CENTRE FOR EMEA BANKING, FINANCE & ECONOMICS

Adaptive Continuous time Markov Chain Approximation Model to General Jump-Diffusions

Mario Cerrato
Chia Chun Lo
Konstantinos Skindilias

Working Paper Series

No 28/11

Adaptive Continuous time Markov Chain Approximation Model to General Jump-Diffusions

Mario Cerrato[a], Chia Chun Lo[b], Konstantinos Skindilias[c]

[a] *University of Glasgow Business School, Adam Smith Building, Glasgow, UK*

[b] *Faculty of Business Administration, University of Macau, Macau, China*

[c] *London Metropolitan Business School, London Metropolitan University, London, UK*

Abstract

We propose a non-equidistant Q rate matrix formula and an adaptive numerical algorithm for a continuous time Markov chain to approximate jump-diffusions with affine or non-affine functional specifications. Our approach also accommodates state-dependent jump intensity and jump distribution, a flexibility that is very hard to achieve with other numerical methods. The Kolmogorov–Smirnov test shows that the proposed Markov chain transition density converges to the one given by the likelihood expansion formula as in Ait-Sahalia (2008). We provide numerical examples for European stock option pricing in Black and Scholes (1973), Merton (1976) and Kou (2002).

Keywords: Markov Chains, Diffusion Approximation, Transition Density, Jump-Diffusion Approximation, Option Pricing

JEL **Classification**: C60, G10, G12, G13

Email addresses: Mario.Cerrato@glasgow.ac.uk (Mario Cerrato), ChiaChunLo@umac.mo (Chia Chun Lo), K.Skindilias@londonmet.ac.uk (Konstantinos Skindilias)

1. Introduction

The approximation of stochastic processes has become a very important part of mathematical finance. In this paper, we propose a methodology to approximate general jump-diffusion processes using a continuous time Markov chain. The idea behind the methodology is to construct a rate generator (or Q matrix) which derives the evolution of the Markov chain in continuous time such that the chain is locally consistent with the underlying process (see Kushner and Dupuis, 2001; Piccioni, 1987).

So far the alternative methods for approximating transition densities are numerically evaluating partial differential equations as in Lindstrom (2007), Markov chain Monte Carlo simulations as in Phillips and Yu (2009), Kalogeropoulos et al. (2010) and Stramer et al. (2010), using a discrete lattice method as in Jensen and Poulsen (2002) or by the likelihood expansion approach as in Ait-Sahalia (2002), Bakshi et al. (2006), Ait-Sahalia (2008) and Preston and Wood (2011).

With respect to the methods above the Markov chain approximation we propose in this paper has several advantages. Firstly, it can accommodate different grid settings depending on the requirements of the application; secondly, it reduces the approximation error by using the proposed adaptive algorithm; thirdly, it is faster than lattice methods, and finally, transition probabilities and derivatives prices are easy to compute.

This paper is organized as follows. Section 2 provides an overview of the continuous time Markov chain. Section 3 outlines the concept of local consistency when approximating diffusion processes. Two examples of an interest rate process (Cox et al., 1985) and a nonlinear drift process illustrate the efficiency and accuracy of our approximation scheme. Section 4 outlines the approximation for a jump-diffusion and discusses the setting for the Q rate matrix under such a scenario. Finally in section 5, option prices using the Markov chain approximation are benchmarked to the Black-Scholes closed form solution, binomial tree method, Merton's and Kou's jump-diffusion models. Section 6 summarizes the results.

2. Overview of Continuous Time Markov Chain

We consider a continuous time Markov chain generated by a transition rate matrix Q. Let Q be a $n \times n$ matrix on a countable set I with non-negative off-diagonal elements and rows that sum to zero. The Q matrix on I is defined as $Q = (q_{ij} : ij \in I)$ and all elements q_{ij} in Q have the following conditions:

$$\begin{aligned} &1. \quad 0 \leq -q_{ii} < \infty \quad \text{for all i} \\ &2. \quad q_{ij} \geq 0 \quad \text{for } i \neq j \\ &3. \quad \sum_{j \in I} q_{ij} = 0 \quad \text{for all i.} \end{aligned} \tag{1}$$

The transition probability matrix $\mathbf{P}$ for a Markov chain is generated by the rate matrix Q though the Kolmogorov equation

$$\mathbf{P}^{'}(t) = \mathbf{P}(t)Q. \tag{2}$$

It can be shown that for a homogeneous continuous time Markov chain there exists a rate generator matrix Q such that the transition probability

$$\mathbf{P}(t) = e^{tQ} = \sum_{k=0}^{\infty} \frac{(tQ)^k}{k!}, \tag{3}$$

is the unique solution to the Chapman-Kolmogorov equation.

3. Markov Chain Approximation

Consider a continuous time Markov chain taking values on a finite n-point grid $G = \{x_1, \cdots, x_n\}$. For simplicity we assume an equally spaced chain with grid size $\triangle x$. However this is not necessary. The Markov chain transition probability over time t is defined by

$$P_{i,j}(t) \equiv \text{Prob}(x(t) = x_j | x(0) = x_i). \tag{4}$$

The approximation of diffusion processes using continuous time Markov chains has been investigated and discussed in the literature, see for example Kushner (1990), Kushner and Dupuis (2001), Rogers and Yousaf (2002) and Chourdakis (2004). The idea behind this approach is that, a continuous time Markov chain is constructed by controlling its transition probability structure over time such

that the sample paths of the resulting Markov chain will approximate the sample paths of the true diffusion process with specified drift and diffusion parameters. To achieve this goal the so called local consistency condition needs to be satisfied. The condition implies that the first and second moments of the continuous time Markov chain should (at least asymptotically) match the instantaneous drift and diffusion of the true process. More specifically, in a diffusion process with the following specifications

$$dX(t) = \mu\left(X(t);\theta\right)dt + \sigma\left(X(t);\theta\right)dW(t), \tag{5}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ are the instantaneous drift and diffusion respectively; θ denotes the unknown parameter set; and $W(t)$ is Brownian motion, the local consistency argument states that it is sufficient for the Markov chain to capture the instantaneous drift and volatility of $X(t)$ if the following two conditions are satisfied:

$$\begin{aligned} \mathrm{E}(\triangle x; x(0) = x_i) &= \mu(\cdot)dt + \circ(\triangle x) \\ \mathrm{Var}(\triangle x; x(0) = x_i) &= \sigma^2(\cdot)dt + \circ(\triangle x). \end{aligned} \tag{6}$$

Piccioni (1987) provides an approximation formula for the rate matrix Q with elements given by:

$$\begin{aligned} q_{i,i-1} &= \frac{1}{2\triangle x^2}\sigma^2\left(x_i\right) + \frac{1}{\triangle x}\mu^-(x_i) \\ q_{i,i} &= -\frac{1}{\triangle x^2}\sigma^2\left(x_i\right) - \frac{1}{\triangle x}\left|\mu(x_i)\right| \\ q_{i,i+1} &= \frac{1}{2\triangle x^2}\sigma^2\left(x_i\right) + \frac{1}{\triangle x}\mu^+(x_i) \\ q_{i,j} &= 0 \qquad \forall j \neq i,\ i-1,\ i+1, \end{aligned} \tag{7}$$

where $\mu^{\pm} = \max\{\pm\mu, 0\}$.

