+1}^+)$, (y_{t+1}, ϕ_{t+1}) and $(y_{t+1}^-, \phi_{t+1}^-)$ (see Exhibit 1), where $y_{t+1} = y_t$ and $y_{t+1}^+ - y_t = y_t - y_{t+1}^- = \eta\sqrt{\Delta t}$, with probabilities:

$$\begin{cases} p_+(y_t, \phi_t, t) &= \frac{1}{2\eta^2} + \frac{m(y_t, \phi_t, t)\sqrt{\Delta t}}{2\eta} \\ p_0(y_t, \phi_t, t) &= 1 - \frac{1}{\eta^2} \\ p_-(y_t, \phi_t, t) &= \frac{1}{2\eta^2} - \frac{m(y_t, \phi_t, t)\sqrt{\Delta t}}{2\eta} \end{cases} \tag{6}$$

The variable η can be fixed at any value larger than 1. It controls the partition size between two consecutive y values on the lattice. It also determines the probabilities. For example, if

$\eta = \sqrt{3/2}$, the probability of a horizontal jump is one third.

Exhibit 1 Here

Since the process for ϕ is locally deterministic (see equation (5)), the values ϕ_{t+1}^+, ϕ_{t+1} and ϕ_{t+1}^- are equal and fully determined by current state variables (y_t, ϕ_t). That is:

$$\phi_{t+1}^+ = \phi_{t+1} = \phi_{t+1}^- = n(y_t, \phi_t, t)\triangle t + \phi_t \qquad (7)$$

Let $\phi_{t+1}^* = n(y_t, \phi_t, t)\triangle t + \phi_t$ represent this common successor value. Because both $m(y, \phi, t)$ and $n(y, \phi, t)$ are linear in ϕ, we see that p_+, p_0, p_- and ϕ_{t+1}^* are all linear with respect to ϕ_t.

In summary, then, our lattice consists of a set of path reconnecting "nodes" for the y values. For a given node, at each date, the number of possible ϕ values equals the number of paths to the node. We will refer to this lattice as the bushy lattice.

Our first phase of the algorithm is to replace the second state variable, namely the path statistic, with a new state variable that can be used to create a new lattice that is not bushy. Hopefully, this can be done without a significant loss of information. Towards, this goal, we shall replace the second state variable, Φ_t, by its conditional expectation. Specifically, let

$$\bar{\phi}_t(y_t) = E_t[\Phi_t | Y_t = y_t]$$

represent the conditional expectation. Further, we denote $\bar{\phi}_t = \bar{\phi}_t(y_t)$ if its meaning in the context is clear, and we shall refer to the lattice of values generated by $(y_t, \bar{\phi}_t)$ as the reduced lattice. The idea here is that, conditional on $Y_t = y_t$, the distribution of Φ_t values is concentrated around its expectation, and deviations from this expected value will have small effects on pricing. This intuition is formalized later on, where convergence issues are addressed.

It turns out that computing these exact conditional expectations, and hence setting up the reduced lattice, can be done without explicitly requiring information from the original bushy lattice! Moreover, given these values up to date t, the exact probability of reaching a particular node at date $t+1$ can also be computed exactly, without considering all the individual paths that lead to the node in the bushy lattice. These results are summarized in Proposition 1 below.

Proposition 1

(i) The exact probability of reaching an individual node at time $t+1$, in the bushy tree, ie. $P(Y_{t+1} = y_{t+1})$, can be computed given information on the reduced tree up to date t. Specifically:

$$P(Y_{t+1} = y_{t+1}) = \sum_{y_t} P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \bar{\phi}_t)P(Y_t = y_t) \qquad (8)$$

where $P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \bar{\phi}_t)$ is given by (6) if $|y_{t+1} - y_t| \le \eta\sqrt{\triangle t}$ and 0 otherwise;

6

(ii) The exact conditional expectations on the bushy tree at time $t + 1$, $\bar{\phi}_{t+1}(y_{t+1})$, can be computed recursively from information on the reduced tree at date t. Specifically:

$$\bar{\phi}_{t+1}(y_{t+1}) = \sum_{y_t} \phi^*_{t+1}(y_t, \bar{\phi}_t(y_t), t) P(Y_t = y_t | Y_{t+1} = y_{t+1}) \tag{9}$$

where

$$P(Y_t = y_t | Y_{t+1} = y_{t+1}) = \frac{P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \bar{\phi}_t) P(Y_t = y_t)}{P(Y_{t+1} = y_{t+1})}, \tag{10}$$

and $\phi^*_{t+1}(y_t, \phi_t, t)$ is given by (7).

Proof : see Appendix

Equation $(8) - (10)$ are exactly the natural ways to update information if we only had one Φ_t (that is $E[\Phi_t | Y_t]$) associated with each node. Equation (8) says the probability of reaching node y_{t+1} comes from contributing nodes $(y_t, \bar{\phi}_t(y_t))$ at time t; equation (10) is the portion of the probability at node y_{t+1} that comes from node $(y_t, \bar{\phi}_t(y_t))$; and equation (9) says the mean value of Φ_{t+1} at node y_{t+1} is the weighted average of Φ_{t+1} values that come from nodes $(y_t, \bar{\phi}_t(y_t))$. Therefore, by using equation $(8) - (10)$, one can use the reduced lattice to establish exactly the same results that we would obtain using the full bushy tree.

