Modelling Share Prices as a Random Walk on a Markov Chain

A Thesis presented for the degree of
Doctor of Philosophy

by

Rukiye Samci Karadeniz

Supervisor: Prof Sergey Utev
Department of Mathematics
University of Leicester
England

November 2016

Abstract

In the financial area, a simple but also realistic means of modelling real data is very important. Several approaches are considered to model and analyse the data presented herein. We start by considering a random walk on an additive functional of a discrete time Markov chain perturbed by Gaussian noise as a model for the data as working with a continuous time model is more convenient for option prices. Therefore, we consider the renowned (and open) embedding problem for Markov chains: not every discrete time Markov chain has an underlying continuous time Markov chain. One of the main goals of this research is to analyse whether the discrete time model permits extension or embedding to the continuous time model. In addition, the volatility of share price data is estimated and analysed by the same procedure as for share price processes. This part of the research is an extensive case study on the embedding problem for financial data and its volatility.

Another approach to modelling share price data is to consider a random walk on the lamplighter group. Specifically, we model data as a Markov chain with a hidden random walk on the lamplighter group $\mathbb{Z}_3$ and on the tensor product of groups $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. The lamplighter group has a specific structure where the hidden information is actually explicit. We assume that the positions of the lamplighters are known, but we do not know the status of the lamps. This is referred to as a hidden random walk on the lamplighter group. A biased random walk is constructed to fit the data. Monte Carlo simulations are used to find the best fit for smallest trace norm difference of the transition matrices for the tensor product of the original transition matrices from the (appropriately split) data.

In addition, splitting data is a key method for both our first and second models. The tensor product structure comes from the split of the data. This requires us to deal with the missing data. We apply a variety of statistical techniques such as Expectation- Maximization Algorithm and Machine Learning Algorithm (C4.5).

In this work we also analyse the quantum data and compute option prices for the binomial model via quantum data.

Acknowledgements

First, I would like to thank my supervisor Prof. Sergey Utev. He helped me a lot with his enthusiasm, knowledge, and amazing motivation for this thesis. He inspired and guided me with his experience and great personality during the whole of my PhD studies.

I also want to thank all the members of staff of the mathematics department at the University of Leicester for their professional support and for sharing their knowledge through departmental meetings and seminars. They were very helpful in developing my skills and increasing my knowledge in the field.

I must also thank all my colleagues in the mathematics department at the University of Leicester for their friendly and freely given aid. The approachable interactions with them always motivate me in this harmonious environment.

Additionally, I would like to thank my husband, who was also one of my proofreaders, for his encouragement and motivation. I must express my profound thanks to my mother, father and sisters, Elif Erdogan, Maside Samci and Rumeysa Samci for their patience and understanding during my PhD studies. Special thanks go to my nieces Zeynep and Betul for their love. I would also like to extend my thanks to my close friends, Seda Ozdemir, Tayyibe Erten, Deniz Sevinc and Hilal Kaya, for their continued support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is based on the analysis of arbitrarily chosen groups of share prices with relatively small data sizes (around 250 closing prices for each group). The share price data were chosen arbitrarily from the Internet for BP day-by-day closing share prices for four different financial years, (April to April) 2009-2010, 2010-2011, 2011-2012, 2012-2013. We work with the log transformed data, i.e., logged data. The aim of this thesis is to construct and analyse several models for the logged data:

(1) First model. We analyse daily share price data and construct our first model-additive functional of a discrete Markov chain perturbed by Gaussian noise. Therefore, we develop an additive functional of a discrete Markov chain perturbed by a Gaussian noise model for the share price adjusted closing data in Chapter 2. The intention was to find the model which on the one hand is not too complicated, and on the other hand fits the data.

(2) Embedding problem. Our model is discrete time. However, financial derivatives are better understood for continuous time models. Furthermore, there are several other reasons why we considered the junction between the discrete time model and the continuous time model [75], [84], [98]. Hence, one of main goals of this research is to treat the discrete time model as a continuous time model. Therefore, we analyse a so-called embedding problem, which is roughly equivalent to asking whether the discrete time Markov chain can be treated as a continuous time Markov chain observed at discrete times.

In the financial area, there are not only continuous time models, such as the

Black-Scholes and Levy process models, but also some discrete-models, such as the Cox-Rubinstein model.

If the model is converted to a continuous time model (embeddable), this means that the result (data) is observable each time. This is a plausible and important methodology in the financial area.

If the model is a continuous time model, many existing formulae (such as option pricing) are applicable to the model.

Additionally, for these analysis, the share prices of twenty different companies for four different financial years are considered to check embbedability.

Overall, this part of research represents an extensive case study on the embedding problem for financial data and its volatility. It demonstrates a real-world financial application of the importance of the embedding problem.

(3) We also model the logged data as a Markov chain with a hidden random walk on the so-called lamplighter group, which are wreath products of groups. Specifically, the hidden random walk is constructed on the lamplighter group $\mathbb{Z}_3$ and on the tensor product of the groups $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. Also, a biased random walk (as introduced [72]) is constructed to fit the data.

(4) Finally, we compute the option price for a binomial model via quantum data. In particular, we consider the quantum binomial model. The real novelty is the analysis of the quantum data. We estimate the parameters of the quantum two-step binomial market. We then find option prices for a multi-step quantum binomial market.

Furthermore, to be consistent, we apply the same approach to the volatility process as to the share price process. We were motivated partly by [35] and regime switching GARCH-type models (e.g. [19]). It is well known that in the classical Black-Scholes model (geometric Brownian motion model), the no arbitrage option price depends solely on interest rate and volatility [35]. However, several studies on the empirical estimation of the volatility show that the Black-Scholes model does not provide a sufficiently good fit to the data. Although many models have been constructed to incorporate the volatility variability [89], research in this area is still ongoing.

In the following section, we present some terminology that will be useful for a better understanding of this work. Then, in the literature review in Section 1.2, the embedding problem, random walk on groups and lamplighter group are examined. Thereafter the algebraic and statistical methods used to derive our models are introduced, respectively, in Section 1.2.2 and Section 1.2.3. Finally, overall thesis results are presented in Section 1.3, and the outline of the thesis is given in the last section of this chapter.

1.1 Basic Facts and Preliminaries

This section is devoted to terminology related to Markov chains and random walks on groups and the embedding problem. Also, some statistical terminology and the definitions of specified norms are given. We introduce the basic tools which are useful to a better understanding of this research.

1.1.1 Trace Norm

Definition 1.1.1 (Trace). *In linear algebra, the sum of the element on the main diagonal is defined as the trace of an n-by-n square matrix A.*

$$\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

where a_{nn} is the entry on the nth row and the nth column of A.

The trace of a matrix is the sum of the eigenvalues and is invariant with respect to a change in the basis. Also, a trace is defined only for only square matrices. The matrix A with complex numbers $A = (a_{ij})$ is called a self-adjoint matrix if $a_{ij} = \bar{a}_{ji}$. When A has real numbers, its self-adjoint is also symmetric. We say that a self-adjoint matrix A is non-negative if A has non-negative eigenvalues [86].

The trace also plays a central role in the distribution of quadratic forms. Regulation via the trace norm (sum of singular values) is a well known method to estimate low rank rectangular matrices [20].

Consider the problem of approximating a noisy target matrix Y with another matrix X [60]. This problem frequently arises in practice. A common general scheme for solving such problems is to select a matrix X that minimizes some combination of the complexity of X and the discrepancy between X and Y. The crux here is the choice of the measure of complexity for X and the measure of discrepancy between X and Y. The most common method used to measure the complexity of a matrix is the use of the rank [56]. The trace norm is the convex envelope of the rank over the unit ball of the spectral norm [51], [83]. Nowadays, the trace norm, rank norm and maximum norm are used as an alternative method to measure complexity with strong connections to maximum-margin linear classification [57].

Let us examine the properties of the trace [86]:

- If $A \leq B$, then $\mathrm{tr}(A) \leq \mathrm{tr}(B)$.
  That is, $B - A$ has non-negative elements. This is also true when $B - A$ is a self-adjoint, non-negative matrix.

- The trace is a linear mapping where A and B are square matrices and c is a constant:

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B),$$

$$\mathrm{tr}(cA) = c\,\mathrm{tr}(A).$$

- A matrix and its transpose have the same trace:

$$\mathrm{tr}(A) = \mathrm{tr}(A^{\intercal}).$$

- **Trace of a product:** The trace of a product is the sum of entry wise product of elements:

$$\mathrm{tr}(A^{\intercal}B) = \mathrm{tr}(AB^{\intercal}) = \mathrm{tr}(B^{\intercal}A) = \mathrm{tr}(BA^{\intercal}) = \sum_{i,j} a_{ij}b_{ij}.$$

- The matrices in the trace of a product can be switched where A is an $m \times n$ and B is a $n \times m$

$$\mathrm{tr}(AB) = \sum_{i=1}^{m}(ab_{ii}) = \sum_{i=1}^{m}\sum_{j=1}^{n}(a_{ij})(b_{ji}) = \sum_{j=1}^{n}\sum_{i=1}^{m}(b_{ji})(a_{ij}) = \sum_{j=1}^{m}(ba_{jj})$$

$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$

- **Cyclic property:** The trace is invariant under cyclic permutations:

$$\operatorname{tr}(ABCD) = \operatorname{tr}(BCDA) = \operatorname{tr}(CDAB) = \operatorname{tr}(DABC).$$

  Arbitrary permutations are not allowed ($\operatorname{tr}(ABC) \neq \operatorname{tr}(ACB)$).

  However, if the matrices are symmetric, the following permutation is allowed:

  $\operatorname{tr}(ABC) = \operatorname{tr}(A^\intercal B^\intercal C^\intercal) = \operatorname{tr}(A^\intercal (CB)^\intercal) = \operatorname{tr}((CB)^\intercal A^\intercal) = \operatorname{tr}((ACB)^\intercal) = \operatorname{tr}(ACB)$.

- If A is a square matrix with real or complex entries and $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A,

$$\operatorname{tr}(A) = \sum_i \lambda_i.$$

  This follows from the fact that A is always similar to its Jordan form, an upper triangular matrix having $\lambda_1, \dots, \lambda_n$ on the main diagonal. In contrast, the determinant of A is the product of its eigenvalues:

$$\det(A) = \prod_i \lambda_i.$$

  More generally,

$$\operatorname{tr}(A^k) = \sum_i \lambda_i^k.$$

Proposition 1.1.1. *Let $A_i, i = 1, \dots, n$, 2×2 matrices and $A_i = A_i^\intercal \geq 0$. A_i is associated if* $\mathbb{E} \operatorname{tr}(A_1 \dots A_n) \geq \operatorname{tr}(\mathbb{E}(A_1) \dots \operatorname{tr}(\mathbb{E}(A_n)$.

Proof:

$A = (a_{ij})$ then $\mathbb{E}[A] = (\mathbb{E}[a_{ij}])$ is the matrix with expectation of elements.
$A_i \geq 0$ and $A = A_i^\intercal$. So,

$$\mathbb{E}[A_1 \dots A_n] \geq \mathbb{E}[A_1] \dots \mathbb{E}[A_n]$$

and taking the trace of both sides:

$$\operatorname{tr}(\mathbb{E}[A_1 \dots A_n]) \geq \operatorname{tr}(\mathbb{E}[A_1] \dots \mathbb{E}[A_n])$$

then use the linearity of the expectation:

$$\mathbb{E}[\mathrm{tr}(A_1 \dots A_n)] \geq \mathrm{tr}(\mathbb{E}(A_1) \dots \mathrm{tr}(\mathbb{E}(A_n).$$

1.1.2 Tensor Product

The term "tensor product" refers to another way of constructing a big vector space out of two (or more) smaller vector spaces. It is also called the Kronecker Product.

Let A and B be the matrices,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \cdots & \ddots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \text{ and } B = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \cdots & \ddots & \cdots \\ b_{m1} & b_{m2} & \cdots & b_{mm} \end{pmatrix}$$

and their tensor product is [6]

$$A \otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1m} & a_{12}b_{11} & \cdots & a_{1n}b_{11} & \cdots & a_{1n}b_{1m} \\ a_{11}b_{21} & a_{11}b_{22} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{11}b_{31} & a_{11}b_{32} & a_{11}b_{33} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \vdots & \cdot & \cdot & \ddots & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{11}b_{m1} & \cdot & \cdot & a_{11}b_{1m} & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{21}b_{11} & \cdot & \cdot & \cdot & a_{22}b_{11} & \cdot & \cdot & \cdot & \cdot \\ \vdots & \cdot & \cdot & \cdot & \cdot & a_{22}b_{1m} & \cdot & \cdot & \cdot \\ a_{31}b_{11} & \cdot & \cdot & \cdot & \cdot & \cdot & a_{31}b_{11} & \cdot & \cdot \\ \vdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \ddots & \cdot \\ a_{n1}b_{m1} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & a_{nn}b_{1m} \end{pmatrix}$$

Proposition 1.1.2. *$A \otimes B$ is a stochastic matrix if A, B are stochastic matrices.*

A stochastic matrix has non-negative elements where the sum of the elements in each row equals 1.

Let $A = (a_{ij}), \quad a_{ij} \geq 0, \quad j = \{1, \dots, n\}$ and $B = (b_{km}), \quad b_{km} \geq 0, \quad m = \{1, \dots, l\}$ be stochastic matrices:

$\sum_j a_{ij} = 1, \quad \sum_m b_{km} = 1$

We begin by taking their tensor product:

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & \cdots \\ a_{21}B & \cdots & \cdots \\ \vdots & \ddots & \cdots \\ a_{n1}B & \cdots & a_{ij}B \end{pmatrix}$$

$$(A \otimes B)_{ik,jm} = a_{ij}b_{km} \geq 0$$

$$\begin{aligned} \sum_{j,m}(A \otimes B)_{ik,jm} &= \sum_m \sum_j a_{ij}b_{km} \\ &= (\sum_j a_{ij})(\sum_m b_{km}) \\ &= 1 \end{aligned}$$

Specifically, consider the 2×2 matrices:

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$\Rightarrow \quad A \otimes B = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \end{pmatrix}$$

If A is a stochastic matrix:

$$a_{11} + a_{12} = 1 \tag{1.1.1}$$

$$a_{21} + a_{22} = 1 \tag{1.1.2}$$

and if B is a stochastic matrix:

$$b_{11} + b_{12} = 1 \tag{1.1.3}$$

$$b_{21} + b_{22} = 1 \tag{1.1.4}$$

Let us examine the first row of the tensor product of the matrices $A \otimes B$ to show that it is a stochastic matrix:

$$a_{11}b_{11} + a_{11}b_{12} + a_{12}b_{11} + a_{12}b_{12} = a_{11}(b_{11} + b_{12}) + a_{12}(b_{11} + b_{12}) = 1 \tag{1.1.5}$$

1.1.1 and 1.1.3 are used to show the equation 1.1.5 is equal to 1. Further, it is obvious that the sum of the each rows is equal to 1. Therefore, the tensor product of the matrices $A \otimes B$ is a stochastic matrix.

Proposition 1.1.3. *Tensor products are not symmetric, in general.*

$$A \otimes B \neq B \otimes A$$

Proposition 1.1.4. *Let X and Y be independent Markov chains with transition matrices P_X and P_Y, respectively. Then, the Markov chain $Z = (X, Y)$ has a transition matrix $P = P_X \otimes P_Y$. Moreover, assume that $P_X = e^{Q_X}$, $P_Y = e^{Q_Y}$ are embeddable. Then, so is $P = e^Q$ with*

$$Q = Q_X \otimes I + I \otimes Q_Y.$$

Proof:

Although the proof is known (for example, it follows from the properties of the Kronecker sum): $e^A \otimes e^B = e^{A \otimes B}$ see e.g. [74]), we give the proof for the reader's convenience.

- Taylor series of exponential function: $e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots$
- A is a $n \times n$ matrix: $e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!}A^k$ where I is the $n \times n$ identity matrix and $A^0 = I$.
- $e^{(tA)} = I + tA + \frac{t^2}{2!}A^2 + \frac{t^3}{3!}A^3 + \cdots$.

Let $P_X^t = e^{tQ_X}$ and $P_Y^t = e^{tQ_Y}$. Then, use the tensor product linearity and a Taylor series of the exponential function of matrices:

$$\begin{aligned}
\lim_{t\downarrow 0} \frac{P_X^t \otimes P_Y^t - \widetilde{I}}{t} &= \lim_{t\downarrow 0} \frac{e^{tQ_X} \otimes e^{tQ_Y} - \widetilde{I}}{t} \\
&= \lim_{t\downarrow 0} \frac{1}{t}((I + tQ_X + O(t^2)) \otimes (I + tQ_Y + O(t^2)) - I) \\
&= \lim_{t\downarrow 0} \frac{1}{t}(I \otimes I + tI \otimes Q_Y + tQ_X \otimes I + O(t^2) - \widetilde{I}) \\
&= I \otimes Q_Y + Q_X \otimes I \\
&= Q_X \otimes I + I \otimes Q_Y
\end{aligned}$$

where $I \otimes I = \widetilde{I}$.

Example: Let $P_X = \begin{pmatrix} 0.5 & 0.5 \\ 0.1429 & 0.8571 \end{pmatrix}$ and $P_Y = \begin{pmatrix} 0.8750 & 0.1250 \\ 0 & 1 \end{pmatrix}$.

Then, the Q matrices are computed through the algebraic approach of the embedding problem (see Section 3.3 for details).

$$Q_X = \begin{pmatrix} -0.8009 & 0.8009 \\ 0.2289 & -0.2289 \end{pmatrix}, \quad Q_Y = \begin{pmatrix} -0.1335 & 0.1335 \\ 0 & 0 \end{pmatrix}$$

$$Q_X \otimes I = \begin{pmatrix} -0.8009 & 0 & 0.8009 & 0 \\ 0 & -0.8009 & 0 & 0.8009 \\ 0.2289 & 0 & -0.2289 & 0 \\ 0 & 0.2289 & 0 & -0.2289 \end{pmatrix}$$

$$I \otimes Q_Y = \begin{pmatrix} -0.1335 & 0.1335 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -0.1335 & 0.1335 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence,

$$Q_X \otimes I + I \otimes Q_Y == \begin{pmatrix} -0.9343 & 0.1335 & 0.8009 & 0 \\ 0 & -0.8009 & 0 & 0.8009 \\ 0.2289 & 0 & -0.3624 & 0.1335 \\ 0 & 0.2289 & 0 & -0.2289 \end{pmatrix}$$

Also,

$$P_X \otimes P_Y = \begin{pmatrix} 0.4375 & 0.0625 & 0.4375 & 0.0625 \\ 0 & 0.5 & 0 & 0.5 \\ 0.1250 & 0.0179 & 0.75 & 0.1071 \\ 0 & 0.1429 & 0 & 0.8571 \end{pmatrix}$$

$$P = e^Q \Rightarrow Q = \begin{pmatrix} -0.9343 & 0.1335 & 0.8009 & 0 \\ 0 & -0.8009 & 0 & 0.8009 \\ 0.2289 & 0 & -0.3623 & 0.1335 \\ 0 & 0.2289 & 0 & -0.2289 \end{pmatrix}$$

Therefore,

$$Q = Q_X \otimes I + I \otimes Q_Y.$$

Notice that by Proposition 1.1.4, the tensor product $P = \otimes_{i=1}^n P_i$ of many small embeddable matrices. P_i is embeddable; however, P is a large matrix for which the embedding property is much harder to verify, in general, since there is no necessary and sufficient condition. It would be interesting to compare the embeddability as $n \to \infty$

Trace of the Tensor Product

Proposition 1.1.5.

$$\operatorname{tr}(A \otimes B) = \operatorname{tr}(B \otimes A) = \operatorname{tr}(A)\operatorname{tr}(B).$$

Proof:

$$\begin{aligned} \operatorname{tr}(A \otimes B) &= (a_{11}b_{11} + a_{11}b_{22} + \cdots + a_{11}b_{1m}) + \cdots + (a_{nn}b_{11} + \cdots + a_{nn}b_{1m}) \\ &= a_{11}(b_{11} + \cdots + b_{1m}) + a_{22}(b_{11} + \cdots + b_{1m}) + \cdots + a_{nn}(b_{11} + \cdots + b_{1m}) \\ &= (a_{11} + a_{22} + \cdots + a_{nn})(b_{11} + \cdots + b_{1m}) \\ &= \operatorname{tr}(A)\operatorname{tr}(B). \end{aligned}$$

Proposition 1.1.6. *Let* $A, B \geq 0$. *Then,* $\operatorname{tr}(A \otimes B) \geq 0$ *and* $\operatorname{tr}(A \otimes A) \geq 0$.

Proof: Apply $A = T^{-1}DT$ and the triangle rule $\mathrm{tr}(ABC) = \mathrm{tr}(BCA)$:
When $A \geq 0$, $\mathrm{tr}(A) = \mathrm{tr}(D)$, all eigenvalues are non-negative. So,

$$A, B \geq 0 \Rightarrow \mathrm{tr}(A), \mathrm{tr}(B) \geq 0$$

Therefore,

$$\mathrm{tr}(A \otimes B) = \mathrm{tr}(A)\,\mathrm{tr}(B) \geq 0 \text{ and } \mathrm{tr}(A \otimes A) = (\mathrm{tr}(A))^2 \geq 0$$

where $A, B \geq 0$.

Proposition 1.1.7. *Let $A, B, C \geq 0$. Then,*

$$\mathrm{tr}(A \otimes B \otimes C) \geq 0.$$

Proof of this proposition is clear by using $\mathrm{tr}(ABC) = \mathrm{tr}(A)\,\mathrm{tr}(B)\,\mathrm{tr}(C)$.

Proposition 1.1.8. *Let $A_1, \cdots, A_{2^r}$ are 2×2, self-adjoint ($A = A^*$ (equal to transposed matrix)) and non-negative matrices (with non-negative eigenvalues). Then,*

$$\mathrm{tr}((A_1 \otimes A_2) \otimes \cdots (A_{2^r-3} \otimes A_{2^r-2})(A_{2^r-1} \otimes A_{2^r})) \geq 0$$

1.1.3 Markov chains

A Markov chain is defined by a countable state space S and a transition matrix P. Also, an initial position (at time 0) has to be specified. Also, the probability of moving from i to j is represented by P_{ij}. Consider a sequence of random variables $(X_n) \geq 0$ that takes on a finite number of possible values on a countable state space S; it can be said that the process is in state i at time n if $X_n = i$. Then, the state is changed from state i to the next state j with a fixed probability P_{ij}. This stochastic process is called a Markov Chain if it has a Markov Property

$P[X_{n+1} = j | X_n = i; X_{n-1} = i_{n-1}, , X_1 = i, X_0 = i_0] = P(X_{n+1} = j | X_n = i) = P_{ij}$

for all states $i_0, i_1, , i_{n-1}, i, j \quad$ and all $\quad n \geq 0$.

In order to model the random variables X_N, one has to find a suitable probability space on which the random position after n steps can be described as the n-th random variable of a Markov chain. The usual choice of the probability space is the trajectory space $\Omega = S^{\mathbb{Z}^+}$ equipped with the product σ-algebra arising from the discrete one on S. An element $\omega = (x_0, x_1, x_2,)$ of Ω represents a possible evolution (trajectory), that is, a possible sequence of points visited one after the other by the Markov chain. Then, X_n is the n-th projection from Ω to S. This describes the Markov chain starting at i, when Ω is equipped with the probability measure given via the Kolmogorov extension theorem by

$$\mathbb{P}_x[X_0 = x_0, X_1 = x_1, , X_n = x_n] = \delta_x(x_0 p(x_0, x_1))p(x_{n-1}, x_n)$$

where $\delta_x(y) = 1$ if $x = y$ and $p(.,.)$ is the transition matrix.

The associated expectation is denoted by $\mathbb{E}_x$. Also, the n-step transition probability, which is the probability of getting from i to j in n steps, is represented as below:

$$P_{ij}^{(n)} = \mathbb{P}_i[X_n = j],$$

Also, the (i, j)-entry of the matrix power P^n, with $P^0 = I$ (I is the identity matrix over S) [96].

Definition 1.1.2. *A Markov chain is irreducible if for every $i, j \in S$, there is some $n \in \mathbb{N}$ such that $P_{ij}^{(n)} > 0$ [96].*

This means that every state $j \in S$ can be reached from every other state $i \in S$ with a positive probability. In this research, we shall always require that the state space is infinite and all states communicate, i.e., the Markov chain is irreducible.

1.1.4 Random Walk on Graphs

Random walk on the n-cycle. Let's take modulo n remainder set $\Omega = \mathbb{Z}_n = \{0, 1, \cdots, n-1\}$. And transition matrix is

$$P(x,y) = \begin{cases} 1/2 & \text{if } y \equiv x+1 \text{ (mod n)}, \\ 1/2 & \text{if } y \equiv x-1 \text{ (mod n)}, \\ 0 & \text{otherwise.} \end{cases}$$

The random walk on the n-cycle is the correlated Markov chain (X_t). The walker can go one step clockwise or one step the other way at each cycle. The random walk on the n-cycle is a simple case of an important type of Markov chain [12].

Graph. A graph $G = (V, E)$ consists of a vertex set V and edge set E, where the elements of E are unordered pairs of vertices: $E \subset \{\{x, y\} : x, y \in V, x \neq y\}$. V is considered to be a set of dots where two dots x and y are joined by a line if, and only if, $\{x, y\}$ is an element of the edge set E. When $\{x, y\} \in E$, y is a neighbour of x (also, x is a neighbour of y), and this relation is represented by $x \sim y$. The degree of a vertex x ($deg(x)$) is the number of neighbours of x [4].

A simple random walk on a given graph $G = (V, E)$ can be defined as being the Markov chain with a state space V and a transition matrix

$$P(x, y) = \begin{cases} \frac{1}{deg(x)} & \text{if } y \sim x, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, when the chain is at vertex x, it checks all the neighbours of x, chooses one uniformly at random, and moves to the chosen vertex [12].

A finite graph is a graph $G = (V, E)$ such that V and E are finite sets. A graph is referred to as being connected if every pair of distinct vertices in the graph is connected [4].

The graph G is referred to as being locally finite if every vertex has a finite degree. Furthermore, G has a bounded geometry if it is connected with bounded vertex degrees [96].

Example: A random walk on a graph $G = (V, E)$ such that $V = \{1, 2, 3\}$ and $E = \{1 \leftrightarrow 2, 1 \leftrightarrow 3\}$

$P(1 \to 2) = P(1, 2) = \frac{1}{2}, \quad d_1 = 2$

$P(1 \to 3) = P(1, 3) = \frac{1}{2}, \quad d_2 = 2$

$P(2 \to 1) = P(2, 1) = 1, \quad d_3 = 1$

$P(3 \to 1) = P(3, 1) = 1 \quad d_4 = 1.$

1.1.5 Random Walk on Groups

Groups. A group consists of set G with a binary operation $\circ$ on G satisfying the following axioms [29]:

(i) *Closure.* For all $a, b \in G$, we have $a \circ b \in G$.

(ii) *Associativity.* For all $a, b, c \in G$, we have $(a \circ b) \circ c = a \circ (b \circ c)$.

(iii) *Identity.* There is an element $e \in G$ satisfying $e \circ a = a \circ e = a$ for all $a \in G$.

(iv) *Inverse.* For all $a \in G$, there is an element $a^* \in G$ satisfying $a \circ a^* = a^* \circ a = e$ (where e is as in the identity element and is unique).

A semi group is a set G_S with a binary operation $\circ$ that satisfies the axioms (i) and (ii) (closure and associativity), but which does not necessarily satisfy the axioms (iii) and (iv) (identity and inverse).

We can define a random walk on a group G with increment distribution μ (which is a probability measure on group $(G, \circ)$) as follows: it is a Markov chain with state space G, and which moves by multiplying the current state on the left by a random element of G selected according to μ. Equivalently, the transition matrix P of this chain has entries [12]

$$P(g, h_g) = \mu(h).$$

If G is a group and S is a subset of elements, then S generates G as a semi group if every element of G can be expressed as a product of elements from S. A group G is finitely generated if it has a finite generating set.

Let us introduce the Cayley graphs to relate random walks on groups with random walks on graphs, that is, graphs that encode the structure of discrete groups. Suppose that the group G is finitely generated, and let S be a symmetric set of generators of G. The Cayley graph $\Gamma = (G, S)$ of G with respect to the generating set S has vertex set G, and two vertices $x, y \in G$ are connected by an edge if, and only if, $x^{-1}y \in S$. This graph is connected locally, finite, and regular (all vertices have the same degree $\mid S \mid$). Notice that Cayley graphs are transitive in the sense that they look the same from every vertex. If $e \in S$, then $\Gamma = (G, S)$ has a loop at each vertex [96].

Branching Tree. A branching tree on a group τ with group operation $\circ$ is defined recursively via generations $\mathbb{G}_n$

$$\mathbb{G}_n = \{g \in \tau : g = a \circ s, a \in \mathbb{G}_{n-1}, s \in S\}.$$

$\mathbb{G}_0 = S$ is the group generator. Let $U \equiv \cup_{n=0}^{\infty} \mathbb{G}_n$. We say that $(a, b) \in E$ if $a, b \in U$ and there is a path $b = a \circ s_1 \circ \cdots \circ s_k$, where each $s_i = S$ [29].

1.2 Literature Review

1.2.1 Embedding problem

Let P be an $(n \times n)$ estimated transition matrix associated with the logged data. The question is whether P can have a representation $P = e^Q$, where Q matrix is transition rate matrix. This is referred to as an embedding problem. In here, the matrix Q is the generator, the real generator, intensity matrix or true generator; otherwise, matrix Q is neither a true generator nor an exact generator.

Elfving first proposed the embedding problem which is also known as Elfving's problem [22]. He gave certain associated necessary conditions, in particular observing that the eigenvalues of P must satisfy two conditions: (i) no eigenvalue other than unity can have unit modulus (and so P must be aperiodic); (ii) every negative eigenvalue must have an even (algebraic) multiplicity.

The problem had been revived by Chung in the more general context of countable Markov chains [47]. From the form of the problem, we easily get the basic result that a chain with a non-singular transition matrix P can be embedded in a continuous time Markov process with a measurable transition function if, and only if, there exists a real matrix P (which is actually Q-matrix) such that $P = e^Q$.

This is not easy to apply for general matrices, although it provides a method of determining whether a particular matrix is embeddable. It may, however, be used to deal with the exceedingly simple case $n = 2$.

The only complete solution is known for $n = 2$ (Kendall, unpublished; formulated in [33]), which states that a stochastic (2×2) matrix P is embeddable if, and only if, $\det(P) > 0$, and $tr(P) > 1$, where $tr(.)$ is the usual trace, i.e., the sum of the diagonal elements of a matrix. However, we shall see later that when $n > 2$, no simple necessary and sufficient condition of this sort is possible. In addition, his propositions show that the set of embeddable matrices has a complex geometry (except when $n = 2$), consisting of $n^2 - n + 1$ different parts, of which at most one can be expressed in algebraic form. It therefore seems unlikely that any explicit characterisation of this geometry can be given.

Runnenburg first considered the case $n = 3$ for the embedding problem in his thesis then he obtained a necessary condition for embeddability [41]. Runnenburg states that the Elfving's problem can only be solved (P be a embeddable stochastic matrix) if all eigenvalues belong to a defined region (H_n).

Cuthbert considered the Jordan canonical form to find conditions for embeddability and gave the simplification of the three-state case [39]. Also, Johansen obtained a more explicit criterion with the expression for the logarithm of a 3×3 stochastic matrix P as a linear combination of powers of P [85]. Johansen left an open question which is considered in [64] and some characterizations of embeddability with a negative eigenvalues for 3×3 stochastic matrices are found.

Later, Culver introduced conditions for the existence and uniqueness of the real logarithm of a matrix [94]. Several of the theoretical results will be applied in the case study in Chapter 3.

Previously, the embedding problem was treated as a problem of pure mathematics. Then, in the 1990s, the problem, which applies to rating transition matrices, received increasing attention in the financial mathematics literature [75].

Notice that many authors in the financial literature assume that the solution to the embedding problem is positive, that is the logged data comes from the continuous time Markov process, and estimate its generator [73], [52]. For example, in credit rating an approximate generator is obtained by assuming that the probability for one rating to make more than one transition in one year is small.

Since the work of Jarrow et al., the use of credit rating transition matrices in credit risk modelling has received increasing attention [73]. For example, Kijima and Komoribayashi provide an improvement on the estimation procedure in [73], [53];

Belkin, Suchower, and Forest Jr. propose a one-factor Markov process to model credit rating transitions [3]; Kijima, from a technical perspective, explains how a Markov chain model can lead to known empirical regularities such as memory in rating changes and long-term reversion of rating [52]. In a useful note, Lando shows how a transition matrix can be used to value credit derivatives such as a default swap [11].