This approximation scheme guarantees a well defined transition probability matrix $\mathbf{P}$ but yields an approximation error $\triangle x\left|\mu(x_i)\right|$ in matching the second moment. To see this, consider a CIR process

$$dX(t) = \alpha(\beta - X(t))dt + \sigma\sqrt{X(t)}dW(t), \tag{8}$$

where α, β and σ are the speed of mean reversion, long-term mean level and volatility respectively. Consider the following parameter values $\alpha = 0.5$, $\beta = 0.06$ and $\sigma = 0.05$, We construct Markov chains based on 19-, 37-, 91- and 181-point grids, equidistantly spanning the interval $[1\%, 10\%]$. Figure 1 presents true drift: $\alpha(\beta - X)$ and diffusion: $\sigma\sqrt{X}$, as well as Markov chain drift: $\mathrm{E}(\triangle x)$ and diffusion: $\sqrt{\mathrm{Var}(\triangle x)}$ together. The Markov chain method yields a perfect fit to the true drift but not for the diffusion. This may be due to the mean-reverting property of the CIR process. This example shows that the Piccioni's scheme fails to produce a good approximation for these kind of processes and furthermore it also shows that the robustness of Piccioni's approximation scheme heavily depends on the length $\triangle x$.

We propose a simple adaptive algorithm along with a general formula for the rate elements. Our approach can accommodate different grid settings for different applications. Let $\triangle x_d$ and $\triangle x_u$ represent the down and up movements from current state x_i. That is $x_{i,i-1} = x_i - \triangle x_d$ and $x_{i,i+1} = x_i + \triangle x_u$. The general formula for the rate elements is then given by,

$$\begin{aligned}
q_{i,i-1} &= \frac{1}{2 \triangle x_d}\mu^-(x_i) + \frac{\sigma^2(x_i) - (\triangle x_d \times \mu^-(x_i) + \triangle x_u \times \mu^+(x_i))}{\triangle x_d(\triangle x_d + \triangle x_u)} \\
q_{i,i} &= -q_{i,i-1} - q_{i,i+1} \\
q_{i,i+1} &= \frac{1}{2 \triangle x_u}\mu^+(x_i) + \frac{\sigma^2(x_i) - (\triangle x_d \times \mu^-(x_i) + \triangle x_u \times \mu^+(x_i))}{\triangle x_u(\triangle x_d + \triangle x_u)} \\
q_{i,j} &= 0 \qquad \forall j \neq i,\ i-1,\ i+1.
\end{aligned} \tag{9}$$

As a result, our formula satisfies the local consistency condition and furthermore non-negative transition probabilities are also guaranteed by imposing the additional constraint

$$\sup_{x_i}\left(\frac{(\triangle x_d \times \mu^-(x_i) + \triangle x_u \times \mu^+(x_i))}{\sigma^2(x_i)}\right) \leq 1. \tag{10}$$

A feature of the adaptive algorithm is that it adjusts the grid size steps $\triangle x_d$ and $\triangle x_u$. The idea behind our algorithm is that when the constraint above is violated it automatically adjusts the grid spacing $\triangle x_i^*$. To show an application of our approximation scheme, we start with a 19-point grid to approximate the CIR process. Figure 2a shows the results. Although the initial grid spacing is not very large as a result of the 19-point grid, our algorithm automatically adjusts the distance between states. Finally a 31-point grid is resolved to satisfy local consistency and achieve a perfect fit to the true drift and diffusion. We also test our methodology with a nonlinear drift process

having the following specification

$$dX(t) = (a_0 + a_1 X(t) + a_2 X(t)^2 + a_3 X(t)^{-1})dt + (b_0 + b_1 X(t) + b_2 X(t)^{b_3})dW, \tag{11}$$

with $a_0 = -0.027, a_1 = 0.709, a_2 = -5.27, a_3 = 0.00056, b_0 = 0, b_1 = 0, b_2 = 0.734, b_3 = 1.311$. These parameters are taken from Takamizawa (2008). Figure 2b shows the results. With an only 73-point grid our adaptive model yields an excellent approximations. It is worth mentioning that more than 700 grid points would have been required to achieve similar results under the Piccioni's scheme. Further, an appealing feature is that the end result is a grid that is denser where required and coarser elsewhere on the grid.

3.1. Transition Probability

Jensen and Poulsen (2002) show that the likelihood expansion approach outperforms other numerical techniques in terms of speed and accuracy when approximating diffusion processes with known transition densities. It is interesting to compare Markov chain approximations and likelihood expansion methods in approximating diffusion with unknown transition density as in the case of nonlinear drift diffusion presented in (11). The likelihood expansion approach transforms the original diffusion into a new one that is closer to normal and builds on it in order to obtain an approximation for the transition density.

Following (11), one can transform x into a unit diffusion process $y = g(x)$. Suppose that $g(x) = \int^x \frac{dy}{\sigma(y(t);\theta)}$ is well defined. After using Ito's lemma, we obtain

$$y_t = y_0 + W(t) + \int_0^t ds \left(\frac{\mu(g^{-1}(y))}{\sigma(g^{-1}(y))} - \frac{1}{2}\sigma'(g^{-1}(y);\theta) \right).$$

The diffusion x_t is reducible if and only if there exists a one-to-one of x_t into y_t. In the case when y_t is not close to normal density, one can standardize the process as $z_t = (y_t - y_0)/\sqrt{\triangle}$. An explicit transition density function for z_t can be obtained using, for example, Hermite expansions. After that one can use the Jacobian formula for the inverted change of variables to approximate the transition density of x_t from the transition density of y_t and z_t. We consider the general likelihood expansion formula of irreducible process as in Ait-Sahalia (2008), and the Matlab symbolic toolbox to obtain the density of x_t in (11)

$$P_{x_0,x}(\triangle) = \frac{e^{c_0 + c_1 \triangle + c_{-1}/\triangle}}{\sqrt{2\pi\triangle}\,(b_0 + b_1 x + b_2 x^{b_3})}, \tag{12}$$

where

$$
\begin{aligned}
c_{-1} \quad = \quad & \frac{\left(b_1 + b_2\, b_3\, {x_0}^{b_3-1}\right)\,(x-x_0)^3}{2\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^3} - \frac{(x-x_0)^2}{2\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^2} - \\
& \frac{\left(11\left(b_1 + b_2\, b_3\, {x_0}^{b_3-1}\right)^2 - 4\, b_2\, b_3\, {x_0}^{b_3-2}\,(b_3-1)\,\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)\right)\,(x-x_0)^4}{24\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^4}
\end{aligned}
$$

$$
\begin{aligned}
c_0 \quad = \quad & \frac{(x-x_0)\,\left(2\, a_0 + \frac{2\, a_3}{x_0} + 2\, x_0\,(a_1 + a_2\, x_0) - \left(b_1 + b_2\, b_3\, {x_0}^{b_3-1}\right)\,\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)\right)}{2\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^2} - \\
& \frac{(x-x_0)^2}{4\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^3}\left(\left(4\, b_1 + 4\, b_2\, b_3\, {x_0}^{b_3-1}\right)\,\left(a_0 + \frac{a_3}{x_0} + x_0\,(a_1 + a_2\, x_0) - \left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)\right.\right. \\
& \left.\left.\left(2\, a_1 + 4\, a_2\, x_0 - \frac{2\, a_3}{{x_0}^2} + \left(b_1 + b_2\, b_3\, {x_0}^{b_3-1}\right)^2\right) + b_2\, b_3\, {x_0}^{b_3-2}\,(b_3-1)\,\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^2\right)\right.
\end{aligned}
$$

$$
\begin{aligned}
c_1 \quad = \quad & \frac{-1}{8\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^2}\left(4\left(a_0 + \frac{a_3}{x_0} + x_0\,(a_1 + a_2\, x_0)\right)^2 + \left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)^2\right. \\
& \left(4\, a_1 + 8\, a_2\, x_0 - \frac{4\, a_3}{{x_0}^2} + \left(b_1 + b_2\, b_3\, {x_0}^{b_3-1}\right)^2 - 2\, b_2\, b_3\, {x_0}^{b_3-2}\,(b_3-1)\,\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)\right) - \\
& \left.\left(8\, b_1 + 8\, b_2\, b_3\, {x_0}^{b_3-1}\right)\,\left(a_0 + \frac{a_3}{x_0} + x_0\,(a_1 + a_2\, x_0)\right)\,\left(b_0 + b_1\, x_0 + b_2\, {x_0}^{b_3}\right)\right).
\end{aligned}
$$