The key to this proof relies on the fact that the transition probabilities $p_+(Y_t, \Phi_t, t)$, $p_0(Y_t, \Phi_t, t)$, $p_-(Y_t, \Phi_t, t)$ and the Φ evolution function are all linear in Φ_t. As a result, this updating scheme can work for higher conditional moments $E[\Phi_t^k | Y_t]$. Hence, the distribution of the random variable $\Phi_t | Y_t = y_t$, can, if necessary, be better approximated using higher moments as additional state variables. As we shall see, however, in this application, such higher moments are unnecessary.

2.2   Construction of Option Prices on the Reduced Lattice

In this section, we see how option prices can be approximated on the reconnecting tree $(y_t, \bar{\phi}_t)$ we have just constructed. Given the state variables at any node, one can reconstruct the term structure, and at the expiration date, boundary prices for the option contract can be established. We will use dynamic programming, rolling back on the tree, to get the option price at date 0. Because we are losing information by keeping only the conditional expectation of Φ, we can't expect to get an exact option value by applying the traditional backward induction. We will, however, start from the usual one period backward recursive approach, and make modifications as we proceed.

Since (y_t, ϕ_t) are state variables, the option value at the expiration date T, say, can be expressed as a function $g(y_T, \phi_T)$. If we had computed all $g(y_{t+1}, \phi_{t+1})$ at time $t + 1$, the backward induction at time t, on the node (y_t, ϕ_t) shown on the left panel of Exhibit 1 is

$$g(y_t, \phi_t) = e^{-r_t \triangle t} [p_+ g(y_{t+1}^+, \phi_{t+1}^+) + p_0 g(y_{t+1}, \phi_{t+1}) + p_- g(y_{t+1}^-, \phi_{t+1}^-)] \tag{11}$$

where $r_t = h(y_t)$; p_+, p_0, p_- are given by (6) and all the successor values of the second state variable, namely ϕ_{t+1}^+, ϕ_{t+1} and ϕ_{t+1}^- are equal to a common value, ϕ_{t+1}^*, as given by (7).

Unfortunately, on our lattice, the option values at the successor nodes are not available. But recall that we used (y_{t+1}, ϕ_{t+1}^1), (y_{t+1}, ϕ_{t+1}^2) and (y_{t+1}, ϕ_{t+1}^3), which came from $(y_t^+, \bar{\phi}_t^+)$, $(y_t, \bar{\phi}_t)$ and $(y_t^-, \bar{\phi}_t^-)$ respectively (see Exhibit 1), to compute $E[\phi_{t+1}|y_{t+1}]$.

If we had computed the option value at these three points at time $t+1$, we could use quadratic interpolation/extrapolation to approximate the date t option value $g(y_{t+1}, \phi_{t+1}^*)$. In particular,

$$
\begin{aligned}
\hat{g}(y_{t+1}, \phi_{t+1}^*) \;=\; & \frac{(\phi_{t+1}^* - \phi_{t+1}^2)(\phi_{t+1}^* - \phi_{t+1}^3)}{(\phi_{t+1}^1 - \phi_{t+1}^2)(\phi_{t+1}^1 - \phi_{t+1}^3)} g(y_{t+1}, \phi_{t+1}^1) + \\
& \frac{(\phi_{t+1}^* - \phi_{t+1}^1)(\phi_{t+1}^* - \phi_{t+1}^3)}{(\phi_{t+1}^2 - \phi_{t+1}^1)(\phi_{t+1}^2 - \phi_{t+1}^3)} g(y_{t+1}, \phi_{t+1}^2) + \\
& \frac{(\phi_{t+1}^* - \phi_{t+1}^1)(\phi_{t+1}^* - \phi_{t+1}^2)}{(\phi_{t+1}^3 - \phi_{t+1}^1)(\phi_{t+1}^3 - \phi_{t+1}^2)} g(y_{t+1}, \phi_{t+1}^3)
\end{aligned}
\tag{12}
$$

Putting these together, we would, at time t, have :

$$
g(y_t, \phi_t) = e^{-r_t \triangle t}[p_+ \hat{g}(y_{t+1}^+, \phi_{t+1}^*) + p_0 \hat{g}(y_{t+1}, \phi_{t+1}^*) + p_- \hat{g}(y_{t+1}^-, \phi_{t+1}^*)], \tag{13}
$$

where $r_t = h(y_t)$; p_+, p_0, p_- are given by (6); ϕ_{t+1}^* is from equation (7) and $\hat{g}(y_{t+1}^+, \phi_{t+1}^*)$, $\hat{g}(y_{t+1}, \phi_{t+1}^*)$ and $\hat{g}(y_{t+1}^-, \phi_{t+1}^*)$ are approximation option values at time $t+1$ given by (12).