If a generator exists, it need not be unique. The first published example of the uniqueness of a generator is given in [37]. Since then, many researchers have considered uniqueness [39], [14], [5]. Notably, Israel et al. present a simple method for finding a generator. They also identify a number of sufficient conditions under which this method works, and discuss some further results on the existence and uniqueness of generators. In addition, they develop an algorithm for searching a valid generator when the simple method fails [75].

Bladt et al. applied the EM algorithm and a Markov chain Monte Carlo (MCMC) procedure to the embedding problem for Markov chains [50]. In this paper, they discuss the problems related to the estimation of maximum likelihood of the intensity matrix based on a discretely sampled Markov jump process and demonstrate that the maximum likelihood estimator can be found either by the EM algorithm or by an MCMC procedure. It is possible that the maximum likelihood estimator does not exist, but this problem can be overcome by using a penalized likelihood function or an MCMC estimator with a suitable prior.

1.2.2 Algebraic Methods

In the thesis, a trace is applied to analyse the error in the embedding problem. Also, it is useful as an expectation of a quantum probability. In addition, is used for computing relative distance between two tensor products of stochastic matrices to find best fit. The trace of a matrix is the sum of the eigenvalues and it is invariant with respect to a change of the basis.

Additionally, tensor product is used while considering a random walk on lamplighter group and embedding problem in order to compute the estimated transition matrices of the splitting data (data is split into two parts “no small jump or small jump" and "no big jump or big jump"). Also, the tensor product plays an important role in quantum mechanics. The phrase "tensor product" refers to another way of constructing a large vector space out of two (or more) smaller vector spaces. It is also called the Kronecker Product. In mathematics, the tensor product, denoted by $\otimes$, may be applied in different contexts to vectors, matrices, vector spaces, and algebra, among many other structures or objects. In general, we use the tensor products of matrices.

In this research, we consider a random walk on a directed Cayley graph with directed graphs [96]. Specifically, let $G = (V, E)$ be a directed graph, a simple random walk on G is a Markov chain with a state space V and transition matrix $P(x, y) = 1/deg(x)$, if $x \to y$ (i.e., x is connected to y) and $P(x, y) = 0$ otherwise. A random walk on a Cayley graph is considered a branching random walk or a connected path on the branching tree starting at the origin and having a single child at each generation. Specifically, we model data as a Markov chain with a hidden random walk on a group. The hidden random walk is constructed on the lamplighter group $\mathbb{Z}_3$ and on the tensor product of groups $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. The lamplighter group has a specific structure where the hidden information is actually explicit. We assume that the positions of the lamplighters are known, but we do not know the status of the lamps. This is referred to as a hidden random walk on the lamplighter group. The biased random walk (as introduced in [72]) is constructed to fit the data.

Many researchers have studied these kinds of processes, in particular the spectral analysis of the lamplighter random walk on an infinite path is considered in [96]. Also, the finite case has been treated and analysed via the probabilistic techniques used for lamplighter processes on a complete graph and on the discrete circle in [61].

1.2.3 Statistical Methods

The tensor product structure comes from the splitting of the data into the "no jump", "small jump" and "big jump" groups and matching into the "no small jump-small jump" and "no big jump-big jump" groups. Then, this requires to deal with the missing data. Splitting data as $(2\times2)(2\times2)$ helps to find the hedging to compute the option price.

Particularly, the transformed data Z is observable data which is considered to be a hidden pair (X, Y) to construct the tensor product structure. The key point of this structure is Z, which is the maximum of the pair.

$$Z = \max(X, Y)$$

Z represents the "no jump", "small jump" and "big jump" group which is split into two groups. Therefore, X represents the "no small jump-small jump" group and Y represents the "no big jump-big jump" group. Also, Y ("no big jump-big jump" group) is a complete dataset whilst X ("no small jump-small jump" group) has missing values. Therefore, this requires to deal with the missing data.

The missing data is a significant issue in this case. Firstly, we present the missing data and treatment methods. Then we introduce and apply the Expectation-Maximization Algorithm as the parameter estimation method and Machine Learning Algorithm (C4.5) as the imputation method in order to treat the missing values [2], [40].

Missing Data Treatment

There is an important problem in data mining where the data is incomplete or a certain amount of the data is missing or wrongly collected. Many methods are applied to deal with this in various areas of research.

In statistics, missing data (values) occur when no data value is gathered for the variable in an observation. Missing data occur because of a non-response in which no information is provided for one or more parts, or indeed for their whole. In order to decide how to treat the missing data, it is useful to know why the data is missing.

Missing data occur in research in economics, sociology, and political science because governments report critical statistics in an incomplete form [88]. Occasionally, missing data can be a result of errors by the researcher (such as collecting data by mistake, [26]).

Missing data can be categorized into various different types as per the generating forms, as follows: missing completely at random, missing at random, and missing not at random [46].

Several techniques are considered to treat the missing data. Many of them, such as case substitution, were developed to deal with missing data in sample surveys, and have some drawbacks when applied in the data mining context. The others, such as replacement of missing values by the attribute mean or mode, are very naive and should be carefully used to avoid the introduction of bias. In the literature, the appropriate methods are divided into the following three categories: (i) ignoring and discarding data, (ii) parameter estimation and (iii) imputation [67].

The imputation technique is mostly associated with machine learning. Machine learning systems are sophisticated systems used for both unsupervised and supervised machine learning, which include a generating model to predict values to alternate the missing data. These predictive models rely on known information from the dataset. When the observed data include information that is beneficial to the prediction of missing data, and the imputation method considers the information to maintain high precision.

Most common learning algorithms, such as multilayer perceptron (MLP, a type of artificial neural network [70]), k nearest neighbours (KNN, ex: the option is defined by the majority of the options of neighbours, [63]), self-organising maps (SOM, a type of artificial neural network [18]) and decision tree (DT, a tree like a graph [76]) algorithms are considered to deal with the missing data in different research areas. Also, they are applied for the same problem domains to examine robustness after imputation [36]. MLP has been found to be useful in the prediction of missing data for the dataset associated with thyroid disease [70]. Also, decision tree algorithms are applied to deal with incomplete industrial datasets [44]. KNN outperforms for handling the absent values in DNA micro arrays [63]. To treat the missing data the KNN algorithm can be used as an imputation method via the C4.5 and CN2 algorithms [23]. Furthermore, the SOM algorithm is applied in various research fields to treat missing data via the imputation method [18], [66], [92].

In addition, Rahman and Davis examined the performance of machine learning methods such as the imputation method to treat missing values [54]. They then compared their results with traditional mean/mode imputation. All the machine learning methods (FURIA, [27], decision tree [76], KNN [13], K-mean clustering [62]) outperformed the statistical method (Mean/Mode [69]) according to their experimental results.

Moreover, two well-known machine learning methods, Autoclass (Bayesian unsupervised learning method, [65]) and C4.5 (decision tree-based supervised learning method, [40]), are applied in order to complete data with missing values [44]. If the learning algorithm has a set of training examples and each example can be seen as a pair, namely the input object and a desired output value (class), the algorithm is considered a supervised learning algorithm. The algorithm analyses the training set and builds a classifier that must be able to correctly classify both training and test examples. A test example is an input object, and the algorithm must predict an appropriate output value (the example must be assigned to a class).

The C4.5 algorithm constructs classifiers that are one of the main tools in data mining. This kind of algorithm has an input and an output. In this algorithm, input is a collection of cases. Each collection of cases belongs to one of a small number of classes, and a fixed set of attributes describe their values. Also, output is a classifier. The classifier exactly predicts the class to which a new case belongs.

The C4.5 algorithm produces classifiers signified as decision trees [40]. Let us consider the details of decision trees with a given set of cases (T). Firstly, the C4.5 algorithm generates an initial tree using a divide-and-conquer algorithm, such as:

- If all the cases in T are from the same class, the tree is a leaf marked as the same as the most-repeated class in the set of cases.

- Otherwise, it needs to choose a test depending on one attribute with two or more outcomes. This test is run for the root of the tree. The each outcome of the test is one branch of the tree. The set of cases T is separated into corresponding $S_1, S_2, \ldots$ based on the outcome for each case. Then the same procedure is followed for each subset.

In the last step of the algorithm, there are many tests to consider. C4.5 considers two searching principles in order to rank possible tests: information gain and gain ratio. Information gain reduces to the smallest value of the total entropy of the subsets T_i but it is biased if there are abundant outcomes; the gain ratio separates information gained by the information provided via the outcomes of the test.

The format of the test outcomes are determined via the types of attributes: numeric or nominal. If there is a numeric attribute A, the threshold θ is found by sorting the set of the cases T on the values of A ($\{A \leq h, A > h\}$) and selecting the split between successive values that maximizes the principle above. An attribute A with discrete values has by default one outcome for each value, but an option allows the values to be grouped into subsets with one outcome for each subset.

Moreover, most researchers consider missing data to be one of the most important statistical research issues. Statistical methods are considered to treat the missing data, e.g., maximum likelihood techniques. The EM algorithm is one of the maximum likelihood techniques used to analyse data with missing values [2]. This is an important technique in parameter estimation in the instance of missing data in comparison to imputation or the filling-in of missing values. Statistical imputation, a less broadly researched area compared to statistical analysis with missing data, encompasses methods (mean imputation, regression imputation, hot-deck imputation).

1.2.4 Motivation via Share Price Modelling

In econometrics and actuarial literature, share price modelling is based on the log transformation, Wilkie model, the Granger approach of co-integration and the linear time series approach [93], [1] [10], [78]. There are several known models, as follows:

Let S_t be the share price at time t and ϵ_t be the residual at time t.

Linear regression model: $S_t = S_0 + bt + \epsilon_t$;

Linear regression model with general intercept: $S_t = a + bt + \epsilon_t$;

Log-linear regression model: $logS_t = logS_0 + bt + \epsilon'_t$ where ϵ'_t is the residual in the logscale at time t;

Log-linear regression model with general intercept: $logS_t = a + bt + \epsilon'_t$;

Difference model: $Z_t = S_t - S_{t-1} = \mu + \eta_t$;

Difference model on logscale: $Z'_t = logS_t - logS_{t-1} = \mu' + \eta'_t$;

In this research, we want to construct a realistic and uncomplicated model, so we perform log transformation and regression analysis. In particular, we apply a log-linear regression model and construct the Markov chains for the residuals (in Chapter 2).

Furthermore, we review several well-known share price models. Generally, option pricing methods are based on Brownian Motion or Levy processes assumptions.

The Black-Scholes model is as a solution to the stochastic differential equation:

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t \tag{1.2.6}$$

where S_t is the share price process [16]. This is based on modelling a share price process as a geometric Brownian motion. S_t is known to have continuous sample paths.

The Black-Scholes-Merton model,

$$\frac{dS_t}{S_t} = \mu\, dt + a\, dN_t \tag{1.2.7}$$

which is a continuous time model of share prices, is used to model share prices as per the stochastic differential equation where N_t is a Poisson process [77].

There is another continuous time model used to incorporate possible jumps,

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t + \; d\{\sum_{j=1}^{N_t} \tilde{\varsigma}_j\} \tag{1.2.8}$$

where $\tilde{\varsigma}_j$ are iid and N_t is a Poisson process with a rate λ [89]. Therefore,

$$\sum_{j=1}^{N_t} \tilde{\varsigma}_j$$

is a compound Poisson process independent of the Brownian motion. Also, the model gives the Black-Scholes model.

The well-known continuous time models for returns are considered above. In the economic literature, discrete time models are known to be more convenient, and which are generally based on regression analysis and the time series approach.

The geometric binomial model derived by Cox-Ross-Rubinstein converts the Black-Scholes-Merton model into a discrete binary tree of prices [30]. The binomial model is:

$$S_n = S_0 Y_1 \dots Y_n \tag{1.2.9}$$

where Y_i are iid variables with two values. Option valuation is computed via application of the risk neutrality assumption over the life of the option, as the price of the underlying instrument evolves. The model defines the security price increasing with a probability p and going down with probability $q = 1 - p$. Further, it does not depend on the past.

In addition, several discrete time models are presented for option pricing with arbitrage possibility [34]. The homogeneous Markov model in discrete time, the homogeneous semi-Markov model in continuous time, and the non-homogeneous semi-Markov model in continuous time are introduced in [34].

1.2.5 Quantum Finance

Quantum finance is an interdisciplinary research area. In this area, theories and methods are developed and combined by physicists and economists in order to solve problems in finance.

One recent trend in the growing area of quantitative finance is to apply techniques borrowed from quantum physics to price derivatives. Principal amongst these are path integrals, which were originally developed to describe the interactions of elementary particles. Path integrals are essentially a means of adding together probabilities, and date. Feynman used path-integral methods to reformulate the rules of quantum physics [80]. Dash applied Feynman path integrals to financial modelling and later many academics began to investigate ways in which path integrals could be applied to the financial markets [42]. The reason was simple: the value of a financial derivative depends on the “path” followed by the underlying asset. One of the most effective ways to demonstrate this is to consider a type of derivative called an option, of which there are two types. A call option is a financial contract issued by one party to another that gives the buyer the right, but not the obligation, to buy the underlying asset for a specified strike price at a future maturity date. The seller or underwriter, normally working for a bank, charges the buyer a premium for the option up front. If the value of the underlying asset is higher than the strike price when the option matures, the buyer will presumably exercise their right to buy the asset at the lower price and then sell it as its present value, resulting in an instant profit less the premium paid to buy the option in the first place. Also, a put option is a similar contract that gives the buyer the right to sell the underlying asset at its maturity for an agreed strike price.

Pricing an option is a complex mathematical problem that involves diffusion processes such as Brownian motion, such as with the random movement of pollen grains suspended in a liquid . Due to the unpredictable behaviour of the underlying assets, the derivatives markets are similar.

The idea of developing a mathematical model for option pricing dates back to 1900 when Bachelier proposed a stochastic process to depict the evolution of stock prices. Later, Fischer Black and Myron Scholes, with Merton, transformed derivatives pricing by developing a pioneering formula for evaluating non-dividend paying stock options [16].

A path integral description of the Black-Scholes model was developed by Baaquie for quantum pricing [7]. Then, a quantum mechanical version of the Black-Scholes equation was derived by Baaquie in order to describe the price of a simple, non-dividend paying option.

In quantum mechanics,the state of a physical system can be represented by a wave function (arbitrary function), $|\psi>$, and the expectation value of an observable that is described by an operator A that is given by the "inner product" $<\psi|A|\psi>$,where $|\psi>$ is the Hermitian conjugate of the wave function. In the financial world the value of an option at a certain time, t, can by analogy be interpreted as the inner product $<f|x>$, where f is the option price and x is the price of the underlying asset.

The evolution of the option value with time, $f(t)$, can be written as $|f(t)>= e^{tH}|f(0)$, where H is the appropriate differential operator or Hamiltonian and $f(0)$ is the value of the option at $t = 0$. The path integral for the option then models the stochastic process followed by the price of the underlying asset, in the same way that the Feynman path integral for, say, an electron takes into account all its possible trajectories. Using simple boundary conditions for the value of the option at certain times, a self-consistent quantum system for the price of an option can be determined.

Other researchers, such as Ilinski, have taken a slightly different path integral method which depends on quantum electrodynamics (QED) [45]. The financial market is formulated using the QED model. Particles with positive and negative charges, corresponding to securities and debts, respectively, interact quantum mechanically with each other through electromagnetic fields.

On the other hand, other models such as Hull-White and Cox-Ingersoll-Ross, have successfully used the same method in the classical setting with interest rate derivatives [28], [31]. Khrennikov builds on the work of Haven and further bolsters the idea that the market efficiency assumption made by the Black-Scholes-Merton equation may not be appropriate [43], [15]. To support this idea, Khrennikov builds on a framework of contextual probabilities using agents as a means of overcoming any criticism of applying quantum theory to finance. Accardi and Boukas again quantize the BlackScholesMerton equation, but in this case they also consider the underlying stock to undergo both Brownian and Poisson processes [49].

Chen presents a quantum binomial options pricing model, simply abbreviated as the quantum binomial model [102]. Chen's quantum binomial options pricing model is to existing quantum finance models what the Cox-Ross-Rubinstein classical binomial options pricing model was to the Black-Scholes-Merton model: a discretized and simpler version of the same result. These simplifications make the respective theories not only easier to analyse but also easier to implement computationally.

1.3 Results

As stated, the thesis is an extensive case study of the financial data on

(i) General stochastic modelling of financial data,

(ii) The embedding problem,

(iii) Modelling of financial data as a random walk on the lamplighter group,

(iv) Treating the data as quantum data and fitting to the quantum binomial market.

General stochastic modelling of financial data.

In this thesis, stochastic modelling of share prices is considered. Particularly, we construct our first model-additive functional of a discrete time Markov chain perturbed by Gaussian noise.

First, we introduce the real data. The data is modelled by the log-linear regression method that is popular in the actuarial literature [87]. After that, Markov chains are constructed for residuals. We define the states of the Markov chains and estimate their transition probabilities.

First the data is considered as a 2×2-state Markov chain ("stay" or "jump"). Second, a 3×3-state Markov chain ("stay", "small jump" or "big jump") is considered. Third, the data is split into two parts (no small jump or small jump and no big jump or big jump) and the tensor product of a 2×2-state Markov chain is considered.

Moreover, the tensor product structure comes from the split of the data into "no jump", "small jump" and "no big jump" groups and matching into the "no small jump-small jump" and "no big jump-big jump" groups. Then, this requires to deal with the missing data. The missing data is a significant issue in the third case. In Section 2.2, we study missing data. In the literature, the methods are divided into the following three categories: (i) ignoring and discarding data, (ii) parameter estimation, and (iii) imputation [67]. In order to treat the missing values, we apply the Expectation-Maximization Algorithm [2] as the parameter estimation method and the C4.5 machine learning algorithm [40] as the imputation method.

Note that our missing data comes from the construction, which can be seen as a side-effect. However, by adding this structure, we greatly simplify the number of parameters that need to be estimated. For example, the transition matrix in the lamplighter group on $\mathbb{Z}_4$ will have 64×64 parameters, but on $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ this is only 4×4. It is traditional in such research that by loosing certain features we can gain new ones.

The approach of approximation/modelling data via the tensor product is new, as far as we are aware. Because of this, the associated problems, such as the embedding problem for the tensor product and the lamplighter group construction for the tensor product, are also new.

Although the idea of exploiting the lamplighter group for such modelling was examined in [99], the tensor product in this research is entirely new.

Embedding problem.

We first analyse the problem via algebraic and perturbation approaches. For the embedding the results for the case study were as follows (see Chapter 3 for the relevant definitions and notations):

(i) *Share prices, algebraic approach.* According to the algebraic approach, the exact generators do not exist for any constructed transition matrices or for any transition matrices ((2×2), (3×3), $(2\times 2)\otimes(2\times 2)$) in any financial year considered (2009-2010, 2010-2011,2011-2012);

(ii) *Volatility, algebraic approach.* According to the algebraic approach, the exact generator exists for several cases for volatility (2×2) transition matrices, and for several cases for volatility $(2 \times 2) \otimes (2 \times 2)$; however, the exact generator does not exist for volatility (3×3) transition matrices for the entire dataset.

(iii) *Share prices, volatility, perturbation approach.*

(2×2) *case.* According to the perturbation approach, the exact generators exist for the slightly perturbed (2×2) transition matrices for the entire financial data and volatilities. The chosen parameter is $\delta = 0.1$.

(3×3) *case.* However, in this case of the perturbation parameter $\delta = 0.1$, none of the perturbed transition matrices have an exact generator. This suggests that considering a larger number of states only makes things worse.

$(2\times 2)\otimes(2\times 2)$ *case.* Surprisingly, for the small perturbation parameter $\delta = 0.1$, in roughly half of cases the perturbed transition matrices have an exact generator.

Overall, this study shows that using a continuous time model for volatility is more stable than the original share prices. In addition, considering a small number of carefully chosen states is more reliable.

Furthermore, the share prices of the twenty different companies for four different financial years (2009-2010, 2010-2011, 2011-2012, 2012-2013) are considered in the 3-by-3 case of the embedding problem via an algebraic approach, indicating the general conclusion that for most data, the Markov chains are not embeddable.

As a result, in general we could not embed the discrete time Markov chains in the continuous time Markov chain. This means that the model we considered would be more appropriately treated as a discrete time model.

Modelling share prices via the random walk on the lamplighter group.

We consider a model as a Markov chain with a hidden random walk on a group. The hidden random walk is constructed on the lamplighter group $\mathbb{Z}_3$ and on the tensor product of groups $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. The lamplighter group has a specific structure where the hidden information is actually explicit. We assume that the positions of the lamplighters are known, but we do not know the status of the lamps. This is referred to as a hidden random walk on the lamplighter group. Also, the biased random walk is constructed to fit the data.

(i) For the randomly chosen datasets, the α-biased random walk on the lamplighter group and $\alpha - \lambda$- biased random walk, as defined in Section 4.1.2, provide significantly better fits to the data. The smallest trace norm value is around 0.02.

(ii) The α-biased random walk on the tensor product of the lamplighter group and $\alpha - \lambda$- biased random walk, as defined in Section 4.1.3, provide a significantly better fit to the data compared with other models. The smallest trace norm value is around 0.01.

(iii) The random walk on the tensor product of the lamplighter group gives a better approximation than the random walk on the lamplighter group.

(iv) Two different generators are chosen randomly for each case, and they produce similar results (sensitivity).

(v) Two different methods (EM and machine learning) are used to deal with the missing data, and they also yield close results (robustness).

(vi) Results are almost identical for share prices and their volatility.

Quantum Finance.

Although the quantum pricing model appears in [102], the analysis of the quantum-type data and its application to option pricing is new.

We analyse the quantum data and compute the option price for a binomial model via the quantum data (see Chapter 5 for the relevant definitions and notations):

(i) We need to work with eigenvalues of the Hermitian operator. More exactly, we observe the eigenvalues of the operator $H^{\otimes n}$.

(ii) The original density matrix ρ is irrelevant to the computation of option price for the binomial model via quantum data; we only need the transformed density matrix $\tilde{\rho}$ for computation. Thus, the main issue is to estimate high "u" and low "d" jumps from $H^{\otimes n}$.

(ii) It appears that several famous statistics in statistical mechanics (Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac) used in the machine learning-type algorithm produce close results. These statistics do not greatly affect the estimation of the parameters "u" and "d".

Although the estimates of model parameters are not generally justified by the traditional goodness of fit, we estimate trace norms to check which models better fit the data.

1.4 Structure of the Thesis

This thesis is composed of six chapters and is organised as follows:

In Chapter 1, we introduce terminology related to the Markov chains and random walks on graphs and groups. Also, we introduce our research and review the related literature. Then overall thesis results are presented in Section 1.3.

In Chapter 2, we consider the stochastic modelling of share prices. At the beginning of this chapter, we construct our first model - an additive functional of a discrete time Markov chain perturbed by Gaussian noise. We then introduce details of statistical methods used to derive the model. In Section 2.2 we discuss the tensor product structure which arises from the splitting of the data into "no jump", "small jump" and "no big jump" groups and matching into the "no small jump-small jump" and "no big jump-big jump" groups. This requires us to deal with missing data. Therefore, we introduce and apply the Expectation-Maximization Algorithm and C4.5 machine learning algorithm . Moreover, in the last section of this chapter, we present a model to estimate the volatility of the share price data. This volatility is analysed via the same procedure as for the share prices process in order to be consistent.

In Chapter 3, we review the embedding problem. Then, "iff" and "necessary" conditions of the embeddability are considered. In Section 3.3 we apply the embedding problem to real-world data to examine whether the transition matrices of the share price and its volatility are embeddable into continuous time Markov chains. First, the embedding problem is considered via an algebraic approach. Then, the perturbation approach is used to consider the embedding problem if the algebraic approach cannot solve the problem exactly. Also, random matrices are considered. The results of these approaches are presented and compared.

In Chapter 4, the data are modelled as a Markov chain with a hidden random walk on lamplighter group. The hidden random walk is constructed on the lamplighter group on $\mathbb{Z}_3$ and biased random walks are considered in Section 4.1.2. Also, in Section 4.1.3 we construct the random walk and biased random walks via the tensor product of the lamplighter groups on $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. At the end of this chapter, we present all the simulated transition matrices by constructing the model as the simple random walk and biased random walks on the lamplighter group to find the best fit to the estimated transition matrices by maximum likelihood in order to compare we calculate the trace error (norm) between the simulated matrices and the one based on the data (share price and its volatility). The trace norm values are demonstrated and compared for all cases in Section 4.2.

In Chapter 5, we compute option prices for the binomial model via quantum data. First, we review the different statistics for both quantum and classical methods. Also, we introduce classical and quantum binomial models. We consider the data derived from the two-step binomial model. Then, we estimate the parameters of the quantum model in order to analyse the quantum data. Finally, we compute option price based on these parameters in a quantum multi-step binomial market.

In Chapter 6, our results are summarised to conclude this research. Finally, we briefly discuss our future research and ideas.

Chapter 2

Stochastic Modelling and Fit to Logged Data

This chapter is devoted to the stochastic modelling of share prices. In this chapter, we have two main goals: the first is to construct our first model-additive functional of a discrete time Markov chain perturbed by Gaussian noise; the second is to consider details of the statistical methods used to derive the model.

We have reviewed several well-known models for share prices in order to choose our model in Section 1.2.4. In this chapter, we introduce real-world data. The data is first modelled by the log-linear regression method that is popular in the actuarial literature [87]. After that, Markov chains are constructed for residuals. We define the states of the Markov chains and estimate their transition probabilities.

First, the data is considered to be a 2×2-state Markov Chain ("stay" or "jump"). Second, a 3×3-state Markov chain ("stay", "small jump" or "big jump") is considered. Third, the data is split into two parts (no small jump or small jump and no big jump or big jump) and a tensor product of a 2×2-state Markov Chain is considered.

Moreover, the tensor product structure comes from the split of the data into "no jump", "small jump" and "no big jump" groups and matching into the "no small jump-small jump" and "no big jump-big jump" groups. This then requires us to deal with the missing data. The missing data is a significant issue in the third case. In Section 2.2 we considered missing data. In the literature, the methods used in this regard are divided into the following three categories: (i) ignoring and discarding data, (ii) parameter estimation, and (iii) imputation [67]. In order to treat the missing values, we apply the Expectation-Maximization Algorithm [2] as the parameter estimation method and the C4.5 machine learning algorithm [40] as the imputation method.

Note that our missing data comes from the construction, which can be seen as an associated side-effect. However, by adding this structure, we greatly simplify the number of parameters that need to be estimated. For example, the transition matrix in the lamplighter group on $\mathbb{Z}_4$ will have 64×64 parameters, whilst on $\mathbb{Z}_2 \otimes \mathbb{Z}_2$ this is only 4×4. It is traditional in research that by loosing or neglecting certain features, we can gain new ones.

Furthermore, in the last section of this chapter, we introduce a model to estimate the volatility of share price data. This volatility is analysed using the same procedure as for the share prices process.

2.1 Establishing Model on the Data

2.1.1 Data

The data used in this research consists of a share price dataset from "British Petroleum (London)". British Petroleum, commonly known as BP is one of the world's largest energy corporations. The company is vertically integrated and operates in the oil and gas industry, including exploration and production, refining, distribution and marketing, petrochemicals, power generation and trading. It also has renewable energy activities in bio-fuels and wind power [9].

The stock price data were chosen arbitrarily from the internet for BP's day-by-day closing share prices for four different financial years, 2009-2010, 2010-2011, 2011-2012, 2012-2013 (April to April). The datasets were obtained randomly from the website *http://uk.finance.yahoo.com*. The datasets are adjusted daily for the closing values (for cash dividends, the value of the dividend is deducted from the last closing sale price of the share) of BP's share prices for one financial year (from April 2012 to March 2013) which is presented in Figure 2.1. (See Appendix for the other years.)

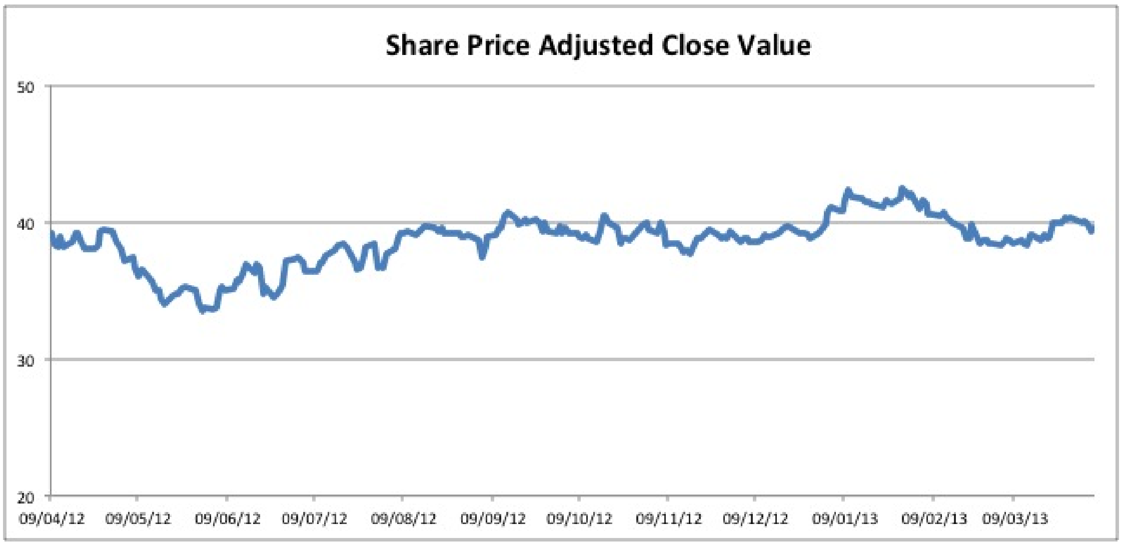


Figure 2.1: BP Share Prices Chart between April 2012 and April 2013: Showing the daily change of BP share prices during the period from April 2012 to April 2013

2.1.2 Establishing the Model

Our first target model is

$$logS_t = \log S_1 + bt + \Sigma_{i=1}^{t}(M_i + \sigma\eta_i) \tag{2.1.1}$$

where S_t is the daily share price process, S_1 is the initial value of the share price, t is the daily unit in a financial year, b is the slope of the share price, M_i is modelled by the Markov chain, σ is the volatility of the residual, η_i are independent and identically distributed Gaussian random variables, and $i = \{1, 2, \cdots, t\}$.

This model is based on the idea of modelling the data as an additive functional of a discrete time Markov chain perturbed by Gaussian noise. For option pricing, it is much easier to work with a geometric random walk (i.e., $S_n = S_0 e^{Y_1+\cdots+Y_n}$ where Y_i are iid). However, the 3×3 transition matrices constructed for all the groups within our data do not fit the geometric random walk model.

Let us consider the methods used to derive and analyse our first model:

In the literature, there are several econometric and actuarial transformations of data such as the linear model, linear model with general intercept, log-linear model, and the difference model.

First we choose the log-linear model as an initial step to transform our data. This is a popular choice for modelling data in the actuarial literature [87]. We denote a constant from simple linear regression by a, and b is the slope of the share price.

Let us model the data (S_t) by our choice of log-linear regression:

$$\log S_t = a + bt + \epsilon_t^{'}$$

where ϵ_t is the residual at time t. Also, $\epsilon^{'}$ is the residual in the log-scale. Here, $a = \log S_1$ and b is the slope of the share price, which is calculated by regression for each data.

We assume that

$$Z_t = \epsilon_t^{'} - \epsilon_{t-1}^{'}$$

are iid errors and $N(0,1)$-Gaussian. Hence, by considering this we derive a simple version of the model (2.1.1)

$$S_t = S_0 \exp{(a + bt + \sum_{k=1}^{t} Z_k)}.$$

Now, our aims are to construct the Markov chains on the transformed data (Z_t) and to estimate the transition matrices of the Markov chains.

2.1.3 Estimating Transition Matrices

In this section, errors will not be iid. Two- and three-state Markov chains are considered.We will later also consider the tensor product of two-state Markov chains.

We choose two states because we know the necessary conditions to solve the embedding problem for two states. There are not any known necessary conditions to find the exact solution for the embedding problem of three states. We choose three states to avoid overcomplicated calculations and to capture the behaviour of the data. Moreover, we use the tensor product of the Markov chains because tensor products maintain the independence of each such tensor. We discuss the tensor product and its properties in the Section 1.1.2.