As the density of x_t is unknown, we use as a benchmark "True" density (eq. (12)) and evaluate the "True" cumulative density function (CDF) numerically by using Gaussian quadrature. Markov chain approximation for the CDF is computed directly from the transition probability matrix generated by the methodology. We consider the same parameters used earlier and set $x_0 = 4\%$ and $\triangle = 1/52$ (weekly). Markov chains are constructed with 47, 67, 81 and 101 grid points respectively. Figure 3 shows the results. The Markov chain methodology proposed in this paper yields good approximations to the "True" density. To further investigate if the difference between the two densities is significant, the Kolmogorov–Smirnov test is performed. The null hypothesis underlying the test is $\mathrm{H}_0 : P_x^{\mathrm{MCA}} = P_x^{\mathrm{True}}$ and the Kolmogorov–Smirnov statistic is given by $D_n = \sup_x |\mathrm{CDF}_n^{\mathrm{MCA}}(x) - \mathrm{CDF}^{\mathrm{True}}(x)|$. The results in Table 1 indicate that even with a coarse Markov chain (41 grid points), the null hypothesis can not be rejected at a 99% confidence level.

4. Local Consistency to a Jump-Diffusion

Our methodology can also be extended to a general jump-diffusion process as for example

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dW(t) + \eta(X(t))dP(X(t)), \tag{13}$$

where $dP(X(t))$ is the differential Poisson process with state dependent intensity $\lambda(X)$. The coefficient $\eta(X)$ determines state dependent jump size at an event with a generalized probability density $\phi_\zeta(v)$. The jump part might be defined by a stochastic integral of the Poisson random measure $\mathcal{P}(dt, dq)$ on the Poisson mark space ζ

$$\eta(X(t))dP(t) = \int_\zeta \eta(v)\mathcal{P}(dt, dv),$$

where $E\{\mathcal{P}(dt, dv)\} = \lambda dt \phi_\zeta(v)dv$. In this case, following Kushner and DiMasi (1978), Medhi (1994) and Hanson (2007), the local consistent rate matrix elements for the jump part Q^{Jump} are given by

$$\begin{aligned}
q_{j,i} &= \lambda(x_i)\phi(x_i; \zeta(x_i) \cap (x_j - x_i - \Delta x/2, x_j - x_i + \Delta x/2]), \quad \text{for} \quad j \neq 1, i, n \\
q_{1,i} &= \lambda(x_i)\phi(x_i; \zeta(x_1) \cap (-\infty, x_1 + \Delta x/2]) \\
q_{1,N_s} &= \lambda(x_i)\phi(x_i; \zeta(x_1) \cap (x_{N_s} - \Delta x/2, \infty)) \\
q_{i,i} &= -\sum_{j \neq i} q_{j,i}.
\end{aligned} \tag{14}$$

Since the Brownian motion and Poisson process in the model above are independent, the Markov chain will be locally consistent to (13), if the rate generator Q^{JD} takes the following form

$$Q^{\text{JD}} = Q^{\text{Diffusion}} + Q^{\text{Jump}}, \tag{15}$$

where $Q^{\text{Diffusion}}$ is obtained from our adaptive approximation scheme in (9) and Algorithm 1, and Q^{Jump} is obtained from (14).

5. MCA Option Pricing Model

In this section we present and discuss a few applications of our methodology to option pricing. Let S denote the share price and consider the risk neutral probability measure $\mathbb{Q}$. The payoff of

a derivative written on S with a maturity T is defined by $C_T(S_T)$. The no-arbitrage argument implies that the fair value of a derivative at t is given by

$$C_t(S_t) = e^{-r(T-t)} E^{\mathbb{Q}} \left\{ C_T(S_T) \mid \mathcal{F}_t \right\}. \tag{16}$$

In order to compute option price $C_t(S_t)$ using our proposed Markov chain approach, we define the share price grid G and let the initial share price S_t lie on the i-th element of the grid $G = [x_1, \ldots, x_i = S_t, \ldots, x_n]$, where $x_1 = S_{min}$ and $x_n = S_{max}$. The payoff of this derivative is defined by $\text{Payoff}(T) = [C_T(x_1), \ldots, C_T(x_i), \ldots, C_T(x_n)]$. The Markov chain approach proposed in this paper computes the derivative price at t as

$$C_t^{MC} = e^{-r(T-t)} \sum_j P_{i,j}(T-t) \times \text{Payoff}_j(T). \tag{17}$$

Using the Ito's lemma the share price $S(t)$ under risk neutral probability measure $\mathbb{Q}$ is given by

$$dS(t) = S(t)\left(r + \lambda(t)(1 - E(V))\right) + S(t)\sigma dW^{\mathbb{Q}} + S(t) d\left(\sum_{j=1}^{N(t)} (V_j - 1) \right), \tag{18}$$

where $N(t)$ is a Poisson process with rate $\lambda(t)$ and $\{V_j\}$ is a sequence of independent identically distributed non-negative random variables. The Poisson process $N(t)$ and Brownian motion W are assumed to be independent. We consider the following three models

- Black and Scholes (1973)
  1. $\lambda(t) = 0$
- Merton (1976)
  1. $\lambda(t) = \lambda$.
  2. V has a log-normal distribution with a mean μ_J and variance parameter σ_J.
  3. $E(V) = \exp\left\{ \mu_J + \frac{\sigma_J^2}{2} \right\}$.
- Kou (2002)
  1. $\lambda(t) = \lambda$.
  2. $Y = \ln(V)$ has an asymmetric double exponential distribution with the density

$$f_Y(y) = p\eta_1 e^{-\eta_1 y} \mathbf{1}_{y \geq 0} + q\eta_2 e^{-\eta_2 y} \mathbf{1}_{y < 0},$$

where $p, q \geq 0$ and $p + q = 1$, represent the probability of upward and downward jumps.

3. $E(V) = E(e^Y) = q\frac{\eta_2}{\eta_1+1} + p\frac{\eta_1}{\eta_1-1}$, $\eta_1 > 1$, $\eta_2 > 0$.

Using (9) the Markov chain method computes the $Q^{\text{Diffusion}}$ by setting $\mu(\cdot) = S(t)\left(r + \lambda(t)(1 - E(V))\right)$ and $\sigma(\cdot) = S(t)\sigma$. For the jump part we consider Q^{Jump} for $d \ln S(t) = \left(\sum_{j=1}^{N(t)} \ln V_j\right) = \eta(\cdot)dP$. Note that this monotone transformation has no effect on the values of rate elements. The Q^{Jump} is then obtained by (14). It is worth to mention that Merton and Kou let $\eta(.)$ have a law of normal and double exponential distribution. The Markov chain method we proposed can accommodate any arbitrary jump amplitude distribution. The transition probability in (17) is obtained from (3) with $Q = Q^{\text{Diffusion}} + Q^{\text{Jump}}$.

