Characteristic Functions and Option Valuation in a Markov Chain Market

Robert J. Elliott[a,], Chuin Ching Liew[b], Tak Kuen Siu[b]

[a]Haskayne School of Business, University of Calgary, Canada; School of Mathematical Sciences, University of Adelaide, South Australia, Australia.
[b]Department of Applied Finance and Actuarial Studies, Faculty of Business and Economics, Macquarie University, Sydney, Australia.

Abstract

We introduce an approach for valuing some path-dependent options in a discrete-time Markov chain market based on the characteristic function of a vector of occupation times of the chain. A pricing kernel is introduced and analytical formulae for the prices of Asian options and occupation time call options are derived.

Keywords: Markov Chain Market, Occupation Times, Characteristic Functions, Asian Options, Occupation Time Derivatives.

1. Introduction

The valuation of options has long been an important issue in financial economics. The history of this problem may be traced to the early work of Bachelier (1900) [1], where movements of share prices were modeled by an arithmetic Brownian motion and an option valuation formula derived. This important piece of work was re-discovered by Paul Samuelson in the 1960s and re-generated interest in option valuation. The pioneering works of Black and Scholes (1973) [2] and Merton (1973) [21] provided a solution to option valuation and hedging. Under the geometric Brownian motion assumption for the price process of the underlying share price, the assumption

Email addresses: `relliott@ucalgary.ca` (Robert J. Elliott), `kennylcc@gmail.com` (Chuin Ching Liew), `ktksiu2005@gmail.com` (Tak Kuen Siu)

Corresponding Author: R.J. Elliott; Tel: +1 403 220 5540, Fax: +1 403 770 8104

of a perfect market and the no-arbitrage assumption, they derived a closed-form pricing formula for the price of a standard European call option. Since the works of Black and Scholes (1973) [2] and Merton (1973) [21], there has been tremendous growth in both academic and practical research on option valuation and hedging, as well as the related trading activities of derivative securities in global financial markets. Coincidentally, the Chicago Board of Trade (CBOT) started trading standardized call option contracts in 1973.

A key insight of Black-Scholes-Merton option pricing theory is the use of risk-neutral valuation, where the appreciation rate of the underlying share is replaced by the risk-free rate of interest and the pricing is then accomplished in the risk-neutral world. This procedure becomes transparent in the discrete-time binomial option valuation model introduced by Cox, Ross and Rubinstein (1976) [10][3]. Besides giving a transparent relationship between risk-neutral valuation and no arbitrage, the binomial, or CRR, option valuation model also provides a simple and efficient numerical scheme to approximate option prices in a continuous-time model. Assuming that the share price takes one of two possible values in each period may not be accurate enough to describe real-world movements of share prices, so more complicated tree structures for option valuation have been proposed in the literature. Boyle (1986) [3] proposed a trinomial lattice model. As an extension of the binomial model, the trinomial lattice model assumes that the price of the underlying share over each time period may take one of the three possible values. Indeed, the trinomial lattice model motivated by a finite-difference numerical scheme for solving partial differential equations. He (1990) [18] proposed a multi-nominal option valuation model, while preserving completeness of the market.

Discrete-time Markov chain models provide an important class of asset price models. They have been considered by authors such as Pliska (1997) [23], Norberg (2005) [22] and van der Hoek and Elliott (2010) [28]. Some related models include Song et al. (2010) [25] for a multivariate Markov chain asset price models and Valakevicius (2009) [27] for a continuous-time Markov chain asset price models. One of the key motives for considering Markov chain asset price models is that discrete-time Markov chain can provide a reasonably approximations to continuous-time diffusion processes. Indeed, Markov chain asset price models may include binomial and trinomial asset

[3] Indeed, this idea of binomial asset price model was originated by William Sharpe.

price models as particular cases. The valuation of some exotic options may be more simple in a discrete-time Markov chain asset price model.

In this paper, we introduce a characteristic function approach for the valuation of some path-dependent options, such as Asian options and occupation time options, in a Markov chain market, where uncertainty is modeled by a discrete-time, finite-state, Markov chain. A characteristic function of a vector of occupation times of the chain over different states is defined, which is the key tool for valuing the options. We also discuss the issue of selecting a pricing kernel in such a Markov chain market. Analytical formulae for the prices of Asian options and occupation time call options are then derived.

This paper is organized as follows. Section 2 presents the Markov chain market and the price dynamics. In Section 3, we derive the characteristic function of a vector of occupation times in different states of the underlying Markov chain. Section 4 discusses the choice of a pricing kernel in the Markov chain market. We derive analytical pricing formulae for an arithmetic Asian option, a geometric Asian option and an occupation time call option using the characteristic function derived in Section 3. The final section gives a summary of the paper.

2. A Markov Chain Market Model

In this section, we present a discrete-time Markov chain market model, where the randomness of the price process of a share is modeled by a discrete-time, finite-state, time-homogeneous, Markov chain. The Markov chain asset price model considered here includes the trinomial asset price model as a special case as explained later in this section. Indeed, similar models were discussed in some recent work such as Valakevicius (2009) [27], Song et al. (2010) [25] and van der Hoek and Elliott (2010) [28].

We consider a complete probability space $(\Omega, \mathcal{F}, P)$, where P is a real-world probability measure. Let $\mathcal{T} := \{0, 1, 2, \ldots, T\}$ be the time parameter set, where T is a finite positive integer. Indeed, one may consider an infinite time parameter set. However, for our purpose, a finite time parameter set is enough. We suppose that the risk-free interest rate is a constant $r \in (0, 1)$.