Actually, Equation (13) is used to compute option prices at $g(y_t, \phi_t^1)$, $g(y_t, \phi_t^2)$ and $g(y_t, \phi_t^3)$ at node $(y_t, \bar{\phi}_t)$. Then the backward induction continues until we get the option price $g(y_0, \phi_0)$ at time 0.

The following proposition gives the convergence property of this algorithm.

Proposition 2

If the function $\bar{\phi}_t(y) = E[\Phi_t|Y_t = y]$ is uniformly Lipschitz in y, i.e., there exist a $M_1 > 0$, s.t.

$$|\bar{\phi}_t(y_1) - \bar{\phi}_t(y_2)| \le M_1|y_1 - y_2|, \;\; 0 < t < T$$

and $|\frac{\partial^i g(y,\phi)}{\partial \phi^i}| \le M_2$, $i = 1, 2, 3$, for some $M_2 > 0$, then the option price $g(y_0, \phi_0)$ computed by recursively applying (12)-(13) converges to the true option value as $\triangle t \to 0$.

Proof : see Appendix.

The bounded derivatives condition is the regular sufficient condition for any lattice model to converge to a continuous model. Since we are working around the conditional expectation $E[\Phi_t|Y_t]$ and using quadratic approximation, the uniformly Lipschitz condition ensures that the approximation error be small enough so that it won't distract the convergence property. Intuitively, this condition is quite weak and should hold in most practical cases.

The modification of the above backward induction to value American option is straightforward. In particular, the price given in equation (13) has to be compared to the exercise value of the claim, with the higher of the two values recorded. The convergence for American claims follows from Proposition 2.

3 An Illustrative Example

To illustrate the algorithm, we consider a three period example. The initial term structure is flat at 6%. The volatility structure is given by $\sigma_f(t,T) = \sigma r(t)^\gamma e^{-\kappa(T-t)}$ where $\sigma = 20\%$, $\gamma = 1$ and $\kappa = 2\%$. The time partition is one year. The value for η is set at $\sqrt{3/2} = 1.2275$, implying the probability of interest rates staying at their current level over a time increment is $1/3$ (see equation (6)). Since $\sigma_f(t,t) = \sigma r(t)$, from equation (3) we have $Y(t) = \frac{ln(r(t))}{\sigma}$. Hence $y_0 = -14.067$ and the gap between adjacent y-values is $\eta\sqrt{\Delta t} = 1.2275$. Given the y-lattice, the interest rate lattice can be recovered as $r_t = e^{\sigma y_t}$. Figure 2 shows the interest rate nodes over three years.

At each "node" of Exhibit 2 we record a matrix of values. In the last row of the matrix, the values of the conditional expectation, $\bar{\phi}$, and the probability of reaching this interest rate node, are recorded. The total probability of reaching a particular node come from the contributing nodes that preceed, but are connected to this node. These contributing probabilities are shown in the second column. For example, consider the interest node $r = 6\%$ in year 2. The probability of reaching this node is 0.3381. There are three paths that lead to this node. The first path is from the upper node in period 1, and this path contributes 0.1109 of the probability; the second path is the horizontal path, and it contributes 0.1111; the final path is through the node $r = 4.70\%$ in year 1, and it contributes the remaining 0.1160. The first column, other than the last row, keeps track of the successor ϕ values from the proceeding connected nodes. For example, in year 2, with the interest rate at 6%, there are three ϕ values recorded. The top value represents the successor value from the $\bar{\phi}$ value in the top node in period 1. Similarly, the second and third values represent the successor values from the $\bar{\phi}$ values in the middle and bottom nodes in the first year. The computations in the lattice proceed forward in time from the original node to maturity.

Insert Exhibit 2 Here

Phase 1 of the algorithm is complete. We have the interest rate nodes and the relevant values of ϕ at each node, for which option prices will be computed.

The second phase begins by recoding the terminal value of the claim at the expiration date, which in our case is year three. In particular, we consider the pricing of a five year discount bond

9

with face value $1000. The strike price is set at the current forward price of $718.44. For each setting of the state variables, at the expiration date, the entire term structure can be computed analytically. In particular, from Ritchken and Sankarasubramanian (1995) we have at date t, contingent on the state variables $(r(t), \phi(t))$, the price of a bond with payout $1 at date T, is given by:

$$P(t,T) = \frac{P(0,T)}{P(0,t)} e^{-\beta(t,T)(r(t)-f(0,t))-\beta^2(t,T)\phi(t)/2}$$

where $\beta(t,T) = \frac{1}{\kappa}(1 - e^{-\kappa(T-t)})$. Given the bond price at expiration and the strike price, the terminal option prices at expiry can be computed. Given these boundary values, we now use backward recursion and interpolation to compute the option prices in previous years. In Exhibit 3, the values of the underlying discount bond and the option price at expiry are shown next to the ϕ values at each interest rate node.