5.1. Numerical Results

- Black-Sholes

Table 2 compares European call option prices using the proposed Markov chain approximation, Black and Scholes (1973) model and binomial tree of Cox and Ross (1976). The parameters we have considered are initial spot price $S_0 = 100$, maturity $T = 0.5$ years, risk-free rate $r = 5\%$ and volatility $\sigma = 25\%$. We set $S_{min} = 20$ and $S_{max} = 200$ and consider 13 different strike prices ranging from 70 to 130. For both, the Markov chain and binomial tree methods, the same number of scenarios at maturity T are considered. Therefore pricing errors and computational time of these two methods are comparable. The most evident advantage of the Markov chain over any tree method is that the tree state space is tied up to the discretization in time (as well as the degrees of freedom associated with the numerical procedure). Three different applications of 46, 91 and 181 states are considered in this example. The pricing error is measured by the sum of squared errors, $\text{SSE} = \sum_i \left(\text{ModelPrice}_\text{i} - \text{BS}_i\right)^2$. The results indicate that the Markov chain method outperforms the binomial tree method in terms of speed and accuracy in all cases. Figure 4 compares rates of convergence to the benchmark Black-Scholes price with respect to the grid size. The Markov chain estimator of the option price approaches faster to the benchmark price than binomial tree estimator.

- Merton

For this case, we set $S_0 = 100$, $T = 1$, risk free rate $r = 0$, and other parameters as given in Andersen and Andreasen (2000). That is the diffusion volatility $\sigma = 17.65\%$, the jump intensity $\lambda = 8.9\%$, the mean of jump amplitude $\mu_J = -88.98\%$ and the jump amplitude volatility $\sigma_J = 45.05\%$. We consider a 181-point grid for the Markov chain with $S_{\min} = 60$ to $S_{\max} = 240$. Figure 5a shows the option prices using the Merton's closed form formula and the ones given by our methodology for 13 different strikes K ranging from 70 to 130. The results shows that the Markov chain method produces a very good fit for the call option prices from deep out-of-money to deep in-the-money strikes and relative errors are all less than 0.1%. Figures 5b and 5c show hedge parameter estimation (delta and gamma) when using our method as well as the Merton's formula. The Markov chain can produce accurate values of delta and gamma across all strike prices. The values of vega are also reported in Figure 5d. Our methodology produces somehow less accurate estimates of vega when deep in-the-money options are considered. Overall these results are satisfactory.

- Kou

We now consider a further example. The parameters for the jump component are taken from Kou (2002) and those are $\eta_1 = 10$, $\eta_2 = 5$, $\lambda = 1$ and $p = 0.4$. The remaining parameters for the diffusion part are: $r = 5\%$ and $\sigma = 0.16$. We price call and put options with $S_0 = 100$ and $K = 98$, for maturities up to half a year. The Markov chain uses a 161-point grid with $S_{\min} = 20$ to $S_{\max} = 180$. Figure 6 shows the option prices computed by Kou's formula and those using the Markov chain approximation proposed in this paper. In all the cases the relative percentage error are all less than 0.5%.

6. Conclusion

We proposed a general Markov chain approximation formula and an adaptive algorithm for determining the grid space of a finite state Markov chain in continuous time. Our approach not only satisfies the local consistency conditions but also guarantees a well defined transition probability matrix. Our numerical examples (on affine and non-affine cases) show that the adaptive algorithm proposed in this paper produces an excellent approximation to the underlying process. Additionally to this, we also show how a generator matrix to the jump part for an arbitrary type of distribution

can be obtained. We have considered different numerical examples: the Black and Scholes (1973) model, the log-normal jump-diffusion of Merton (1976) and double exponential jump-diffusion of Kou (2002). We note that our Markov chain method can capture all the salient features of these models. Our methodology can easily accommodate state dependent volatility, intensity, and jump amplitude distribution while the same degree of flexibility is not easily achievable when using tree methods. The pricing of high-dimensional American options is on the agenda for future research.

References

Ait-Sahalia, Y., 2002. Maximum-likelihood estimation of discretely-sampled diffusions: A closed-form approximation approach. Econometrica 70 (1), 223–262.

Ait-Sahalia, Y., 2008. Closed-form likelihood expansions for multivariate diffusions. Annals of Statistics 36 (2), 906–937.

Andersen, L., Andreasen, J., 2000. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. Review of Derivatives Research 4, 231–262.

Bakshi, G., Ju, N., Ou-Yang, H., 2006. Estimation of continuous-time models with an application to equity volatility dynamics. Journal of Financial Economics 82 (1), 227–249.

Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. Journal of Political Economy 81 (3), 637–654.

Chourdakis, K., 2004. Non-affine option pricing. Journal of Derivatives 11, 10–25.

Cox, J. C., Ingersoll, E., Ross, S. A., 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385–407.

Cox, R., Ross, S. A., 1976. The valuation of options for alternative stochastic processes. Journal of Financial Economics 3 (1), 145–166.

Hanson, F. B., 2007. Applied Stochastic Processes and Control for Jump-Diffusions. SIAM, US.

Jensen, B., Poulsen, R., 2002. Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets. Journal of Derivatives 9 (4), 18–32.

Kalogeropoulos, K., Roberts, G. O., Dellaportas, P., 2010. Inference for stochastic volatility models using time change transformations. Annals of Statistics 38 (2), 784–870.

Kou, S. G., 2002. A jump-diffusion model for option pricing. Management Science 48 (8), 1086–1101.

Kushner, H. J., 1990. Numerical methods for stochastic control problems in continuous time. SIAM Journal of Control and Optimisation 28 (5), 999–1048.

Kushner, H. J., DiMasi, G., 1978. Approximations for functionals and optimal control problems on jump-diffusion processes. J. Math. Anal. Applics. 63, 772–800.

Kushner, H. J., Dupuis, P. G., 2001. Numerical methods for stochastic control problems in continuous time, 2nd Edition. Stochastic Modelling and Applied Probability. Springer-Verlag, London, UK.

Lindstrom, E., 2007. Estimating parameters in diffusion processes using an approximate maximum likelihood approach. Annals of Operations Research 151 (1), 269–288.

Medhi, J., 1994. Stochastic Processes. John Wiley and Sons, US.

Merton, R. C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3 (1-2), 125–144.

Phillips, P. C. B., Yu, J., 2009. Simulation-based estimation of contingent-claims prices. Review of Financial Studies 22 (9), 3669–3705.

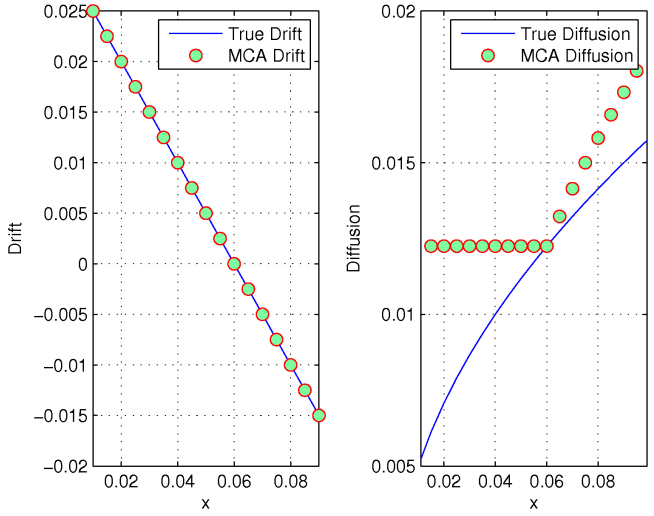
Piccioni, M., 1987. Convergence of implicit discretisation schemes for linear differential equations with applicaiton to filtering. In: Prato, G., Tubaro, L. (Eds.), Stochastic PDE's and Applications, Lecture Notes in Mathematics 1236. Springer-Verlag.