To describe uncertainty or randomness in the Markov chain market, we consider a discrete-time, N-state, time-homogeneous Markov chain $\{\mathbf{X}_t\}_{t \in \mathcal{T}}$. Following the convention in Elliott et al. (1994) [12], we identify the state space of the chain $\{\mathbf{X}_t\}_{t \in \mathcal{T}}$ with the canonical state space given by the set of standard unit vectors in $\mathbb{R}^N$:

$$\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\}.$$

Here, for each $i = 1, 2, \cdots, N$, $\mathbf{e}_i$ is the unit vector in $\mathbb{R}^N$ with one as the i-th element and zeros elsewhere. That is, $\mathbf{e}_i := (0, \ldots, 1, \ldots, 0)'$ with $\mathbf{x}'$ the transpose of a vector $\mathbf{x}$.

To describe the probability law of the chain, we define the following transition probabilities and transition matrix:

$$a_{ji} = P(\mathbf{X}_{t+1} = \mathbf{e}_j \mid \mathbf{X}_t = \mathbf{e}_i),$$
$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ a_{21} & a_{22} & \cdots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{NN} \end{pmatrix}.$$

Define the martingale increment process $\{\mathbf{V}_t\}$ by

$$\mathbf{V}_{t+1} := \mathbf{X}_{t+1} - \mathbf{A}\mathbf{X}_t$$

so

$$\mathrm{E}[\mathbf{V}_{t+1} \mid \mathcal{F}_t] = \mathbf{0} \in \mathbb{R}^N.$$

Here $\{\mathcal{F}_t\}_{t \in \mathcal{T}}$ is the natural filtration generated by the Markov chain.

We now define a share price process $\{S_t\}_{t \in \mathcal{T}}$ by assuming that it can only take values from a finite set of values $\mathcal{S} = \{s_1, s_2, \ldots, s_N\} \subset [0, \infty)$. Write

$$\mathbf{s} := (s_1, s_2, \ldots, s_N)'.$$

Without loss of generality, we suppose that $0 \leq s_1 < s_2 < \cdots < s_N$. Then in our model, the share price process $\{S_t\}$ is governed by the Markov chain $\{\mathbf{X}_t\}$ by means of the definition:

$$S_t = \langle \mathbf{s}, \mathbf{X}_t \rangle.$$

Consequently, the price process $\{S_t\}$ is, again, a discrete-time, finite-state Markov chain. Here, $\langle \cdot, \cdot \rangle$ is the scalar product.

It is not difficult to see that a finite time trinomial asset price model is a particular case of our Markov chain market model as illustrated in the following example.

Example 2.1. *Consider a three-period trinomial asset price model, where the share price process is modeled as:*

$$S_t = \begin{cases} e^u S_{t-1} & \textit{if the share price rises;} \\ S_{t-1} & \textit{if the share price remains the same;}\,, \\ e^{-u} S_{t-1} & \textit{if the share price falls,} \end{cases} \qquad S_0 = s$$

with (real-world) probability that the share price rises, remains and falls in the next period being p_u, p_m and p_d, respectively.
In this case, the state space of the share price is given by:

$$\mathcal{S} = \{e^{-3u}s, e^{-2u}s, e^{-u}s, s, e^u s, e^{2u}s, e^{3u}s\}$$

and the transition matrix is

$$\mathbf{A} = \begin{pmatrix} 1 & p_u & 0 & 0 & 0 & 0 & 0 \\ 0 & p_m & p_u & 0 & 0 & 0 & 0 \\ 0 & p_d & p_m & p_u & 0 & 0 & 0 \\ 0 & 0 & p_d & p_m & p_u & 0 & 0 \\ 0 & 0 & 0 & p_d & p_m & p_u & 0 \\ 0 & 0 & 0 & 0 & p_d & p_m & 0 \\ 0 & 0 & 0 & 0 & 0 & p_d & 1 \end{pmatrix}.$$

Consequently, the three-period trinomial asset price model is a special case of the Markov chain market model.

Similarly, the binomial asset price model is a special case of our model as well. When the number of time periods increases, the number of states increase. In fact, for a T-period trinomial or binomial model, the corresponding state space has $2T + 1$ elements.

In the existing literature on Markov regime-switching models, the evolution of the state of an economy over time is usually modeled by a finite-state Markov chain, (see, for example, Buffington and Elliott (2002) [5]). Indeed, the state of an economy in a country, or a region, may be reflected in the major share indices in that country, or that region. There is a saying that a share index of a region is a thermometer of the economy in that region. For example, one may derive some idea about the American, Japanese, UK and Hong Kong economies by looking at the S&P 500 index, the Nikkei 225 index, the FTSE 100 and the Heng Seng Index, respectively. Consequently, since the share price process $\{S_t\}$ in our model is assumed to follow a finite-state Markov chain, it may be considered as a proxy for the price process of a share index.

3. Characteristic Function of Occupation Times

In this section, we derive the characteristic function of a vector of occupation times in different states of the Markov chain $\{\mathbf{X}_t\}_{t\in\mathcal{T}}$. Indeed, the characteristic function can be regarded as the Discrete Fourier Transform (DFT) of the joint probability mass function for the vector of occupation times. Consequently, using the Inverse DFT (IDFT), the joint probability mass function can be recovered. The characteristic function derived here is a generalization of the one derived in Elliott and Tsoi (2005) [13], where a two-state Markov chain was considered. A continuous version of the characteristic function was derived in Buffington and Elliott (2002) [5].