Insert Exhibit 3 Here

Exhibit 3 also shows the values of the option prices on the lattice. For example, consider the computations at the 6% interest rate node in year two, where the state variable $\phi = 2.8224$. The probabilities of an up, horizontal, and down jump, computed using equation (6) are 0.3021, 0.3333, and 0.3646 respectively. The successor ϕ value is 4.1495. Going to the next period, where interest rates could be 7.67%, 6%, or 4.70%, the option prices need to be identified. The interpolated values at the three successor values are 0, 0, and 43.71 respectively. Given the risk neutral probabilities and successor option values, the option price of 15.01 can be computed. Through repeated backward recursions, the date 0 value of this claim is $23.27.

Unlike the original algorithm of Li, Ritchken and Sankarasubramanian [1995], this algorithm only requires at most three values of ϕ at each interest rate node. This is in contrast to their algorithm in which more than 25 values of the ϕ values at each node are typically necessary. With just three values of the auxiliary variable at each node, the overall computational effort is about 3 times slower than computing prices of options when there are no auxiliary variables.

4 Computational Experience

To illustrate the convergence behavior we consider the prices of European and American claims in a model where the initial yield curve is flat at 6%, and the volatility structure of forward rates is given by

$$\sigma_f(t,T) = \sigma r(t)^\gamma e^{-\kappa(T-t)}.$$

Exhibit 4 shows the convergence property of the proposed algorithm. Here we take $\gamma = 0.5$, $\sigma = 0.02$ and $\kappa = 0.01$. Two claims are used in the experiment: one is a 5-year call option on

a 10 year discount bond and the other is a 1 year call option on a 3 month spot rate. Both European and American style contracts are considered. Simulation prices for European options are generated using $100,000$ paths with 500 time partitions for each path, and the resulting 95% confidence intervals for the price provided a benchmark for establishing the precision of our model prices.

<center>Exhibit 4 Here</center>

Exhibit 4 shows that European prices generated by our algorithm are consistent with the simulated prices.

Exhibit 5 tests the model with all kinds of different parameters. We consider a 6 month European call option on a 3 month spot rate. Strike rate are set at 95%, 100% and 105% of at the money strike. The parameters are the elasticity value, γ, the mean reversion parameter, κ and the initial volatility of the spot rate, σ_0, where $\sigma_0 = \sigma r_0^\gamma$. These values were varied over an extremely large range of values compared to their levels that are encountered in practice. The prices produced by the algorithm are compared to prices generated using Monte Carlo simulation.

<center>Exhibit 5 Here</center>

$20,000$ sample paths are used to generate the simulation prices. Of the 144 cases reported, there are 5 cases where the interval did not contain the price produced by the algorithm, which is somewhat consistent with the stated 95% confidence. Indeed, all the lattice prices differed from the simulation prices by no more than one half a basis point, which in practice, is already well within the bid-ask spread.

The construction of the tree takes a little additional work beyond the regular trinomial model but is about the same as the forward scan procedure used to find the minimum and maximum values of ϕ at each node in the LRS model. The prices produced by this algorithm are computed in a fraction of the time it takes to compute prices by the LRS model. The computational effort is about the same as the LRS algorithm with the number of partitions for the auxiliary variable at each node, K, equaling 3. As K in the LRS model increases, the computational time increases at a rate much greater than the rate of increase in K. Indeed, as the number of partitions, K, increases, significant time is spent in identifying the successor ϕ values at which option prices need to be interpolated. $K \geq 25$ is a typical value recommended by Li, Ritchken and Sankarasubramanian [1995].

<center>11</center>

5 NGARCH and Bivariate Stochastic Volatility Option Models

Our algorithm can be used in other applications involving auxiliary variables. In this section we consider using the algorithm to price options under a generalized GARCH process considered by Duan [1995]. The conditional variance process of the limiting diffusion for this model is a process which includes that used by Hull and White [1987] in their stochastic volatility option pricing model. In this application the primary variable will be the logarithmic stock price, and the auxiliary state variable will be the local volatility, the value of which is determined by the path up to the current date.

Partition each time period ("day") into "trading periods" of width Δt. Label these trading periods by consecutive integers starting from the beginning of the current period, period 0. Let y_t be the logarithmic price at the beginning of the tth trading period, and let ϕ_t be the variance for this trading period. Assume y_0 and ϕ_0 are given. The dynamics of the return generating process are

$$
\begin{align}
y_{t+1} &= y_t + (r_f + \lambda\sqrt{\phi_t} - \phi_t/2)\Delta t + \sqrt{\phi_t}\sqrt{\Delta t}v_{t+1} \tag{14}\\
\phi_{t+1} &= \phi_t + \beta_0\Delta t + \phi_t[\beta_1 + \beta_2 q - 1]\Delta t + \phi_t\beta_2\sqrt{\Delta t}[(v_{t+1} - c)^2 - q], \tag{15}
\end{align}
$$

where $q = (1 + c^2)$, λ is the market price of risk, and $\{v_t \mid t = 0, 1, 2, \ldots\}$ is a sequence of independent standard normal random variables. Notice that when $\Delta t = 1$ this equation reduces to the NGARCH model considered by Duan [1995]. Duan [1997] shows through an equilibrium argument that a local risk neutralized valuation relationship holds, in which options can be priced as if the economy were risk neutral under the pricing measure

$$
\begin{align}
y_{t+1} &= y_t + (r_f - \phi_t/2)\Delta t + \sqrt{\phi_t}\sqrt{\Delta t}\epsilon_{t+1} \tag{16}\\
\phi_{t+1} &= \phi_t + \beta_0\Delta t + \phi_t[\beta_1 + \beta_2 q - 1]\Delta t + \phi_t\beta_2\sqrt{\Delta t}[(\epsilon_{t+1} - c - \lambda\sqrt{\Delta t})^2 - q], \tag{17}
\end{align}
$$

where $\{\epsilon_t \mid t = 0, 1, 2, \ldots\}$ is a sequence of independent standard normal random variables.