Preston, S. P., Wood, A., 2011. Approximation of transition densities of stochastic differential equations by saddlepoint methods applied to small-time ito-taylor sample-path expansions. Statistics and Computing 38 (2), 1–13.

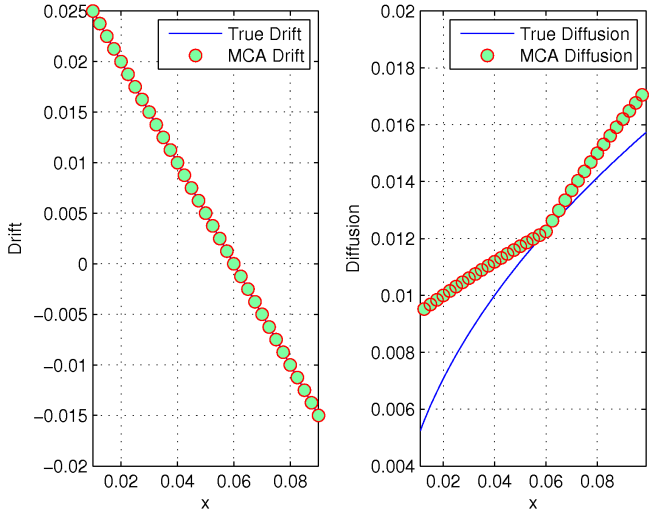
Rogers, L., Yousaf, F., 2002. Markov chains and the potential approach to modelling interest rates and exchange rates. In: H. Geman, D. M., Pliska, S., Vorst, T. (Eds.), Mathematical Finance: Bachelier Congress 2000 (Paris). Springer Finance, London, UK, pp. 375–406.

Stramer, O., Bognar, M., Schneider, P., 2010. Bayesian inference for discretely sampled markov processes with closed-form likelihood expansions. Journal of Financial Econometrics 8 (4), 450–480.

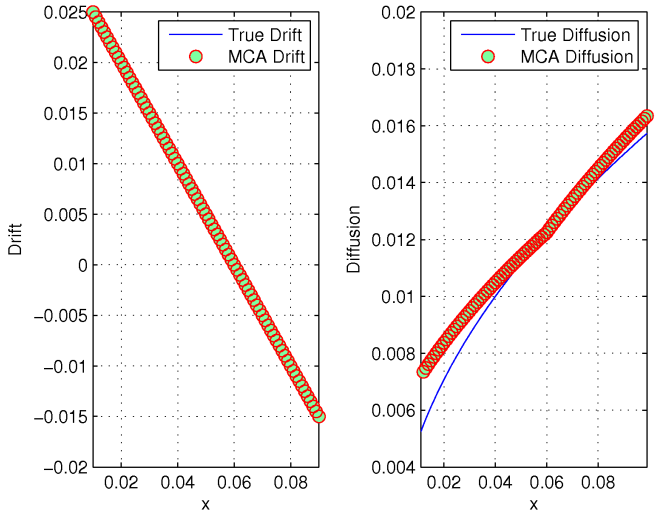
Takamizawa, H., 2008. Is nonlinear drift implied by the short-end of the term structure ? Review of Financial Studies 21 (1), 311–346.



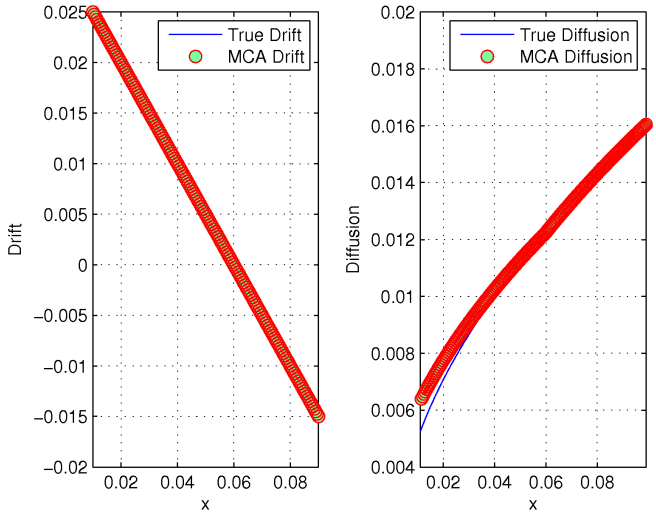
(a) 19-point grid (Piccioni)



(b) 37-point grid (Piccioni)

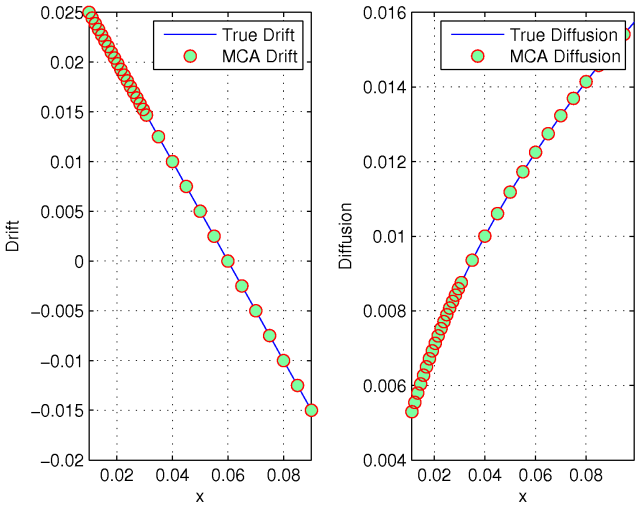


(c) 91-point grid (Piccioni)

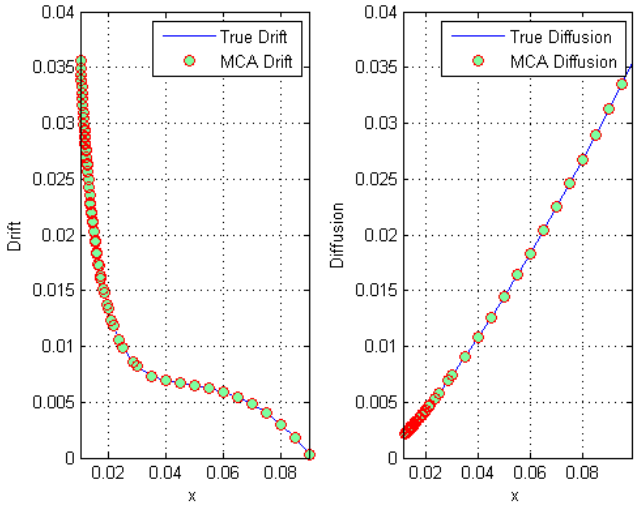


(d) 181-point grid (Piccioni)

Figure 1: Approximation to the CIR process $dX = \alpha(\beta - X)dt + \sigma\sqrt{X}dW$, under Piccioni's scheme. Solid line: True drift and diffusion for the CIR model. Dotted line: Approximated drift and diffusion using Piccioni's formula. The parameter set is $\{\alpha, \beta, \sigma\} = \{0.5, 0.06, 0.05\}$.