Firstly, for each $i = 1, 2, \cdots, N$, write $J^i(t, T)$ for the occupation time of the chain in state $\mathbf{e}_i$ from time t to T. That is,

$$J^i(t, T) := \sum_{k=t}^{T} \langle \mathbf{X}_k, \mathbf{e}_i \rangle \ .$$

Then we define a vector of occupation times as follows:

$$\mathbf{J}(t, T) := \begin{pmatrix} J^1(t, T) \\ J^2(t, T) \\ \vdots \\ J^N(t, T) \end{pmatrix} \in D^N_{t,T} \ ,$$

where $D_{t,T} := \{0, 1, \cdots, T - t + 1\}$ and $D^N_{t,T}$ is the N-folded product of $D_{t,T}$.

By definition,

$$\sum_{k=1}^{N} J^k(t, T) = T - t + 1,$$

which implies that $\mathbf{J}(t, T)$ can actually only take values from a subset $D'_{t,T}$ of $D^N_{t,T}$, where

$$D'_{t,T} = \{\mathbf{v} \in D^N_{t,T} : \langle \mathbf{v}, \mathbf{1} \rangle = T - t + 1\}.$$

Ross (1976) [24] shows that the number of non-negative integer solutions to the equation

$$v_1 + \cdots + v_n = r$$

is equivalent to the number of ways for placing r identical objects into n distinct boxes, i.e. $\binom{r+n-1}{n-1}$. So, the number of elements in $D'_{t,T}$ is $\binom{T+N-t}{N-1}$.

Let $\mathbf{u} := (u_1, u_2, \cdots, u_N)' \in D_{t,T}^N$. Then the conditional characteristic function of the random vector $\mathbf{J}(t,T)$ is defined as:

$$\begin{aligned}\phi_{\mathbf{J}(t,T)}(\mathbf{u}) &:= \mathrm{E}\left[\exp\left\{-\frac{2\pi i}{T-t+2}\langle \mathbf{u}, \mathbf{J}(t,T)\rangle\right\} \;\Big|\; \mathcal{F}_t\right] \\ &= \mathrm{E}\left[\exp\left\{-\frac{2\pi i}{T-t+2}\sum_{k=t}^{T}\langle \mathbf{X}_k, \mathbf{u}\rangle\right\} \;\Big|\; \mathbf{X}_t\right] .\end{aligned}$$

Here E is expectation under P. The last equality follows from the Markov property.

Note that, for each $k = 1, 2, \cdots, N$, the random variable $J^k(t,T)$ is discrete. Consequently, the conditional characteristic function $\phi_{\mathbf{J}(t,T)}$ is the (multi-dimensional) Discrete Fourier Transform (DFT) of the conditional joint probability mass function for the vector of random variables $\mathbf{J}(t,T)$ given $\mathcal{F}_t$. If $p_{t,T}(\mathbf{j})$ is the conditional probability mass function of $\mathbf{J}(t,T)$, we can suppose, without loss of generality, that the domain of p is $D_{t,T}^N$. Then

$$\phi_{\mathbf{J}(t,T)}(\mathbf{u}) = \sum_{\mathbf{j}\in D_{t,T}^N} \exp\left\{-\frac{2\pi i}{T-t+2}\langle \mathbf{u}, \mathbf{j}\rangle\right\} p_{t,T}(\mathbf{j}) .$$

This formula enables us to recover the conditional joint probability mass function of $\mathbf{J}(t,T)$ from the conditional characteristic function via inversion. We shall employ the IDFT instead of the standard Inverse Fourier Transform. To use the Inverse Discrete Fourier Transform (IDFT) in the multi-dimensional case, we require the extra coefficient $-\frac{2\pi i}{T-t+2}$.

Remark. As noted before that $\mathbf{J}(t,T)$ only takes values from $D'_{t,T} \subset D_{t,T}^N$, so we would expect $p_{t,T}(\mathbf{j}) = 0$ if $\mathbf{j} \notin D'_{t,T}$.

The following theorem gives a compact formula for the conditional characteristic function.

Theorem 3.1. *Let $\mathbf{u} = (u_1, u_2, \cdots, u_N)' \in D_{t,T}^N$ Then the conditional characteristic function of $\mathbf{J}(t,T)$ associated with $\mathbf{u}$ given $\mathcal{F}_t$ under P is:*

$$\phi_{\mathbf{J}(t,T)}(\mathbf{u}) := \left\langle e^{-\frac{2\pi i}{T-t+2}\langle \mathbf{X}_t, \mathbf{u}\rangle} (\mathbf{B}(\mathbf{u})\mathbf{A})^{T-t}\mathbf{X}_t, \mathbf{1}\right\rangle , \tag{1}$$

where $\mathbf{1} = (1, 1, \cdots, 1)' \in \mathbb{R}^N$; $\mathbf{B}(\mathbf{u}) := \mathrm{diag}(e^{-\frac{2\pi i}{T-t+2}u_1}, \cdots, e^{-\frac{2\pi i}{T-t+2}u_N})$, the $N \times N$ diagonal matrix with diagonal elements $(e^{-\frac{2\pi i}{T-t+2}u_1}, \cdots, e^{-\frac{2\pi i}{T-t+2}u_N})$.