Duan [1997] shows that the limiting diffusion of the process for the price of the underlying security in equations (14) and (15) is

$$
\begin{align}
dy_t &= (r_f + \lambda\sqrt{\phi_t} - \phi_t/2)dt + \sqrt{\phi_t}dW_1(t) \tag{18}\\
d\phi_t &= [\beta_0 + (\beta_1 + \beta_2 q - 1)\phi_t]dt - 2c\beta_2\phi_t dW_1(t) + \sqrt{2}\beta_2\phi_t dW_2(t) \tag{19}
\end{align}
$$

where $dW_1(t)$ and $dW_2(t)$ are independent Wiener processes. The limiting diffusion under the locally risk neutralized measure is shown to be

$$
\begin{align}
dy_t &= (r_f - \phi_t/2)dt + \sqrt{\phi_t}dZ_1(t) \tag{20}\\
d\phi_t &= [\beta_0 + (\beta_1 + \beta_2 q - 1 + 2\lambda\beta_2 c)\phi_t]dt - 2c\beta_2\phi_t dZ_1(t) + \sqrt{2}\beta_2\phi_t dZ_2(t) \tag{21}
\end{align}
$$

where $dZ_1(t)$ and $dZ_2(t)$ are independent Wiener processes.

Ritchken and Trevor [1999] provide a lattice based algorithm for pricing such contracts where the primary state variable is the logarithmic price and the auxiliary state variable is the path dependent local volatility. In their algorithm a grid of values at each node is used to proxy for the local volatility. The grid size at a particular node is based on the maximum and minimum local volatility levels that are feasible for paths that end up at the given node.

To build a trinomial lattice for process (16)-(17), we follow Ritchken and Trevor [1999] by setting the grid size for y as $\eta\sqrt{\phi_0}\sqrt{\triangle t}$. The transition probability at node (y_t, ϕ_t) is given by,

$$
\begin{cases}
p_+(y_t, \phi_t, t) &= \frac{\phi_t}{2\phi_0\eta^2} + \frac{(r_f - \phi_t/2)\sqrt{\phi_0\triangle t}}{2\eta\phi_0} \\
p_0(y_t, \phi_t, t) &= 1 - \frac{\phi_t}{\phi_0\eta^2} \\
p_-(y_t, \phi_t, t) &= \frac{\phi_t}{2\phi_0\eta^2} - \frac{(r_f - \phi_t/2)\sqrt{\phi_0\triangle t}}{2\eta\phi_0}
\end{cases}
\tag{22}
$$

These probabilities are linear in ϕ_t, which is required for Proposition 1. Unfortunately, from node (y_t, ϕ_t) to node (y_{t+1}, ϕ_{t+1}), the update for ϕ_t is given by,

$$\phi_{t+1} = f(\phi_t, y_t, y_{t+1}) = f_1(\phi_t, y_t, y_{t+1}) + f_2(\phi_t, y_t, y_{t+1})$$

where $f_1(\phi_t, y_t, y_{y+t})$ is some linear function of ϕ_t, but

$$f_2(\phi_t, y_t, y_{t+1}) = \beta_2(\frac{1}{4}\triangle t^{3/2}\phi_t^2 - \frac{1}{2}c^*\phi_t^{3/2} - 2c^*(y_{t+1} - y_t - r_f\sqrt{\triangle t})\sqrt{\triangle t}\phi_t^{1/2})$$

with $c^* = c - \lambda\sqrt{\triangle t}$, which is not affine in ϕ_t. Therefore Proposition 1 is not applicable. Indeed,

$$E[\phi_{t+1}|y_{t+1}] = E[f(\phi_t, y_t, y_{t+1})|y_{t+1}] \neq \text{weighted average of } f(E[\phi_t|y_t], y_t, y_{t+1}).$$

Since Proposition 1 is not valid, computing the exact conditional expectation of the auxiliary state variable at each node cannot be accomplished by simple recursions. However, the function $f(\phi_t, y_t, y_{t+1})$ is close to linear in the sense that the weighted average of $f(E[\phi_t|y_t], y_t, y_{t+1})$ provides an excellent approximation to $E[\phi_{t+1}|y_{t+1}]$. Hence, as a heuristic, we take as a reduced form lattice the sequence of points $(y_t, \tilde{\phi}_t(y_t))$ where $\tilde{\phi}_t(y_t) = \text{weighted average of } f(\tilde{\phi}_{t-1}(y_{t-1}), y_{t-1}, y_t)$. While the option prices produced by this algorithm are not guaranteed by Proposition 2, as we shall see from numerical experiments, the prices serve as good approximations.