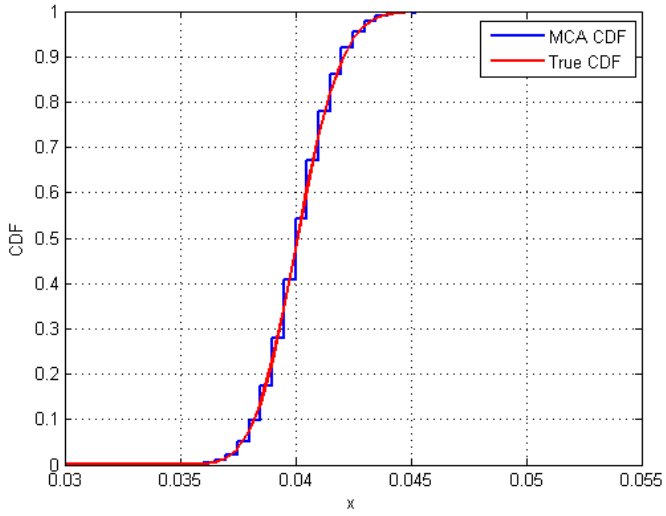


(a) CIR 31 grid points (Adaptive)

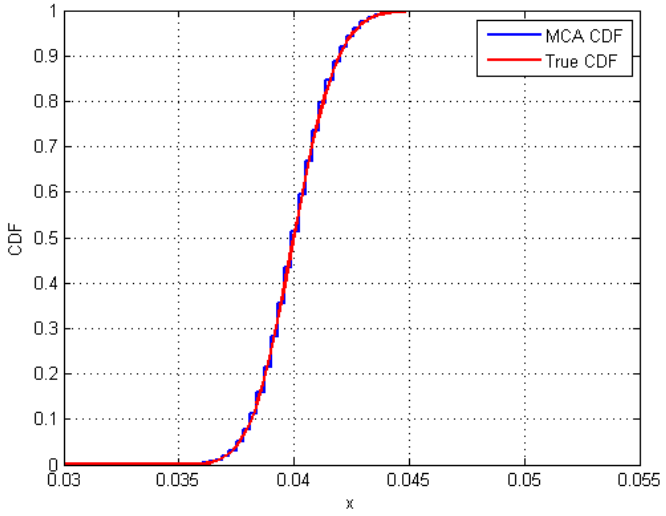


(b) Nonlinear drift 73 grid points (Adaptive)

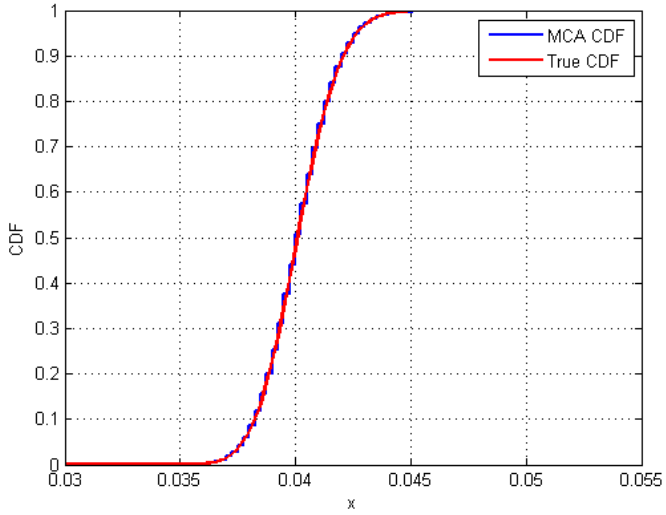
Figure 2: The adaptive algorithm as an approximation to the CIR process $dX = \alpha(\beta - X)dt + \sigma\sqrt{X}dW$ and the nonlinear drift process $dX = (a_0 + a_1X + a_2X^2 + a_3X^{-1})dt + (b_0 + b_1X + b_2X^{b_3})dW$, started initially with a 19-point grid and resolved to a 31 and 73-point grid respectively. The resulting non-equidistant grid is denser for low values of X. Parameter set: $\{\alpha, \beta, \sigma\} = \{0.5, 0.06, 0.05\}$, and $\{a_0, a_1, a_2, a_3, b_0, b_1, b_2, b_3\} = \{-0.027, 0.709, -5.27, 0.00056, 0, 0, 0.734, 1.311\}$ for the two processes respectively.



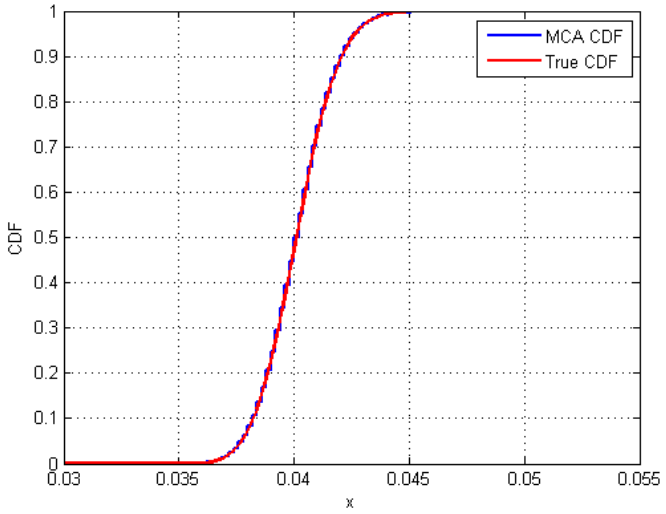
(a) 41-point grid (MCA)



(b) 67-point grid (MCA)

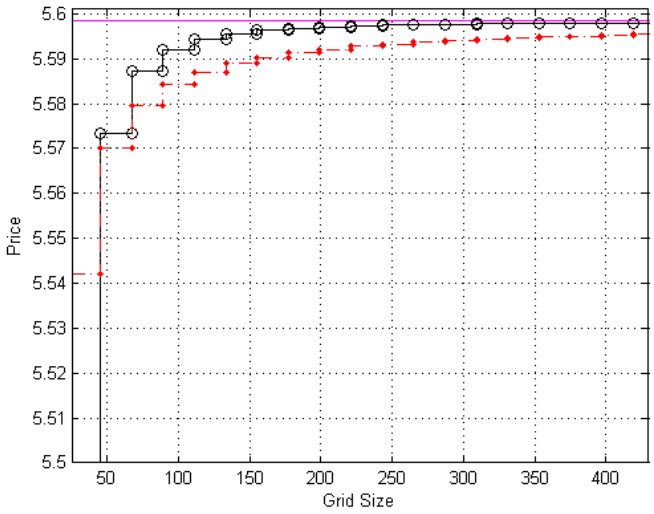


(c) 81-point grid (MCA)

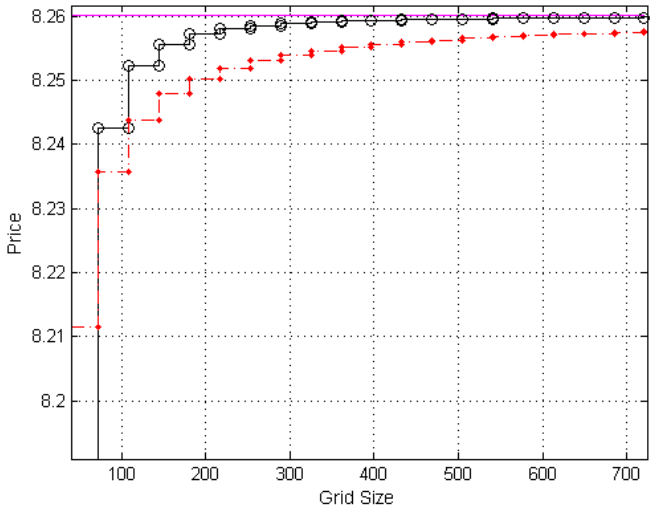


(d) 101-point grid (MCA)