Proof. For notational convenience, write $c = -\frac{2\pi}{T-t+2}$. For $n \geq t$, consider the $\mathbb{R}^N$-valued process defined by:

$$\mathbf{Z}_n = \mathbf{X}_n \exp\left\{ ci \sum_{k=t}^{n} \langle \mathbf{X}_k, \mathbf{u} \rangle \right\} .$$

Now, put $u_N = 0$. Then

$$\begin{aligned}
\mathbf{Z}_T &= \mathbf{X}_T \exp\left\{ ci \sum_{k=t}^{T} \langle \mathbf{X}_k, \mathbf{u} \rangle \right\} \\
&= \mathbf{X}_T \exp\left\{ ci \langle \mathbf{X}_T, \mathbf{u} \rangle \right\} \exp\left\{ ci \sum_{k=t}^{T-1} \langle \mathbf{X}_k, \mathbf{u} \rangle \right\} \\
&= \sum_{j=1}^{N} \left(e^{ciu_j} \langle \mathbf{X}_T, \mathbf{e}_j \rangle \mathbf{e}_j \right) \exp\left\{ ci \sum_{k=t}^{T-1} \langle \mathbf{X}_k, \mathbf{u} \rangle \right\} \\
&= \sum_{j=1}^{N} \left(e^{ciu_j} \langle \mathbf{A}\mathbf{X}_{T-1} + \mathbf{V}_T, \mathbf{e}_j \rangle \mathbf{e}_j \right) \exp\left\{ ci \sum_{k=t}^{T-1} \langle \mathbf{X}_k, \mathbf{u} \rangle \right\} \\
&= \left[\sum_{j=1}^{N} \left(e^{ciu_j} \langle \mathbf{A}\mathbf{X}_{T-1}, \mathbf{e}_j \rangle \mathbf{e}_j \right) + \sum_{j=1}^{N} \left(e^{ciu_j} \langle \mathbf{V}_T, \mathbf{e}_j \rangle \mathbf{e}_j \right) \right] \exp\left\{ ci \sum_{k=t}^{T-1} \langle \mathbf{X}_k, \mathbf{u} \rangle \right\} .
\end{aligned}$$

Conditioning both sides on $\mathcal{F}_{T-1}$ gives:

$$\begin{aligned}
\mathrm{E}[\mathbf{Z}_T \mid \mathcal{F}_{T-1}] &= \sum_{j=1}^{N} \left(e^{ciu_j} \langle \mathbf{A}\mathbf{X}_{T-1}, \mathbf{e}_j \rangle \mathbf{e}_j \right) \exp\left\{ ci \sum_{k=t}^{T-1} \langle \mathbf{X}_k, \mathbf{u} \rangle \right\} \\
&= \sum_{j=1}^{N} \left(e^{ciu_j} \langle \mathbf{A}\mathbf{Z}_{T-1}, \mathbf{e}_j \rangle \mathbf{e}_j \right) \\
&= \mathbf{B}(\mathbf{u})\mathbf{A}\mathbf{Z}_{T-1} .
\end{aligned}$$

Consequently,

$$\mathrm{E}[\mathbf{Z}_T \mid \mathcal{F}_t] = (\mathbf{B}(\mathbf{u})\mathbf{A})^{T-t}\mathbf{Z}_t .$$

Noting that $\langle \mathbf{X}_T, \mathbf{1}\rangle = 1$,

$$\begin{aligned}\langle \mathrm{E}[\mathbf{Z}_T \mid \mathcal{F}_t], \mathbf{1}\rangle &= \left\langle \mathrm{E}\left[\exp\left\{ci\sum_{k=t}^{T}\langle \mathbf{X}_k, \mathbf{u}\rangle\right\}\mathbf{X}_T \;\middle|\; \mathcal{F}_t\right], \mathbf{1}\right\rangle \\ &= \mathrm{E}\left[\exp\left\{ci\sum_{k=t}^{T}\langle \mathbf{X}_k, \mathbf{u}\rangle\right\} \;\middle|\; \mathcal{F}_t\right] .\end{aligned}$$

Hence the result follows. □

We now derive the characteristic function of the occupation time of the Markov chain $\mathbf{X}$. For each $k = 1, 2, \ldots, N$, let $p^k_{t,T}$ be the conditional probability mass function of the random variable $J^k(t, T)$ given $\mathcal{F}_t$. Write $\phi_{J^k(t,T)}$ for the conditional characteristic function of $J^k(t, T)$ given $\mathcal{F}_t$, which is given by:

$$\begin{aligned}\phi_{J^k(t,T)}(u) &:= \mathrm{E}\left[\exp\left\{-\frac{2\pi i}{T-t+2}uJ^k(t,T)\right\} \;\middle|\; \mathbf{X}_t\right] \\ &= \sum_{j=0}^{T} e^{-\frac{2\pi i}{T-t+2}uj} p^k_{t,T}(j) .\end{aligned}$$

This is the univariate DFT of $p^k_{t,T} : D_{t,T} \to [0, 1]$. Then the following corollary gives a compact formula for the conditional characteristic function $\phi_{J^k(t,T)}(u)$.

Corollary 3.1. *For each $u \in D_{t,T}$ and $k = 1, 2, \cdots, N$, the conditional characteristic function of $J^k(t, T)$ given $\mathcal{F}_t$ under P evaluated at u is given by:*

$$\phi_{J^k(t,T)}(u) = \left\langle e^{-\frac{2\pi i}{T-t+2}u\langle \mathbf{X}_t, \mathbf{e}_k\rangle}(\mathbf{B}_2(u)\mathbf{A})^{T-t}\mathbf{X}_t, \mathbf{1}\right\rangle . \tag{2}$$

Here $\mathbf{B}_2(u) := \mathrm{diag}(1, 1, \cdots, e^{-\frac{2\pi i}{T-t+2}u}, \cdots, 1)$, the $(N \times N)$ diagonal matrix with diagonal elements being $(1, 1, \cdots, e^{-\frac{2\pi i}{T-t+2}u}, \cdots, 1)$, where the term $e^{-\frac{2\pi i}{T-t+2}u}$ is at the k-th position.