Exhibit 6 shows the convergence of at the money European call option prices with expiration dates ranging from 2 to 200 days that are obtained using the algorithm. The number of time partitions for each day are shown in the columns. The prices obtained using the algorithm are indicated. For example, consider the 20 day contract. If a lattice is constructed with one trading opportunity per day, the price is \$1.869. If the partition is refined to 5 trading opportunities per day, then the price is \$1.867. The "true" price obtained using Monte Carlo simulation of

13

the bivariate diffusion process given in equations (20) and (21) is, with 95% confidence, in the interval [$1.864, $1.867]. The confidence intervals are based on simulations of 50,000 paths, using Black Scholes as a control variate and antithetic variables as variance reduction techniques. The simulations also used 30 time partitions in each day.

Exhibit 6 Here

From the Exhibit we see that the algorithm produces very accurate results, with convergence occurring using 3 revisions a day. Moreover, in contrast to Ritchken and Trevor [1999], where at each node in the lattice, a grid size of over 20 is recommended, here, in our lattice, at each node, there are only 3 option prices stored. The computational times for these contracts are therefore an order of magnitude lower than Ritchken and Trevor. Notice that the simulations are of a bivariate diffusion, whereas the lattice prices are based off our simple reduced lattice.

The final experiment we conducted was to investigate the performance of the algorithm for pricing contracts over different strike prices. Exhibit 7 shows the results. The data used for these experiments is taken from Ritchken and Trevor, so that comparisons can be made with their algorithm.

Exhibit 7 Here

For each contract the 95% confidence intervals for the true bivariate diffusion option price is displayed. Notice that the intervals, constructed using 50,000 simulations, antithetic variables and the Black Scholes model as a control variate produces very tight confidence intervals. In spite of this, almost all the contracts fell in the intervals. This suggests that the reduced form lattice is a useful model for pricing options in a GARCH framework.

6   Conclusion

There are many problems in financial economics, where valuations require carrying an auxiliary state variable. Oftentimes these auxiliary state variables arise based on the structure of the particular financial contract. For example, there are a wide range of mortgage contracts, and exotic option contracts that involve path dependent statistics that need to be carried along. In other problems, such as in implementing Markovian HJM models, or in GARCH option pricing models, the auxiliary state variable arises directly in the underlying dynamics of the process. Current numerical procedures for computing prices of claims under these processes require very fine partitions for the auxiliary state variable for each level of the primary state variable. Examples of such algorithms include Hull and White [1993b], Li, Ritchken and Sankarasubramanian [1995]

and Ritchken and Trevor [1999]. In this article we have developed a new algorithm which is much more efficient than the existing methods. In addition we have identified conditions under which the algorithm is guaranteed to converge to accurate prices. Examples of Markovian HJM models were provided to demonstrate the efficiency. In addition, the approach is useful for pricing options under GARCH and stochastic volatility processes.

It would be desirable to extend the models here to handle problems with multiple auxiliary state variables. Solving such problems is important especially if one is interested in implementing HJM models with a richer volatility structure than that described by Ritchken and Sankarasubramanian [1995]. For example, Chiarella and Kwon [2000] have identified classes of volatility structures for forward rates, that include the models of Ritchken and Sankarasubramanian as special cases, and are Markov with respect to several state variables, most of which are path statistics. It remains for future research to establish algorithms that can efficiently handle such structures. It also remains for future research to investigate how our models can be extended to handle problems with multiple stochastic drivers.

References

Abramowitz, M., and I. Stegun, "Handbook of Mathematical Functions", *Applied Mathematics Series*,National Bureau of Standards, 55 (1964).

Andersen, L. and J. Andreasen, "Volatility Skews and Extensions to the LIBOR Market Model", *Applied Mathematical Finance*, 7 (2000),1-32.

Barraquand, J., and T. Pudet,"Pricing of American Path Dependent Contingent Claims",*Journal of Finance 6* (1996), 17-51.

Boyle, P., Broadie, M., and P. Glasserman, " Monte Carlo Methods for Security Pricing",*Journal of Economic Dynamics & Control 21* (1997), 1267-1321.

Brace, A., Gatarek, D., and Musiela, M. "The Market Model of Interest Rate Dynamics", *Mathematical Finance 7* (1997), 127–155.

Chiarella, C., and N. El-Hassan, "Evaluation of Derivative Security Prices in the Heath-Jarrow-Morton Framework as Path Integrals using Fast Fourier Transform Techniques", *The Journal of Financial Engineering 6* (1997), 121–147.

Chiarella, C., and O. Kwon, "Forward Rate Dependent Markovian Transformations of the Heath-Jarrow-Morton Term Structure Model", *Working Paper* (2000), University of Technology Sydney.

Cox, J., J. Ingersoll, and S. Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53 (1985) 363-384.