We model Z_i by $Z_i = M_i + \sigma\eta_i$ where η_i is iid, $N(0,1)$ and M_i are Markov chains. We discretise the error Z_t for all the data and then estimate the transition matrices by maximum-likelihood estimation (MLE).

Let us construct the Markov chain on the transformed data Z_t.

Then, θ_j is defined such that:

$$\theta_1 = \frac{a * (Ma - Mi)}{n}$$
$$\theta_2 = \frac{(Ma - Mi)}{1 + a}$$

where a is a fixed parameter, Ma (maximum jump) is the largest value of the transformed data and Mi (minimum jump) is the minimum value of the transformed data. We undertook some preliminary research as to the choice of a. Eventually, we decided to choose the same a for all models for the comparison reasons. We aim to find a with the largest embeddable proportion.

Specifically, a three-state Markov Chain is chosen in this research to avoid over-complicated calculations whilst still being representative of the data's behaviour.

$$Z_t < \theta_1 \qquad (\text{“no jump”}, 0, M_i = 0)$$
$$\theta_1 \leq Z_t < \theta_2 \qquad (\text{“small jump”}, s, M_i = 1)$$
$$Z_t > \theta_2 \qquad (\text{“big jump”}, b, M_i = 2)$$

Hence the value of the Markov chain is defined on the transformed data for each data $Z_t, t = \{1, \ldots, n\}$ as $M_j = Z_t, j = \{1, 2, 3\}$.

By abuse of notation, the Markov chain M_i will have states: “no jumps”, “small jump” and “big jump”.

Notice that M_i does not represent the approximate changes in Z_i. This simplified labelling is convenient and sufficient in our research.

Also, a two-state Markov Chain is chosen:

$$Z_t < \theta_2 \qquad (\text{“stay”}, 0, M_i = 0)$$
$$Z_t > \theta_2 \qquad (\text{“ jump”}, s, M_i = 1)$$

Then, the value of the Markov chain is defined on the transformed data for each data $Z_t, t = \{1, \ldots, n\}$ as $M_j = Z_t, j = \{1, 2\}$.

Then, we estimate the transition matrices on the state space S via a well-known MLE method.

To estimate the transition probabilities p_{ij},

$$p_{ij} = P\{X_{n=1} = j | X_0 = i\}, \quad i, j \in S$$

We use MLE estimators and write the log likelihood function:

$$\log p = \sum_{i,j} n_{ij} \log p_{ij} \quad \text{and} \quad \sum_{j} p_{ij} = 1$$

$$\hat{p}_{ij} = \frac{n_{ij}}{\sum_{j=1}^{m} n_{ij}}$$

where n_{ij} is the number of times that the process has been observed to go from state i directly to state j and $\hat{p}_{ij}$ is the MLE estimate of p_{ij}.

Notice that p_{ij} does not have a specific distribution. The joint likelihood is defined by

$$\begin{aligned} L &= L(x_1 \ldots, x_n) = \prod_{t=1}^{n-1} p_{x_t, x_{t+1}} \\ &= \prod_{i,j \in S} p_{i,j}^{\nu_{i,j}} \end{aligned}$$

where S is the state space and $\nu_{i,j}$ is the number of observed links $i \to j$ in the sample $x_1, \ldots, x_n$.

To derive $\hat{p}_{i,j}$ we notice that the log likelihood is

$$\log L = \sum_{i,j} \nu_{i,j} \log p_{i,j}$$

In addition we need to incorporate the constrains

$$\Sigma_j p_{i,j} = 1 \ , \ j \in S$$

Consider the Lagrangian that incorporates the constrains

$$\begin{aligned} S &= \log L - \sum_{i} \lambda_i \Big(\sum_{j} p_{i,j} - 1 \Big) \\ &= \sum_{i,j} \nu_{i,j} \log p_{i,j} - \sum_{i} \lambda_i \Big(\sum_{j} p_{i,j} - 1 \Big) \end{aligned}$$

Then,

$$0 = \frac{\partial S}{\partial p_{ij}} = \frac{\nu_{i,j}}{p_{i,j}} - \lambda_i$$

Together with the constrains

$$\begin{aligned} \nu_{i,j} &= \lambda_i p_{i,j}, \sum_j p_{i,j} = 1, \\ \Longrightarrow \nu_{i,j} &= \lambda_i p_{i,j}, \sum_j \nu_{i,j} = \lambda_i, \\ \Longrightarrow \hat{p}_{i,j} &= \frac{\nu_{i,j}}{\sum_j \nu_{i,j}}. \end{aligned}$$

that is, the MLE estimate of $p_{i,j}$ is the proportion of the number of links $i \to j$ over the number of all links from i.

In Table 2.1, estimated two-state and three-state transition matrices are illustrated in the second and third columns, respectively. The rows of the two-state transition matrices are close, and the case looks to be independent. However, unlike the previous case, in the 3×3 case the rows are very different. This shows that the original process is not a homogeneous random walk with independent increments (a similar result was obtained in [98] for the 5x5 case).

2.2 Splitting the Data

The tensor product structure arises from splitting the data into "no jump", "small jump" and "big jump" groups and matching into the "no small jump-small jump" and "no big jump-big jump" groups. This then requires us to deal with the missing data. Splitting data as $(2 \times 2)(2 \times 2)$ helps us to find the hedging to compute the option price.

In particular, the transformed data Z is observable data which is considered a hidden pair (X, Y) needed to construct the tensor product structure. The key point of this structure is Z, which is the maximum of the pair.

$$Z = \max(X, Y)$$

Z represents the "no jump", "small jump" and "big jump" group which is split into two groups. Therefore, X represents the "no small jump-small jump" group and Y represents the "no big jump-big jump" group.

Cases	2×2 transition matrices	3×3 transition matrices
BP(2009-2010)	$\begin{pmatrix} 0.8918 & 0.1082 \\ 0.8929 & 0.1071 \end{pmatrix}$	$\begin{pmatrix} 0.2051 & 0.7436 & 0.0513 \\ 0.1429 & 0.7551 & 0.1020 \\ 0.1667 & 0.7500 & 0.0833 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.8597 & 0.1403 \\ 0.8611 & 0.1389 \end{pmatrix}$	$\begin{pmatrix} 0.4381 & 0.4762 & 0.0857 \\ 0.4310 & 0.4310 & 0.1379 \\ 0.2500 & 0.4444 & 0.3056 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.8098 & 0.1902 \\ 0.8125 & 0.1875 \end{pmatrix}$	$\begin{pmatrix} 0.0667 & 0.8667 & 0.0667 \\ 0.1140 & 0.7668 & 0.1192 \\ 0.1667 & 0.6667 & 0.1667 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.9103 & 0.0897 \\ 0.9167 & 0.0833 \end{pmatrix}$	$\begin{pmatrix} 0.2162 & 0.6757 & 0.1081 \\ 0.1327 & 0.7704 & 0.0969 \\ 0.1200 & 0.8000 & 0.0800 \end{pmatrix}$

Table 2.1: Estimated transition matrices by MLE

Let us consider a basic example to clarify the tensor product structure:

Let $Z = \{s, s, s, 0, b, s, 0, 0, b, s, b, 0, s, 0\}$ be transformed data where 0 is "no jump", s is "small jump" and b is "big jump" and, the data is split as follows:

$$X = \{s, s, s, 0, ?, s, 0, 0, ?, s, ?, 0, s, 0\}$$

$$Y = \{\hat{0}, \hat{0}, \hat{0}, \hat{0}, b, \hat{0}, \hat{0}, \hat{0}, b, \hat{0}, b, \hat{0}, \hat{0}, \hat{0}\}$$

where 0 is "no small jump", s is "small jump" and $\hat{0}$ is "no big jump", b is "big jump". Also "?" represents the missing values.

Briefly, Y ("no big jump-big jump" group) is a complete dataset and X ("no small jump-small jump" group) has missing values. Therefore, this requires us to deal with the missing data.

The type of missing data and treatment methods used to deal with the missing data are reviewed in the following section. Then, details of our selected methods, the C4.5 and EM algorithms, are presented. Also, the results (estimating transition matrices after dealing with the missing values) are illustrated at the end of this section.

2.2.1 Missing Data

Little and Rubin consider the missing data randomness to be one of three main types as follows [46]:

- Missing completely at random (MCAR): This is the highest level of randomness. It occurs when the probability of an instance (case) having a missing value for an attribute does not depend on either the known values or the missing data. In this level of randomness, any missing data treatment method can be applied without risk of introducing bias to the data.
- Missing at random (MAR): When the probability of an instance having a missing value for an attribute may depend on the known values, but not on the value of the missing data itself;
- Not missing at random (NMAR): When the probability of an instance of having a missing value for an attribute could depend on the value of that attribute.

Several techniques are considered to treat the missing data. Many of them, such as case substitution, were developed for dealing with missing data in sample surveys, and have some drawbacks when applied to the data mining context. The others, such as replacement of missing values by the attribute mean or mode, are very naive and should be used with due caution to avoid insertion of bias [46].

Specifically, these techniques to handle the missing data treatment can be divided into three categories [46]:

- Ignoring and discarding data: There are two main ways to discard data with missing values. The first is known as complete case analysis. It is available in all statistical packages and is the default method in many programs. This method consists of discarding all instances (cases) with missing data. The second method is known as discarding instances and/or attributes. This method consists of determining the extent of the missing data in each instance and attribute, and deleting the instances and/or attributes with high levels of missing data. Before deleting any attribute, it is necessary to evaluate its relevance to the analysis. Unfortunately, relevant attributes should be kept even when they have a high number of missing values. Both methods, complete case analysis and discarding instances and/or attributes, should be applied only if missing data are MCAR, because missing data that are not MCAR have non-random elements that can bias the results.

- Parameter estimation: Maximum likelihood procedures are used to estimate the parameters of a model defined for the complete data. Maximum likelihood procedures that use variants of the Expectation-Maximization algorithm [2] can handle parameter estimation in the presence of missing data.

- Imputation: Imputation is a class of procedures that are intended to fill in missing values with estimated ones. The objective is to employ known relationships that can be identified in the valid values of the dataset to assist to estimate the missing values.

Specifically in this research, we apply the Expectation-Maximization Algorithm [2] as the parameter estimation method and C4.5 machine learning algorithm [40] as the imputation method in order to handle the missing data. C4.5 and EM are two of the 10 data mining algorithms discussed in the IEEE International Conference on Data Mining (ICDM) in December 2006 [100]. These algorithms are the most significant data mining algorithms used within the research community. We start with descriptions of these algorithms and apply them to treat our missing data.

C4.5 Algorithm

The C4.5 algorithm was developed by Quinlan and includes modelling missing variables with the supervised induction of a decision tree-based classifier [40], [44]. This method estimates the possible value for the attribute of interest.

The classifier used by C4.5 is a decision tree which is built from root to leaves by respecting Occams Razor, which state that given two equally likely solutions to a given problem, the simpler is more likely to be correct (i.e., we should choose the simpler solution).

In order to handle the missing data with machine learning, there are three main points that need to be addressed:

- Selecting the best candidate test.
- Dealing with instances that have missing values on the test variable.
- Proceeding with an instance that has to be tested against a variable which has a missing value for the given instance.

Entropy is a common means of measuring impurity:

$$Entropy = \sum_i -p_i \log_2 p_i$$

where p_i is the probability of class i. It is computed as the proportion of class i in the set. Entropy comes from information theory. The higher the entropy, the greater the information content. We want to determine which attribute in a given set of training feature vectors is most useful for discriminating between the classes to be learned. Information gain tells us how important a given attribute of the feature vectors is. We will use this to decide the ordering of attributes in the nodes of a decision tree.

The C4.5 algorithm [40] uses the InfoGain or GainRatio tests. InfoGain is the gain in entropy after a test splits data. It may be noted that for missing labels on test columns, such instances will not produce any gain in information. Therefore, if there are missing data and p is the fraction of instances with complete non-missing values, the InfoGain can be computed as:

$$\begin{aligned} InfoGain &= p * (OldEntropy - NewEntropy) + (1-p) * 0 \\ &= p * (OldEntropy - NewEntropy). \end{aligned}$$

After a test is selected, the instances are split into two or multiple groups (one group for each node). To choose the node for the instance with the missing data on the test variable, C4.5 proposes to send all instances with missing values to all child nodes, but with their weight being equal to the proportion of instances from that child node to the total non-missing instances. For example, consider a test column "Movement" with 15 instances having *up*, 35 instances having *down* and three instances having missing value. Then, the three missing instances will be sent to both "up" and "down" child nodes, with weights 15/50 for *up* and 35/50 for *down* nodes. At the prediction step, to deal with the instance of missing values on the test variable, all the possibilities are searched but with the weights. A prediction is made for each possible subnode. Finally, the class with the largest density value is chosen for the prediction. Let us now consider the details of the C4.5 algorithm.

Briefly, the C4.5 algorithm follows these steps:

Step 1: Check if algorithm satisfies termination criteria

Step 2: Computer information-theoretic criteria for all attributes

Step 3: Choose best attribute according to the information-theoretic criteria

Step 4: Create a decision node based on the best attribute in step 3

Step 5: Induce (i.e., split) the dataset based on the newly created decision node in step 4

Step 6: For all sub-datasets in step 5, call the C4.5 algorithm to get a sub-tree (recursive call)

Step 7: Attach the tree obtained in step 6 to the decision node in step 4

Step 8: Return tree

In this research, the information-theoretic criteria is chosen as the InfoGain test. Tree growing is terminated when all the instances covered by a specific branch are pure.

In other words, C4.5 follows a probabilistic procedure to deal missing values in the training data and test data. Each case of the training data is denoted by a weight w_i having output O_i for the value of an attribute. If the output is known and has a value O_i, then $w_i = 1$. Otherwise, other outcomes are denoted by a weight is equal to 0. If the output is missing, the weight of any output O_j for that attribute is the relative frequency of that output among all training cases whose outputs for this attribute are known. The same procedure is used for the test data. In the case of the training data, the target attribute cannot be a missing variable.

C4.5 directly exploits the predicted missing attribute values by using the values of the target attribute (for discrete value attributes) to predict for test data. For this reason, the training data should have a specified value for the target attribute. C4.5 does not deal with continuous variables as target classes. To handle this, for continuous variables intervals would be used on the real line as classes. C4.5 needs disconnected training data for all the candidate attributes in order to make any predictions.

EM Algorithm

The Expectation Maximization (EM) algorithm is a well-known iterative algorithm used for parameter estimation by maximum likelihood to deal with datasets with missing or incomplete random variables [2], [91].

Let us summarize the idea of the EM algorithm with a simple example which is about the outside temperature from your window for each of the 24 hours of the day. This temperature is represented by $x \in \mathbb{R}^{24}$ and is dependent on season $\theta \in \{summer, fall, winter, spring\}$, where the seasonal temperature distribution $p(x|\theta)$ is known. But what if we could only measure the average temperature $y = \bar{x}$ for some day and we would like to estimate what the season θ is? Specifically, the maximum likelihood estimate of θ is considered; that is, the value $\hat{\theta}$ that maximizes $p(y|\theta)$. The EM algorithm iteratively alternates between making guesses about the complete data x, and finding the θ that maximizes $p(x|\theta)$ over θ. Hence, the EM algorithm attempts to find the maximum likelihood estimate of θ given [the data, measurement or observation] y.

Each iteration of the algorithm usually includes two steps:

- The expectation step (E-step): The missing data are estimated given the observed data and current estimate of the model parameters

- The maximization step (M-step): The likelihood function is maximized under the assumption that the missing data are known.

The algorithm is run until the change of the estimated parameter reaches the chosen threshold.

Let us analyse the steps to EM, breaking down the usual two-step description into a five-step description [55]:

Step 1: Let be $m = 0$ and make an initial estimate $\theta(m)$ for θ.

Step 2: Given the observed data y and pretending for the moment that your current guess $\theta^{(m)}$ is correct, formulate the conditional probability distribution $p(x|y, \theta^{(m)})$ for the complete data x.

Step 3: Using the conditional probability distribution $p(x|y, \theta^{(m)})$ calculated in Step 2, form the conditional expected log-likelihood, which is called the Q function:

$$\begin{aligned} Q = (\theta|\theta^{(m)}) &= \int_{\chi(y)} \log p(x|\theta) p(x|y, \theta^{(m)}) dx \\ &= E_{X|y,\theta^{(m)}}[\log p(X|\theta)], \end{aligned}$$

Step 4: Find the θ that maximizes the Q function; the result is our new estimate $\theta^{(m+1)}$.

Step 5: Let be $m := m+1$ and go back to Step 2. (The EM algorithm does not specify a termination criterion; standard criteria are to iterate until the estimate stops changing: $||\theta^{(m+1)} - \theta^{(m)}|| < \epsilon$ for some $\epsilon > 0$, or to iterate until the log-likelihood $l(\theta) = \log p(y|\theta)$ stops changing: $|l(\theta^{(m+1)} - l(\theta^{(m)})| < \epsilon$ for some $\epsilon > 0$.)

The classical description of the EM algorithm has only two steps. The above Steps 2 and 3 combined are referred to as the E-step for expectation, and Step 4 is called the M-step for maximization.

2.2.2 Results

In this section, the tensor product structure is considered, which arises from the splitting of the data into ‘the ‘no jump”, “small jump” and “no big jump” groups and matching into the “no small jump- small jump” and “no big jump-big jump” groups. Then, this requires us to deal with the missing data. After splitting, we treat the missing data via the C4.5 and EM algorithms. Therefore, we have two independent datasets, namely the “no small jump- small jump” and “no big jump-big jump”. We assume that the transformed data are as $Z = (X, Y)$. X is for “no small jump- small jump” and Y is for “no big jump- big jump” groups. Also, the Y

Cases	P_X	P_Y
BP(2009-2010)	$\begin{pmatrix} 0.2419 & 0.7581 \\ 0.2449 & 0.7551 \end{pmatrix}$	$\begin{pmatrix} 0.9060 & 0.0940 \\ 0.9167 & 0.0833 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.5319 & 0.4681 \\ 0.5652 & 0.4348 \end{pmatrix}$	$\begin{pmatrix} 0.8864 & 0.1136 \\ 0.6944 & 0.3056 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.2333 & 0.7667 \\ 0.2344 & 0.7656 \end{pmatrix}$	$\begin{pmatrix} 0.7293 & 0.2707 \\ 0.6761 & 0.3239 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.2787 & 0.7213 \\ 0.2296 & 0.7704 \end{pmatrix}$	$\begin{pmatrix} 0.9009 & 0.0991 \\ 0.9200 & 0.0800 \end{pmatrix}$

Table 2.2: Transition matrices via EM Algorithm.

Cases	P_X	P_Y
BP(2009-2010)	$\begin{pmatrix} 0.2072 & 0.7928 \\ 0.3729 & 0.6271 \end{pmatrix}$	$\begin{pmatrix} 0.9060 & 0.0940 \\ 0.9167 & 0.0833 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.3489 & 0.6511 \\ 0.6504 & 0.3496 \end{pmatrix}$	$\begin{pmatrix} 0.8864 & 0.1136 \\ 0.6944 & 0.3056 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.0463 & 0.9537 \\ 0.3221 & 0.6779 \end{pmatrix}$	$\begin{pmatrix} 0.7293 & 0.2707 \\ 0.6761 & 0.3239 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1332 & 0.8668 \\ 0.3146 & 0.6854 \end{pmatrix}$	$\begin{pmatrix} 0.9009 & 0.0991 \\ 0.9200 & 0.0800 \end{pmatrix}$

Table 2.3: Transition matrices via Machine Learning.

Cases	2×2 transition matrices	3×3 transition matrices
BP(2009-2010)	$\begin{pmatrix} 0.8190 & 0.1810 \\ 0.7755 & 0.2245 \end{pmatrix}$	$\begin{pmatrix} 0.1176 & 0.8529 & 0.0294 \\ 0.1295 & 0.8705 & 0 \\ 0 & 1.0000 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.9593 & 0.0407 \\ 0.9091 & 0.0909 \end{pmatrix}$	$\begin{pmatrix} 0.2807 & 0.7018 & 0.0175 \\ 0.1960 & 0.8040 & 0 \\ 1.0000 & 0 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.8159 & 0.1841 \\ 0.7115 & 0.2885 \end{pmatrix}$	$\begin{pmatrix} 0.0357 & 0.9643 & 0 \\ 0.1161 & 0.8795 & 0.0045 \\ 0 & 1.0000 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.8186 & 0.1814 \\ 0.8837 & 0.1163 \end{pmatrix}$	$\begin{pmatrix} 0.0968 & 0.8387 & 0.0645 \\ 0.1200 & 0.8800 & 0 \\ 0 & 1.0000 & 0 \end{pmatrix}$

Table 2.4: Estimated transition matrices for volatilities

("no big jump-big jump group) is a complete dataset and the X ("no small jump-small jump group) has missing values. Therefore, only this requires us to deal with the missing data. After dealing with these missing values via the selected algorithms, the appropriate transition matrices can be estimated. P_X is the estimated transition matrix of X and P_Y is the estimated transition matrix of Y. Table 2.2 and Table 2.3 illustrate the estimated transition matrices for the Markov chains.

2.3 Volatility

There are many methods that can be used to estimate volatility, and we apply the simplified maximum likelihood estimator.

The suggested model is:

$$dS_t = \mu S_t \, dt + \sigma S_t \, dB_t \tag{2.3.2}$$

where S_t is the daily share price, t is the daily unit in a financial year, μ is the mean, σ is the volatility and $\{B_t, t \geq 0\}$ is standard Brownian motion [71].

Cases	P_X	P_Y
BP(2009-2010)	$\begin{pmatrix} 0.1471 & 0.8529 \\ 0.1434 & 0.8566 \end{pmatrix}$	$\begin{pmatrix} 0.9680 & 0.0320 \\ 1.0000 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.2954 & 0.7046 \\ 0.2054 & 0.7946 \end{pmatrix}$	$\begin{pmatrix} 0.9841 & 0.0159 \\ 1.0000 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.0596 & 0.9404 \\ 0.1289 & 0.8711 \end{pmatrix}$	$\begin{pmatrix} 0.9755 & 0.0245 \\ 0.8571 & 0.1429 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1094 & 0.8906 \\ 0.1408 & 0.8592 \end{pmatrix}$	$\begin{pmatrix} 0.9595 & 0.0405 \\ 1.0000 & 0 \end{pmatrix}$

Table 2.5: Transition matrices via EM Algorithm for volatilities.

The solution to the Black-Scholes equation (2.3.2) is a geometric Brownian motion $S_t = S_0 e^{at+\sigma B_t}$, where $a = \mu - \sigma^2/2$. The solution is found from the Ito formula

$$df(t, S_t) = f'_s\, dS_t + f'_t\, dt + \frac{1}{2} f''_{ss}(\, dS_t)^2 \tag{2.3.3}$$

which is applied to $f(t, S_t) = lnS_t$. More exactly,

$$dlnS_t = \frac{1}{S_t}\, dS_t - \frac{1}{S_t^2}(\, dS_t)^2, \tag{2.3.4}$$

where $(\, dS_t)^2 = S_t^2 \sigma^2 (\, dB_t)^2 = S_t^2 \sigma^2\, dt$ and by integration gives the solution.

$$OP(f(S_T)|t, x) = e^{-\rho(T-t)} - E[f(x) e^{N(a(T-t), \sigma^2(T-t))}] \tag{2.3.5}$$

is at time t the option price of the option claim where $N(.)$ is a standard normal cumulative distribution function, $a = \rho - \sigma^2/2$, ρ is the interest rate and σ the volatility. It is well known that in the classical Black-Scholes model, the no arbitrage option price depends on the interest rate ρ and volatility σ but not on drift μ. However, several studies on empirical estimation of volatility show that the Black-Scholes model does not provide a sufficiently good fit to data. However, many models have been constructed to incorporate the volatility variability [89].

In this research, the volatility of the share price data is estimated and analysed by the same procedure as for the share prices process. In order to estimate the volatility, we set:

$$\hat{\sigma}_t^* = \sqrt{\frac{1}{d-1}\sum_{j=t-d+1}^{t}(S_j - \bar{S}_t)^2}$$

where $\hat{\sigma}_t^*$ is the estimated volatility and

$$\bar{S}_t = \frac{1}{d}\sum_{j=t-d+1}^{t} S_j$$

$\hat{\sigma}_t^*$ estimates the random volatility σS_t in the Black-Scholes model. So,

$$\hat{\sigma} = \frac{1}{S_t}\hat{\sigma}_t^* = \frac{1}{S_t}\sqrt{\frac{1}{d-1}\sum_{j=t-d+1}^{t}(S_j - \bar{S}_t)^2}.$$

After the estimation of the volatility we follow the same procedure (i.e., Markov chains are constructed and the transition matrices are estimated.) for the estimated volatility, such as share prices. (For details see Section 2.1.3.) We model $\hat{\sigma}_t^*$by $\hat{\sigma}_t^* = M_i + i\eta_i$ where η_i is iid, $N(0,1)$ and M_i are Markov chains. Then, we discretise the error for all the data and then we estimate the transition matrices by maximum-likelihood estimation (MLE). The Markov chain is constructed on the estimated volatility $\hat{\sigma}_t^*$ as the transformed share price data Z_t in Section 2.1.3. Specifically, two- and three-state Markov Chains are chosen for the volatility process, the same as for the share price process.

The estimated transition matrices for two and three states are illustrated for volatilities, respectively, in second and third columns in Table 2.4.

The tensor product structure is also considered regarding the volatility process. Therefore, we need to follow the same splitting procedure (see Section 2.2) for the volatility process. Then, we need to deal with the missing data after splitting the data. After that, the transition matrices and their tensor products are estimated via the C4.5 and EM algorithms. P_X is the estimated transition matrix of X and P_Y is the estimated transition matrix of Y via these algorithms. Table 2.5 and Table 2.6 illustrate the estimated transition matrices for the Markov chains. Note that as in Table 2.4, Y is a complete dataset and so there is no error, i.e., P_Y are identical. By contrast, P_X has missing data and estimates of P_X are different for different methods.

Cases	P_X	P_Y
BP(2009-2010)	$\begin{pmatrix} 0.0932 & 0.9068 \\ 0.1378 & 0.8622 \end{pmatrix}$	$\begin{pmatrix} 0.9680 & 0.0320 \\ 1.0000 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.2285 & 0.7715 \\ 0.1793 & 0.8207 \end{pmatrix}$	$\begin{pmatrix} 0.9841 & 0.0159 \\ 1.0000 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.1280 & 0.8720 \\ 0.1257 & 0.8743 \end{pmatrix}$	$\begin{pmatrix} 0.9755 & 0.0245 \\ 0.8571 & 0.1429 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1155 & 0.8845 \\ 0.0494 & 0.9506 \end{pmatrix}$	$\begin{pmatrix} 0.9595 & 0.0405 \\ 1.0000 & 0 \end{pmatrix}$

Table 2.6: Transition matrices via Machine Learning for volatilities.

2.4 Conclusion

In this chapter, our first model was developed for our share price data namely that of an additive functional of a Markov chain perturbed by Gaussian noise. Also, several statistical techniques are considered, which were subsequently used to build the model. In particular, missing data was incurred when we split the data into two groups (the “no small jump-small jump” and “no big jump-big jump” groups). Therefore, missing data treatment methods were considered. In particular, the C4.5 algorithm and EM algorithm were applied to handle our missing data. Transition matrices for the Markov chains are estimated for these two groups of data after dealing with the missing data, the results of which were then presented. In addition, the volatility of the share price data was estimated and analysed by the same procedure as for the share prices process.

Chapter 3

Embedding Problem

In this chapter we present the embedding problem and check the embeddability of the estimated transition matrices for our dataset. It appears from our research that the embedding of the discrete time Markov chains into the continuous time Markov chains seems to be an even a bigger problem than the independence assumption. Most of the results presented in this chapter are submitted in [82].

3.1 The origin of the problem

The embedding problem raised in the pure mathematics area at the beginning. Since continuous time Markov chains are easier to handle in option pricing, it is desirable to treat discrete time Markov chains as continuous time Markov chains as observed at discrete times. Let P be an $n \times n$ real matrix with non-negative entries and with row-sums 1, which is a time-homogeneous Markov transition matrix. The question is how to find generator Q, an $n \times n$ real matrix with non-negative off-diagonal entries and with row sums of 0, such that $e^Q = P$.

Every discrete time Markov chain does not have an underlying continuous time chain, and the necessary and sufficient conditions for this to be the case are unknown. The conditions (real, non negative off-diagonal elements; zero row sums) must be satisfied by a matrix to be a generator, which are known. Also, if Q is a generator then all matrices $P = e^Q$ are stochastic matrices. However, the exact conditions a stochastic matrix P must satisfy so that it can be written as $P = e^Q$ with Q a generator are not known. The subset of $n \times n$ embeddable matrices within the set of all $n \times n$ stochastic matrices has a very complicated geometrical structure (except if $n = 2$).

We give a real-world application of the importance of the embedding problem in this part of the research.

In finance, one of the main research problems has became credit risk modelling and credit derivatives pricing. The transition matrix is a milestone for this problem. Given that empirically estimated matrices are mostly for a one-year period, there is a need to recover a matrix generator so that one can obtain a transition matrix for any arbitrary period of time, as often dictated by a valuation problem such as pricing a default swap [75]. In this research, we are interested in identifying conditions under which a true generator does or does not exist with real datasets.

We summarize how to estimate the transition rates for Markov chains: Suppose that $\{C_t\}, t \in \mathbb{R}_+$, is a time homogeneous, continuous time Markov chain with a finite state space $\kappa = \{1, \ldots, K\}$ under the same probability measure Q. Also, there is a given transition matrix for C corresponding to time $t = 1$, which is denoted by $P = [p_{ij}(1)]_{1 \leq i,j \leq K}$ for every $i, j = 1, \ldots, K$ and every $t \in \mathbb{R}_+$ we have

$$p_{ij}(1) = Q\{C_1 = j | C_0 = i\} = Q\{C_{t+1} = j | C_t = i\}$$

The embedding problem for C relative to Q can be stated as follows:

Find $K \times K$ matrix $\hat{\wedge}$ with non-negative off-diagonal entries and with all rows summing to 0, such that $e^{\hat{\wedge}} = P$ [90]. More clearly, $\hat{\wedge} = [\widehat{\lambda}_{ij}]_{1 \leq i,j \leq K}$, where $\hat{\lambda}_{ij} \geq 0$ for every $i, j = 1, \ldots, K$ with $i \neq j$ and

$$\hat{\wedge}_{ij} = -\sum_{j \neq i} \hat{\lambda}_{ij}$$

where $\hat{\lambda}_{ij}$ is the estimated transition matrix.

The embedding problem is an NP-hard [90] which is finding a $K \times K$ matrix $\hat{\Lambda}$ with non-negative off-diagonal entries and with all rows summing to 0, such that $e^{\hat{\Lambda}} = P$.

An NP-hard problem is a problem that is related to an NP-hard (non-deterministic polynomial-time hard) in computational complexity theory. This a class of problems that, informally, are "at least as hard as the hardest problems in NP" [90]. Cubitt et al. prove that any computationally efficient method of determining which dynamical equations are consistent with a set of measurement data would solve the P vs. NP problem.

3.2 An overview of known facts

In order to analyse the embedding problem, first the matrix logarithm is introduced, and then the "iff" and "necessary" conditions of the embeddability are considered in this part.

Matrix Logarithm:

Let us begin with the complex logarithm:

For $z \in \mathbb{C} \setminus \{0\}$ a logarithm of z is any $w \in \mathbb{C}$ such that $e^w = z$.

$$z = |z|e^{i\phi} \quad \Rightarrow \quad w = \log_k z = \log|z| + i(\phi + 2k\pi), \quad k \in \mathbb{Z}.$$

And the principal value (branch) of the logarithm:

$$z = |z|e^{i\phi}, |z| > 0, -\pi < \phi < \pi \quad \Rightarrow \quad \log z = \log|z| + i\phi.$$

After that, we review the primary and non-primary matrix functions. Non-primary matrix functions play a key role in the embedding problem. Let $A \in M_n$ where M_n is a class of $n \times n$ matrix. A has the Jordan canonical form $A = SJS^{-1}$.