Proof. Note that $\langle \mathbf{v}, \mathbf{J}(t, T)\rangle = uJ^k(t, T)$ if $\mathbf{v} = (0, 0, \cdots, u, \cdots, 0) \in \mathbb{R}^N$, where the term "$u$" appear at the k-th position. The result then follows directly from Theorem 3.1. □

Using the IDFT, we can invert the conditional characteristic functions to find the conditional probability mass functions for $\mathbf{J}(t,T)$ and $J^k(t,T)$ given $\mathcal{F}_t$ under P. We give the formulae in the sequel.

For the conditional probability mass function of $\mathbf{J}(t,T)$ given $\mathcal{F}_t$ under P,

$$p_{t,T}(\mathbf{j}) = \frac{1}{(T-t+2)^N} \sum_{\mathbf{u} \in D_{t,T}^N} e^{\frac{2\pi i}{T-t+2}\langle \mathbf{u}, \mathbf{j} \rangle} \phi_{\mathbf{J}(t,T)}(\mathbf{u}) \ , \tag{3}$$

and for the conditional probability function of $J^k(t,T)$, $k = 1, 2, \ldots, N$ given $\mathcal{F}_t$ under P,

$$p_{t,T}^k(j) = \frac{1}{T-t+2} \sum_{u=0}^{T-t} e^{\frac{2\pi i}{T-t+2}uj} \phi_{J^k(t,T)}(u). \tag{4}$$

Relevant details of the DFT and the IDFT can be found in Frazier (1999) [14] and Stein and Shakarchi (2003) [26].

For the case of $p_{t,T}^k(j)$, a simpler method is available. Let

$$\psi_{J^k(t,T)}(u) = \mathrm{E}[\exp\{iuJ^k(t,T)\} \mid \mathbf{X}_t] \ .$$

Then it can be shown that

$$\psi_{J^k(t,T)}(u) = \left\langle e^{iu\langle \mathbf{X}_t, \mathbf{e}_k \rangle} (\mathbf{B}_3(u)\mathbf{A})^{T-t} \mathbf{X}_t, \mathbf{1} \right\rangle \ , \tag{5}$$

where $\mathbf{B}_3(u) := \mathrm{diag}(1, \cdots, 1, e^{iu}, 1, \cdots, 1)$, the $N \times N$ diagonal matrix with diagonal elements $(1, \cdots, 1, e^{iu}, 1, \cdots, 1)$; the term e^{iu} is at the k-th position.

Consequently,

$$p_{t,T}^k(j) = \frac{1}{2\pi} \int_0^{2\pi} e^{ijx} \psi_{J^k(t,T)}(x) \mathrm{d}x.$$

A similar account can be found in Elliott and Tsoi (2005) [13], where the special case of a two-state Markov chain was presented.

4. Selection of a Pricing Kernel

In this section, we first present a measure change for the Markov chain and then discuss a method to determine a pricing kernel, or measure, based on the measure change. In particular, the transition matrix of the chain under the selected pricing measure is characterized.

4.1. A Measure Change

Suppose $\mathbf{C} := (c_{ji})_{i,j=1,2,\cdots,N}$ is an $N \times N$ matrix with real-valued entities such that $0 \leq c_{ji} \leq 1$, for all $i, j = 1, 2, \cdots, N$ and

$$\sum_{j=1}^{N} c_{jk} = 1 \ .$$

for all $k = 1, 2, \cdots N$. Consequently, this matrix $\mathbf{C}$ can be a candidate of a transition probability matrix of the Markov chain.

Define, for each $l = 1, 2, \cdots, T$,

$$\lambda_l := \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{c_{ji}}{a_{ji}} \langle \mathbf{X}_l, \mathbf{e}_j \rangle \langle \mathbf{X}_{l-1}, \mathbf{e}_i \rangle \ .$$

Here we assume that $a_{ji} > 0$, for each $i, j = 1, 2, \cdots, N$, so that λ_l is well-defined.

Consider an $\{\mathcal{F}_t\}$-adapted process $\{\Lambda_t\}_{t \in \mathcal{T}}$ defined by:

$$\Lambda_t := \prod_{k=1}^{t} \lambda_k; \quad \Lambda_0 = 1 \ .$$

Then we can define a probability measure $Q \sim P$ by putting:

$$\left. \frac{\mathrm{d}Q}{\mathrm{d}P} \right|_{\mathcal{F}_t} := \Lambda_t \ ,$$

for all $t \in \mathcal{T}$.

The following lemma is standard. We state the result without giving the proof.

Lemma 4.1. *$\{\Lambda_t\}_{t \in \mathcal{T}}$ is an $(\{\mathcal{F}_t\}, P)$-martingale.*

The next proposition gives the dynamics of the chain $\{\mathbf{X}_t\}$ under the new measure Q. This result is also standard and can be found in Elliott et al. (1994) [12]. So we state the result here without giving the proof.

Proposition 4.1. *Under Q, $\{\mathbf{X}_t\}_{t \in \mathcal{T}}$ is a Markov chain with transition probability matrix $\mathbf{C}$.*

4.2. Risk Neutral Transition Matrix

To determine a price for a contingent claim in the Markov chain market, we need to determine a transition matrix under a pricing measure, say a measure Q of the form introduced in Proposition 4.1.

The fundamental theorem of asset pricing by Harrison and Kreps (1979) [15] and Harrison and Pliska (1981, 1983) [16], [17] states that the absence of arbitrage opportunities is "essentially" equivalent to the existence of an equivalent martingale pricing measure under which discounted price processes are martingales. In our model, this martingale condition is equivalent to saying that if Q is an equivalent martingale measure, then

$$S_{t-1} = \mathrm{E}^Q[e^{-r} S_t \mid \mathcal{F}_{t-1}] \ , \quad t = 1, 2, \cdots, T \ . \tag{6}$$

Here E^Q is expectation under Q.