Duan, J., "The GARCH Option Pricing Model". *Mathematical Finance 5* 1995), 13–32.

Duan, J., "Augmented GARCH(p,q) Process and its Diffusion Limit", *Journal of Econometrics* 79 (1997), 97-127.

Heath, D., Jarrow, R., and A. Morton, "Bond Pricing and the Term Structure of Interest Rates : A New Methodology for Contingent Claims Valuation", *Econometrica 60* (1992),77–105.

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

Hull, J., and A. White, "Bond Option Pricing Based on a Model for the Evolution of Bond Prices", *Advances Futures Options Research 6* (1993), 1–13.

Hull, J., and A.White, "Efficient Procedures for Valuing European and American Path Dependent Options", *Journal of Derivatives 1* (1993b),21–31.

Hull, J., and A. White, "Forward Rate Volatilities, Swap Rate Volatilities, and the Implementation of the Libor Market Model", *Journal of Fixed Income* ,10,(2000), 46-62.

Inui, K., and M. Kijima, "A Markovian Framework in Multi-Factor Heath-Jarrow-Morton models". *Journal of Financial and Quantitative Analysis 3* (1999), 423–440.

Kamrad, B. and P. Ritchken, "Multinomial Approximating Models for Options with k State Variables," *Management Science* 37 (1991),1640-1653.

Li, A., Ritchken, P., and L. Sankarasubramanian, "Lattice Model for Pricing American Interest Rate Claims". *Journal of Finance 50* (1995),719–737.

Longstaff, F., and E. Schwartz, "Valuing American Options by Simulation: A Simple Least-Squares Approach", *The Review of Financial Studies*, 14 (2001),113-147

Ritchken, P., and L. Sankarasubramanian, "Volatility Structures of Forward Rates and the Dynamics of the Term Structure", *Mathematical Finance 5* (1995),55–72.

Ritchken, P., Sankarasubramanian, L., and A. Vijh, "The Valuation of Path Dependent Contracts on the Average", *Management Science 39*, (1993), 1202–1213.

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

Appendix

Proof of Proposition 1

First, consider equation (8). We have:

$$P(Y_{t+1} = y_{t+1}) = \sum_{(y_t, \phi_t)} P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \phi_t) P(Y_t = y_t, \Phi_t = \phi_t)$$

$$= \sum_{y_t} \sum_{\phi_t | y_t} P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \phi_t) P(\Phi_t = \phi_t | Y_t = y_t) P(Y_t = y_t)$$

But $P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \phi_t)$ is given by (6) if $|y_{t+1} - y_t| \le \lambda \sqrt{\triangle t}$ and 0 otherwise, which is linear with respect to ϕ_t. Therefore

$$\sum_{\phi_t | y_t} P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \phi_t) P(\Phi_t = \phi_t | Y_t = y_t)$$

$$= P(Y_{t+1} = y_{t+1} | Y_t = y_t, \sum_{\phi_t | y_t} (\Phi_t = \phi_t) P(\Phi_t = \phi_t | Y_t = y_t))$$

$$= P(Y_{t+1} = y_{t+1} | Y_t = y_t, E[\Phi_t | Y_t = y_t])$$

$$= P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \bar{\phi}_t) \tag{23}$$

This gives equation (8).

For equation (10), we have

$$P(Y_t = y_t | Y_{t+1} = y_{t+1}) = \frac{P(Y_{t+1} = y_{t+1}, Y_t = y_t)}{P(Y_{t+1} = y_{t+1})}$$

$$= \frac{\sum_{\phi_t | y_t} P(Y_{t+1} = y_{t+1}, Y_t = y_t, \Phi_t = \phi_t)}{P(Y_{t+1} = y_{t+1})}$$

$$= \frac{\sum_{\phi_t | y_t} P(Y_{t+1} = y_{t+1} | Y_t = y_t, \Phi_t = \phi_t) P(\Phi_t = \phi_t | Y_t = y_t) P(Y_t = y_t)}{P(Y_{t+1} = y_{t+1})}$$

Using (23) again, we get equation (10).

Finally, for equation (9)

$$\bar{\phi}_{t+1}(y_{t+1}) = E[\Phi_{t+1} | Y_{t+1} = y_{t+1}]$$

$$= \sum_{(y_t, \phi_t)} \phi^*_{t+1}(y_t, \phi_t, t) P(Y_t = y_t, \Phi_t = \phi_t | Y_{t+1} = y_{t+1})$$

$$= \sum_{(y_t, \phi_t)} \phi^*_{t+1}(y_t, \phi_t, t) P(\Phi_t = \phi_t | Y_t = y_t, Y_{t+1} = y_{t+1}) P(Y_t = y_t, | Y_{t+1} = y_{t+1})$$

Since Φ_t is a path statistic only up to time t,

$$P(\Phi_t = \phi_t | Y_t = y_t, Y_{t+1} = y_{t+1}) = P(\Phi_t = \phi_t | Y_t = y_t)$$