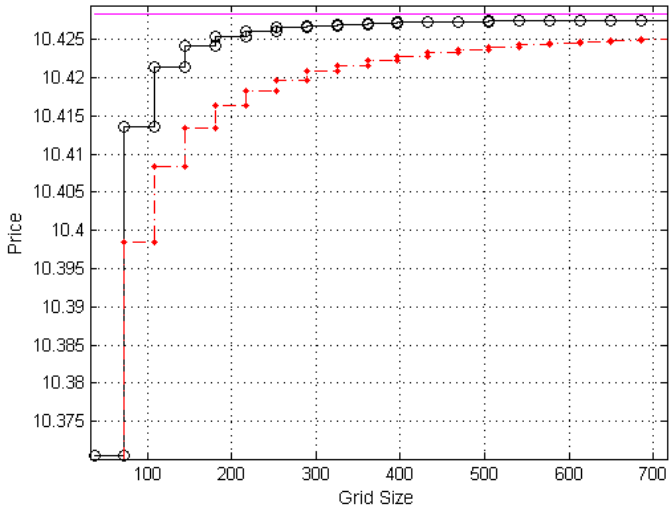
Figure 3: MCA to the nonlinear drift process $dX = (a_0 + a_1 X + a_2 X^2 + a_3 X^{-1})dt + (b_0 + b_1 X + b_2 X^{b_3})dW$. Solid line: True CDF evaluated by likelihood expansion formula in Ait-Sahalia (2008). Stairstep: CDF generated by MCA.



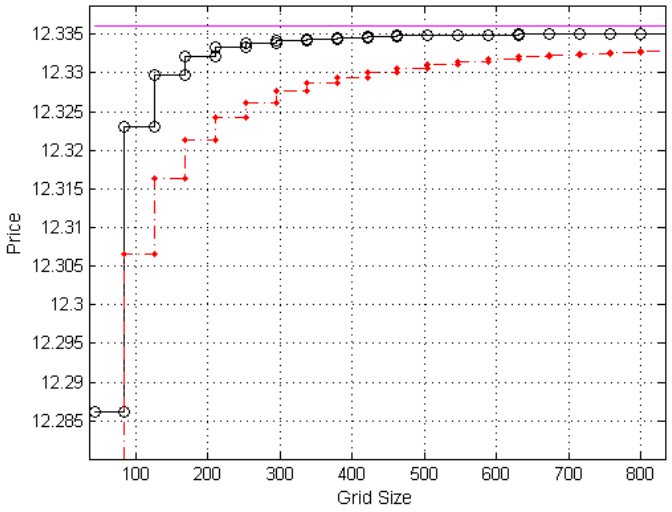
(a) T=0.25 year



(b) T=0.5 year

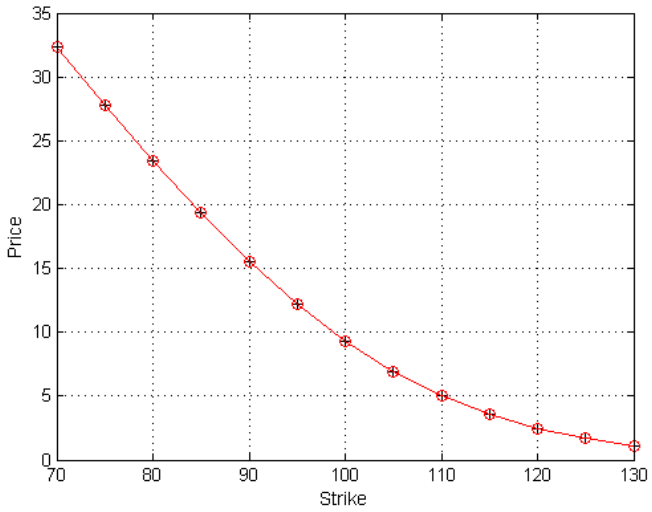


(c) T=0.75 year

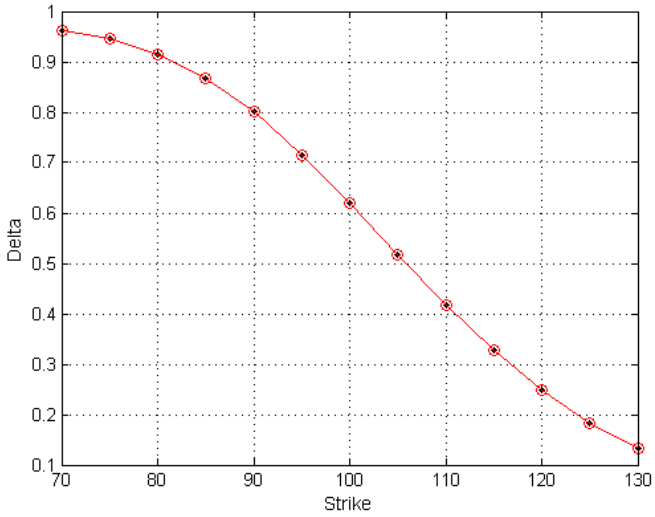


(d) T=1 year

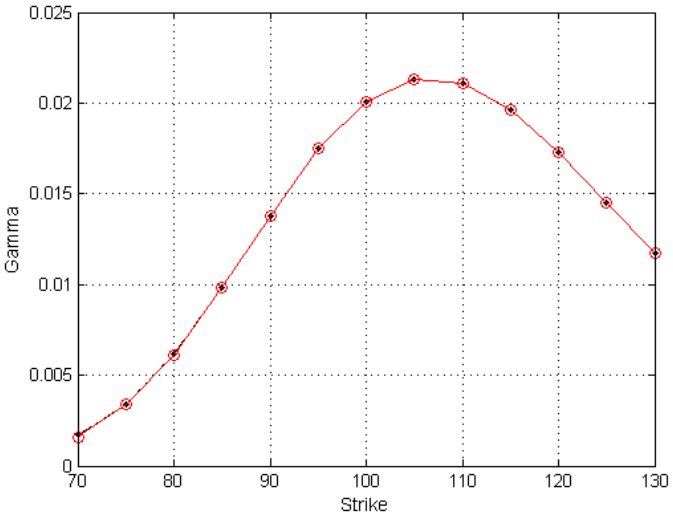
Figure 4: Option price convergence to the benchmark Black-Scholes price as a function of grid size (and space). MCA (circle), CRR tree (dots) and benchmark Black-Scholes (solid). Parameters: $\{S0, K, \sigma, r\} = \{100, 100, 0.25, 0.05\}$