The following proposition gives the martingale condition, or the martingale restriction, in the Markov chain market model.

Proposition 4.2 (Martingale Condition)**.** *Suppose Q is an equivalent measure of the form introduced in Proposition 4.1 so that under Q, $\mathbf{X}$ is a Markov chain with transition matrix $\mathbf{C}$. Then Q is a martingale measure if*

$$\langle \mathbf{s}, (e^{-r}\mathbf{C} - \mathbf{I})\mathbf{e}_k \rangle = 0 \ , \tag{7}$$

for all $k = 1, 2, \ldots, N$.

Proof. Using (6) and the Markov property,

$$S_{t-1} = \mathrm{E}^Q[e^{-r} S_t \mid \mathcal{F}_{t-1}].$$

That is:

$$\begin{aligned} \langle \mathbf{s}, \mathbf{X}_{t-1} \rangle &= \mathrm{E}^Q[e^{-r} \langle \mathbf{s}, \mathbf{X}_t \rangle \mid \mathcal{F}_{t-1}] \\ &= \langle \mathbf{s}, e^{-r} \mathrm{E}^Q[\mathbf{X}_t \mid \mathcal{F}_{t-1}] \rangle \\ &= \langle \mathbf{s}, e^{-r} \mathrm{E}^Q[\mathbf{X}_t \mid \mathbf{X}_{t-1}] \rangle \\ &= \langle \mathbf{s}, e^{-r} \mathbf{C}\mathbf{X}_{t-1} \rangle. \end{aligned}$$

This means that

$$\langle \mathbf{s}, (e^{-r}\mathbf{C} - \mathbf{I})\mathbf{X}_{t-1} \rangle = 0$$

or

$$\langle \mathbf{s}, (e^{-r}\mathbf{C} - \mathbf{I})\mathbf{e}_k \rangle = 0$$

for all $k = 1, 2, \ldots, N$. $\square$

In the case of $N = 2$, the "risk-neutral" transition probability matrix $\mathbf{C}$ can be determined uniquely as

$$\mathbf{C} = \left(\begin{array}{cc} \alpha & 1-\beta \\ 1-\alpha & \beta \end{array} \right) ,$$

where

$$\alpha = \frac{e^r s_1 - s_2}{s_1 - s_2}, \qquad \beta = \frac{s_1 - e^r s_2}{s_1 - s_2}.$$

However, in the case where $N > 2$, the uniqueness is not guaranteed. Therefore, we need to determine a "risk-neutral" transition probability matrix $\mathbf{C}$ subject to some additional conditions. One possible way to determine $\mathbf{C}$ is by calibration to option prices data. The idea would be to determine $\mathbf{C}$ so that the sum of squared deviation of the actual option prices from the theoretical ones is minimized. In practice, this calibration exercise is usually done using price data of simple options, such as standard European-style call or put options.

Suppose the vector of share prices $\mathbf{s}$, the current share price S_0 and the interest rate r are given. Then the theoretical price of a vanilla European call option in the Markov chain market model is:

$$\begin{aligned} c(K, \mathbf{C}, \tau) &= e^{-r\tau} \mathrm{E}^Q[\max\{S_\tau - K, 0\} \mid \mathbf{X}_0] \\ &= e^{-r\tau} \sum_{i=1}^{N} \sum_{j=1}^{N} (s_j - K)^+ (\mathbf{C}^\tau)_{ji} \langle \mathbf{X}_0, \mathbf{e}_i \rangle. \end{aligned}$$

Here K is the strike price and $\tau \in \mathcal{T}$ is the time-to-maturity.
Then a risk-neutral transition probability matrix $\mathbf{C} := (c_{ji})_{i,j=1,2,\cdots,N}$ can be determined using the following conditions:

1. $0 \leq c_{ji} \leq 1$ for all $i, j = 1, 2, \cdots, N$;
2. $\sum_{j=1}^{N} c_{jk} = 1$ for all $k = 1, 2, \cdots N$;
3. $\langle \mathbf{s}, (e^{-r}\mathbf{C} - \mathbf{I})\mathbf{e}_k \rangle = 0$ for all $k \in \mathcal{K}$;
4. $\sum_{j=1}^{L} \sum_{i=1}^{M} |c(K_i, \mathbf{C}, \tau_j) - c_{\mathrm{market}}(K_i, \tau_j)|^2$ is minimized, for given strike prices $K_1, K_2, \cdots, K_M$ and maturities $\tau_1, \tau_2, \cdots, \tau_L$.

Note that by imposing Condition 4, we are using the minimum square deviation price calibration to select a risk-neutral measure. A similar account can be found in Cont and Tankov (2006) [8] and references therein, where the least-square calibration was used to select a risk-neutral measure to price options under an exponential Lévy model, (see Section 3, Problem 2, therein). The basic idea of the calibration is to select a set of risk-neutral parameters in the model dynamics so as to minimize the discrepancy between the theoretical prices implied by the model and the observed market option prices. The calibration idea in Cont and Tankov (2006) [8] could be applied to our Markov chain framework. This would provide a topic for further research. Carr and Cousot (2011) [7] mentioned that discrete-time, finite-state, Markov chain asset price models are among the very few models which are arbitrage free and which can be calibrated to a finite number of observed option quotes. Indeed, the issues of arbitrage-free and exact calibration to option quotes go hand-in-hand. Inspired by the results developed in Carr and Madan (2005) [6], Buehler (2006) [4], Cousot (2007) [9] and Davis and Hobson (2007) [11] established, independently, the existence result of an arbitrage-free Markov chain asset price model which can calibrated exactly to the data under certain explicit conditions.