We have

$$\bar{\phi}_{t+1}(y_{t+1}) = \sum_{y_t} \sum_{\phi_t|y_t} \phi^*_{t+1}(y_t, \phi_t, t) P(\Phi_t = \phi_t | Y_t = y_t) P(Y_t = y_t, |Y_{t+1} = y_{t+1})$$

According to equation (7), $\phi^*_{t+1}(y_t, \phi_t, t)$ is linear with respect to ϕ_t, therefore

$$\begin{aligned}
\bar{\phi}_{t+1}(y_{t+1}) &= \sum_{y_t} \phi^*_{t+1}(y_t, \sum_{\phi_t|y_t} \phi_t P(\Phi_t = \phi_t | Y_t = y_t), t) P(Y_t = y_t, |Y_{t+1} = y_{t+1}) \\
&= \sum_{y_t} \phi^*_{t+1}(y_t, E(\Phi_t | Y_t = y_t), t) P(Y_t = y_t | Y_{t+1} = y_{t+1}) \\
&= \sum_{y_t} \phi^*_{t+1}(y_t, \bar{\phi}_t(y_t), t) P(Y_t = y_t | Y_{t+1} = y_{t+1})
\end{aligned}$$

This proves Proposition 1.

Proof of Proposition 2

Using the same arbitrage arguments as in the regular case, it is shown that option price $g(y, \phi, t)$, in absence of arbitrage, satisfies the following partial differential equation :

$$\frac{1}{2}\frac{\partial^2 g}{\partial y^2} + m(y, \phi, t)\frac{\partial g}{\partial y} + n(y, \phi, t)\frac{\partial g}{\partial \phi} + \frac{\partial g}{\partial t} - r_t g = 0 \tag{24}$$

The trinomial backward induction equation (11) is actually the finite difference approximation of the above PDE, with truncation error $O[(\triangle t)^{\frac{3}{2}}]$. The total effect of this truncation error is at most $N \times O[(\triangle t)^{\frac{3}{2}}] = \frac{T}{\triangle t}O[(\triangle t)^{\frac{3}{2}}] = O(\sqrt{\triangle t})$, which says the price converges as $\triangle t \to 0$.

We use equation (12)-(13) instead of (11) in backward induction. The additional error we introduce is the error of the quadratic approximation (12). In order to keep convergence property, this error should have the same magnitude as truncation error $O[(\triangle t)^{\frac{3}{2}}]$. But the inerpolation/extrapolation error in a quadratic approximation like (12) is given by (see Abramowitz and Stegun (1964))

$$\frac{1}{3!}\frac{\partial^3 g(y, \xi)}{\partial \phi^3}(\phi^*_{t+1} - \phi^1_{t+1})(\phi^*_{t+1} - \phi^2_{t+1})(\phi^*_{t+1} - \phi^3_{t+1})$$

It remains to show

$$|\phi^*_{t+1} - \phi^i_{t+1}| = O(\sqrt{\triangle t}), \ i = 1, 2, 3 \tag{25}$$

By carefully following the tree construction procedure (8)-(10) and using the Lipschitz condition we can show

$$|\phi^i_{t+1} - \phi^j_{t+1}| = O(\sqrt{\triangle t}), \ i, j = 1, 2, 3 \tag{26}$$

This is also true at time t and therefore

$$|\phi^i_t - \bar{\phi}_t| = O(\sqrt{\triangle t}), \ i = 1, 2, 3 \tag{27}$$

19

We only care $\phi^*_{t+1} = \phi^*_{t+1}(y_t, \phi^i_t, t)$, $i = 1, 2, 3$. Since $\phi^*_{t+1}(y_t, \bar{\phi}_t, t)$ is one of ϕ^j_{t+1}, $j = 1, 2, 3$, without loss of generality, say $\phi^1_{t+1} = \phi^*_{t+1}(y_t, \bar{\phi}_t, t)$, then by (7) and (27)

$$
\begin{aligned}
|\phi^*_{t+1} - \phi^1_{t+1}| &= |n(y_t, \phi^i_t, t)\triangle t + \phi^i_t - n(y_t, \bar{\phi}_t, t)\triangle t - \bar{\phi}_t| \\
&\leq O(\triangle t) + |\phi^i_t - \bar{\phi}_t| = O(\sqrt{\triangle t})
\end{aligned}
$$

This together with (26) gives (25), which is what we want.

Endnotes

[1] This is not the only approach. Chiarella and El-Hassan [1997], for example, look at such a problem and evaluate prices as path integrals using fast Fourier transform techniques.

[2] For example, it might be possible to identify the maximum and minimum value of the auxiliary state variable to each node. Then this range can be split up into buckets of equal width.

[3] For a review of the literature see Boyle, Broadie, and Glasserman [1997], and Longstaff and Schwartz [2000].

[4] For example, see Andersen and Andreasen [1997], Brace, Gatarek, and Musiela [1997], and Hull and White [2000].

[5] Innui and Kijima [1999] and Chiarella and Kwon [2000], by adding additional state variables, have further generalized the class of volatility structures that permit Markovian representations.

[6] For a discussion on this trinomial lattice see Kamrad and Ritchken [1991].