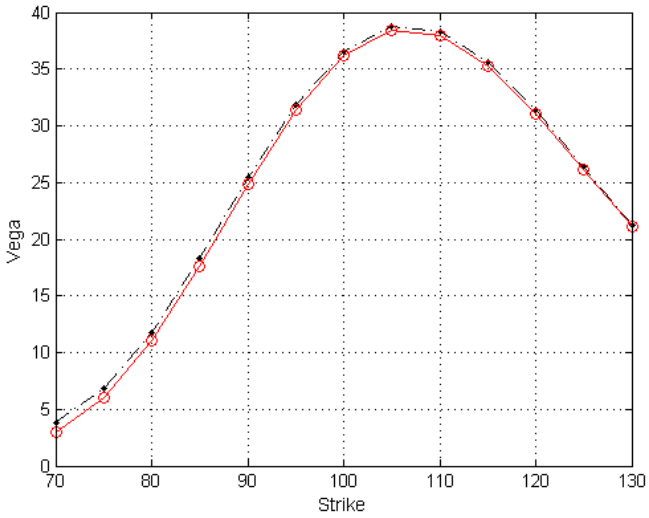
(a) Call Option Prices



(b) Delta

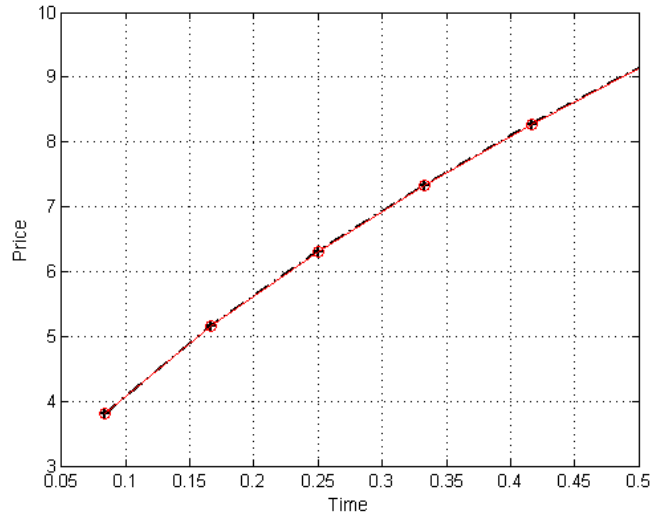


(c) Gamma

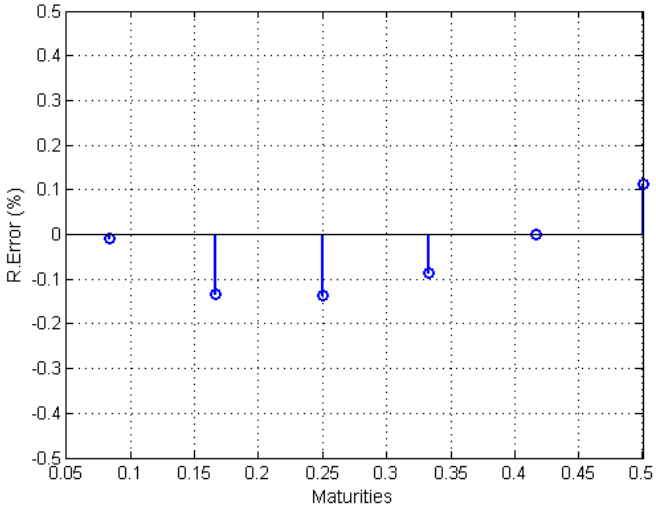


(d) Vega

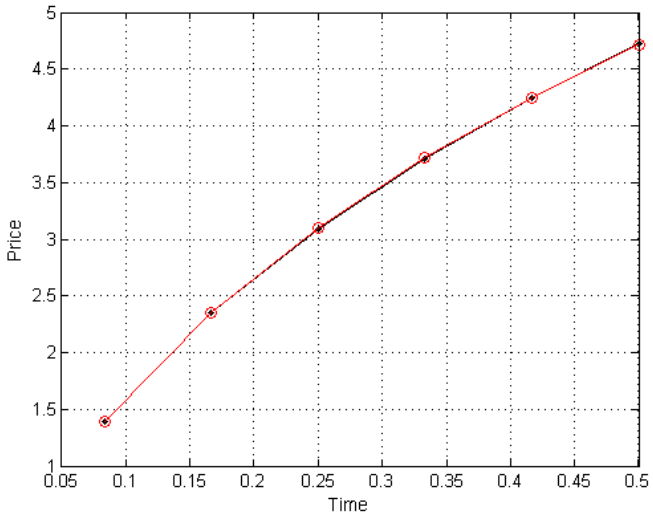
Figure 5: MCA approximation to the Merton jump-diffusion. The figures plot MCA estimates (circle) and Merton closed form solution values (dot).



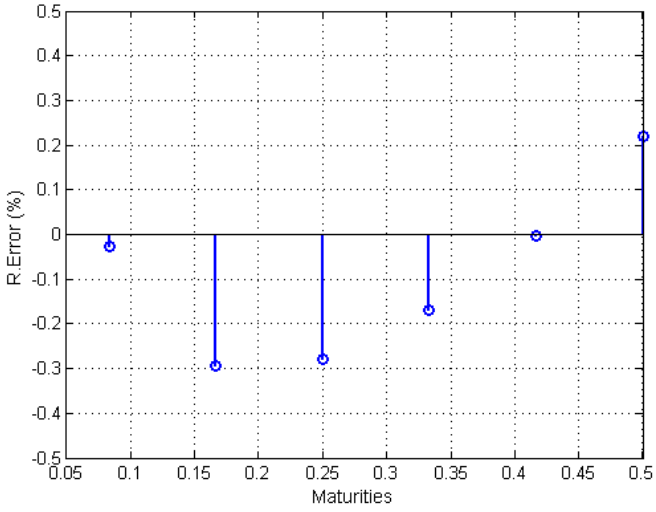
(a) Call Prices



(b) Call Price Relative Error



(c) Put Prices



(d) Put Price Relative Error

Figure 6: MCA approximation to the Kou jump-diffusion. Option prices from the MCA (circle) and Kou's closed form solution (dot).

Algorithm 1 Adaptive Local Consistency Algorithm for Grid Setting

Given an initial grid $G = [x_1, \cdots x_n]$
Set a threshold $\hbar$ for the minimum of grid size
find $x \in G$ violating 10 ***While*** x is not empty and $\triangle x \geq \hbar$ ***Do***
Select x_{max}
Adjust the grid size around x_{max} by $x_{max} - \triangle x$ and $x_{max} + \triangle x$
Update new states in G and check constraint
End While

Grid Point	41	67	81	101
Kolmogorov–Smirnov statisticD_n	0.0675	0.0401	0.0332	0.0265
Critical Value ($\alpha = 0.01$)	0.2546	0.199	0.1811	0.1622

Table 1: Kolmogorov–Smirnov Test

.

		Node: 46	Grid: 46	Node: 91	Grid: 91	Node: 181	Grid: 181
Strike	BS	Bino	MCA	Bino	MCA	Bino	MCA
70	31.8082	31.8057	31.8159	31.8079	31.8082	31.8080	31.8081
75	27.0844	27.0747	27.0909	27.0842	27.0871	27.0825	27.0841
80	22.5415	22.5445	22.5302	22.5339	22.5386	22.5433	22.5407
85	18.2885	18.3040	18.2908	18.2945	18.2908	18.2878	18.2870
90	14.4371	14.4627	14.4427	14.4385	14.4285	14.4297	14.4349
95	11.0775	11.1107	11.0678	11.0919	11.0778	11.0723	11.0748
100	8.2600	8.2974	8.2149	8.2405	8.24875	8.2503	8.2571
105	5.9885	6.0230	5.9771	6.0008	5.9883	5.9879	5.9858
110	4.2258	4.2449	4.2272	4.2409	4.2172	4.2326	4.2236
115	2.9065	2.8957	2.9042	2.9103	2.9075	2.9095	2.9048
120	1.9517	1.9192	1.9360	1.9422	1.9477	1.9456	1.9506
125	1.2817	1.2811	1.2858	1.2637	1.2838	1.2835	1.2811
130	0.8247	0.8262	0.8342	0.8102	0.8239	0.8178	0.8244
	SSE	0.00623	0.00288	0.00172	0.000319	0.000329	0.00004
	CPU Time (Second)	0.00113	0.000504	0.00315	0.000750	0.01076	0.00789

Table 2: European call option prices from Black-Scholes(BS), binomial tree (Bino) and Markov chain approximation (MCA) are presented together with parameters $\{S_0, T, r, \sigma\} = \{100, 0.5, 0.05, 0.25\}$. Three different number of share price scenarios, 46, 91 and 181 are considered for the binomial tree and Markov chain approximation model.