If $f(t)$ is a scalar-valued function of a complex variable t such that $f(\lambda)$ is defined for each eigenvalue λ of A. $f(t)$ is $(k-1)$-times differentiable at each λ, where k is the algebraic multiplicity of λ. Then

$$f(A) = Sf(J)S^{-1}$$

$f(J)$ is the direct sum of $f(J_\lambda)$, and $f(J_\lambda)$ has entries defined in terms of derivatives of $f(t)$ and is evaluated at λ for each Jordan block (Jordan block over a ring, whose identities are zero and one) is a matrix composed of 0 elements everywhere except for the diagonal, which is filled with a fixed element, and for the superdiagonal, which is composed of ones [86]). Let us consider an example:

$$J = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \Longrightarrow f(J) = \begin{pmatrix} f(\lambda) & f'(\lambda) \\ 0 & f(\lambda) \end{pmatrix}$$

$f(A)$ is the primary matrix function associated with the stem function $f(t)$.

Also, an arbitrary matrix function is not a primary matrix function if it cannot be defined by way of a stem function.

In addition we need the power series:

$$\log I + A = A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \ldots, \quad \rho(A) < 1$$

$$\log A = (A - I) - \frac{(A-I)^2}{2} + \frac{(A-I)^3}{3} - \frac{(A-I)^4}{4} + \ldots, \quad \rho(A - I) < 1$$

Theorem 3.2.1. *Let $A \in \mathbb{R}^{n \times n}$.*

- *A has a real logarithm if, and only if, A has an even number of Jordan blocks of each size for every negative eigenvalue.*
- *If A has any negative eigenvalues then no primary logarithm is real [59].*

Therefore, if P has a negative eigenvalue of odd algebraic multiplicity, P is not embeddable.

Let us check the Theorem 3.2.1 on our dataset (see Table 2.1 for the estimated transition matrices in Section 2.1.3):

For the 2009-2010 financial year, $(1, 0.0425, 0.0010)$ are the eigenvalues of the estimated transition matrix P. This has a real logarithm such that:

$$Q = \begin{pmatrix} -1.3696 & 4.4204 & -3.0508 \\ 0.0169 & -1.5218 & 1.5049 \\ 2.1554 & 4.9955 & -7.1509 \end{pmatrix}$$

For 2010-2011, $(1, -0.0387, 0.2135)$ are the eigenvalues of the estimated transition matrix P. Then, no primary logarithm is real.

For 2011-2012, $(1, 0.0001 + 0.0228i, 0.0001 - 0.0228i)$ are the eigenvalues of the estimated transition matrix P. Then, no primary logarithm is real.

For 2012-2013, $(1, 0.0821, -0.0155)$ are the eigenvalues of the estimated transition matrix P. Then, no primary logarithm is real.

Theorem 3.2.2. *For $A \in \mathbb{C}^{n\times n}$ with no eigenvalues on $\mathbb{R}^- = (-\infty, 0]$ there is a unique logarithm X of A (the principal logarithm, $\log A$), all of whose eigenvalues lie in the strip*

$$\{z : -\pi < Im(z) < \pi\}.$$

If A is real then $\log A$ is real [59].

Proposition 3.2.1. *If A has real distinct positive eigenvalues then $\log A$ is the only real logarithm of A [5].*

Iff conditions:

First, the iff conditions are a well-known practical condition for 2×2 matrices [33]:

Proposition 3.2.2. *A stochastic 2×2 matrix P is embeddable if, and only if,*

$$det(P) > 0,$$

or equivalently

$$tr(P) > 1.$$

Here $tr(.)$ is used to denote the sum of the diagonal elements. Traces and their properties are considered in Section 1.1.1.

Proposition 3.2.3. *A stochastic matrix P is embeddable if, and only if, P is infinitely divisible [33].*

$\forall m \in \mathbb{N}$ there exists a stochastic matrix Q_m such that

$$P = Q_m^m$$

which leads to the stochastic roots of stochastic matrices and links to discrete Markov chains.

Let us continue with the criterion for embeddability for 3×3 matrices.

Proposition 3.2.4. *P is a 3×3 stochastic matrix and the eigenvalues of P are $1, \lambda_1, \lambda_2$ such as $|\lambda_1|, |\lambda_2| \leq 1$. Assume that one of the following conditions is satisfied:*

- *λ_1 and λ_2 are positive.*

- *λ_1 and λ_2 are complex conjugate.*

- *$\lambda \equiv \lambda_1 = \lambda_2 < 0$ and P is diagonalizable. (There exists a real logarithm with two 1×1 Jordan blocks for λ.)*

Then P is embeddable [33].

Let us check the Proposition 3.2.4 on our dataset (see Table 2.1 for the estimated transition matrices in Section 2.1.3):

For the 2009-2010 financial year, $(1, \lambda_1 = 0.0425, \lambda_2 = 0.0010)$ are the eigenvalues of the estimated transition matrix P. This also satisfies the first condition of the proposition.

However, for other financial years (2010-2011, 2011-2012, 2012-2013), none of the conditions of the proposition are satisfied.

Proposition 3.2.5. *P is a 3×3 stochastic matrix, det$(P > 0)$, with eigenvalues $1, \gamma e^{i\delta}, \gamma e^{-i\delta}$, where $0 < \gamma < 1$, and $0 \leqslant \delta \leqslant \pi$ and P has a diagonal Jordan form. Suppose that P has the spectral decomposition*

$$P = X diag\{1, \gamma e^{i\delta}, \gamma e^{-i\delta}\} X^{-1}.$$

Then,

- *- if $\delta = 0$, P is embeddable if, and only if,*

$$X diag\{0, \log \gamma, \log \gamma\} X^{-1} \text{ is a } Q\text{- matrix;}$$

- *- if $0 < \delta < \pi$, P is embeddable if, and only if,*

$$X diag\{0, \log \gamma + i\delta, \log \gamma - i\delta\} X^{-1} \text{ is a } Q\text{- matrix;}$$

  or

$$X diag\{0, \log \gamma + i(\delta - 2\pi), \log \gamma - i(\delta - 2\pi)\} X^{-1} \text{ is a } Q\text{- matrix;}$$

- *- if $\delta = \pi$ it is necessary for P to be embeddable that [38]*

$$P = X diag\{0, \log \gamma, \log \gamma\} X^{-1} \text{ is a } Q\text{- matrix.}$$

Let us check the Proposition 3.2.5 on our dataset (see Table 2.1 for the estimated transition matrices in Section 2.1.3):

For the 2009-2010 financial year, $(1, 0.0010, 0.0425)$ are the eigenvalues of the estimated transition matrix P, and $det(P) > 0$. Let us compute γ:

$(1, \gamma e^{i\delta}, \gamma e^{-i\delta}) \Rightarrow \quad \gamma e^{i\delta} = 0.0010$ and $\gamma e^{-i\delta} = 0.0425$

So,

$\gamma = 0.00652 \Rightarrow \quad 0 < \delta < \pi$

The second condition of the proposition is satisfied. Then, we suppose that $P = X diag\{1, \gamma e^{i\delta}, \gamma e^{-i\delta}\} X^{-1}$, where

$$X = \begin{pmatrix} 1 & 0.7803 & 4.0800 \\ 1 & -0.2831 & -0.9613 \\ 1 & 1 & 1 \end{pmatrix}$$

Therefore, $Q = X diag\{0, \log \gamma + i\delta, \log \gamma - i\delta\} X^{-1}$ and P is embeddable.

For all the other financial years (2010-2011, 2011-2012, 2012-2013) except 2009-2010, the conditions of the proposition are not satisfied.

Proposition 3.2.6. *Let P be a 3×3 stochastic matrix with distinct eigenvalues $(1, \lambda_1, \lambda_2)$.*

- *If λ_1 and λ_2 are positive then P can be embeddable if, and only if,*

$$p_{ij}^{(2)} \leq p_{ij} \frac{(\lambda_2^2 - 1) \ln \lambda_1 - (\lambda_1^2 - 1) \ln \lambda_2}{(\lambda_2 - 1) \ln \lambda_1 - (\lambda_1 - 1) \ln \lambda_2}, \quad i \neq j.$$

- *If eigenvalues are $(1, \lambda, \lambda)$, $0 < \lambda < 1$ then P can be embeddable if, and only if, [85]*

$$p_{ij}^{(2)} \leq p_{ij} \frac{\lambda^2 \ln \lambda^2 - \lambda^2 + 1}{\lambda \ln \lambda - \lambda + 1}, \quad i \neq j.$$

Here, $p_{ij}^{(2)}$ denotes the (i, j)th element of P^2 [85].

Proposition 3.2.7. *Let P be a 3×3 stochastic matrix with eigenvalues $(1, \lambda_1, \lambda_2)$. If $\lambda_1 = e^{(\alpha + i\beta)}$, $\lambda_2 = e^{(\alpha - i\beta)}$, $0 < \beta < \pi$, then P can be embeddable if, and only if,*

$$p_{ij}^{(2)}(\beta(e^{\alpha} \cos \beta - 1) - \alpha e^{\alpha} \sin \beta) \geq p_{ij}(\beta(e^{2\alpha} \cos 2\beta - 1) - \alpha e^{2\alpha} \sin 2\beta)$$

or

$$p_{ij}^{(2)}((\beta - 2\pi)(e^{\alpha} \cos \beta - 1) - \alpha e^{\alpha} \sin \beta) \geq p_{ij}((\beta - 2\pi)(e^{2\alpha} \cos 2\beta - 1) - \alpha e^{2\alpha} \sin 2\beta)$$

where $i \neq j$ [85].

Proposition 3.2.8. *P is a 3×3 embeddable matrix with a negative eigenvalue λ, then,*

- $P = Udiag(1, \lambda, \lambda)U^{-1}$
- $P = P^{\infty} + \lambda(I - P^{\infty})$,

where

$$P^{\infty} = \lim_{n} P^{n} = Udiag(1, 0, 0)U^{-1},$$

P^{∞} is a stochastic matrix with identical rows and non-zero elements [64].

To understand this proposition, let us focus on the matrices:

$$P = P(\lambda, P^{\infty}) = P^{\infty} + \lambda(I - P^{\infty})$$

where

$$P^{\infty} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix},$$

$$p_1 + p_2 + p_3 = 1, \quad p_i > 0, i = 1, 2, 3,$$

and

$$\lambda < 0, \quad \lambda - \min_{i} \frac{p_i}{1 - p_i}.$$

Proposition 3.2.9. *A stochastic matrix $P(\lambda, P^{\infty})$ is embeddable if, and only if, there exists a stochastic Q matrix such as $Q^2 = P(\lambda, P^{\infty})$ and one of the following conditions holds [64]:*

$$q_{ij} \geq (1 + \frac{\sqrt{|\lambda|}}{\pi} \ln|\lambda|)p_i,$$

$$q_{ij} \leq (1 - \frac{\sqrt{|\lambda|}}{3\pi} \ln|\lambda|)p_i,$$

for all $i \neq j$.

Necessary conditions:

The embedding problem is introduced in [22], and the certain necessary conditions are given by Elfving. Let us start the necessary conditions of the embeddability:

Proposition 3.2.10. *P is a stochastic matrix with distinct eigenvalues. If P is embeddable, then*

- *P is non-singular*
- *P has no simple negative roots*
- *no eigenvalue λ of P is such that $|\lambda| = 1$ other than $\lambda = 1$.*

Proposition 3.2.11. *If a stochastic matrix $P \in \mathbb{R}^{n\times n}$ is embeddable [33]*

- *if $P^k = (p_{ij}^{(k)})$, then $p_{ij} = 0 \Rightarrow p_{ij}^{(k)} = 0, \quad k = 2, 3, \ldots$*
- *$det(P) > 0$.*

The following proposition is needed to meet Runnenburg's necessary condition for embeddability.

Proposition 3.2.12. *The eigenvalues $\mu_1, \ldots, \mu_n$ of an intensity matrix Q, for $n \geq 3$, satisfy [41]*

$$(\frac{1}{2} + \frac{1}{n})\pi \leq \mu_j \leq (\frac{3}{2} - \frac{1}{n})\pi.$$

Only a finitely many admissible logarithms have to be investigated while searching for a generator for P.

Proposition 3.2.13. *$\lambda_j(t)$ is eigenvalue of a stochastic matrix P. If P is embeddable ($P(t) = e^{Qt}$) at least one eigenvalue lies on the boundary curve of the region H_n ($\lambda_j(t) \in H_n$). H_n is a heart-shaped region in the complex plane contained in the circle, and is symmetric with respect to the real axis with a boundary for $Imz \geq 0$ given by the curve [41]*

$$x(t) + iy(t), \quad 0 \leq t \leq \frac{\pi}{\sin\frac{2\pi}{n}}$$

$$x(t) = [exp(-t + t\cos\frac{2\pi}{n})]\cos(t\sin\frac{2\pi}{n})$$
$$y(t) = [exp(-t + t\cos\frac{2\pi}{n})]\sin(t\sin\frac{2\pi}{n})$$

The following proposition is proven in [14] from the general theory of Markov chains.

Proposition 3.2.14. *If P is embeddable and $p_{ij} > 0, p_{j,k} > 0$, then $p_{ik} > 0$ [14].*

Proposition 3.2.15. *If P is embeddable then $\prod_i p_{ii} \geq det(P)$ [24].*

Proposition 3.2.16. *If P is embeddable and Q is the generator, then $\Lambda(Q) \subseteq \{z : |z - \log det(P)| \leq -\log det(P)\}$ [39].*

Proposition 3.2.17. *If $P \neq I$ is embeddable and*

- *there exist distinct indices i_1, i_2 such that for all k, $p_{i_1 k} = 0$ then $p_{i_2 k} = 0$,*
- *there exist distinct indices j_1, j_2 such that for all k, $p_{kj_1} = 0$ then $p_{kj_2k} = 0$ [8].*

3.2.1 Computing the candidate generator

Israel et al. suggested the use of the following representation

$$Q = \sum_{n=1}^{\infty} (-1)^n \frac{(P-I)^n}{n}$$

which is equal to $Q = T^{-1}(\log J)T$ under the Jordan representation provided the series exist. Motivated by the embedding problem, we assume that $P_{ij} = P_t$ is the transition probability of the continuous time Markov chain [75].

They also stated the following:

Theorem 3.2.3. *Let P be an $n \times n$ Markov transition matrix and suppose that $S < 1$ (where $S = \|P - I\|$ is the norm of the matrix $P - I$ (e.g., the trace norm))*

Then the series

$$Q = (P - I) - \frac{(P-I)^2}{2} + \frac{(P-I)^3}{3} - \frac{(P-I)^4}{4} + \ldots$$

converges geometrically quickly and gives rise to an $n \times n$ matrix Q having row-sums 0, such that $e^Q = P$ exactly [75].

The problem has been extensively studied since the complete solution is known only for a state space with two elements, or $n = 2$ [33]. The solution for a 2×2 stochastic matrix can be found via a so-called ergodic representation. Let

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$$

so

$$P^t = P^\infty + \lambda^t(1 - P^\infty) \text{ where}$$

$$\lambda = \det(P) \text{ and } P^\infty = \begin{pmatrix} \pi_1 & \pi_2 \\ \pi_1 & \pi_2 \end{pmatrix}$$

so

$$P^t = e^{tQ} = \begin{pmatrix} \pi_1 + \pi_2\lambda^t & \pi_2 - \pi_2\lambda^t \\ \pi_2 - \pi_1\lambda^t & \pi_1 + \pi_1\lambda^t \end{pmatrix}$$

$$\Longrightarrow Q = \lim_{t\to 0}(P^t - I)/t = \begin{pmatrix} \pi_2\log(\lambda) & -\pi_2\log(\lambda) \\ -\pi_1\log(\lambda) & \pi_1\log(\lambda) \end{pmatrix}$$

So the necessary and sufficient condition for the existence of a continuous Markov chain or generator Q such that $P = e^Q$ is $\det(P) = \lambda > 0$.

3.3 Results and Comparisons

In this section, we consider the dataset with our models. First the embedding problem is considered via an algebraic approach. Then a perturbation approach is used to consider the embedding problem if it cannot be solved exactly via the algebraic approach. Also, random matrices are considered for comparison.

3.3.1 Embedding problem with our dataset

This section is devoted to an explanation of the algebraic and perturbation approaches and their applications to our estimated transition matrices. Let us start with algebraic approach:

3.3.2 Algebraic Approach

Motivated by Israel et al.,

$$Q = \sum_{n=1}^{\infty}(-1)^n \frac{(P-I)^n}{n}$$

which is equal to $Q = T^{-1}(\log J)/T$ under the Jordan representation, provided the series exist. Motivated by the embedding problem, we assume that $P_{ij} = P_t$ is the transition probability of the continuous time Markov chain.

They also emphasise that, even if the series fail to converge or converge to a matrix Q that is not a valid generator, this does not preclude the possibility that a valid generator for P still exists [75]. Here we choose $t = 1$, then

$$e^Q = P$$

from the definition

$$e^Q = \sum_{j=0}^{\infty} \frac{Q^j}{j} \text{ and } P = TDT^{-1}$$

so

$$e^Q = P = TDT^{-1} \text{ and } Q = T(\log D)T^{-1}$$

We analyse how this works with our dataset after an explanation of the perturbation approach in Section 3.3.4.

3.3.3 Perturbation Approach

Perturbation theory is used to find an approximate solution to a problem which can not be solved exactly by starting from the exact solution of a related problem $Q = \sum_{n=1}^{\infty}(-1)^n \frac{(P-I)^n}{n}$. This is exactly the case with our data. In our case, we find that the algebraic approach does not give an exact solution to the transition matrices.

Since our transition matrices are found up to the associated statistical error, it is reasonable to introduce a perturbation to find the generators if the perturbed transition matrices $\hat{P}$ are embeddable:

$$\hat{P} \longrightarrow (1-\delta)P + \delta B \longrightarrow Q$$

where B is embeddable.

P are the estimated transition probabilities of the Markov chains, and δ is a parameter.

So, we are looking for suitable B matrices and a minimum value of the parameter δ:

$$\delta(B) = \hat{\delta} = \min\{\delta > 0 : (1-\delta)P + \delta B \text{ is embeddable}\}$$

Example.

B, δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$B_1 = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$	N	N	N	N	N	N	N	N	N
$B_2 = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$	Y	Y	Y	Y	Y	Y	Y	Y	Y
$B_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 3.1: Pertubated transition matrix might be embeddable.

In particular, for 2×2 transition matrices (we used 2×2 transition matrices because we know the necessary and sufficient conditions to be embeddable), we define

$$G_{2\times 2} = \{B: \quad B \quad \text{embeddable}\}$$

$$= \left\{ B = \begin{pmatrix} a & \bar{a} \\ \bar{b} & b \end{pmatrix} \quad : ab \geqslant \bar{a}\bar{b}, \qquad 0 \leqslant a, \bar{a} \leqslant 1, 0 \leqslant b, \bar{b} \leqslant 1 \right\}$$

where $\bar{a} = 1 - a$ and $\bar{b} = 1 - b$. Find with a given P transition matrix:

$$\delta_P = \min_B \min_\delta \{\delta > 0: \quad (I - \delta)P + \delta B, \text{ embeddable}\}$$

Here, $\min_B$ is the value of the embeddability not over matrices.

Let us show how this works with our data:

We do not consider the perturbation approach for all B matrices in $G_{2\times 2}$. Rather, we consider this just for$G \subseteq G_{2\times 2}$

$$G = \left\{ B_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} B_2 = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix} B_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

and we get the table for $\delta \in (0, 1)$. To explain the choice of the set G we note that B_3 is theoretically the best, B_1 is theoretically the worst and B_2 is stable and good in numerical implementations. In addition, we note that B_1 is not exactly embeddable.

In Table 3.1, Y shows that the pertubated transition matrix $\hat{P}$ is embeddable ($det(\hat{P}) \geqslant 0$) with the B matrices and the parameter δ. Also, N shows that the pertubated transition matrices $\hat{P}$ are not embeddable $(det(\hat{P}) < 0)$ with the B matrices and the parameter δ. As a result, $\hat{P}$ is embeddable with B_2 and B_3 for all values of δ.

After that, we decided to check the perturbation with these two matrices with 0.01 grids instead of 0.1 in $(0, 1)$ to find the value of the parameter δ. Then, we found the transition matrix is embeddable when $\delta > 0.03$, pertubated with B_2. Also, it is embeddable when $\delta > 0.01$, pertubated with the B_3 matrix. After that, we consider 0.001 grids to find the minimum value of the parameter δ more specifically. We computed that the pertubated transition matrix $\hat{P}$ is embeddable when $\delta > 0.010$.

$\hat{P} = (1-\delta)P + \delta B_3$ is embeddable when $\delta > 0.010$, which is the result of empirical calculations.

Now, we will consider the situation theoretically:

P is not an embeddable transition matrix, such that

$$P = \begin{pmatrix} a & \bar{a} \\ \bar{b} & b \end{pmatrix}$$

$$det(P) < 0 \Rightarrow \bar{a} = 1 - a \text{ and } \bar{b} = 1 - b \Rightarrow a + b < 1$$

Also, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is an embeddable matrix:

$$\hat{P} = (1 - \delta)P + \delta B$$

We want to compute δ when $\hat{P}$ is embeddable. To find δ,

$$\hat{P} = (1 - \delta)P + \delta B = (1 - \delta) \begin{pmatrix} a & \bar{a} \\ \bar{b} & b \end{pmatrix} + \delta \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} (1 - \delta)a + \delta & (1 - \delta)\bar{a} \\ (1 - \delta)\bar{b} & (1 - \delta)b + \delta \end{pmatrix}$$

$$det(\hat{P}) = (1 - \delta)(a + b - 2) + 1$$

If $\hat{P}$ is embeddable, the determinant of the pertubated transition matrix should be equal to or greater than zero. P is a transition matrix, so $a+b$ cannot be greater than

2 ($0 \leqslant a+b \leqslant 2$). Also, $a+b$ should be less than 1 because it is not embeddable. Therefore, the parameter δ is found as follows:

$$\delta \geqslant 1 + \frac{1}{a+b-2}$$

Example. For matrix B_2, a is equal to 0.6705 and b is equal to 0.2892 in our estimated 2-by-2 transition matrix. The parameter computed as a result of the theoretical approach is:

$$\delta \geqslant 0.0387$$

Consequently, we can say that our previous empirical calculation supports theoretical approaches.

Now, let us compute minimum value of the parameter δ with an arbitrary embeddable B matrix:

$$B = \begin{pmatrix} x & \bar{x} \\ \bar{y} & y \end{pmatrix} \qquad xy \geqslant \bar{x}\bar{y} \Rightarrow x+y \geqslant 1$$

Also, the transition matrix as above,

$$P = \begin{pmatrix} a & \bar{a} \\ \bar{b} & b \end{pmatrix}$$

$$\delta = \delta_P = min_B \min \{\delta > 0 : (1-\delta)P + \delta B \text{ is embeddable}\}$$

To find δ,

$$\hat{P} = (1-\delta)P + \delta = (1-\delta)\begin{pmatrix} a & \bar{a} \\ \bar{b} & b \end{pmatrix} + \delta \begin{pmatrix} x & \bar{x} \\ \bar{y} & y \end{pmatrix}$$

$$= \begin{pmatrix} (1-\delta)a + \delta x & (1-\delta)\bar{a} + \delta\bar{x} \\ (1-\delta)\bar{b} + \delta\bar{y} & (1-\delta)b + \delta y \end{pmatrix}$$

$$det(\hat{P}) = a + b - 1 + \delta(x + y - a - b)$$

We know that if $\hat{P}$ is embeddable, the determinant of the pertubated transition matrix should be equal to or greater than zero.

$$det(\hat{P}) \geqslant 0 \qquad \Rightarrow det(\hat{P}) = a + b - 1 + \delta(x + y - a - b) \geqslant 0$$

$$\delta \geqslant \frac{1-a-b}{x+y-a-b}$$

As a result, we compute δ such as the above where $x+y \geqslant 1$ and $x+y-a-b \geqslant 0$ for 2×2 matrices. If we consider with our dataset, $a = 0.6705$ and $b = 0.2892$. Then, the parameters compute such that:

$$\delta \geqslant \frac{0.0403}{x+y-0.9597}$$

Furthermore, we take embeddable matrices B as empirical calculations and calculate δ for each B matrix with Matlab. The perturbation of our transition matrix P is embeddable if δ is greater than or equal to 1 with the B_1 matrix. It is embeddable if δ is greater than or equal to 0.0314 with matrix B_2. Finally, the transition matrix is pertubated by the B_3 identity matrix which is embeddable when δ is greater than or equal to 0.0107. As our aim is to compute $min_B \delta_B$ when the transition matrix is not embeddable ($a+b<1$), B_3 is more suitable for this purpose than the others.

Lemma 3.3.1. *If P_1 and P_2 are embeddable matrices, then $\frac{P_1+P_2}{2}$ is also an embeddable matrix.*

Proof. We know the condition of embedding for 2×2 matrices. Let,

$$P_1 = \begin{pmatrix} a & \bar{a} \\ \bar{b} & b \end{pmatrix} \qquad det(P_1) > 0, a+b > 1$$

$$P_2 = \begin{pmatrix} x & \bar{x} \\ \bar{y} & y \end{pmatrix} \qquad det(P_2) > 0, x+y > 1$$

So,

$$\frac{P_1+P_2}{2} = \frac{1}{2} \begin{pmatrix} a+x & \bar{a}+\bar{x} \\ \bar{b}+\bar{y} & b+y \end{pmatrix}$$

$$det(\frac{P_1+P_2}{2}) = a+b+x+y-2$$

Also, we know $a+b>1, \quad x+y>1$

$$a+b+x+y>2$$

Then,

$$det(\frac{P_1 + P_2}{2}) = a + b + x + y - 2 > 0$$

This satisfies the condition for embedding 2×2 matrices. Therefore, we can say that if P_1 and P_2 are embeddable matrices, then $\frac{P_1+P_2}{2}$ is also an embeddable matrix for 2×2 matrices. □

3.3.4 Applications of these approaches

In this section, the modelled share prices and volatility estimated from our real dataset are considered in three cases in order to apply these approaches.

Case 1: The data is considered to be a 2×2-state Markov Chain (stay or jump) because the existence condition of the generator is known.

Case 2: A 3×3 Markov chain ("stay", "small jump" or "big jump") is considered.

Case 3: The data is split into two parts ("no small jump" or "small jump" and "no big jump" or "big jump") and their tensor products are considered.

For the first and second cases, transition matrices for the Markov Chains are estimated by MLE. However, in the last case, the transition matrices of the two split Markov chains are estimated via the EM algorithm and C4.5 machine learning algorithm in order to deal with missing values.

After estimating the transition matrices, their tensor product is computed. This transition matrix is the transition matrix for Case 3.

Finally, Matlab is used to find the eigenvalues of the transition probability matrices in order to obtain the generator Q matrices.

The estimated transition matrices for two and three states are illustrated in Table 2.1 in Section 2.1.3. Also, the transition matrices and their tensor products estimated via the EM algorithm and C4.5 are illustrated, respectively, in Table 2.2 and Table 2.3 in Section 2.2.2.

BP daily share prices for the 2009-2010 financial year are considered as an example with all details. The results for other financial years are illustrated in tables in each respective case.

Moreover, the share prices of the twenty different companies for four different financial years (2009-2010, 2010-2011, 2011-2012, 2012-2013) are considered in the

Cases	Q matrices
BP(2009-2010)	$\begin{pmatrix} -0.7363+0.3395i & 0.7363-0.3395i \\ 6.0761-2.8020i & -6.0761+2.8020i \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -0.9207+0.4401i & 0.9207-0.4401i \\ 5.6506-2.7014i & -5.6506+2.7014i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -1.1219+0.5959i & 1.1219-0.5959i \\ 4.7926-2.5457i & -4.7926+2.5457i \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -0.4502+0.2800i & 0.4502-0.2800i \\ 4.6012-2.8616i & 4.6012+2.8616i \end{pmatrix}$

Table 3.2: Q matrices of the 2-by-2 transition matrices for all financial years

3-by-3 case of the embedding problem via an algebraic approach in order to give a general conclusion.

Case 1 ***(***2×2***)***

Algebraic Approach:

P is the transition matrix for the 2009-2010 financial year:

$$P_t = \begin{pmatrix} 0.8918 & 0.1082 \\ 0.8929 & 0.1071 \end{pmatrix}$$

According to the existence condition for 2×2 stochastic transition matrices, $det(P) = -0.0011 < 0$, hence the exact generator does not exist. Applying the algebraic approach: According to the function $e^Q = P = TDT^{-1}$, the eigenvalue (D) and eigenvector of P should be found

$$Q = T(\log D)T^{-1} = \begin{pmatrix} -0.7363+0.3395i & 0.7363-0.3395i \\ 6.0761-2.8020i & -6.0761+2.8020i \end{pmatrix}$$

The matrix is even not the real matrix.

Table 3.2 shows that the exact generators do not exist for the other transition matrices for all financial years for the share price process. Moreover, Table 3.3 shows that the exact generators exist for the volatilities for all financial years except the 2012-2013 financial year. Therefore, we only consider the perturbation approach for this financial year for volatilities in this case, and for all the estimated transition matrices for the share price process.

Cases	Q matrices
BP(2009-2010)	$\begin{pmatrix} -0.5932 & 0.5932 \\ 2.5418 & -2.5418 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -0.1282 & 0.1282 \\ 2.8635 & -2.8635 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -0.4645 & 0.4645 \\ 1.7951 & -1.7951 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -0.4653+0.5351i & 0.4653-0.5351i \\ 2.2666-2.6065i & -2.2666+2.6065i \end{pmatrix}$

Table 3.3: Q matrices of the 2-by-2 transition matrices for volatilities

Perturbation Approach:

P is the same transition matrix for the 2009-2010 financial year with the algebraic approach:

$$P_t = \begin{pmatrix} 0.8918 & 0.1082 \\ 0.8929 & 0.1071 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$$

which is a fixed embeddable matrix. Also δ is a parameter; we choose a low level, $\delta = 0.1$, to say that if $\delta(\hat{P}) < \delta = 0.1$, then $\hat{P}$ is embeddable.

Let us compute the perturbed transition matrix by:

$$\hat{P} \longrightarrow (1-\delta)P + \delta B \longrightarrow Q$$

$$\hat{P} = \begin{pmatrix} 0.8693 & 0.1307 \\ 0.8369 & 0.1631 \end{pmatrix}$$

and $det(\hat{P}) = 0.0323$. So, the exact generator exists:

$$Q = T(\log D)T^{-1} = \begin{pmatrix} -0.4633 & 0.4633 \\ 2.9663 & -2.9663 \end{pmatrix}$$

In addition, Table 3.4 shows the perturbed transition matrices (first column of the table) and their Q matrices (second column of the table) for all other transition

Cases	$\hat{P}$ perturbed matrices	Q matrices
BP(2010-2011)	$\begin{pmatrix} 0.8404 & 0.1596 \\ 0.8083 & 0.1917 \end{pmatrix}$	$\begin{pmatrix} -0.5671 & 0.5671 \\ 2.8718 & -2.8718 \end{pmatrix}$
BP(2011-2012	$\begin{pmatrix} 0.7955 & 0.2045 \\ 0.7646 & 0.2354 \end{pmatrix}$	$\begin{pmatrix} -0.7337 & 0.7337 \\ 2.7433 & -2.7433 \end{pmatrix}$
BP(2012-2013	$\begin{pmatrix} 0.8859 & 0.1141 \\ 0.8584 & 0.1416 \end{pmatrix}$	$\begin{pmatrix} -0.4216 & 0.4216 \\ 3.1719 & -3.1719 \end{pmatrix}$

Table 3.4: 2-by-2 perturbed transition matrices and Q matrices for the 2010-2011, 2011-2012, and 2012-2013 financial years.

matrices for the three financial years (2010-2011, 2011-2012, 2012-2013). So, the exact generators exist for all our perturbed transition matrices with the chosen parameter $\delta = 0.1$ and embeddable B (fixed) matrix ($B = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$).

According to the algebraic approach, the exact generators exists for volatilities for all the financial years except the 2012-2013 financial year. Therefore, we only consider the perturbation approach for the volatilities for this financial year in this case. So,

$$\hat{P}_v = \begin{pmatrix} 0.7928 & 0.2072 \\ 0.7901 & 0.2099 \end{pmatrix}$$

where $\hat{P}_v$ is the perturbed transition matrix and $det(\hat{P}_v) = 0.0026$ for the 2012-2013 financial year volatilities. So, the exact generator exists:

$$Q_v = T(\log D)T^{-1} = \begin{pmatrix} -1.2341 & 1.2341 \\ 4.7053 & -4.7053 \end{pmatrix}$$

with the chosen parameter $\delta = 0.17$ (observe $\delta_3 = 0.5$ and we choose a low level $\tilde{\delta_3} = 0.17$) and embeddable B (fixed) matrix. However, the perturbed transition matrix is not embeddable with the parameter $\delta = 0.1$ and embeddable B (fixed) matrix ($B = B_2 = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$).