5. Derivatives Pricing

In Section 3, the conditional characteristic function of the vector of occupation times of a Markov chain in different states was obtained. The conditional joint probability mass function of the occupation times can then be determined from this conditional characteristic function. This result can be applied for derivatives pricing. In particular, the conditional joint mass probability function may be used for option valuation in regime-switching markets. Further, several derivative securities involving occupation times of share prices, such as cumulative options (see Jeanblanc et al. (2009) [19]), step options, switch options, day-in/day-out options, occupation time options (see Linetsky (1999) [20]), require the conditional probability mass function to determine their fair prices. In the sequel, we illustrate the use of the conditional joint probability mass function to price an arithmetic Asian option, a geometric Asian option and an occupation time call option.

5.1. Asian Options

Consider a fixed-strike arithmetic Asian option with maturity at time T, strike price K and discrete monitoring at time points $0, 1, \ldots, T$. Denote the arithmetic average of the share prices up to time T as $A_1(T)$. Then

$$\begin{aligned} A_1(T, \mathbf{J}(0,T)) &= \frac{1}{T+1} \sum_{t=0}^{T} S_t \\ &= \frac{1}{T+1} \sum_{k=1}^{N} s_k J^k(0,T) \\ &= \frac{\langle \mathbf{s}, \mathbf{J}(0,T) \rangle}{T+1}. \end{aligned}$$

Denote the geometric average of the share prices up to time T as $A_2(T)$. Then

$$\begin{aligned} A_2(T, \mathbf{J}(0,T)) &= (S_0 S_1 S_2 \cdots S_T)^{1/(T+1)} \\ &= \exp\left\{ \frac{1}{T+1} \sum_{t=0}^{T} \ln S_t \right\} \\ &= \exp\left\{ \frac{1}{T+1} \sum_{k=1}^{N} J^k(0,T) \ln s_k \right\} \\ &= \exp\left\{ \frac{\langle \ln \mathbf{s}, \mathbf{J}(0,T) \rangle}{T+1} \right\}, \end{aligned}$$

where $\ln \mathbf{s} = (\ln s_1, \ln s_2, \cdots, \ln s_N)'$.

The payoff of the Asian call option at time T is given by:

$$\max\{A_i(T) - K, 0\}$$

where $i = 1, 2$, depending on the type of averaging of the option.

A price of the Asian option can then be determined as:

$$\begin{aligned} p_{\text{asian}} &= \mathrm{E}^Q[e^{-rT} \max\{A_i - K, 0\} \mid \mathbf{X}_0] \\ &= e^{-rT} \sum_{\mathbf{j} \in D'_{t,T}} \max\{A_i(T, \mathbf{j}) - K, 0\} Q(\mathbf{J}(0,T) = \mathbf{j} \mid \mathbf{X}_0) \ , \end{aligned}$$

where

$$\begin{aligned}
&Q(\mathbf{J}(0,T) = \mathbf{j} \mid \mathbf{X}_0) \\
&= \frac{1}{(T+2)^N} \sum_{\mathbf{u} \in D_{0,T}^N} e^{\frac{2\pi i}{T+2}\langle \mathbf{u}, \mathbf{j} \rangle} \langle e^{-\frac{2\pi i}{T+2}\langle \mathbf{X}_0, \mathbf{u} \rangle} (\mathbf{B}(\mathbf{u})\mathbf{C})^T \mathbf{X}_0, \mathbf{1} \rangle \\
&= \frac{1}{(T+2)^N} \sum_{\mathbf{u} \in D_{0,T}^N} e^{\frac{2\pi i}{T+2}[(j_1 - x_1)u_1 + \cdots + (j_N - x_N)u_N]} \langle (\mathbf{B}(\mathbf{u})\mathbf{C})^T \mathbf{X}_0, \mathbf{1} \rangle \ ,
\end{aligned}$$

following Theorem 3.1, if $\mathbf{X}_0 := (x_1, \cdots, x_N)$; $\mathbf{j} := (j_1, \cdots, j_N)$; $\mathbf{u} := (u_1, \cdots, u_N)$.

The analytical formula we have for the Asian option gives the exact price rather than an approximation.

5.2. Occupation Time Options

An occupation time call option, as mentioned in Linetsky (1999) [20], is an option on the occupation time whose the payoff is given by:

$$\max\{\tau_B^- - \alpha T, 0\} \ ,$$

where τ_B^- is the amount of time that the share price is lower than the fixed barrier level B, and α is a fixed constant. The payoff of an occupation time put option can be defined similarly.

Suppose $s_k \le B < s_{k+1}$, for some $k = 1, 2, \ldots, N-1$. Note that the cases where $B < s_1$ and where $s_N \le B$ are trivial, so we do not discuss them. Then,

$$\begin{aligned}
\tau_B^- &= J^1(0,T) + J^2(0,T) + \cdots + J^k(0,T) \\
&= \langle \mathbf{J}(0,T), \mathbf{1}_k \rangle \ ,
\end{aligned}$$

where $\mathbf{1}_k = (1, 1, \cdots, \underbrace{1}_{k^{\text{th}}}, 0, \cdots, 0) \in \mathbb{R}^N$.

The characteristic function of τ_B^- is given below, as a corollary of Theorem 3.1. The proof is omitted.

Corollary 5.1. *The characteristic function of τ_B^- is given by:*

$$\begin{aligned}
\psi_{\tau_B^-}(u) &:= \mathrm{E}[e^{iu\tau_B^-} \mid \mathbf{X}_0] \\
&= \left\langle e^{iu\langle \mathbf{X}_t, \mathbf{e}_k \rangle} (\mathbf{B}_4(u)\mathbf{C}^*)^{T-t} \mathbf{X}_t, \mathbf{1} \right\rangle
\end{aligned}$$

where $\mathbf{B}_4(u) = \mathrm{diag}(e^{iu}, e^{iu}, \cdots, \underbrace{e^{iu}}_{k^{\text{th}}}, 1, \cdots, 1)$.