Random Search:

In this part, random matrices were used to find minimum value of the parameter δ in the perturbation approach with the fixed matrix $B = B_2$. First, we choose four iid variables and the P matrix is constructed such that:

$$P_t = \begin{pmatrix} \frac{\xi_1'}{\xi_1'+\xi_2'} & \frac{\xi_2'}{\xi_1'+\xi_2'} \\ \frac{\xi_1''}{\xi_1''+\xi_2''} & \frac{\xi_2''}{\xi_1''+\xi_2''} \end{pmatrix}$$

where ξ_i' and ξ_i'' are the iid for all $i = 1, 2$. Also, all the random variables are normally distributed ($\xi_i \sim NE(1)$). Then, perturbation is considered for this random matrix with the same B matrix as the previous part to find δ.

All these steps are repeated 10^5 times in order to find the minimum value of the parameter δ, which is 0.2. The aim is to find practical $\delta = \delta_n$ (such as confidence interval) such that for the observed error $\delta(\hat{P}) >> \delta$ we can say that the matrix $\hat{P}$ is not embeddable (observe $\delta = 0.2$).

Case 2 (3×3)

Algebraic Approach:

P is the transition matrix for the 2009-2010 financial year:

$$P_t = \begin{pmatrix} 0.2051 & 0.7436 & 0.0513 \\ 0.1429 & 0.7551 & 0.1020 \\ 0.1667 & 0.7500 & 0.0833 \end{pmatrix}$$

We do not have any necessary and sufficient conditions for 3×3 stochastic transition matrices, $det(P) > 0$. Using the algebraic approach: According to the function $e^Q = P = TDT^{-1}$, we wish to find the eigenvalue (D) and eigenvector (T) of P

$$Q = T(\log D)T^{-1} =$$

$$\begin{pmatrix} -1.7094 & 3.7052 & -2.0156 \\ 0.1432 & -1.2451 & 1.1117 \\ 1.6950 & 3.9188 & -5.6713 \end{pmatrix}$$

The matrix is a real matrix. However there is a negative value on the off-diagonal, so the exact generator does not exist.

Also, Table 3.5 shows that the exact generators do not exist for the transition matrices for the two financial years 2009-2010, and 2011-2012. Instead, we consider the perturbation approach for these years.

Cases	Q matrices
BP(2010-2011)	$\begin{pmatrix} -1.5678+1.2048i & 1.5172+1.5096i & 0.0507+0.3047i \\ 1.3732-1.3663i & -1.7774+1.7121i & 0.4040-0.3456i \\ 0.1481+0.8886i & 1.3020-1.1135i & -1.4500+0.2248i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -6.6596 & 9.7921 & -3.1315 \\ 0.3780 & -0.8546 & 0.4766 \\ 3.9983 & -3.9513 & -0.0473 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -2.0995-0.0793i & 1.5290+0.6971i & 0.5704-0.6178i \\ 0.3762-0.0330i & -0.7550+0.2900i & 0.3788-0.2571i \\ 0.1588+0.3761i & 3.6563-3.3070i & -3.8152+2.9308i \end{pmatrix}$

Table 3.5: Q matrices of the 3-by-3 transition matrices for the 2010-2011, 2011-2012, 2012-2013 financial years

Cases	Q matrices
BP(2009-2010)	$\begin{pmatrix} -2.5570 & 1.8508 & 0.7062 \\ 0.3887 & -0.2968 & -0.0919 \\ -3.1243 & 5.8412 & -2.7169 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -1.5839+0.8545i & 1.5372-0.6754i & 0.0466-0.1790i \\ 0.4293-0.1886i & -0.4273+0.1491i & -0.0020+0.0395i \\ 2.6579-10.2045i & -0.4079+8.0665i & -2.2500+2.1380i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -2.4491+2.5892i & 2.2517-2.7298i & 0.1974+0.1406i \\ 0.2710-0.3286i & -0.2741+0.346i & 0.0031-0.0178i \\ 5.3235+3.7917 & 0.6971-3.9976i & -6.0206+0.2059i \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -2.3304 & 1.2007 & 1.1298 \\ 0.3269 & -0.1968 & -0.1301 \\ -2.0164 & 4.3503 & -2.3339 \end{pmatrix}$

Table 3.6: Q matrices of the 3-by-3 transition matrices for volatilities

Moreover, Table 3.6 illustrates that the exact generators do not exist for all other transition matrices for the three financial years 2010-2011, 2011-2012, 2012-2013 either for the volatility case. Therefore, we consider the perturbation approach for all the estimated transition matrices.

Perturbation Approach:

P is the same transition matrix for the 2009-2010 financial year found with the algebraic approach:

$$P_t = \begin{pmatrix} 0.2051 & 0.7436 & 0.0513 \\ 0.1429 & 0.7551 & 0.1020 \\ 0.1667 & 0.7500 & 0.0833 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

which is a fixed embeddable matrix. Also δ is a parameter which we choose as $\delta = 0.1$. Let us compute the pertubated transition matrix by:

$$\hat{P} \longrightarrow (1-\delta)P + \delta B \longrightarrow Q$$

$$\hat{P} = \begin{pmatrix} 0.2178 & 0.7026 & 0.0795 \\ 0.1619 & 0.7129 & 0.1251 \\ 0.1834 & 0.7083 & 0.1083 \end{pmatrix}$$

and $det(\hat{P}) = 3.6279 \times 10^{-5}$. The matrix is real but not the true generator (negative value on the off-diagonal).

$$Q = T(\log D)T^{-1} = \begin{pmatrix} -1.4852 & 4.2201 & -2.7351 \\ 0.0164 & -1.8104 & 1.7940 \\ 2.1430 & 4.7858 & -6.9286 \end{pmatrix}$$

Table 3.7 shows the perturbed transition matrices and Table 3.8 shows their Q matrices for all the other transition matrices for the three financial years 2010-2011, 2011-2012, and 2012-2013. So, the exact generators do not exist for all our perturbed

Cases	$\hat{P}$ perturbed matrices
BP(2010-2011)	$\begin{pmatrix} 0.4276 & 0.4619 & 0.1105 \\ 0.4212 & 0.4212 & 0.1574 \\ 0.2583 & 0.4333 & 0.3084 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.0934 & 0.8134 & 0.0934 \\ 0.1359 & 0.7235 & 0.1406 \\ 0.1834 & 0.6334 & 0.1834 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.2279 & 0.6415 & 0.1306 \\ 0.1528 & 0.7267 & 0.1205 \\ 0.1413 & 0.7533 & 0.1053 \end{pmatrix}$

Table 3.7: 3-by-3 pertubated transition matrices

Cases	Q matrices
BP(2010-2011)	$\begin{pmatrix} -1.6362+1.1813i & 1.5290-1.4803i & 0.1073+0.2988i \\ 1.4101-1.3897i & -1.8709+1.7414i & 0.4606-0.3515i \\ 0.1851+0.8652i & 1.3138-1.0842i & -1.4988+0.2189i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -6.6737 & 9.7014 & -3.0268 \\ 0.4695 & -1.0512 & 0.5816 \\ 4.0899 & -4.0428 & -0.0474 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -2.1324+0.0880i & 1.4546-0.7734i & 0.6778+0.6854i \\ 0.4486+0.0417i & -0.9347-0.3663i & 0.4861+0.3246i \\ 0.2313-0.3674i & 3.5819+3.2307i & -3.8132-2.8632i \end{pmatrix}$

Table 3.8: Q matrices of the 3-by-3 pertubated transition matrices

Cases	Q matrices
BP(2009-2010)	$\begin{pmatrix} -2.4820 & 1.6971 & 0.7846 \\ 0.5701 & -0.5565 & -0.0136 \\ -2.9452 & 5.6879 & -2.7430 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -1.6598+1.1501i & 1.5356-0.9094i & 0.1242-0.2408i \\ 0.4598+0.1075i & -0.5354-0.0850i & 0.0755-0.0225i \\ 2.6931-9.9196i & -0.4138+7.8432i & -2.2793+2.076i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -2.6580-2.4010i & 2.2622+2.5320i & -0.3957-0.1313i \\ 0.1679+0.5156i & -0.3682-0.5437i & 0.2004+0.0282i \\ 5.2096-3.6014i & 0.7119+3.7980i & -5.9215-0.1969i \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -2.2958 & 1.1101 & 1.1856 \\ 0.4668 & -0.3929 & -0.0739 \\ -1.8766 & 4.2600 & -2.3834 \end{pmatrix}$

Table 3.9: Q matrices of the 3-by-3 pertubated transition matrices for volatilities

transition matrices with the chosen parameter, $\delta = 0.1$, and the embeddable B (fixed) matrix.

The perturbation approach is then considered for the transition matrices for the volatility case. However, the exact generators exist for all two of the perturbed transition matrices with the chosen parameter $\delta = 0.1$ and embeddable B (fixed) matrix. Q matrices of the 3-by-3 pertubated transition matrices are illustrated in Table 3.9 for all financial years. In Cases 1 and 4, the matrices are real but not embeddable. (Pertubated transition matrices are illustrated in Table B.1.)

Random Search:

In this section, random matrices were used to find the minimum value of the parameter δ in the perturbation approach. First, we choose nine iid variables and the P matrix is constructed such that:

$$P_t = \begin{pmatrix} \frac{\xi_1'}{\xi_1'+\xi_2'+\xi_3'} & \frac{\xi_2'}{\xi_1'+\xi_2'+\xi_3'} & \frac{\xi_3'}{\xi_1'+\xi_2'+\xi_3'} \\ \frac{\xi_1''}{\xi_1''+\xi_2''+\xi_3''} & \frac{\xi_2''}{\xi_1''+\xi_2''+\xi_3''} & \frac{\xi_3''}{\xi_1''+\xi_2''+\xi_3''} \\ \frac{\xi_1'''}{\xi_1'''+\xi_2'''+\xi_3'''} & \frac{\xi_2'''}{\xi_1'''+\xi_2'''+\xi_3'''} & \frac{\xi_3'''}{\xi_1'''+\xi_2'''+\xi_3'''} \end{pmatrix}$$

where ξ_i', ξ_i'' and ξ_i''' are iid for all $i = 1, 2, 3$. Also, all the random variables normally distributed ($\xi_i \sim NE(1)$). Then, perturbation is considered for this random matrix with the same B matrix as previous to find δ. All these steps are repeated 10^5 times in order to find the minimum value of the parameter δ, which is 0.5.

Case 3 ($(2 \times 2) \otimes (2 \times 2)$***)***

Algebraic Approach:

The data is split into two groups ($Z = (X, Y)$). In this part we will consider P_X to be the transition matrix for the X group of the data via the EM and C4.5 (machine learning) algorithms to treat the missing values. However, there is no missing data in the Y group of the data. Therefore, the P_Y transition matrix is same for all processes.

P_X is the transition matrix of the X part of the split data for the 2009-2010 financial year via the EM algorithm:

$$P_X = \begin{pmatrix} 0.2419 & 0.7581 \\ 0.2449 & 0.7551 \end{pmatrix}$$

According the existence condition of 2×2 stochastic transition matrices $det(P) = -0.0030 < 0$, so the exact generator does not exist. Applying the algebraic approach:

According to the function $e^Q = P = TDT^{-1}$, we wish to find the eigenvalue (D) and eigenvector of P

$$Q_X = T(\log D)T^{-1} = \begin{pmatrix} -4.3907 + 2.3745i & 4.3907 - 2.3745i \\ 1.4184 - 0.7671i & -1.4184 + 0.7671i \end{pmatrix}$$

The matrix is even not a real matrix.

Moreover, Table 3.10 shows that the exact generators do not exist any other transition matrices (P_X via EM algorithm for the three financial years (2010-2011, 2011-2012, 2012-2013). And, Table 3.11 shows that the exact generators do not exist any other transition matrices (P_X via C4.5 algorithm for our dataset).

Also, P_Y transition matrices are same in the same financial for EM algorithm and C4.5 algorithm processes. P_Y is the transition matrix of Y part of the splitted data for 2009-2010 financial year:

$$P_Y = \begin{pmatrix} 0.9060 & 0.0940 \\ 0.9167 & 0.0833 \end{pmatrix}$$

Cases	Q matrices (P_X via EM)
BP(2009-2010)	$\begin{pmatrix} -4.3907+2.3745i & 4.3907-2.3745i \\ 1.4184-0.7671i & -1.4184+0.7671i \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -1.5412+1.4234i & 1.5412-1.4234i \\ 1.8610-1.7184i & -1.8610+1.7184i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -5.2174+2.4060i & 5.2174-2.4060i \\ 1.5951-0.7356i & -1.5951+0.7356i \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -2.2862 & 2.2862 \\ 0.7277 & 0.7277 \end{pmatrix}$

Table 3.10: Q_X matrices via the EM algorithm for the 2010-2011, 2011-2012, 2012-2013 financial years

Cases	Q matrices (P_X via EM)
BP(2009-2010)	$\begin{pmatrix} -1.2225+2.2366i & 1.2225-2.2366i \\ 0.5750-1.0050i & -0.5750+1.0050i \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -0.5998+1.57726i & 0.5998-1.57726i \\ 0.5992-1.5700i & -0.5992+1.5700i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -0.9629+2.3484i & 0.9629-2.3484i \\ 0.3252-0.7932i & 0.3252+0.7932i \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -1.2525+2.3050i & 1.2525-2.3050i \\ 0.4546-0.8366i & -0.4546+0.8366i \end{pmatrix}$

Table 3.11: Q_X matrices via the C4.5 algorithm for the 2010-2011, 2011-2012, 2012-2013 financial years

Cases	Q_Y matrices
BP(2009-2010)	$\begin{pmatrix} -0.4220+0.2922i & 0.4220-0.2922i \\ 4.1155-2.8494i & -4.1155+2.8494i \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -0.2320 & 0.2320 \\ 1.4182 & -1.4182 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -0.8388 & 0.8388 \\ 2.0949 & -2.0949 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -0.3849+0.3055i & 0.3849-0.3055i \\ 3.5732-2.8361i & -3.5732+2.8361i \end{pmatrix}$

Table 3.12: Q_Y matrices

According the existence condition of 2×2 stochastic transition matrix, $det(P) = -0.0107 < 0$, so the exact generator does not exist. Applying algebraic approach:

According to the function $e^Q = P = TDT^{-1}$, find the eigenvalue (D) and eigenvector of P

$$Q_Y = \begin{pmatrix} -0.4220+0.2922i & 0.4220-0.2922i \\ 4.1155-2.8494i & -4.1155+2.8494i \end{pmatrix}$$

The matrix is not a real matrix, so the exact generator does not exist.

There are exact generators for the 2010-2011, and 2011-2012 financial years. However, generators do not exist exist for the 2009-2010 and 2012-2013 financial years. These are presented in Table 3.12.

Now, let us find the generator for the tensor product of the transition matrices from splitting the data.

$$P = P_X \otimes P_Y = \begin{pmatrix} 0.2192 & 0.0227 & 0.6868 & 0.0713 \\ 0.2218 & 0.0202 & 0.6949 & 0.0632 \\ 0.2219 & 0.0230 & 0.6841 & 0.0710 \\ 0.2245 & 0.0204 & 0.6922 & 0.0629 \end{pmatrix}$$

P is the tensor product of the transition matrices where P_X is the transition matrix of the X group via the EM algorithm and P_Y is the transition matrix of the Y group for the 2009-2010 financial year. Table 2.2 and Table 2.3 illustrate the estimated transition matrices for the Markov chains via these algorithms.

Also, in Section 2.3, Table 2.5 and Table 2.6 illustrate that the tensor product of the estimated transition matrices for the volatility case where P_X is estimated via the EM and C4.5 algorithms, respectively.

We do not have any necessary and sufficient conditions for 4×4 stochastic transition matrices, $det(P) = 2.1871 \times 10^{-9} < 0$. Using the properties of the tensor product:

$$Q = I \otimes Q_Y + Q_X \otimes I =$$

$$\begin{pmatrix} -4.8127 + 2.6667i & 0.4220 - 0.2922i & 4.3907 - 2.3745i & 0 \\ 4.1155 - 2.8494i & -8.5062 + 5.2239i & 0 & 4.3907 - 2.3745i \\ 1.4184 - 0.7671i & 0 & -1.8404 + 1.0593i & 0.4220 - 0.2922i \\ 0 & 1.4184 - 0.7671i & 4.1155 - 2.8494i & -5.5339 + 3.6165i \end{pmatrix}$$

As before, the matrix is even not a real matrix, so the exact generator does not exist. Also, the Q matrix of the tensor product of the transition matrices for the other financial years are illustrated in Table 3.13, where P_X are estimated by the EM algorithm. Table 3.14 shows the Q matrix of the tensor product of the transition matrix for our dataset where P_X are estimated by the C4.5 algorithm.

Moreover, for the volatility case, the exact generators do not exist for all transition matrices (P_X via the EM and C4.5 algorithms for our dataset), which are illustrated in Table B.2. Also,the P_Y transition matrices are the same for the same financial years using both the EM and C4.5 algorithm processes (see Table B.3)). The Q matrix of the tensor product of the transition matrices for the other financial years are illustrated in Table 3.15 where P_X is estimated by the EM algorithm for the estimated volatilities. Table 3.16 shows the Q matrix of the tensor product of the transition matrix for our dataset, where P_X is estimated by the C4.5 algorithm for the estimated volatilities. There are no real generators for the tensor products of the transition matrices, therefore we need to consider the perturbation approach for all the tensor products of the estimated transition matrices in both situations (share price and volatility).

Perturbation Approach:

Q matrices
$$\begin{pmatrix} -1.7740+1.4231i & 0.2327+0.0001i & 2.5419-1.4231i & -0.0006-0.0001i \\ 1.4214-0.0003i & -2.9626+1.4235i & -0.0022-0.0004i & 1.5434-1.4228ii \\ 1.8620-1.7185i & -0.0011+0.0001i & -2.0940+1.7185i & 0.2330-0.0001i \\ -0.0036-0.0005i & 1.8645-1.7189 & 1.4212+0.0003i & -3.2822+1.7181i \end{pmatrix}$$
$$\begin{pmatrix} -3.4903+1.7073i & 0.8451-0.0001i & 2.6518-3.7073i & -0.0067+0.0001i \\ 2.0998+0.0014i & -4.7463+1.7081i & -0.0058+0.0014i & 2.6525-1.7081i \\ 2.2270-1.4343i & -0.0057-0.0001i & -3.0655+1.4343i & 0.8441+0.0001i \\ 0.0069+0.0014i & 2.2269-1.4335i & 2.1010-0.0014i & -4.3207+1.4335i \end{pmatrix}$$
$$\begin{pmatrix} -2.6701+0.3029i & 0.3839-0.3029i & 2.2850-0.0004i & 0.0012+0.0004i \\ 3.5896-2.8387i & -5.8758+2.8387i & -0.0079-0.0004i & 2.2941+0.0004i \\ 0.7270+0.0009i & 0.0007-0.0009i & -1.1117+0.3058i & 0.3840-0.3058i \\ -0.0040+0.0009i & 0.7318-0.0009i & 3.5694-2.8358i & -4.2972+2.8358i \end{pmatrix}$$

Table 3.13: Q matrix of the tensor product of the transition matrices (P_X is estimated via the EM algorithm) for the 2010-2011, 2011-2012, 2012-2013 financial years

Q matrices
$\begin{pmatrix} -2.1088 & 2.4259 & -1.3261 & 2.7995 \\ -1.2164 & -1.0241 & -0.1448 & 3.2328 \\ 1.4578 & -0.5469 & 0.4409 & -1.8548 \\ -0.0723 & -0.5639 & 1.3951 & -0.8042 \end{pmatrix}$
$\begin{pmatrix} -0.8320+1.5716i & 0.2321+0.0001i & 0.5999-1.5716i & -0.00061-0.0001i \\ 1.4179+0.0002i & -2.0177+1.5715i & 0.0003-0.0002i & 0.5995-1.5715i \\ 0.5994-1.5700i & -0.0002+0.0001i & -0.8314+1.5700i & 0.2322-0.0001i \\ -0.0006-0.0006i & 0.3258-0.7926i & 2.0955+0.0006i & -2.4207+0.7926i \end{pmatrix}$
$\begin{pmatrix} -1.8029+2.3486i & 0.8401-0.0002i & 0.9642-2.3486i & -0.0013+0.0002i \\ 2.0977-0.0006i & -3.0605+2.3490i & -0.0027+0.0006i & 0.9656-2.3490i \\ 0.3256-0.7929i & -0.0004-0.0002i & -1.1643+0.7929i & 0.8391+0.0002i \\ -0.0006-0.0006i & 0.3258-0.7926i & 2.0955+0.0006i & -2.4207+0.7926i \end{pmatrix}$
$\begin{pmatrix} -1.6372+2.1617i & 0.3847+0.1433i & 1.2523-1.8562i & 0.0002-0.4488i \\ 3.5673+1.3277i & -4.8198+0.9773i & 0.0059-4.1638i & 1.2466+1.8588i \\ 0.4542-0.6740i & 0.0004-0.1625i & -0.8391+0.9795i & 0.3845-0.1429i \\ 0.0037-1.5081i & 0.409+0.6715i & 3.5695-1.3280i & -4.0241+2.1646i \end{pmatrix}$

Table 3.14: Q matrix of the tensor product of the transition matrices (P_X is estimated via the C4.5 algorithm) for the 2009-2010, 2010-2011, 2011-2012, 2012-2013 financial years

Q matrices
$\begin{pmatrix} -4.7475+0.0692i & 0.1067-0.0692i & 4.6408-0.1666i & 0.1666i \\ 3.3353-2.1616i & -7.9761+2.1616i & 5.2058i & 4.6408-5.2058i \\ 0.7868-0.0282i & 0.0282i & -0.8936-0.0692i & 0.1067+0.0692i \\ 0.8826i & 0.7868-0.8826i & 3.3353+2.1616i & -4.1221-2.1616i \end{pmatrix}$
$\begin{pmatrix} -1.9170-0.0268i & 0.0647+0.0268i & 1.8523+0.0759i & -0.0759i \\ 4.0784+1.6910i & -5.9307-1.6910i & -4.7835i & 1.8523+4.7835i \\ 0.5427+0.0222i & -0.0222i & -0.6074+0.0268i & 0.0647-0.0268i \\ -1.4015i & 0.5427+1.4015i & 4.0784-1.6910i & -4.6211+1.6910i \end{pmatrix}$
$\begin{pmatrix} -2.4098-2.6086i & 0.0593-0.1534i & 2.3505+2.6086i & 0.1534i \\ 2.0747-5.3706i & -4.4252+2.6086i & 5.3706i & 2.3505-2.6086i \\ 0.3230+0.3585i & 00211i & -0.3823-0.3585i & 0.0593-0.0211i \\ 0.7381i & 0.3230-0.3585i & 2.0747-0.7381i & -2.3977+0.3585i \end{pmatrix}$
$\begin{pmatrix} -3.1137-2.5906i & 0.1248-0.1222i & 2.9889+2.7129i & 0 \\ 3.0820-3.0194i & -6.0709+0.3065i & 0 & 2.9889+2.7129i \\ 0.4724+0.4287i & 0 & -0.5971-0.3065i & 0.1248-0.1222i \\ 0 & 0.4724+0.4287i & 3.0820-3.0194i & -3.5544+2.5906i \end{pmatrix}$

Table 3.15: Q matrix of the tensor product of the transition matrices for the volatilities (P_X estimated via the EM algorithm)

Q matrices
$$\begin{pmatrix} -2.8065+2.8246i & 0.1067-0.0974i & 2.6998-2.7272i & 0 \\ 3.3353-3.0442i & -6.0351+5.7714i & 0 & 2.6998-2.7272i \\ 0.4103-0.4144i & 0 & -0.5170+0.5118i & 0.1067-0.0974i \\ 0 & 0.4103-0.4144i & 0.3353-3.0442i & -3.7456+3.4586i \end{pmatrix}$$
$$\begin{pmatrix} -2.5087+0.0492i & 0.0648-0.0492i & 2.4439 & 0 \\ 4.0766-3.0924i & -6.5205+3.0924i & 0 & 2.4439 \\ 0.5680 & 0 & -0.6328+0.0492i & 0.0648-0.0492i \\ 0 & 0.5680 & 4.0766-3.0924i & -4.6446+3.0924i \end{pmatrix}$$
$$\begin{pmatrix} -5.3688 & 0.0593 & 5.3095 & 0 \\ 2.0744 & -7.3839 & 0 & 5.3095 \\ 0.7654 & 0 & -0.8247 & 0.0593 \\ 0 & 0.7654 & 2.0744 & -2.8389 \end{pmatrix}$$
$$\begin{pmatrix} -2.6977+0.1223i & 0.1248-0.1223i & 2.5729 & 0 \\ 3.0816-3.0193i & -5.6545+3.0193i & 0 & 2.5729 \\ 0.1437 & 0 & -0.2685+0.1223i & 0.1248-0.1223i \\ 0 & 0.1437 & 3.0816-3.0193i & -3.2253+3.0193i \end{pmatrix}$$

Table 3.16: Q matrix of the tensor product of the transition matrices for the volatilities (P_X estimated via the C4.5 algorithm)

In this subsection we follow the same procedure as for Cases 1 and 2 shown at the beginning of this section. The main difference is that we have two 2×2 transition matrices (P_X, P_Y) for perturbation. Also, we have two different results for P_X because of the missing data treatment.

Therefore, we start with the perturbation of these transition matrices for all financial years. To perturb, we use

$$B = \begin{pmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{pmatrix}$$

which is a fixed embeddable matrix and nearly the best embeddable matrix according to our preliminary calculations. Also, δ is a parameter which we choose as $\delta = 0.1$.

Then,

$$\hat{P} \longrightarrow (1-\delta)P + \delta B \longrightarrow Q$$

We then compute their Q matrices $(\hat{Q}_X, \hat{Q}_Y)$. Table B.4 and Table B.5 show that the pertubated transition matrices $\hat{P_X}$ estimated via the EM and C4.5 algorithms. Also, P_Y is the same for both these algorithms. Table B.6 illustrates the pertubated $\hat{P}_Y$ transition matrices (see Appendix for these tables).

Finally, we use the property of the tensor product to compute the $\hat{Q}$ matrices for the tensor products of the pertubated transition matrices:

$$Q = I \otimes Q_Y + Q_X \otimes I$$

The $\hat{Q}$ matrices of the tensor product of the pertubated transition matrices are illustrated in Table 3.17 and Table 3.18 for all transition matrices for all financial years. Hence, the exact generators do not exist for any of our perturbed transition matrices with the chosen parameter, $\delta = 0.1$, and embeddable B (fixed) matrix.

Furthermore, the perturbation approach is considered for the transition matrices for the volatility case. However, the exact generators do not exist for any perturbed transition matrices with the chosen parameter, $\delta = 0.1$, and the embeddable B (fixed) matrix. The perturbed transition matrices for each method (EM and C4.5) and their Q matrices are illustrated in Table B.7, Table B.8 and Table B.9 in the Appendix. Also, the $\hat{Q}$ matrices of the tensor products of the pertubated transition

Cases	Q matrices
BP(2009-2010)	$\begin{pmatrix} -3.0253 & 0.4520 & 2.5733 & 0 \\ 3.2901 & -5.8634 & 0 & 2.5733 \\ 0.9124 & 0 & -1.3644 & 0.4520 \\ 0 & 0.9124 & 3.2901 & -4.2025 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -2.8674 & 0.2697 & 2.5977 & 0 \\ 1.3095 & -3.9072 & 0 & 2.5977 \\ 3.0972 & 0 & -3.3668 & 0.2697 \\ 0 & 3.0971 & 1.3095 & -4.4066 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -3.3219 & 0.7568 & 2.5651 & 0 \\ 1.7538 & -4.3189 & 0 & 2.5651 \\ 0.8663 & 0 & -1.6231 & 0.7568 \\ 0 & 0.8663 & 1.7538 & -2.6201 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -2.4059 & 0.5139 & 1.8920 & 0 \\ 3.6124 & -5.5044 & 0 & 1.8920 \\ 0.6652 & 0 & -1.1791 & 0.5139 \\ 0 & 0.6652 & 3.6124 & -4.2776 \end{pmatrix}$

Table 3.17: Q matrix of the tensor product for the perturbed transition matrices (P_X is estimated via the EM algorithm) for all financial years

Q matrices
$\begin{pmatrix} -1.8951+2.1028i & 0.4520 & 1.4431-2.1028i & 0 \\ 3.2901 & -4.7332+2.1028i & 0 & 1.4431-2.1028i \\ 0.7129-1.0388i & 0 & -1.1649+1.0388i & 0.4520 \\ 0 & 0.7129-1.0388i & 3.2901 & -4.0030+1.0388i \end{pmatrix}$
$\begin{pmatrix} -0.9878+1.5716i & 0.2697 & 0.7181-1.5716i & 0 \\ 1.3095 & -2.0276+1.5716i & 0 & 0.7181-1.5716i \\ 0.7173-1.5700i & 0 & -0.9870+1.5700i & 0.2697 \\ 0 & 0.7173-1.5700i & 1.3095 & -2.0268+1.5700i \end{pmatrix}$
$\begin{pmatrix} -1.8854+2.3058i & 0.7568 & 1.1286-2.3058i & 0 \\ 1.7538 & -2.8824+2.3058i & 0 & 1.1286-2.3058i \\ 0.4091-0.8358i & 0 & -1.1659+0.8358i & 0.7568 \\ 0 & 0.4091-0.8358i & 1.7538 & -2.1629+0.8358i \end{pmatrix}$
$\begin{pmatrix} -1.9831+2.2617i & 0.5139 & 1.4692-2.2617i & 0 \\ 3.6124 & -5.0815+2.2617i & 0 & 1.4692-2.2617i \\ 0.5716-0.8799i & 0 & -1.0855+0.8799i & 0.5139 \\ 0 & 0.5716-0.8799i & 3.6124 & -4.1840+0.8799i \end{pmatrix}$

Table 3.18: Q matrix of the tensor product for the perturbed transition matrices (P_X is estimated via the C4.5 algorithm) for all financial years

Q matrices
$\begin{pmatrix} -3.0855 & 0.3368 & 2.7487 & 0 \\ 5.0595 & -7.8082 & 0 & 2.7487 \\ 0.573 & 0 & -0.8941 & 0.3368 \\ 0 & 0.5573 & 5.0595 & -5.6168 \end{pmatrix}$
$\begin{pmatrix} -1.8179 & 0.1924 & 1.6255 & 0 \\ 3.7697 & -5.3952 & 0 & 1.6255 \\ 0.5337 & 0 & -0.7261 & 0.1924 \\ 0 & 0.5337 & 3.7697 & -4.3034 \end{pmatrix}$
$\begin{pmatrix} -3.1522+2.6857i & 0.1266 & 3.0256-2.6857i & 0 \\ 1.8402 & -4.8658+2.6857i & 0 & 3.0256-2.6857i \\ 0.5136-0.4559i & 0 & -0.6402+0.4559i & 0.1266 \\ 0 & 0.5136-0.4559i & 1.8402 & -2.3538+0.4559i \end{pmatrix}$
$\begin{pmatrix} -4.8353+0.2185i & 0.4015-0.2185i & 4.4338 & 0 \\ 5.3695-2.9230i & -9.8033+2.9230i & 0 & 4.4338 \\ 0.8500 & 0 & -1.2515+0.2185i & 0.4015-0.2185i \\ 0 & 0.8500 & 5.3695-2.9230i & -6.2195+2.9230i \end{pmatrix}$

Table 3.19: Q matrix of the tensor product for the perturbed transition matrices (P_X is estimated via the EM algorithm) for the volatilities

matrices are illustrated in Table 3.19 and Table 3.20 for all financial years. Hence, the exact generators do not exist for any of the perturbed transition matrices with the chosen parameter, $\delta = 0.1$, and the embeddable B (fixed) matrix.

Random Search:

In this subsection, random matrices are used to find minimum value for the parameter δ for the perturbation approach. First, we choose eight iid variables, P_X, and the P_Y matrix is constructed such that:

$$P_X = \begin{pmatrix} \frac{\xi_1'}{\xi_1'+\xi_2'} & \frac{\xi_2'}{\xi_1'+\xi_2'} \\ \frac{\xi_1''}{\xi_1''+\xi_2''} & \frac{\xi_2''}{\xi_1''+\xi_2''} \end{pmatrix}$$

$$P_Y = \begin{pmatrix} \frac{\eta_1'}{\eta_1'+\eta_2'} & \frac{\eta_2'}{\eta_1'+\eta_2'} \\ \frac{\eta_1''}{\eta_1''+\eta_2''} & \frac{\eta_2''}{\eta_1''+\eta_2''} \end{pmatrix}$$

Q matrices
$\begin{pmatrix} -4.5468+2.6506i & 0.3368 & 4.2100-2.6506i & 0 \\ 5.0595 & -9.2695+2.6506i & 0 & 4.2100-2.6506i \\ 0.7799-0.4910i & 0 & -1.1167+0.4910i & 0.3368 \\ 0 & 0.7799-0.4910i & 5.0595 & -5.8394+0.4910i \end{pmatrix}$
$\begin{pmatrix} -2.2089 & 0.1924 & 2.0165 & 0 \\ 3.7697 & -5.7862 & 0 & 2.0165 \\ 0.5395 & 0 & -0.7319 & 0.1924 \\ 0 & 0.5395 & 3.7697 & -4.3092 \end{pmatrix}$
$\begin{pmatrix} -2.9603 & 01.126 & 2.8337 & 0 \\ 1.8402 & -4.6739 & 0 & 2.8337 \\ 0.5073 & 0 & -0.6339 & 0.1266 \\ 0 & 0.5073 & 1,8402 & -2.3475 \end{pmatrix}$
$\begin{pmatrix} -2.5747+0.2185i & 0.4015-0.2185i & 2.1732 & 0 \\ 5.3695-2.9230i & -7.5427+2.9230i & 0 & 2.1732 \\ 0.2038 & 0 & -0.6053+0.2185i & 0.4015-0.2185i \\ 0 & 0.2038 & 5.3695-2.9230i & -5.5733+2.9230i \end{pmatrix}$

Table 3.20: Q matrix of the tensor product for the perturbed transition matrices (P_X is estimated via the C4.5 algorithm) for the volatilities

where $\xi_i^{'}$, $\xi_i^{''}$ are iid and $\eta_i^{'}$, $\eta_i^{''}$ are iid for all $i = 1, 2$. Also, all the random variables are normally distributed ($\xi_i \sim NE(1)$, $\eta_i \sim NE(1)$). Then, perturbation is considered for this random matrix with the same B matrix as previous to find δ.

All these steps are repeated 10^5 times in Matlab in order to find the minimum value of the parameter δ, which is 0.22.

3.4 More Data Analysis on the Embedding for the Financial Data

In this section, more financial data is tested in the 3-by-3 case of the embedding problem to give a general conclusion. For these analyses, the data is the share prices of the twenty different companies for four different financial years (2009-2010, 2010-2011, 2011-2012, 2012-2013). First, the 3-by-3 transition matrices for these share prices are estimated using MLE and their Q matrices are then computed by the algebraic approach to the embedding problem (see Section 3.3.2 for details). All these matrices are presented in Appendix B. In general, the Q matrices are not real, or are real with negative values on their diagonals. Therefore, the estimated transition matrices are not embeddable except for the transition matrices associated with the Royal Mail (2012-2013), Sainsbury (2012-2013), Unilever (2010-2011), Whitbread (2011-2012), WHSmith (2011-2012), and Wolseley (2010-2011) share prices. Briefly, for most data the Markov chains are not embeddable.

3.5 Conclusion

The intention of this chapter was to analyse whether the discrete time model permits extension or embedding in the continuous time model. If the model is converted to a continuous time model (embeddable), it means that the result (data) is observable each time. This is a plausible and important method in the financial sector. If the model is a continuous time model, many existing formulae (such as option pricing) are applicable to the model. We analyse this problem by applying it to our data.

This part of the research is an extensive case study of the embedding problem for financial data and its volatility. It gives a real financial application that illustrates the importance of the embedding problem. As a result, in general we could not embed the discrete time Markov chain in the continuous time Markov chain. This means that the model we considered should be treated as a discrete time model.

Inspecting the case study, our overall results are as follows:

(i) *Share prices, algebraic approach.* According to the algebraic approach, the exact generators do not exist for any constructed transition matrices for any transition matrix types ((2×2), (3×3), $(2\times 2)\otimes(2\times 2)$) in any financial year considered (2009-2010, 2010-2011, 2011-2012);

(ii) *Volatility, algebraic approach.* According to the algebraic approach, exact generators exist in several cases for the volatility (2×2) transition matrices, and for several cases of volatility for the $(2\times 2)\otimes(2\times 2)$ case; however, the exact generator does not exist for the volatility (3×3) transition matrices for any of the data.

(iii) *Share prices, volatility, perturbation approach.*

(2×2) *case.* According to the perturbation approach, the exact generators exist for the slightly perturbed (2×2) transition matrices for all the financial data and volatilities. The chosen parameter is $\delta = 0.1$.

(3×3) *case.* However, in the case of the perturbation parameter $\delta = 0.1$, none of the perturbed transition matrices have an exact generator. This suggests that if one were to consider a larger number of states, this would only make things worse.

$(2\times 2)\otimes(2\times 2)$ *case.* Surprisingly, for the small perturbation parameter $\delta = 0.1$, roughly half of the perturbed transition matrices have an exact generator.

For the (2×2) case, data often appear as independent observations. This explains why, in most cases, the embedding problem has a negative solution.

Moreover, for these analyses, the share prices of the twenty different companies for four different financial years (2009-2010, 2010-2011, 2011-2012, 2012-2013) are considered for embeddability via the algebraic approach. The overall conclusion is that in most cases the Markov chains are not embeddable.

Overall, this study shows that using a continuous time model for volatility is more stable than the original share prices. In addition, considering a small number of carefully chosen states is generally more reliable.

Chapter 4

Random Walk on The Lamplighter Group

This chapter is based on the analysis of arbitrarily chosen groups of share prices of relatively small data sizes (around 250 closing prices for each group). In addition, their volatility is analysed using the same procedure as for share prices. The research presented here is effectively a continuation of the study in [99].

In particular, by constructing the Markov chain models, we find that traditional models such as the geometric Levy process pricing models do not provide proper fits to this data because the estimated transition matrices for all our datasets have highly variable rows. Moreover, traditional models such as (i) Brownian motion, (ii) random walks with iid increments, (iii) geometric Brownian motion, (iv) geometric Levy processes and geometric random walks, and (v) continuous time homogeneous Markov chain pricing models do not fit the data [99]. The embedding of discrete time Markov chains into continuous time Markov chains seems to be an even bigger problem than the independence assumption. We recall here that the embedding problem is to solve the log matrix problem, i.e., to find the Q matrix such that the stochastic matrix P has a representation $P = e^Q$ (see the previous chapter for details of the embedding problem and its connection with financial data). Due to the fact that interest rates are practically zero, it has become increasingly popular to use random walks as the modelling tool of choice for risky assets. Random walks on the wreath products (which is a specialized product of two groups based on a semi-direct product) are known in the literature as lamplighter random walks because of the intuitive interpretation of such walks in terms of the configuration of lamps (as defined in [58]).

Motivated by the nature of share prices, we discuss several procedures to model risky assets via the random walk on the lamplighter group (or its tensor products). Specifically, we model data as a geometric Markov chain with a hidden random walk on the group [101]. The hidden random walk is constructed on the lamplighter group on $\mathbb{Z}_3$ and on the tensor product of groups $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. The lamplighter group has a specific structure where the hidden information is actually explicit. We assume that the positions of the lamplighters are known, but we do not know the status of the lamps. We refer to this as a hidden random walk on the lamplighter group. Choosing the semi-group generators for the branching random walk requires tedious calculations and is still an open question for future research [21]. To analyse the sensitivity of the generators, we choose at least two different generator sets.

We also construct the biased random walks on the tensor product of the lamplighter group models (as introduced in [72]) to fit the data. Overall, several branching walk models are considered. A Monte Carlo simulation is then applied to find the best fit. The results are then compared with analytic errors computed for the relative distance between two tensor products of random stochastic matrices.

The missing data algorithms (which are considered in Chapter 2) and Monte Carlo simulation are used to find the best fit in the sense of finding the random walk for which the distance between the original matrix and the corresponding 3×3 reduced transition matrix is smallest. In this chapter, as a measure of the fit of the class of stochastic matrices, we consider the smallest trace norm difference between two transition matrices. The fit is relatively good. Moreover, for the randomly chosen data sets, the α-biased random walk on the tensor product of the lamplighter group and $\alpha - \lambda$-biased random walk provide significantly better fits to the data.

Additionally, we follow the same procedures for the volatility process as for the share price data to ensure consistency. Then, all the results are compared with analytic errors. Most of the results presented in this chapter have been reported in [81].

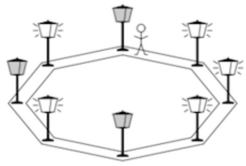


Figure 4.1: Lamplighter group.

4.1 Random Walk on the Group

In this section, we work mainly with directed graphs and have found that branching trees and graphs are particularly useful in the stochastic modelling of the data. We construct the Cayley graph on the lamplighter group by choosing particular group generators. Then, we model the jump part of the shares as a random walk on the associated Cayley graph of the lamplighter group, which is called a random walk on the lamplighter group. Therefore, details of the random walk on the group, graphs and branching trees have been reviewed in Chapter 1.

As stated above, the data jumps are modelled as a geometric Markov chain with a hidden random walk on the lamplighter group on $\mathbb{Z}_3$ and on the tensor product of groups $\mathbb{Z}_2 \otimes \mathbb{Z}_2$.

Let us begin with definition of the lamplighter group:

4.1.1 Lamplighter Group

The lamplighter group G_1 is defined as a semi-direct product, $G_1 := \mathbb{Z} \ltimes \Sigma_{x\in\mathbb{Z}}\mathbb{Z}_2$, with the direct sum of copies of $\mathbb{Z}_2$ indexed by $\mathbb{Z}$ ($\ltimes$ is the semi-direct product); for $m, m' \in \mathbb{Z}$ and $\eta, \eta' \in \Sigma_{x\in\mathbb{Z}}\mathbb{Z}_2$ the group operation is

$$(m, \eta)(m', \eta') := (m + m', \eta \oplus \rho^{-m}\eta')$$

where $\oplus$ is a component-wise addition modulo 2 and ρ is a left shift [72]. The m-move is for the lamplighter and the η-move for the status of lamps (on $\rightarrow$ off and off $\rightarrow$ on), (see Figure 4.1 and [12] for more details).

For example, $P = \{0, 1, 2\}$ is the element and ρ is a left shift. So,

$$\rho P = \{1, 2, 0\}$$

ρ^{-1} is right shift:

$$\rho^{-1}P = \{2, 0, 1\}$$

Also, ρ^{-m} has become an m-steps right shift. If $m = 4$,

$$\rho^{-4}P = \{2, 0, 1\}$$

Additionally, the element $\eta \in \Sigma_{x\in\mathbb{Z}}\mathbb{Z}_2$ is referred to as a configuration and $\eta(k)$ as the bit at k. We further identify $\mathbb{Z}_2$ with $\{0, 1\}$.

The position of the marker is denoted by $M(x)$ in the state x, which is the first component of an element $x = (m, \eta) \in G_1$.

$(1, 0)$,$(-1, 0)$ and $(0, 1_0)$ are the generators of the lamplighter group G_1.

The reason for the name of this group is that a streetlamp at each integer with the configuration η represents which lights are on, namely those where $\eta = 1$ [72].

We also may imagine a lamplighter at the position of the marker. The first two generators of G_1 correspond to the lamplighter taking a step either to the right or to the left (leaving the lights unchanged). The third generator ($(0, 1_0)$) corresponds to flipping the light at the position of the lamplighter (see Figure4.2 and notations [72].

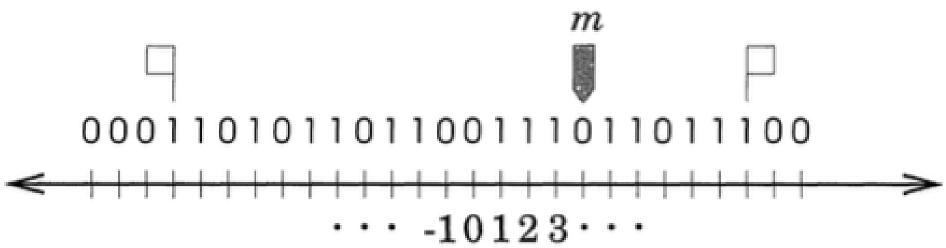


Figure 4.2: An element of the lamplighter group.

4.1.2 Random Walk on the Lamplighter Group

In order to construct a branching-type random walk on the group and apply it to fit the data, we choose a generator set S for the group, generating it as a semi-group that is in general non-symmetric ($x \in S$ but $x^{-1} \notin S$). Then, we construct the branching random walk on the group, and finally we model the hidden Markov chain on the lamplighter group. Branching means that off-springs of x are xy, where $y \in S$.

We begin by considering the lamplighter group on $\mathbb{Z}_3$. There are three positions, 0, 1, and 2, on the group. They refer to: no jump, small jump, and big jump, respectively. Direct analysis shows that the lamplighter group on $\mathbb{Z}_3$ has 24 elements, which are listed in the following set:

$$\begin{gathered}E = \{e1 = (0, (0, 0, 0)), e2 = (0, (0, 0, 1)), e3 = (0, (0, 1, 0)), e4 = (0, (1, 0, 0)),\\ e5 = (0, (0, 1, 1)), e6 = (0, (1, 0, 1)), e7 = (0, (1, 1, 0)), e8 = (0, (1, 1, 1)),\\ e9 = (1, (0, 0, 0)), e10 = (1, (0, 0, 1)), e11 = (1, (0, 1, 0)), e12 = (1, (1, 0, 0)),\\ e13 = (1, (0, 1, 1)), e14 = (1, (1, 0, 1)), e15 = (1, (1, 1, 0)), e16 = (1, (1, 1, 1,)),\\ e17 = (2, (0, 0, 0)), e18 = (2, (0, 0, 1)), e19 = (2, (0, 1, 0)), e20 = (2, (1, 0, 0)),\\ e21 = (2, (0, 1, 1)), e22 = (2, (1, 0, 1)), e23 = (2, (1, 1, 0)), e24 = (2, (1, 1, 1))\}\end{gathered}$$

To examine the sensitivity of the generators, two different generator sets are chosen at random. First, we choose a random set of elements and verified that the set was indeed the generator (as a semi-group). If the set generates the group, the set is chosen as the generator set. Else, we choose another random set and repeat all the steps again until we find two different generator sets. Theoretically, it may appear that for two different generators the results may be qualitatively different. Choosing the "right" generator is still an open question [21].

The two randomly chosen generator sets of the lamplighter group on $\mathbb{Z}_3$ are:

$$\begin{aligned} S_1 &= \{e4 = (0, (1, 0, 0)), e11 = (1, (0, 1, 0))\},\\ S_2 &= \{e10 = (1, (0, 0, 1)), e20 = (2, (1, 0, 0))\} \end{aligned}$$

We explain in detail the next procedure for the first generator set S_1. Then, we compare the results for both choices in Section 4.2.

Let us continue with constructing the simple random walk on the group. The transition probabilities for the simple random walk with d links are defined by

$$w_{ij} = 1/d \text{ if } i \text{ links to } j, w_{ij} = 0 \text{ otherwise}$$

Then, the 24×24 transition matrix W of the simple random walk is defined as follows:

$$w_{ij} = \begin{cases} 1/2 & \text{if} \quad i \quad \text{links to} \quad j,\\ 0 & \text{otherwise} \end{cases}$$

The hidden Markov chain on the lamplighter group is then constructed to model the data. For the hidden part, it is assumed that we know the lamplighter positions, but we do not know the status of the lamps. Hence, the possible positions are observed as follows:

$$\mathbf{0} = \{e1 = (0,(0,0,0)), e2 = (0,(0,0,1)), e3 = (0,(0,1,0)), e4 = (0,(1,0,0)),$$
$$e5 = (0,(0,1,1)), e6 = (0,(1,0,1)), e7 = (0,(1,1,0)), e8 = (0,(1,1,1))\}$$
$$\mathbf{1} = \{e9 = (1,(0,0,0)), e10 = (1,(0,0,1)), e11 = (1,(0,1,0)), e12 = (1,(1,0,0)),$$
$$e13 = (1,(0,1,1)), e14 = (1,(1,0,1)), e15 = (1,(1,1,0)), e16 = (1,(1,1,1,))\}$$
$$\mathbf{2} = \{e17 = (2,(0,0,0)), e18 = (2,(0,0,1)), e19 = (2,(0,1,0)), e20 = (2,(1,0,0)),$$
$$e21 = (2,(0,1,1)), e22 = (2,(1,0,1)), e23 = (2,(1,1,0)), e24 = (2,(1,1,1))\}$$

Finally, we construct the branching tree and then the branching type random walk. There are two methods that can be used to construct the branching tree: the first method is that of choosing one of the elements from the generator set which became the initial point of the tree; then, all the generators are used to generate the branching tree. We need to stop when we receive each element of the group, and so the branching tree is completed. Figure 4.3 shows the generated branching tree via the first method, which illustrates the generated branching tree with the first generator set (S_1), starting from e_4.

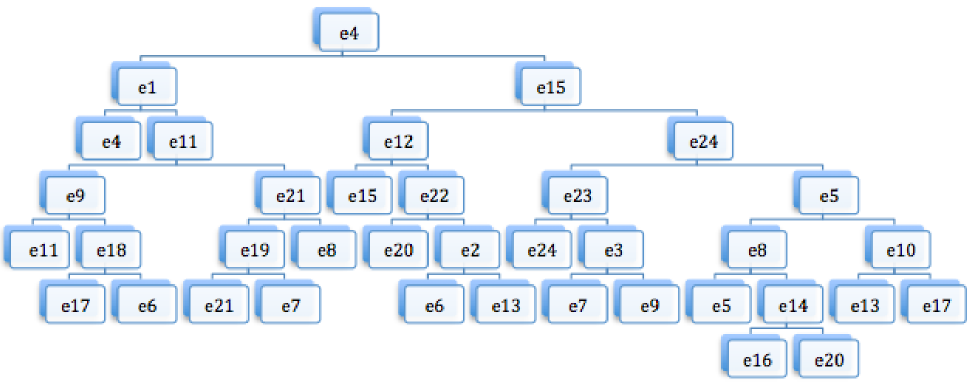


Figure 4.3: The branching tree via method 1.

The second method is to use all the elements of each generator as the initial

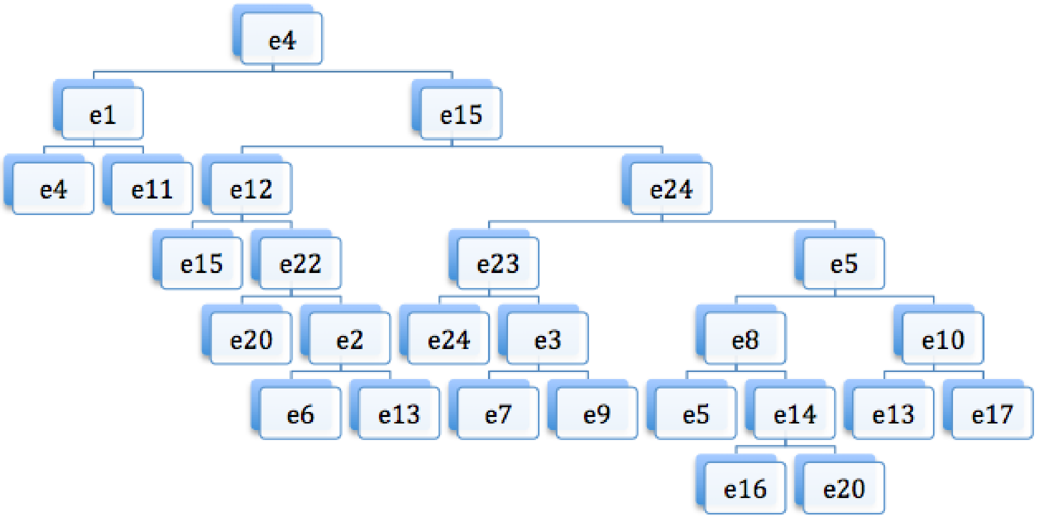


Figure 4.4: The branching tree via method 2: Initial point is $e4$.

points of the tree. Then, all elements of each generator are used to generate the branching tree. As with the first method, we need to stop when we receive each element of the group, at which point the branching tree is completed. Figure 4.4 and Figure 4.5 show the branching trees generated by each generator via the second method. Note that each element of the generator generates the branching tree.

We run two branching trees simultaneously until we have all the elements [17]. Figure 4.3, Figure 4.4 and Figure 4.5 show that both approaches generate the whole group [29].

Based on the original branching walk, we construct a new random walk on group of states. Then, for the new random walk, we find the transition matrix by simulation. The simulation of the model is run 10^5 times to find the transition matrix so as to subsequently find the best fit for the estimated 3×3 transition matrix by the maximum-likelihood estimation (MLE). The answer may be found theoretically, but it seems the random simulation is a considerably more efficient way of finding it.

Biased Random Walk on the Lamplighter Group

In this part, biased random walks on the lamplighter group are considered.

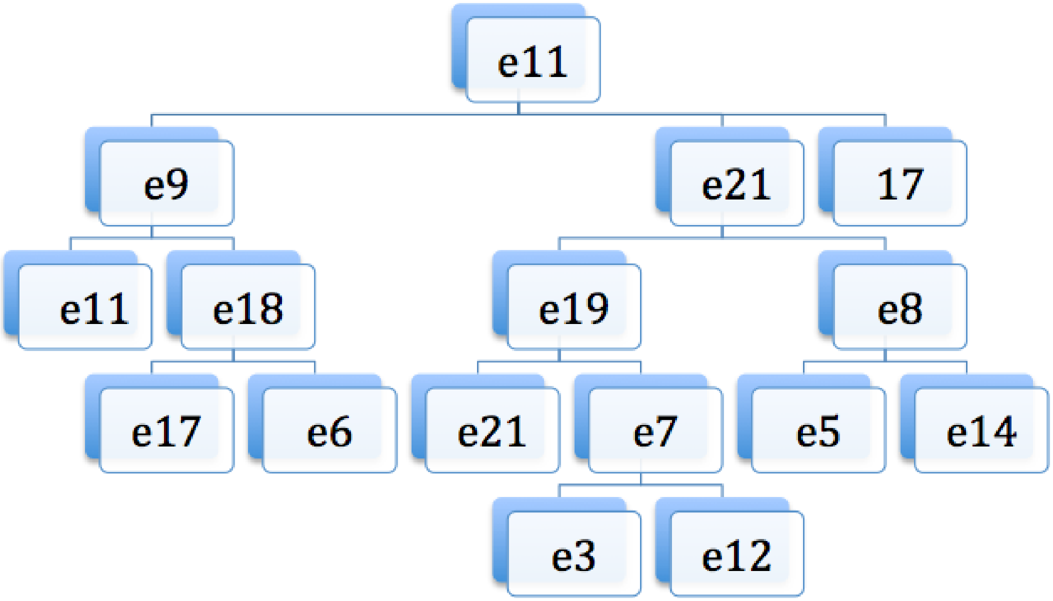


Figure 4.5: The branching tree via method 2: Initial point is $e11$.

The λ- biased random walk on the lamplighter group: Following Lyons, Pemantle, and Peres, for $\lambda > 0$ we define the λ- biased random walk RW_λ on a locally connected finite graph with a distinguished vertex Θ as the time-homogeneous Markov chain $\{X_n; n \geq 0\}$ with the following transition probabilities [72]. The distance from a vertex $|v|$ to Θ is the number of edges on the shortest path joining the two vertexes. Suppose that v is a vertex of the graph. Let $v_1, ..., v_k(k \geq 1$ unless $v = \Theta)$ be the neighbours of v at a distance $|v| - 1$ from Θ and let $u_1, u_2, ..., u_j \quad (j \geq 0)$ be the neighbours of v. Then, the transition probabilities are

$$\begin{aligned} w(v, v_i) &= \frac{\lambda}{(k\lambda + j)} \quad \text{for} \quad i = 1, ..., k; \\ w(v, u_i) &= \frac{1}{(k\lambda + j)} \quad \text{for} \quad i = 1, ..., k. \end{aligned}$$

And,

$$w_{ij} = \begin{cases} 1/d & \text{if there are} \quad d \quad \text{links where} \quad d > 0, \\ 0 & \text{otherwise} \end{cases}$$

when the λ- biased condition for the neighbours of the vertex v is satisfied [72].

To construct the λ- biased random walk for the data, we closely follow the construction procedure for the simple random walk. We work with the same 3×3-

state Markov chain and generator sets S_1, S_2. Moreover, we only treat the case of the first generator set S_1. The results are then analysed for both generator sets (S_1, S_2) in Section 4.2.

Then, the λ- biased random walk is constructed on the lamplighter group via transition probabilities:

$$w(v, v_i) = \frac{\lambda}{(\lambda+1)}\ ,\ w(v, u_i) = \frac{1}{(\lambda+1)} \tag{4.1.1}$$

and,

$$w_{ij} = \begin{cases} 1/2 & \text{if} \quad d = 2, \\ 0 & \text{otherwise} \end{cases} \tag{4.1.2}$$

when the λ- biased condition for the neighbours is satisfied.

Let us take a vertex $v = e_8$ as an example to explain the biased random walk:

In our branching tree the distinguished vertex is $\Theta = e_4$ and neighbours of the vertex are $\{v_1 = e_5, v_2 = e_14\}$. The distance from a vertex $|v|$ to Θ is the number of the edges on a shortest path joining the two vertexes. In this case,

$|v_1| = |v| - 1$

The vertex has a neighbour with shorter distance to the distinguished vertex. Therefore, the biased random walk can be considered here with the transition probabilities:

$w_{8,5} = \frac{1}{\lambda+1}, \quad w_{8,14} = \frac{\lambda}{\lambda+1}$

The transition probabilities of a 24×24 Markov chain are calculated based on the Cayley graph. The transition matrix P_1 consists of two parts.

In the first part all links below have transition probabilities of 1/2, i.e.,

$w_{4,1} = w_{4,15} = w_{1,4} = w_{1,11} = \frac{1}{2}\ ,\ w_{11,9} = w_{11,21} = w_{18,17} = w_{18,6} = \frac{1}{2};$

$w_{17,18} = w_{17,4} = w_{6,2} = w_{6,15} = \frac{1}{2}\ ,\ w_{21,19} = w_{21,8} = w_{19,21} = w_{19,7} = \frac{1}{2};$

$w_{14,16} = w_{14,20} = w_{16,14} = w_{16,23} = \frac{1}{2}\ ,\ w_{15,12} = w_{15,24} = w_{1,4} = w_{1,11} = \frac{1}{2};$

$w_{12,15} = w_{12,22} = w_{20,22} = w_{20,1} = \frac{1}{2}\ ,\ w_{2,6} = w_{2,13} = w_{13,10} = w_{13,19} = \frac{1}{2}$

$w_{24,23} = w_{24,5} = w_{23,24} = w_{23,3} = \frac{1}{2}\ ,\ w_{10,13} = w_{10,17} = \frac{1}{2};$

The second part probabilities are defined by

$w_{9,11} = \frac{1}{\lambda+1}, \quad w_{9,18} = \frac{\lambda}{\lambda+1}\ ,\ w_{7,12} = \frac{1}{\lambda+1}, \quad w_{7,3} = \frac{\lambda}{\lambda+1};$

$w_{8,5} = \frac{1}{\lambda+1}, \quad w_{8,14} = \frac{\lambda}{\lambda+1}\ w_{22,2} = \frac{1}{\lambda+1}, \quad w_{22,20} = \frac{\lambda}{\lambda+1};$

$w_{3,7} = \frac{1}{\lambda+1}, \quad w_{3,9} = \frac{\lambda}{\lambda+1}\ , \ w_{5,8} = \frac{1}{\lambda+1}, \quad w_{5,10} = \frac{\lambda}{\lambda+1}$

Then, the hidden Markov chain on the lamplighter group is constructed to model the data. The hidden part is same as before, with the known lamplighter positions but unknown states of lamps. A Monte Carlo simulation is run 10^5 times to choose the optimal parameter λ to find the transition matrix that allows us to find the best fit for the estimated transition matrices. The value of the parameter λ is chosen by the minimum over grid $(0, 1)$.

The α- biased random walk on the lamplighter group: We consider a slightly perturbed simple random walk on the lamplighter group generated as a semi-group with a non-symmetric set of generators. The approach is similar to the previous cases with the same set up: the same two generator sets (S_1, S_2) are chosen and the same 3×3-state Markov chain is considered as the initial matrix. And again, we only treat the case of the first generator set S_1 with the results then being analysed for both generator sets (S_1, S_2) in Section 4.2.

As before, based on the Cayley graph, we calculate the transition probability of a 24×24-state Markov chain . Notice that $e_1 = 0$ is not in our generators (i.e., $e_1 = 0 \notin S_i$), and so staying at the same position is not allowed in the branching-type random walk. To modify this, the α parameter is introduced and the transition matrix is perturbed by the diagonal matrix. For instance, the 24×24 matrix for the α- biased random walk is:

$$w_{ij} = \begin{cases} \frac{1}{2}(1-\alpha) & \text{if} \quad i \quad \text{links to} \quad j, \\ \alpha & \text{if} \quad i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{4.1.3}$$

The other steps (e.g., hidden part, observations,...) are similar to the previous approaches. A Monte Carlo simulation used to find the best fit using trace norm differences for the estimated transition matrices.

The $\alpha - \lambda$- biased random walk on the lamplighter group: The process is similar to the previous cases.

The transition probabilities for the $\alpha - \lambda$- biased random walk are:

$$w(v, v_i) = (1-\alpha)\frac{\lambda}{(k\lambda + j)})\ ,\ w(v, u_i) = (1-\alpha)\frac{1}{(k\lambda + j)}) \quad \text{for} \quad i = 1, ..., k.$$

v_i and u_i are sites satisfying conditions explained in the "Biased Random Walk on the Lamplighter Group" section.

And

$$w_{ij} = \begin{cases} (1-\alpha)\frac{1}{d} & \text{if} \quad i \quad \text{links to} \quad j, \\ \alpha & \text{if} \quad i = j, \\ 0 & \text{otherwise} \end{cases}$$

when the neighbours of the vertex v satisfy the λ- biased condition.

Specifically for this case, the 24×24 matrix W is:

$$w(v, v_i) = (1-\alpha)\frac{\lambda}{(\lambda+1)} \ , \ w(v, u_i) = (1-\alpha)\frac{1}{(\lambda+1)} \tag{4.1.4}$$

and

$$w_{ij} = \begin{cases} \frac{1}{2}(1-\alpha) & \text{if} \quad i \quad \text{links to} \quad j, \\ \alpha & \text{if} \quad i = j, \\ 0 & \text{otherwise} \end{cases} \tag{4.1.5}$$

when the λ- biased condition for the neighbours is satisfied.

Notice that the 24×24 transition matrix P_2 of the Markov chain is found by

$$P_2 = (1-\alpha)P_1 + \alpha I$$

where I is the identity matrix and P_1 is the transition matrix of the λ-biased case. Finally, the Monte Carlo simulation is run 10^5 times to chose the optimal parameters λ and α (0 to 1) in order to compute the transition matrix used to find the best fit for the original transition matrices. The values of the parameters λ and α are chosen by the minimum over grid $(0, 1)$.

4.1.3 Random Walk on the tensor product of the Lamplighter Group

In this section, we consider the tensor product of the lamplighter groups on $\mathbb{Z}_2$. First, let the group G be the tensor product of two groups $G = G_1 \otimes G_2$. The elements of the group G are pairs of the elements of the groups G_1 and G_2.

$$G = G_1 \otimes G_2 = (a, b), \quad a \in G_1, b \in G_2;$$
$$(a_1, b_1) \otimes (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$$

Consider the lamplighter group on the group G, and particularly $G_1 = G_2 = \mathbb{Z}_2$.

Notice that the elements of the lamplighter group on $\mathbb{Z}_2$:

$$E = (0,(0,0)),(0,(0,1)),(0,(1,0)),(0,(1,1)),(1,(0,0)),(1,(0,1)),(1,(1,0)),(1,(1,1))$$

Then, we can introduce the elements of the tensor product of the lamplighter group on the group G. The tensor product of the lamplighter group has 64 elements because of property of itself and the tensor product. There are four positions 0, 1, 2, and 3 on this group. They refer to differences between the daily-adjusted closing values of the share prices as no small jump, small jump and no big jump, big jump, respectively.