Similarly to the argument in previous section,

$$Q(\tau_B^- = k \mid \mathbf{X}_0) = \frac{1}{2\pi} \int_0^{2\pi} e^{ijx} \psi_{\tau_B^-}(x) \mathrm{d}x.$$

Given a barrier level B and α, the price of the occupation time call in our model is:

$$\begin{aligned} p_{\text{occupation}} &= \mathrm{E}^Q[e^{-rT} \max\{\tau_B^- - \alpha T, 0\} \mid \mathbf{X}_0] \\ &= e^{-rT} \sum_{k=0}^{T+1} \max\{k - \alpha T, 0\} Q(\tau_B^- = k \mid \mathbf{X}_0). \end{aligned}$$

6. Conclusion

We considered the valuation of Asian options and occupation time options in a Markov chain market model, where uncertainty of share price movements is modulated by a discrete-time, finite-state, Markov chain. The characteristic function for occupation times of the chain was derived. It was then used to derive pricing formulae for Asian options and occupation time options. The issue of selecting a pricing kernel in the Markov chain market model was also discussed.

Acknowledgments: The authors would like to thank the referees for helpful comments. Robert J. Elliott and Tak Kuen Siu would like to acknowledge the Discovery Grant from the Australian Research Council (ARC), (Project No.: DP1096243).

[1] L. Bachelier, Théorie de la spéculation, Annales Scientifiques de l'École Normale Supérieure 3 (1900) 21–86.

[2] F. Black, M. Scholes, The pricing of options and corporate liabilities, J. Polit. Econ. 81 (1973) 637–659.

[3] P. Boyle, Option Valuation Using a Three-Jump Process, Int. Options J. 3 (1986) 7–12.

[4] H. Buehler, Expensive martingales. Quant. Finance 6 (2006) 207–218.

[5] J. Buffington, R.J. Elliott, American options with regime switching, Int. J. Theor. and Appl. Finance. 5 (2002) 497–514.

[6] P. Carr, D. Madan, A note on sufficient conditions for no arbitrage, Finance Lett. 2 (2005) 125–130.

[7] P. Carr, L. Cousot, A PDE approach to jump-diffusions, Quant. Finance 11 (2011) 33–52.

[8] R. Cont and P. Tankov, Retrieving Lévy processes from option prices: regularization of an ill-posed inverse problem, SIAM J. Control Optim. 45 (2006) 1–25.

[9] L. Cousot, Conditions on option prices for absence of arbitrage and exact calibration, J. Bank. Finance 31 (2007) 3377–3397.

[10] J.C. Cox, S.A. Ross, M. Rubinstein, Option Pricing: A Simplified Approach, J. Financ. Econ. 7 (1979) 229–263.

[11] M. Davis, D. Hobson, The range of traded option prices, Math. Finance 17 (2007) 1–14.

[12] R.J. Elliott, L. Aggoun, J.B. Moore, Hidden Markov Model: Estimation and Control, Springers-Verlag, 1994.

[13] R.J. Elliott, A.H. Tsoi, Hidden Markov Filter Estimation of the Occurrence Time of an Event in a Financial Market, Stoch. Anal. Appl. 23 (2005) 1165–1177.

[14] M.W. Frazier, An Introduction to Wavelets Through Linear Algebra, Springer, 1999.

[15] J.M. Harrison, D.M. Kreps, Martingales and arbitrage in multiperiod securities markets, J. Econ. Theory 20 (1979) 381–408.

[16] J.M. Harrison, S. Pliska, Martingales and stochastic integrals in the theory of continuous trading, Stoch. Proc. Appl. 11 (1981) 215–260.

[17] J.M. Harrison, S. Pliska, A stochastic calculus model for continuous trading, complete markets, Stoch. Proc. Appl. 15 (1983) 313–316.

[18] H. He, Convergence from Discrete- to Continuous-Time Contingent Claims Prices, Rev. Financ. Stud. 3 (1990) 523–46.

[19] M. Jeanblanc, M. Yor, M. Chesney, Mathematical Methods for Financial Markets, Springer Finance, London, 2009.

[20] V. Linetsky, Step Options, Math. Finance. 9 (1999) 55–96.

[21] R.C. Merton, The theory of rational option pricing, Bell J. Econ. and Manag. Sci. 4 (1973) 141–183.

[22] R. Norberg, Anomalous PDE-s in Markov chains: Domains of validity and numerical solutions, Finance and Stochastics 9 (2005) 519–537.

[23] S.R. Pliska. Introduction to Mathematical Finance, Blackwell Publishing, United States, 1997.

[24] S.M. Ross, A First Course in Probability, fifth ed., Prentice Hall, New York, 1976.

[25] N. Song, W.K. Ching, T.K. Siu, E.S. Fung, M.K. Ng, Option Valuation Under a Multivariate Markov Chain Model, Third International Joint Conference on Computational Science and Optimization (2010) 177–181.

[26] E.M. Stein, R. Shakarchi, Fourier Analysis: An Introduction, Princeton University Press, 2003.

[27] E. Valakevicius, Continuous Time Markov Chain Model of Asset Prices Distribution, Computational Science - ICCS 2009. Lecture Notes on Computer Science 5545 (2009) 598–605.

[28] J. van der Hoek, R.J. Elliott, Asset Pricing Using Finite State Markov Chain Stochastic Discount Factor, *preprint*, 2010.