The set of the elements is:

$E = \{e1 = (0,(0,0,0,0)), e2 = (0,(0,0,0,1)), e3 = (0,(0,0,1,0)), e4 = (0,(0,1,0,0)),$
$e5 = (0,(1,0,0,0)), e6 = (0,(0,0,1,1)), e7 = (0,(0,1,1,0)), e8 = (0,(1,1,0,0)),$
$e9 = (0,(0,1,0,1)), e10 = (0,(1,0,0,1)), e11 = (0,(1,0,1,0)), e12 = (0,(0,1,1,1)),$
$e13 = (0,(1,0,1,1)), e14 = (0,(1,1,0,1)), e15 = (0,(1,1,1,0)), e16 = (0,(1,1,1,1)),$
$e17 = (1,(0,0,0,0)), e18 = (1,(0,0,0,1)), e19 = (1,(0,0,1,0)), e20 = (1,(0,1,0,0)),$
$e21 = (1,(1,0,0,0)), e22 = (1,(0,0,1,1)), e23 = (1,(0,1,1,0)), e24 = (1,(1,1,0,0)),$
$e25 = (1,(0,1,0,1)), e26 = (1,(1,0,0,1)), e27 = (1,(1,0,1,0)), e28 = (1,(0,1,1,1)),$
$e29 = (1,(1,0,1,1)), e30 = (1,(1,1,0,1)), e31 = (1,(1,1,1,0)), e32 = (1,(1,1,1,1)),$
$e33 = (2,(0,0,0,0)), e34 = (2,(0,0,0,1)), e35 = (2,(0,0,1,0)), e36 = (2,(0,1,0,0)),$
$e37 = (2,(1,0,0,0)), e38 = (2,(0,0,1,1)), e39 = (2,(0,1,1,0)), e40 = (2,(1,1,0,0)),$
$e41 = (2,(0,1,0,1)), e42 = (2,(1,0,0,1)), e43 = (2,(1,0,1,0)), e44 = (2,(0,1,1,1)),$
$e45 = (2,(1,0,1,1)), e46 = (2,(1,1,0,1)), e47 = (2,(1,1,1,0)), e48 = (2,(1,1,1,1)),$
$e49 = (3,(0,0,0,0)), e50 = (3,(0,0,0,1)), e51 = (3,(0,0,1,0)), e52 = (3,(0,1,0,0)),$
$e53 = (3,(1,0,0,0)), e54 = (3,(0,0,1,1)), e55 = (3,(0,1,1,0)), e56 = (3,(1,1,0,0)),$
$e57 = (3,(0,1,0,1)), e58 = (3,(1,0,0,1)), e59 = (3,(1,0,1,0)), e60 = (3,(0,1,1,1)),$
$e61 = (3,(1,0,1,1)), e62 = (3,(1,1,0,1)), e63 = (3,(1,1,1,0)), e64 = (0,(1,1,1,1))\}$

As per the first model (see Section 4.1.2), we randomly choose two different generator sets of the tensor product of the lamplighter group $G_1 \otimes G_2$, generating it as a semi-group. The generator sets of the group are chosen as:

$$S_3 = \{e18 = (1, (0, 0, 0, 1)), e35 = (2, (0, 0, 1, 0))\},$$

$$S_4 = \{e36 = (2, (0, 1, 0, 0)), e50 = (3, (0, 0, 0, 1))\}$$

Then, we construct the simple random walk and the biased random walks via the tensor product of the lamplighter groups. We also consider the hidden Markov chain on the group to model the data. Finally, we estimate the transition matrix for the new model. The overall procedure is the same as in the case of the lamplighter group on Z_3, where necessary definitions and explanations can be found in Section 4.1.2.

Let us start by constructing the simple random walk on the group. The transition matrix of the simple random walk is:

$$w_{ij} = \begin{cases} 1/2 & \text{if } i \text{ links to } j, \\ 0 & \text{otherwise} \end{cases}$$

The hidden part and observations are similar to the previous model. Figure 4.6 shows the generated branching tree with the first generator set (S_3).

We run the simulation 10^5 times to find the best fit to the original matrix. The data is split as in Section 2.2. The transition matrices are estimated by the MLE and result in the tensor product of the estimated transition matrices. This tensor product is then used as the original matrix.

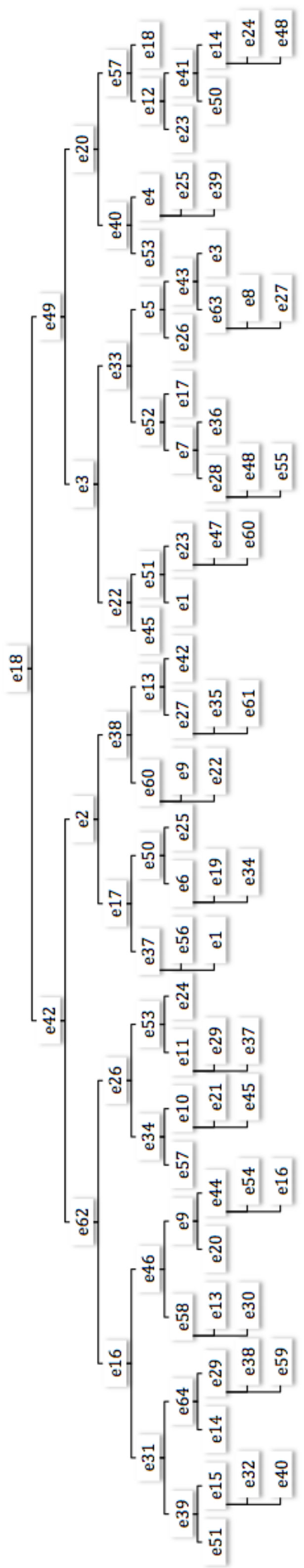


Figure 4.6: The branching tree via S_3.

Biased Random Walk on the tensor product of the Lamplighter Group

In this subsection, we construct the biased random walk via the tensor product of the lamplighter group by choosing optimal parameters. Let us start with the λ-biased random walk:

The λ- biased random walk on the tensor product of the lamplighter group: We consider a slightly perturbed λ- biased random walk on a lamplighter group [72].

Then, the λ- biased random walk is constructed on the lamplighter group via transition probabilities as defined in (4.1.1)-(4.1.2).

The transition probability of a 64×64-state Markov chain is calculated based on the Cayley graph. In particular, the transition probabilities are as below:

$$w_{1,18} = w_{1,35} = w_{2,17} = w_{2,38} = w_{3,22} = w_{3,33} = w_{5,26} = w_{5,43} = \tfrac{1}{2}$$
$$w_{6,19} = w_{6,34} = w_{7,28} = w_{7,36} = w_{8,30} = w_{8,47} = w_{9,20} = w_{9,44} = \tfrac{1}{2}$$
$$w_{10,21} = w_{10,45} = w_{11,29} = w_{11,37} = w_{12,23} = w_{12,41} = w_{13,27} = w_{13,42} = \tfrac{1}{2}$$
$$w_{14,24} = w_{14,48} = w_{15,32} = w_{15,40} = w_{16,31} = w_{16,46} = w_{17,37} = w_{17,50} = \tfrac{1}{2}$$
$$w_{18,42} = w_{18,49} = w_{19,43} = w_{19,54} = w_{21,33} = w_{21,58} = w_{23,47} = w_{23,60} = \tfrac{1}{2}$$
$$w_{24,36} = w_{24,62} = w_{25,46} = w_{25,52} = w_{27,35} = w_{27,61} = w_{28,48} = w_{28,55} = \tfrac{1}{2}$$
$$w_{29,38} = w_{29,59} = w_{30,41} = w_{30,56} = w_{32,44} = w_{32,63} = w_{33,52} = w_{33,5} = \tfrac{1}{2}$$
$$w_{34,57} = w_{34,10} = w_{35,55} = w_{35,11} = w_{36,49} = w_{36,8} = w_{37,56} = w_{37,1} = \tfrac{1}{2}$$
$$w_{38,60} = w_{38,13} = w_{39,51} = w_{39,15} = w_{40,53} = w_{40,4} = w_{42,62} = w_{42,2} = \tfrac{1}{2}$$
$$w_{43,63} = w_{43,3} = w_{44,54} = w_{44,16} = w_{45,64} = w_{45,6} = w_{47,59} = w_{47,7} = \tfrac{1}{2}$$
$$w_{48,61} = w_{48,12} = w_{49,3} = w_{49,20} = w_{51,1} = w_{51,23} = w_{52,7} = w_{52,17} = \tfrac{1}{2}$$
$$w_{53,11} = w_{53,24} = w_{54,2} = w_{54,28} = w_{55,4} = w_{55,19} = w_{56,15} = w_{56,21} = \tfrac{1}{2}$$
$$w_{57,12} = w_{57,18} = w_{58,13} = w_{58,30} = w_{59,5} = w_{59,31} = w_{60,9} = w_{60,22} = \tfrac{1}{2}$$
$$w_{61,10} = w_{61,32} = w_{62,16} = w_{62,26} = w_{63,8} = w_{63,27} = w_{64,14} = w_{64,29} = \tfrac{1}{2}$$
$$w_{31,64} = \tfrac{1}{\lambda+1}, \quad w_{31,39} = \tfrac{\lambda}{\lambda+1}, \; w_{46,58} = \tfrac{1}{\lambda+1}, \quad w_{46,9} = \tfrac{\lambda}{\lambda+1}$$
$$w_{26,53} = \tfrac{1}{\lambda+1}, \quad w_{26,24} = \tfrac{\lambda}{\lambda+1}, \; w_{50,25} = \tfrac{1}{\lambda+1}, \quad w_{50,6} = \tfrac{\lambda}{\lambda+1}$$
$$w_{22,45} = \tfrac{1}{\lambda+1}, \quad w_{22,51} = \tfrac{\lambda}{\lambda+1}, \; w_{20,40} = \tfrac{1}{\lambda+1}, \quad w_{20,57} = \tfrac{\lambda}{\lambda+1}$$
$$w_{4,25} = \tfrac{1}{\lambda+1}, \quad w_{4,39} = \tfrac{\lambda}{\lambda+1}, \; w_{41,14} = \tfrac{1}{\lambda+1}, \quad w_{41,50} = \tfrac{\lambda}{\lambda+1}$$

We use a similar process to the simple random walk on the tensor product of the lamplighter group. Finally the simulation is run 10^5 times, and the optimal

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.6478 & 0.3522 & 0 \\ 0 & 0.7650 & 0.2350 \\ 0.2222 & 0 & 0.7778 \end{pmatrix}$	$\begin{pmatrix} 0.2500 & 0.7500 & 0 \\ 0.0020 & 0.9970 & 0.0010 \\ 0 & 1.0000 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.5000 & 0.5000 & 0 \\ 0 & 0.8571 & 0.1429 \\ 0 & 0 & 1.0000 \end{pmatrix}$	$\begin{pmatrix} 0.3273 & 0.3165 & 0.3561 \\ 0.2414 & 0.4409 & 0.3177 \\ 0.2785 & 0.4399 & 0.2816 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.5584 & 0.4416 & 0 \\ 0 & 0.7258 & 0.2742 \\ 0.2544 & 0 & 0.7456 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.6667 & 0.3333 \\ 0.0010 & 0.9970 & 0.0020 \\ 0.3333 & 0.6667 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.4000 & 0.6000 & 0 \\ 0 & 0.9932 & 0.0068 \\ 0.0036 & 0 & 0.9964 \end{pmatrix}$	$\begin{pmatrix} 0.2500 & 0.5000 & 0.2500 \\ 0.0010 & 0.9950 & 0.0040 \\ 0.2000 & 0.8000 & 0 \end{pmatrix}$

Table 4.1: $\alpha - \lambda$ biased random walk on the lamplighter group.

parameter λ is found that gives the best fit to the original matrix.

The α- biased random walk on the tensor product of the lamplighter group: The 64×64 transition matrix is defined as in (4.1.3).

The other steps (e.g., hidden part, observations,...) are similar to the previous approaches. A Monte Carlo simulation is used to find the best fit using norm differences for the estimated transition matrices.

The α–λ- biased random walk on the tensor product of the lamplighter group: The transition probabilities for the $\alpha - \lambda$- biased random walk on the lamplighter group are defined similar to the first model, more specifically as in (4.1.4) -(4.1.5). The Monte Carlo simulation is run 10^5 times to choose the optimal parameter λ and α (0 to 1) to find the best fit for the estimated transition matrices.

4.2 Results and Comparisons

A branching-type random walk is constructed on the lamplighter group with two different generator sets (S_1, S_2) in Section 4.1.2. Also, a biased random walk is considered on the lamplighter group. 3×3 transition matrices are estimated by the Monte Carlo simulation. We estimate the transition matrices by constructing the model as the simple random walk and biased random walks on the lamplighter group to find the best fit for the estimated transition matrices. Table 2.1 shows the estimated transition matrices by maximum likelihood. Table 4.1 illustrates the transition matrices for a $\alpha - \lambda$ biased random walk on the lamplighter group. (The estimated transition matrices are illustrated in Table C.1, C.2, C.3 for the simple random walk and the other biased random walks on the lamplighter group in the Appendix.)

We calculate the trace error (norm) between the simulated matrices and the original transition matrices in order to compare with the four different random walks on the lamplighter group. Also, we consider the same computation as for another generator set to check sensitivity. The first row of Table 4.2 shows the comparison of trace norm values for all cases of each of the four methods with the first generator S_1. A comparison of trace norm values found for all methods with a second generator S_2 are stated in the second row of the table. These show that the best approximation was given by the α- biased and $\alpha - \lambda$- biased random walks. The smallest norm value is around 0.02. Also, the trace norm errors do not display significant differences between the values for the two different generator sets (S_1, S_2). The results raise an open question as to the assessment of the best generator fit and statistical hypothesis testing of the best fit.

To be consistent, we apply the four different methods with the four different random walks on the lamplighter group and two different generator sets to the volatility in the same procedure as the share price data. The simulated transition matrices are illustrated in Table C.5, Table C.6, Table C.7 and Table C.8 (for these tables, see the Appendix). Again, we calculate the trace error (norm) between the simulated matrices and that based on the data to compare the methods for the volatility processes. The first row of Table 4.3 shows a comparison of the trace norm values for all cases for each of the four methods with the first generator S_1. A comparison of the trace norm values for all methods with the second generator S_2 are reported in the second row of the table. They show that the best approximation

Cases	**Simple RW**	λ **biased**	α **biased**	$\alpha - \lambda$ **biased**
BP(2009-2010)	0.9715	0.8422	0.3077	0.0997
BP(2010-2011)	0.9684	0.8561	0.2952	0.0530
BP(2011-2012)	0.9664	0.8330	0.3140	0.0731
BP(2012-2013)	0.9781	0.9630	0.3136	0.0742
BP(2009-2010)	0.9832	0.8996	0.2224	0.0455
BP(2010-2011)	0.9076	0.8616	0.2649	0.0417
BP(2011-2012)	0.9868	0.9003	0.3426	0.0669
BP(2012-2013)	0.9855	0.9080	0.2992	0.0253

Table 4.2: Norm errors of the random walk on the lamplighter group

was given by the α- biased and $\alpha - \lambda$- biased random walks, which was a somewhat unexpected result. The smallest norm value is around 0.05; additionally, there is no significant difference between the trace error values for the two different generator sets, as per the share price data.

Cases	**Simple RW**	λ **biased**	α **biased**	$\alpha - \lambda$ **biased**
BP(2009-2010)	0.4902	0.4625	0.2453	0.1450
BP(2010-2011)	0.3380	0.2694	0.2798	0.1254
BP(2011-2012)	0.4995	0.4779	0.2215	0.1347
BP(2012-2013)	0.4958	0.4657	0.2828	0.1098
BP(2009-2010)	0.4862	0.3914	0.2217	0.0825
BP(2010-2011)	0.4447	0.3757	0.2507	0.1091
BP(2011-2012)	0.4826	0.4043	0.2215	0.0580
BP(2012-2013)	0.4529	0.3883	0.2064	0.0870

Table 4.3: Norm errors of the random walk on the lamplighter group for volatility.

Moreover, a branching-type random walk is constructed on the tensor product of the lamplighter group with two different generator sets (S_2, S_3) in Section 4.1.3. Also, a biased random walk is considered on the tensor product of the lamplighter group. 2×2 transition matrices estimated by the Monte Carlo simulation and their tensor product is computed. We estimate the transition matrices by constructing the model as the simple random walk and biased random walks on the lamplighter group to find the best fit for the estimated transition matrices. Specifically, we assume that the transformed data are such that $Z = (X, Y)$. X - "no jump", "small jump" and Y - "no big jump", "big jump" groups. P_X is the estimated transition matrix of X, and P_Y is the estimated transition matrix of Y. In order to estimate the transition matrix P_X, we deal with the missing data using both the EM and C4.5 algorithms. Therefore, we estimate two different transition matrices in each case and take their tensor product, as illustrated in Table C.9 and Table C.10 (see Appendix). Moreover, a branching-type random walk is constructed on the tensor product of the lamplighter group with two different generator sets (S_3, S_4) in Section 4.1.3. Also, a biased random walk is considered on the tensor product lamplighter group. Their transition matrices and the estimated transition matrices from Section 4.1.2 are compared with the tensor product of the original transition matrices $(P_X \otimes P_Y)$. The trace norm is applied to find the best fit to the data. To give an idea of the results of these methods, Table C.4 shows the transition matrices for the λ- biased random walk on the tensor product of the lamplighter group with the generator S_3 (see Appendix).

In order to check consistency, we once again apply the same procedure for volatility as we did for share prices. Therefore, a branching-type random walk is constructed on the tensor product of the lamplighter group with two different generator sets (S_3, S_4) for the volatility procedure. Also, a biased random walk is considered on the tensor product lamplighter group. Their transition matrices and the estimated transition matrices are compared with the tensor product of the original transition matrices $(P_X \otimes P_Y)$ (which are illustrated in Table C.11 and Table C.12 in the Appendix). Again, we calculate the trace error (norm) between the simulated matrices and that based on the data to compare the methods for the volatility processes.

Table 4.4 shows the comparison of the norm errors of the random walk on the tensor product of the lamplighter group with two different generator sets (S_3, S_4), where the transition matrices are estimated via the EM algorithm and C4.5 machine

Cases	**Simple RW**	λ **biased**	α **biased**	$\alpha - \lambda$ **biased**
BP(2009-2010)	1.2659	0.9677	0.3767	0.1797
BP(2010-2011)	1.1207	0.8234	0.2334	0.0290
BP(2011-2012)	1.1675	0.8690	0.2934	0.0123
BP(2012-2013)	1.2215	0.9214	0.3400	0.1089
BP(2009-2010)	1.2201	0.9123	0.3001	0.1569
BP(2010-2011)	1.1814	0.8692	0.2771	0.0783
BP(2011-2012)	1.2062	0.8966	0.3166	0.0995
BP(2012-2013)	1.2662	0.9528	0.3531	0.1534
BP(2009-2010)	1.2400	0.9460	0.3495	0.1261
BP(2010-2011)	1.1738	0.8759	0.2082	0.1334
BP(2011-2012)	1.1611	0.8612	0.2050	0.1399
BP(2012-2013)	1.2504	0.9528	0.3347	0.1142
BP(2009-2010)	1.2575	0.9457	0.3138	0.1141
BP(2010-2011)	1.1882	0.8790	0.2684	0.1016
BP(2011-2012)	1.1674	0.8626	0.2424	0.1283
BP(2012-2013)	1.2553	0.9407	0.3201	0.1176

Table 4.4: Norm errors of the random walk on the tensor product of the lamplighter group

learning algorithm. The first and second parts of the table show the results for the generator set S_3, whilst the third and forth parts of the table shows the results for the generator set S_4. The best approximation is again achieved by the α- biased and $\alpha - \lambda$- biased random walks. The smallest norm value is around 0.01; also, there is no significant difference between the values for the two different generator sets (S_3, S_4) and the two different missing value treatment methods (EM and C4.5 algorithm).

In addition, Table 4.5 shows the comparison of the norm errors of the random walk on the tensor product of the lamplighter group using the two different generator sets (S_3 and S_4) with the transition matrices estimated via the EM and C4.5 machine

Cases	**Simple RW**	λ **biased**	α **biased**	$\alpha - \lambda$ **biased**
BP(2009-2010)	0.4008	0.4089	0.2795	0.0880
BP(2010-2011)	0.3869	0.3891	0.2558	0.0698
BP(2011-2012)	0.3918	0.3875	0.2650	0.0296
BP(2012-2013)	0.4063	0.4102	0.2983	0.1008
BP(2009-2010)	0.4201	0.4060	0.2587	0.1086
BP(2010-2011)	0.3875	0.3714	0.2409	0.0672
BP(2011-2012)	0.4127	0.3966	0.2595	0.0855
BP(2012-2013)	0.4268	0.4123	0.2840	0.1061
BP(2009-2010)	0.4088	0.4131	0.2974	0.0843
BP(2010-2011)	0.3933	0.3974	0.2696	0.0637
BP(2011-2012)	0.3877	0.3905	0.2479	0.0486
BP(2012-2013)	0.4172	0.4165	0.2065	0.0651
BP(2009-2010)	0.4331	0.4181	0.2001	0.1120
BP(2010-2011)	0.4024	0.3874	0.2571	0.0692
BP(2011-2012)	0.4012	0.3858	0.2755	0.0410
BP(2012-2013)	0.4206	0.4076	0.2895	0.0566

Table 4.5: Norm errors of the random walk on the tensor product of the lamplighter group for volatilities

learning algorithms for the volatility process. The best approximation is again achieved by the α- biased and $\alpha - \lambda$- biased random walks. The smallest norm value is around 0.01, and there is no significant difference between the values for the two different generator sets and the two different missing value treatment methods. The choice of a "good" (best) generator is an open question.Theoretically it may appear that for two different generator's results may be qualitatively different. For the chosen simple trace norm metric, the trace errors for the different generators do not show any large differences. However, by choosing the weighted norms we may obtain more significant errors.

Briefly, the fit is relatively good. For the randomly chosen datasets, the α-biased random walk on the lamplighter group and $\alpha - \lambda$- biased random walk provide significantly better fits to the data. The smallest trace norm value is around 0.02. Also, the α-biased random walk on the tensor product of the lamplighter group and $\alpha - \lambda$- biased random walk provide significantly better fits to the data compared with other models. The smallest trace norm values is around 0.04. The random walk on the tensor product of the lamplighter group gives a better approximation than the random walk on the lamplighter group. Also, two different generators are chosen randomly for the each case, and each produce similar results. Therefore, this shows the sensitivity. Two different methods (EM and machine learning) are used to deal with the missing data. They yield close results, showing the robustness. In addition, the results are almost same regarding share price and its volatility. Although the method works well in many situations, results were occasionally unsatisfactory, such as with BP (2009-2010). This poses the question as to funding a better approach, e.g., funding a better generator.

For the moment we do not have results showing any direct advantage of the lamplighter construction with respect to the general construction of the Markovian model. However, parametric models are often easier to handle then non-parametric models [97], [32].

Chapter 5

Option Price for Binomial Model via Quantum Data

This chapter is devoted to the analysis of the quantum data. First, we introduce three different statistics, namely the Maxwell-Boltzmann, Bose-Einstein and Fermi-Dirac statistics, for both quantum and classical methods. Then, we present a single-step classical binomial model, and a one-step and multi-step quantum binomial model. Specifically, we estimate the parameters of the quantum two-step binomial market. Then, we find option prices based on these parameters in a quantum binomial market.

5.1 Statistical Mechanics

There are both quantum and classical statistics. The usual classical statistics are statistics on data such as that of moment estimator. Quantum statistics are mostly associated with the probabilities of eigenvalues defined by quantum state, density, ρ, etc. Statistics consider the extent to which particles can be distinguished and how that affects the number of unique states the particles can create. The statistics of distinguishable classical particles is described using Maxwell-Boltzmann statistics. In other words, the configurations of two different particles in two different states are not the same. Extending this to N particles yields the Maxwell-Boltzmann distribution of particles in different states.

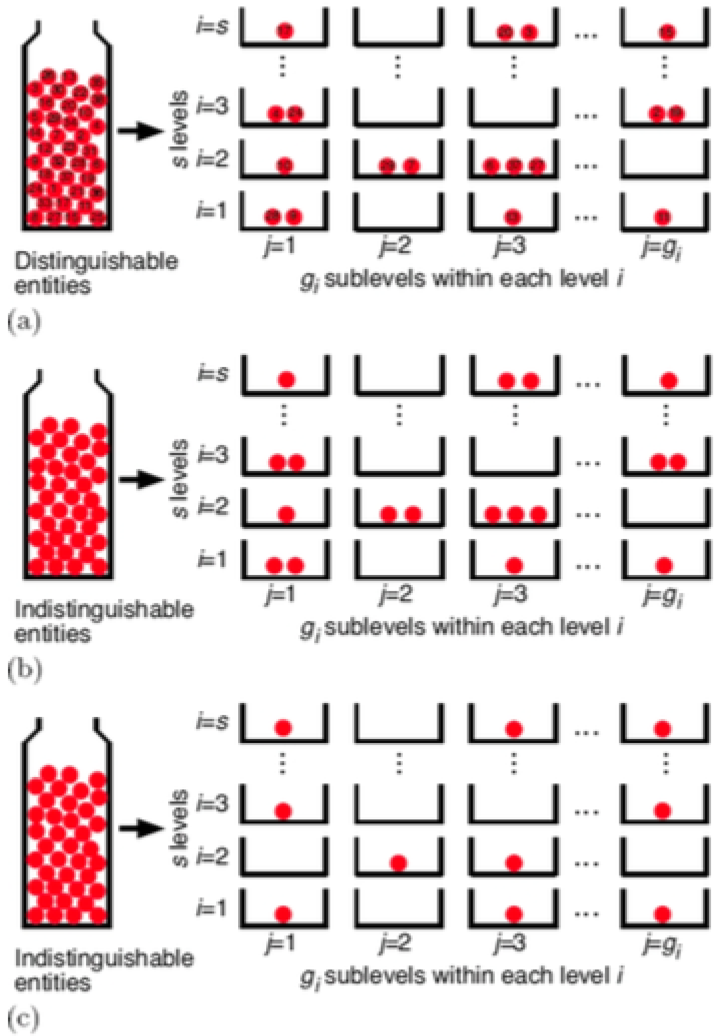


Figure 5.1: Configurations of (a) Maxwell-Boltzmann, (b) Bose-Einstein and (c) Fermi-Dirac ball-in-box models. Note that all N entities are fully allocated.

In addition, the statistics of indistinguishable quantum particles are described using Bose-Einstein and Fermi-Dirac statistics. When particles are indistinguishable, the number of unique states is decreased, as the number of unique configurations is reduced. See Figure 5.1 for a detailed picture of the difference between these statistical models [79]. Also, see further for details of likelihood for each of the statistics.

5.2 Classical and Quantum Binomial Model

In finance theory, the binomial model is a useful and well-known model used for pricing a stock option, and was first proposed by Cox, Ross and Rubinstein [30]. This model converts the Black-Scholes-Merton model into a discrete binary tree model of prices. The binomial tree is constructed between the valuation and expiration dates (number of steps). For each n steps of the binomial model, n new tree nodes are created. These new nodes represent a single, discrete change in the underlying stock price. For example, the term single-step means that only one discrete change in the stock underlying the option. In this research, we mostly consider a two-step quantum binomial model. First of all, we present the single-step classical binomial model. Then, we introduce the option price formula for single-step, two-step and multi-step quantum binomial models with the notation used in [102]. See Appendix C for details of the derivation of the option price formulae for these binomial models.

Single-Step Classical Binomial Model:

A binomial market (B, S) consists of a risk-free bank account B and stock of price S. An arbitrage-free portfolio is:

$$B_1 = B_0(1 + r), S_1 = S_0(1 + R) \tag{5.2.1}$$

where the interest rate r is constant and the volatility rate R takes two values such as:

$$-1 \leq u < r < d \tag{5.2.2}$$

S_1 has two outcomes. $C_u = [S_0(1 + u) - K]^+$ is the price of the call option if there is an upward movement u in the stock price and $C_d = [S_0(1 + d) - K]^+$ is the price of the call option if there is an downward movement d in the stock price and strike price K. Therefore, the formula for the current price of an option C is

$$C = \frac{1}{1+r}[\frac{r-d}{u-d}C_u + \frac{u-r}{u-d}C_d]. \tag{5.2.3}$$

Equivalently,

$$C = \frac{1}{1+r}[q_u C_u + q_d C_d]. \tag{5.2.4}$$

Single-Step Quantum Binomial Model:

Let us consider the single-step binomial model for the price of a call option with a risk-free bank account B, a stock S, and strike price K. An arbitrage-free replicating portfolio is:

$$B_1 = B_0(1+r), S_1 = S_0(I_2 + A) \tag{5.2.5}$$

where A is the quantum operator (or observable), which is a self-adjoint non-negative matrix. Notice that by abuse of notation we denote by S_1 a random variable in the classical case and an operator in the quantum case. Traditionally, saying the observed matrix refers to the observed eigenvalues.

And,

$$H = [S_1 - K]^+ \tag{5.2.6}$$

which takes two values: h_b is the price of the call option if there is an upward movement in the stock of $(1+b)$ and h_a is the price of the call option if there is a downward movement (we follow the notation used in [102] here, where $b = u - 1$ and $a = d - 1$):

$$h_b = [S_0(1+b) - K]^+, h_a = [S_0(1+a) - K]^+ \tag{5.2.7}$$

Hence, the current option value C is

$$C = \frac{1}{1+r}\,\mathrm{tr}[\rho H] = \frac{1}{1+r}[\frac{b-r}{b-a}h_a + \frac{r-a}{b-a}h_b]. \tag{5.2.8}$$

where the density matrix ρ (i.e., a self-adjojnt positive matrix with the trace 1) for all states in the risk-neutral world.

Multi-Step Quantum Binomial Model:

A single-step quantum binomial model is considered in order to derive the $N-$period multi-step model. In the multi-step model, each step is taken using the tensor product (see Chapter 2 for details of the tensor product) of the previous step to build a composite quantum system. Let us consider a call options in the $N-$ period quantum binomial market (B, S). Its payoff is

$$H_N = [S_N - K]^+ \tag{5.2.9}$$

where K is the strike price. And,

$$H_N = [S_N - K]^+ = \sum_{n=0}^{N}[S_0(1+b)^n(1+a)^{N-n} - K][\sum_{|\sigma|} \otimes_{j=1}^{N} |w_{j\sigma} >< w_{j\sigma}|] \tag{5.2.10}$$

In here, $w_{j\sigma} = |1 >= \binom{1}{0}$ means to choose $b = u - 1$ at time j in the tensor product; similarly, $w_{j\sigma} = |0 >= \binom{0}{1}$ means to choose $a = d - 1$

$$f(S_N) = \sum_{n=0}^{N} f(u^n d^{N-n}) \sum_{|\sigma|} |w_{j\sigma} >< w_{j\sigma}|$$

The second sum is over all permutations with a fixed number of chosen u's =n.

The price for the call option in this multi-step quantum binomial pricing can be written as follows:

$$C_0^N = \mathrm{tr}(\rho^{\otimes N}[S_N - K]^+) \tag{5.2.11}$$

Then, this equation is taken and the equivalent of the Cox-Ross-Rubinstein option pricing formula in [102] can be derived as follows:

$$C_0^N = (1+r)^{-N} \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} q^n (1-q)^{N-n} [S_0(1+b)^n(1+a)^{N-n} - K]^+. \tag{5.2.12}$$

where $q = q_u$ and $1 - q = q_d$ are the same as for the classical case.

Moreover, the option price formula is found via Bose-Einstein statistics instead of the classical Maxwell-Boltzmann statistics as follows [102]:

$$C_0^N = \mathrm{tr}(\rho^{\otimes N}[S_N - K]^+) \tag{5.2.13}$$

Also this equation is used to derive a new quantum option pricing formula in [102] as follows:

$$C_0^N = (1+r)^{-N} \sum_{n=0}^{N} \left(\frac{q^n(1-q)^{N-n}}{\sum_{k=0}^{N} q^k (1-q)^{N-k}} [S_0(1+b)^n(1+a)^{N-n} - K]^+\right). \tag{5.2.14}$$

Two-Step Quantum Binomial Model:

Now we present the two-step quantum binomial model, which is the key model in this research. We define the two-step binomial market (B, S) with a risk-free bank account $B = (B_0, B_1, B_2)$ and a stock $S = (S_0, S_1, S_2)$ as follows:

$$B_1 = B_0(1+r)^2, S_2 = S_0 \otimes_{j=1}^{2} (I_2 + A) \tag{5.2.15}$$

where stock price movement is represented by the quantum operator A, which is a Hermitian matrix.

In order to derive the option price formula in a two-step quantum binomial market, let us take $N = 2$ in equation 5.2.11. The option price formula is as follows:

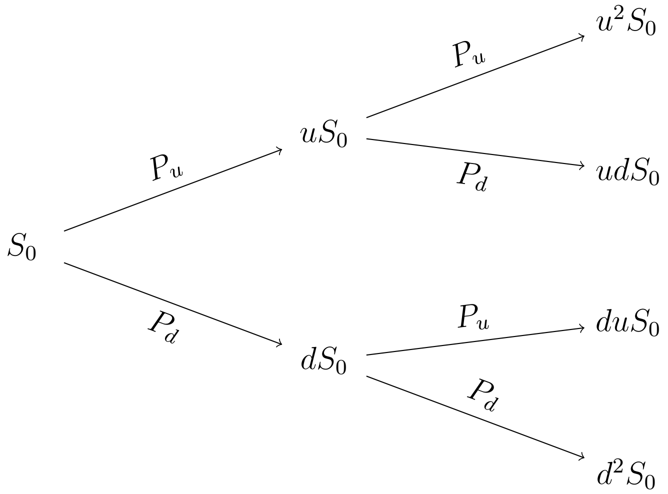
$$C_0^2 = \mathrm{tr}(\rho_j^{\otimes 2}[S_2 - K]^+) \tag{5.2.16}$$

The following formula is equal to (5.2.16), which is derived via classical Maxwell-Boltzmann statistics.

$$C_0^2 = (1+r)^{-2} \sum_{n=0}^{2} \frac{2!}{n!(2-n)!} q^n (1-q)^{2-n} [S_0(1+b)^n (1+a)^{2-n} - K]^+. \tag{5.2.17}$$

$$C_0^2 = \frac{1}{(1+r)^2} (q^2[S_0(1+b)^2 - K]^+ + 2q(1-q)[S_0(1+b)(1+a) - K]^+ \\ + (1-q)^2[S_0(1+a)^2 - K]^+).$$

5.3 Computing Option Price via the Two-Step Quantum Binomial Model



The diagram above illustrates a two-step binomial tree in a classical probability model. $S_0, S_1, \ldots$ is the dataset for which the probabilities P_u, P_d are observed. The estimated value of P_u is the total number of upward movements over the total number of movements. Also, the estimated value of P_d is the total number of downward movements over the total number of movements. Hence, for one step

$$Pf(X) = f(Xu)P_u + f(Xd)P_d.$$

In addition, we need to present a new parameter r, which is the interest rate, in order to consider the same binomial tree in a Cox-Rubinstein binomial market's so-called classical binomial model. In this model, the arbitrage-free martingale probabilities q_u and q_d are considered for computation in a Cox-Rubinstein binomial market. The main purpose where is the estimation of the parameters u and d. Note that P_u and P_d are irrelevant for the option pricing; however, they might be used to estimate the model.

Briefly, the formula of the option price via the classical binomial model is:

$$OP(C) = \frac{\sum_{j=0}^{n} f(S_0 u^j d^{n-j}) \binom{n}{j} q_u^j q_d^{n-j}}{(1+r)^n} \tag{5.3.18}$$

where $uq_u + dq_d = 1 + r$ (the same u and d as defined above) with no arbitrage condition ($d < 1 + r < u$) and $f(S_0 u^j d^{n-j}) = Cu^j d^{n-j}$ where the option price of option claim $C = f(S_n)$. Therefore,

$$q_u = \frac{1 + r - d}{u - d}, q_d = 1 - q_u.$$

Also, in this research we consider the call option ($C = (S_2 - K)^+$). For a two-step binomial model, the option price is:

$$OP = \frac{Cuuq_u^2 + 2Cudq_u q_d + Cddq_d^2}{(1+r)^2}. \tag{5.3.19}$$

In this research, we work with quantum data. Firstly, expressing odd/ known results in a new form may allow for a better qualitative understanding. Additionally, the quantum model is well established [102], [95]. In the classical case, the probabilities P_u, P_d are observed from the data $S_0, S_1, \ldots$. The all difference between the classical case and quantum case is about data. In the quantum case, the data is observed $S_0, S_1, \ldots + errors$ with outliers or missing data. In our research, the observed data is the operator $H^{\otimes 2}$ that has eigenvalues $H^{\otimes 2}$.

Furthermore, the general formula for option price via the quantum binomial model is as follows:

$$OP(f(S_T) = f(S_0 H^{\otimes n}) = \frac{\mathrm{tr}(\rho^{\otimes n} f(S_0 H^{\otimes n}))}{(1+r)^n} \tag{5.3.20}$$

where ρ is density matrix ($\rho \geqslant 0, \rho^* = \rho$) and must have a trace of one. Also,

$$S_n = S_0 H^{\otimes n}, \rho = \hat{\rho}^{\otimes n}, \mathrm{tr}(\rho H) = 1 + r$$

For a one-step quantum binomial model, the option price is:

$$OP(f(S_T)) = OP(f(S_0 H) = \frac{\mathrm{tr}(\rho f(S_0 H))}{(1+r)} \tag{5.3.21}$$

where S_0 is constant and a Hermitian matrix H is the quantum operator, as follows:

$$H = V^* D_H V \text{ and } D_H = \begin{pmatrix} u & o \\ 0 & d \end{pmatrix} \geq 0$$

where V^* is the adjoint of V and V, V^*, D_H are for the diagonal representation.

Lemma 5.3.1. *Let H be a self-adjoint matrix:*

$$g(H) = V^* g(D_H) V.$$

Also, ρ is a density matrix, as follows:

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \geq 0$$

so, $\rho_{11}, \rho_{21} \geq 0$. ρ is a self-adjoint ($\rho^* = \rho$), ρ^* is an adjoint of ρ and $\bar{a}$ is the conjugate:

$$\rho^* = \begin{pmatrix} \rho_{11} & \bar{\rho}_{21} \\ \bar{\rho}_{12} & \rho_{22} \end{pmatrix}$$

so, $\rho_{12} = \rho_{21}$.

Also, the arbitrage-free condition $\mathrm{tr}(\rho H) = 1 + r$ is satisfied such that:

$$\begin{aligned} \mathrm{tr}(\rho H) &= \mathrm{tr}(\rho V^* D_H V) \\ &= \mathrm{tr}((V \rho V^*) D_H) \\ &= \mathrm{tr}(\tilde{\rho} D_H) \end{aligned}$$

if $\tilde{\rho} = V \rho V^*$ is a density matrix.

Lemma 5.3.2. *$\tilde{\rho} = V\rho V^*$ is a density matrix if ρ is the density.*

Proof:

(i) $\tilde{\rho}$ must be a self-adjoint

(ii) $\tilde{\rho}$ must have a trace of one.

$$\tilde{\rho}^* = (V\rho V^*)^* = (V^*)^*\rho V^* = V\rho V^* = \tilde{\rho}$$

and

$$\mathrm{tr}(\tilde{\rho}) = \mathrm{tr}(V\rho V^*) = \mathrm{tr}(V^*V\rho) = \mathrm{tr}(\rho) = 1$$

Therefore, $\tilde{\rho}$ is a density matrix.

Hence,

$$\mathrm{tr}(\tilde{\rho}D_H) = 1 + r$$
$$\mathrm{tr}(\tilde{\rho}) = 1$$

and

$$\tilde{\rho}_{11} + \tilde{\rho}_{22} = 1$$
$$\tilde{\rho}_{11}u + \tilde{\rho}_{22}d = 1 + r.$$

Then, $\tilde{\rho}_{11} = q_u, \tilde{\rho}_{22} = q_d$ and

$$q_u = \frac{1 + r - d}{u - d}, q_d = 1 - q_u.$$

Finally, we obtain the option price formula for the one-step model:

$$OP(f(S_1)) = \frac{f(S_0u)q_u + f(S_0d)q_d}{(1 + r)} \tag{5.3.22}$$

In this research, our aim is to estimate the parameters required to describe a two-step quantum binomial market. Therefore, we derive the option price formula for the two-step quantum binomial model via same procedure used above for the one-step quantum binomial model. For a two-step quantum binomial model, the option price is:

$$OP(f(S_T)) = OP(f(S_0H^{\otimes 2}) = \frac{\mathrm{tr}(\rho^{\otimes 2}f(S_0H^{\otimes 2}))}{(1 + r)^2} \tag{5.3.23}$$

where S_0 is a constant, Hermitian matrix $H^{\otimes 2}$ is a quantum operator and $\rho^{\otimes 2}$ is a density matrix. Also, the arbitrage-free condition $\mathrm{tr}(\rho^{\otimes 2}H^{\otimes 2}) = 1 + r$ is satisfied:

$$\mathrm{tr}(\rho^{\otimes 2}H^{\otimes 2}) = \mathrm{tr}(\tilde{\rho}D_H)$$

and $\tilde{\rho}^{\otimes 2}$ is a density matrix from Lemma 5.3.2. Hence,

$$\mathrm{tr}(\tilde{\rho}^{\otimes 2}D_H^{\otimes 2}) = 1 + r$$
$$\mathrm{tr}(\tilde{\rho}^{\otimes 2}) = 1$$

and diagonal elements of the $\tilde{\rho}^{\otimes 2}$ are as $q_u^2, q_uq_d, q_dq_u, q_d^2$,

$$q_u = \frac{1+r-d}{u-d}, q_d = 1 - q_u.$$

Finally, we obtain the option price formula for the two-step model:

$$OP(f(S_2)) = \frac{f(S_0u^2)q_u^2 + f(S_0ud)q_uq_d + f(S_0du)q_dq_u + f(S_0d^2)q_d^2}{(1+r)^2} \tag{5.3.24}$$

Thus, in general (5.3.20) we obtain the same answer as for the classical case (5.3.18). The general formula of the option price in a quantum binomial market is:

$$OP(f(S_T) = f(S_0H^{\otimes n}) = \frac{\mathrm{tr}(\rho^{\otimes n}f(S_0H^{\otimes n}))}{(1+r)^n} \tag{5.3.25}$$
$$= \frac{\sum_{j=0}^{n} f(S_0u^jd^{n-j})\binom{n}{j}q_u^jq_d^{n-j}}{(1+r)^n} \tag{5.3.26}$$

In conclusion, we need to work with eigenvalues of the Hermitian operator; more specifically, we observe the eigenvalues of the operator $H^{\otimes n}$. The original density matrix ρ is irrelevant to the computation of option price for the binomial model via quantum data. We only need the transformed density matrix $\tilde{\rho}$ for computation [95]. Thus, the main job is to estimate high "u" and low "d" jumps from a set of numbers (eigenvalues of $H^{\otimes n}$).

Briefly, we need to estimate the parameters "u" ,"d" in both the classical and quantum models. The main point is availability. If we have more data, we observe H, (rather than $H^{\otimes 2}$), which is easier.

5.3.1 Estimating the Parameters ("u" and "d") for a Two-Step Quantum Binomial Model

The observable data (given) λ is real data, which is a set of eigenvalues of $H^{\otimes 2}$:

$$\lambda_i = \frac{S_i}{S_{i-1}} \tag{5.3.27}$$

where S_i is the monthly share price process for $i = 1, \dots, N$. Therefore, our dataset is:

$$\mathbf{A} = \{\lambda_1, \dots, \lambda_N\} \tag{5.3.28}$$

which comes from a family such as:

$$\mathbf{B} = \{u^j d^{n-j} : \quad j = \{0, 1, \dots, n\}\}. \tag{5.3.29}$$

In addition we have errors:

$$\lambda_i = u^j d^{n-j} + \epsilon_i \tag{5.3.30}$$

where $j = j(i)$. Now, our goal is one of identifying the parameters "u" and "d". Briefly, the data is statistically "dirty". Even without the presence of errors ϵ_i it is not a straightforward question. In general, the machine learning-type algorithm takes four main steps in order to estimate "u" and "d":

Step 1: Randomly choose a pair (u, d) such that:

$$f_j(u, d) = u^j d^{n-j}$$

where $j = \{1, \dots, n+1\}$.

Step 2: Given λ, find a minimum

$$\epsilon_j = |\lambda - f_j(u, d)| \text{ and } Arg = j$$

for which the minimum $\epsilon = \epsilon_1 \dots \epsilon_{n+1}$ is obtained. Then,

$$A_j = \{\lambda_k \colon j = Arg(\vec{\epsilon})\}.$$

Step 3: Based on Maxwell-Boltzmann statistics, estimate $\hat{p}$, (and so $\hat{q} = 1 - \hat{p}$).

Step 4: Compute u, d by the minimum value of the risk function:

$$\sum_{\lambda_k \in A_j} |\lambda_k - u^j d^{n-j}|)\hat{p}^j \hat{q}^{n-j} \binom{n}{j} \to \min_{u,d}.$$

Specifically, in this research we consider a two-step quantum binomial market.

Case n=2: It is rational to treat fortnight closing data as once observed end of the monthly closing data. In particular, we first randomly choose initial values for u and d that satisfy the following conditions:

$$\begin{cases} d \in (\sqrt{\lambda_{min}}, \sqrt{\lambda_{med}}) \\ d < u \\ u \in (\sqrt{\lambda_{med}}, \sqrt{\lambda_{max}}) \end{cases} \tag{5.3.31}$$

Then, errors are computed to identify classes of the data:

$$\epsilon_{1i} = |\lambda_i - u^2|$$
$$\epsilon_{2i} = |\lambda_i - ud|$$
$$\epsilon_{3i} = |\lambda_i - d^2|$$

If ϵ_{1i} is the minimum error λ_i goes to A_{uu} class, else if ϵ_{2i} is the minimum error λ_i goes to A_{ud} class and else ϵ_{3i} is the minimum error λ_i goes to A_{dd} class.

Second, we compute the probabilities p and $q = (1-p)$ of u and d by likelihood: At the beginning of this chapter, we review three different statistics.

In Maxwell-Boltzmann statistics, $u^j d^{n-j}$ appears with equal chances as a tensor product for each u, whilst each d appears independently with fixed probability.

$$u^j d^{n-j} \to \binom{n}{j} P_u^j P_d^{n-j}$$

where $\binom{n}{j}$ are the places to put u.

Bose-Einstein statistics ignore the placement, and each $u^j d^{n-j}$ counts once.

$$C P_u^j P_d^{n-j}$$

where C is a normaliser.

Fermi-Dirac statistics ignores also the placement. There is no p, q, and one simply puts 1/3 for each event uu, ud, dd. (Also $ud = du$.)

The Maxwell-Boltzmann statistics likelihood function is

$$L = (p^2)^{n_{uu}}(q^2)^{n_{dd}}(2pq)^{n_{ud}} \tag{5.3.32}$$

where n_{uu} is the number of elements in class A_{uu}, n_{dd} is the number of elements in class A_{dd} and n_{ud} is the number of the elements in class A_{ud}.

The Bose-Einstein statistics likelihood function is similar

$$L = C^N(p^2)^{n_{uu}}(q^2)^{n_{dd}}(pq)^{n_{ud}} \tag{5.3.33}$$

where n_{uu} is the number of elements in class A_{uu}, n_{dd} is the number of elements in class A_{dd} and n_{ud} is number of the elements in class A_{ud}. Also, $N = n_{uu} + n_{ud} + n_{dd}$ and

$$C = \frac{1}{1 - pq}$$

where $p^2 + pq + q^2 = p^2 + 2pq + q^2 - pq = 1 - pq$.

Since Fermi-Dirac statistics do not depend on p, q, the likelihood step is not present.

In the last step, we compute u and d via the probabilities computed in the second step.

$$F = p^2(\sum_{\lambda_i \in A_{uu}} |\lambda_i - u^2|) + 2pq(\sum_{\lambda_i \in A_{ud}} |\lambda_i - ud|) + q^2(\sum_{\lambda_i \in A_{dd}} |\lambda_i - d^2|) \tag{5.3.34}$$

We find the minimum value of the risk function F with minimum u and d which satisfy the conditions in (5.3.31). All the steps are repeated until convergence.

5.3.2 Results

In reality, we have always look only at available data, not the mathematical model. In this example, we consider real model is fortnight and available data is monthly share prices. Therefore, dissimilar to the other models, we consider the monthly closing share prices for the model presented in this chapter. The share prices data

Cases	BP(2009-2010)	BP(2010-2011)	BP(2011-2012)	BP(2012-2013)
OP	0.4738	0.5589	0.5062	0.4803

Table 5.1: The option prices for a quantum binomial market for all financial years

were obtained from the Internet for BP monthly closing share prices for four different financial years, namely (April to April) 2009-2010, 2010-2011, 2011-2012, 2012-2013. The real data S_n is transformed to $\lambda_i = \frac{S_i}{S_{i-1}}$. We treat the data as a set of eigenvalues of $H^{\otimes 2}$ (same operator, iid observation). Here we assume that we do not observe the closing prices at the end of each fortnight. $(H \to two \quad weeks, \quad H^{\otimes 2} \to month, \cdots)$

Then, we estimate the parameters "u" and "d" by the machine learning-type algorithm introduced in Section 5.3.1. As a result of this algorithm, the option price is computed using the estimated parameters "u" and "d" in the quantum binomial market, and which are illustrated in Table 5.1. Also, the interest rate is chosen to satisfy a no arbitrage condition. Surprisingly, a 0 interest rate gives negative option prices.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

The main purpose of this thesis was to suggest an alternative means of modelling share prices in the real market, and thereby to gain further understanding of related mathematical methodology in the financial area. We considered a number of problems in several areas which are considered substantial in certain mathematical and financial areas.

In this thesis, we modelled the BP share prices over four different financial years, (April to April) 2009-2010, 2010-2011, 2011-2012, 2012-2013, which were chosen arbitrarily from the internet. For the first and the second model, we worked with day-by-day closing share prices, which were considered after log transformation. However, for the quantum model case, we worked with BP's monthly closing share prices for the same financial years. In addition, the volatility of the share price data was estimated and analysed by the same procedure as for share prices process.

First, the logged data were modelled as an additive functional of a discrete time Markov chain perturbed by Gaussian noise. Applying the real share price data, we assumed our model was a discrete time model.

Second, though our model was discrete time, working with continuous time is easier in theory for a number of reasons. Therefore, we considered the embedding problem, which examined whether a discrete time Markov chain can be treated as a continuous time Markov chain [22]. Although the embedding problem was applied to the financial area by Israel [75], one of the goals of this research was to analyse whether the discrete time model permitted extension or embedding into the continuous time model. If the model could be converted to a continuous time model (i.e., it was indeed embeddable), this would have meant that the result (data) was observable each time. This is a plausible and important method in the financial area. If a model is a continuous time model, many existing formulae (such as option pricing) are then directly applicable to the model.

We analysed this problem by applying it to our data. This part of the research is an extensive case study on the embedding problem for the financial data and its volatility. It illustrates the importance of the embedding problem in a real financial application. As a result, in general we found we could not embed the discrete time Markov chain into the continuous time Markov chain. This means that the model we considered should be treated as a discrete time model only. Overall, this study shows that using a continuous time model for volatility is more stable than the original share prices. In addition, considering a small number of carefully chosen states is more reliable in terms of modelling.

Third, we present a new and simple, yet realistic method of modelling, analysing and considering financial data in our second model. We modelled the logged data as an additive functional of a discrete time Markov chain with a hidden random walk on the so-called lamplighter group, which are wreath products of groups. In particular, the hidden random walk was constructed on the lamplighter group $\mathbb{Z}_3$ and on the tensor product of groups $\mathbb{Z}_2 \otimes \mathbb{Z}_2$. Also, a biased random walk (as introduced in [72]) was constructed to fit the data. After comparison to different random walks with two different generator sets, for the randomly chosen datasets the α-biased random walk on the lamplighter group and $\alpha - \lambda$- biased random walk were shown to provide good fits to the data; in addition, the α-biased random walk on the tensor product of the lamplighter group and $\alpha - \lambda$- biased random walk were also shown to provide significantly better fits to the data compared with other models. The random walk on the tensor product of the lamplighter group gave a better approximation to the data than the random walk on the lamplighter group. Also, two different generators were chosen randomly for each case, and each produced similar results. Therefore, this shows the sensitivity. Two different methods (EM and Machine Learning) are used to deal with the missing data. They yield close results showing the robustness.

Furthermore, splitting data is a key method for both our first and second models. The tensor product structure comes from the splitting of the data into no jump, small jump and no big jump groups and matching it into the no small jump-small jump and no big jump-big jump groups. This then required us to deal with any missing data. Splitting data in a $(2 \times 2)(2 \times 2)$ manner helps to find the hedging to compute the option price, but the missing data is important in any such case. Therefore, the missing data and appropriate treatment methods have been reviewed. Specifically, we used the Expectation- Maximization Algorithm as the parameter estimation method and the C4.5 machine learning algorithm as the imputation method in order to treat the missing values.

Finally, we analysed the quantum data, which was subsequently used to compute the option price. In particular, we considered a binomial model. In finance theory, the binomial model is a useful and well-known model for pricing stock options. This model was used to convert the Black-Scholes-Merton model into a discrete binary tree model of prices. Quantum binomial options pricing model made the respective theories not only easier to analyse but also easier to implement on a computer.

Specifically, we considered a two-step quantum binomial model to estimate the appropriate parameters. We realized that the original density matrix ρ is irrelevant to the computation of option price using a binomial model via quantum data. We only needed to transform the density matrix $\tilde{\rho}$ for computation. Thus, the main requirement was to estimate u and d from $H^{\otimes n}$. Therefore, we estimated these parameters using the binomial quantum model, and subsequently computed option prices based on these parameters in a quantum binomial market.

6.2 Future Work

In theory, working with continuous time is easier than discrete time. The embedding problem plays a significant role in finance. Whilst this thesis contains an extensive case study on the embedding problem regarding financial data and its volatility, more research needs to be done for the embedding problem in the area of finance.

In case 2, for the random search in the embedding problem in Section 3.3.4 in Chapter 3, all steps are repeated 10^5 times in order to find the minimum value of the parameter δ, which was found to be 0.2. The aim is to find practical $\delta = \delta_n$ (such as a confidence interval) such as that for the observed error $\delta(\hat{P}) >> \delta$. We say that the matrix $\hat{P}$ is not embeddable (observe $\delta = 0.2$). We believe this result can be proved theoretically, but at this stage, we do not know how.

Further, the statistical distribution chosen for computation was the Maxwell-Boltzmann classical statistical distribution. The influence of statistical methods such as those of Maxwell-Boltzmann or Bose-Einstein statistics on parameter estimation in quantum binomial market is an interesting, and open, question.

In this research, we use entropy in a machine learning algorithm to treat the missing data. Additionally, application of entropy should be considered for data analysis. Due to the entropy property $H_{XY} = H_X + H_Y$ for independent variables, entropy could be used as an alternative metric for data fitting problems.

Theoretical results for the distance between two stochastic random matrices is another interesting open question.

In addition, the results are similar for share price and its volatility. Although the method works well in many situations, sometimes results are not satisfactory such for as BP (2009-2010). This poses the question as to whether a better approach can be found, e.g., finding a better generator.

Bibliography

[1] A. D. Wilkie (1984), *A Stochastic Investment Model for Actuarial Use*, Transactions of the Faculty of Actuaries, **39**, pp 341–403.

[2] A. P. Dempster, N. M. Laird and D. B. Rubin (1977), *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the royal statistical society. Series B (methodological), **39**, pp 1–38.

[3] B. Belkin, S. Suchower and L. Forest Jr (1998), *A one-parameter representation of credit risk and transition matrices*, Credit Metrics monitor, **1(3)**, pp 46–56.

[4] B. Bollobas (2013), *Modern graph theory*, Springer Science and Business Media, **(Vol. 184)**.

[5] B. Singer and S. Spilerman (1976), *The representation of social processes by Markov models*, American Journal of Sociology, **82**, pp 1–54.

[6] B. E. Baaquie (2013), *The theoretical foundations of quantum mechanics*, Springer Science and Business Media.

[7] B. E. Baaquie, C. Coriano and M. Srikant (2004), *Hamiltonian and potentials in derivative pricing models: exact results and lattice simulations.* Physica A: Statistical Mechanics and its Applications, **334(3)**, pp 531–557.

[8] B. Fuglede (1988), *On the imbedding problem for stochastic and doubly stochastic matrices*, Probability theory and related fields, **80(2)**, pp 241–260.

[9] BP at a glance (2016), http://www.bp.com/en/global/corporate/about-bp/bp-at-a-glance.html.

[10] C. W. Granger (1983), *Co-integrated Variables and Error-Correcting Model*, Unpublished discussion paper, pp 83–13, University of California, San Diego.

[11] D. Lando (2000), *Some elements of rating-based credit risk modeling*, Advanced Fixed-Income Valuation Tools, pp 193–215.

[12] D. A. Levin, Y. Peres and E. L. Wilmer (2009), *Markov chains and mixing times*. American Mathematical Soc.

[13] D. W. Aha, D. Kibler and M. K. Albert (1991), *Instance-based learning algorithms*, Machine learning, **6(1)**, pp 37–66.

[14] E. B. Davies (2010), *Embeddable Markov matrices*, Electronic Journal of Probability, **15**, pp 1474–1486.

[15] E. E. Haven (2002), *A discussion on embedding the BlackScholes option pricing model in a quantum physics setting*, Physica A: Statistical Mechanics and its Applications, **304(3)**, pp 507–524.

[16] F. Black and M. Scholes (1973), *The pricing of options and corporate liabilities*, The journal of political economy, pp 637–654.

[17] F. Celler, C. R. Leedham-Green, S. H. Murray, A. C. Niemeyer and E. A. O'brien (1995), *Generating random elements of a finite group*, Communications in algebra, **23(13)**, pp 4931–4948.

[18] F. Fessant and S. Midenet (2002), *Self-organising map for data imputation and correction in surveys*, Neural Computing and Applications, **10(4)**, pp 300–310.

[19] F. Klaassen (2002), *Improving GARCH volatility forecasts with regime-switching GARCH. In Advances in Markov-Switching Models*, Physica-Verlag HD, pp 223–254.

[20] F. R. Bach (2008), *Consistency of trace norm minimization*, Journal of Machine Learning Research, **9**, pp 1019–1048.

[21] F. R. Chung and R. L. Graham (1997), *Random walks on generating sets for finite groups*, The electronic journal of combinatorics, **4(2)**, pp 14–28.

[22] G. Elfving (1937), *Zur Theorie der Markoffschen Ketten*, Acta societatis Scientiarum Fennicae, **2(8)**.

[23] G. E. Batista and M. C. Monard (2002), *A Study of K-Nearest Neighbour as an Imputation Method*, HIS, **87**, pp 251–260.

[24] G. S. Goodman (1970), *An intrinsic time for non-stationary finite Markov chains*. Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete, **16(3)**, pp 165–180.

[25] H. Kesten (1959), *Symmetric random walks on groups*, Transactions of the American Mathematical Society, **92(2)**, pp 336-354.

[26] H. J. Adr and M. Adr (2008), *Advising on research methods: A consultant's companion*, Johannes van Kessel Publishing, pp 305-332.

[27] J. Hhn and E. Hllermeier (2009) *FURIA: an algorithm for unordered fuzzy rule induction*, Data Mining and Knowledge Discovery, **19(3)**, pp 293–319.

[28] J. Hull (1993), *The pricing of options on interest rate caps and floors using the Hull-White model*, Journal of Financial Engineering, **2(3)**, pp 287–296.

[29] J. Meier (2008), *Groups, graphs and trees: an introduction to the geometry of infinite groups*, Cambridge university press.

[30] J. C. Cox, S. A. Ross and M. Rubinstein (1979), *Option pricing: A simplified approach*, Journal of financial Economics, **7(3)**, pp 229–263.

[31] J. C. Cox, J. E. Ingersoll Jr, S.A. Ross (1985), *A theory of the term structure of interest rates*, Econometrica: Journal of the Econometric Society, pp 385–407.

[32] J. E. Walsh (1962), *Handbook of Nonparametric Statistics*, New York: D.V. Nostrand.

[33] J. F. C. Kingman (1962), *The imbedding problem for finite Markov chains*, Probability Theory and Related Fields, **1(1)**, pp 14–24.

[34] J. Janssen, R. Manca and G. Di Biase (1997), *Markov and semiMarkov option pricing models with arbitrage possibility*, Applied Stochastic Models and Data Analysis, **13(2)**, pp 103–113.

[35] J. M. Harrison and S. R. Pliska (1981), *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic processes and their applications, **11(3)**, pp 215–260.

[36] J. M. Jerez, I. Molina, P. J. Garca-Laencina, E. Alba, N. Ribelles, M. Martn and L. Franco (2010), *Missing data imputation using statistical and machine learning methods in a real breast cancer problem*, Artificial intelligence in medicine, **50(2)**, pp 105–115.

[37] J. M. Speakman (1967), *Two Markov chains with a common skeleton*, Probability Theory and Related Fields, **7(3)**, pp 224–224.

[38] J. R. Cuthbert (1972), *On uniqueness of the logarithm for Markov semi-groups*, Journal of the London Mathematical Society, **2(4)**, pp 623–630.

[39] J. R. Cuthbert (1973), *The logarithm function for finite-state Markov semi-groups*, Journal of the London Mathematical Society, **2(3)**, pp 524–532.

[40] J. R. Quinlan (2014), *C4. 5: programs for machine learning*, Elsevier.

[41] J. T. Runnenburg (1962), *On Elfving's problem of imbedding a time-discrete Markov chain in a time-continuous one for finitely many states I.* Proceedings of the KNAW-Series A, Mathematical Sciences, **65(5)**, pp 536–548.

[42] J. W. Dash (2004), *Quantitative finance and risk management.* World Scientific.

[43] A. Khrennikov (2007), *Classical and quantum randomness and the financial market*, arXiv preprint arXiv:0704.2865.

[44] K. Lakshminarayan, S. A. Harp, R. P. Goldman and T. Samad (1996), *Imputation of Missing Data Using Machine Learning Techniques*, In KDD, pp 140–145.

[45] K. Ilinski (2001), *Physics of finance: Gauge modelling in non-equilibrium pricing*, Wiley.

[46] R. J. Little and D. B. Rubin (2014), *Statistical analysis with missing data*, John Wiley and Sons.

[47] K. L. Chung (1967), *Markov Chains with Stationary Transition Probabilities*, 2d Ed. Springer.

[48] K. R. Parthasarathy (2012), *An introduction to quantum stochastic calculus*, Springer Science and Business Media.

[49] L. Accardi and A. Boukas (2007), *The quantum Black-Scholes equation*, arXiv preprint arXiv:0706.1300.

[50] M. Bladt and M. Srensen (2005), *Statistical inference for discretely observed Markov jump processes*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), **67(3)**, pp 395–410.

[51] M. Fazel, H. Hindi and S. P. Boyd (2001), *A rank minimization heuristic with application to minimum order system approximation*, In American Control Conference, 2001. Proceedings of the 2001, **6**, pp 4734–4739.

[52] M. Kijima (1998), *Monotonicities in a Markov chain model for valuing corporate bonds subject to credit risk*, Mathematical Finance, **8(3)**, pp 229–247.

[53] M. Kijima and K. Komoribayashi (1998), *A Markov chain model for valuing credit risk derivatives*, The Journal of Derivatives, **6(1)**, pp 97–108.

[54] M. M. Rahman and D. N. Davis (2013), *Machine learning-based missing value imputation method for clinical datasets*, In IAENG Transactions on Engineering Technologies, Springer Netherlands, pp 245–257.

[55] M. R. Gupta and Y. Chen (2011), *Theory and use of the EM algorithm*, Now Publishers Inc.

[56] N. Srebro and A. Shraibman (2005), *Rank, trace-norm and max-norm*, In International Conference on Computational Learning Theory, Springer Berlin Heidelberg, pp 545–560.

[57] N. Srebro, J. Rennie and T. S. Jaakkola (2004), *Maximum-margin matrix factorization*, In Advances in neural information processing systems, pp 1329–1336.

[58] N. Varopoulos (1983), *Random walks on soluble groups*, Bull. Sci. Math., **107**, pp 337-344.

[59] N. J. Higham (2008), *Functions of matrices: theory and computation*, Siam.

[60] N. Sahand and M. J. Wainwright (2011), *Estimation of (near) low-rank matrices with noise and high-dimensional scaling*, The Annals of Statistics, pp 1069-1097.

[61] O. Hggstr and J. Jonasson (1997), *Rates of convergence for lamplighter processes*, Stochastic Processes and their Applications, **67(2)**, pp 227–249.

[62] O. Maimon and L. Rokach (2005), *Data mining and knowledge discovery handbook*, New York: Springer.

[63] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani and R. B. Altman (2001), *Missing value estimation methods for DNA microarrays*, Bioinformatics, **17(6)**, pp 520–525.

[64] P. Carette (1995), *Characterizations of Embeddable* 3×3 *Stochastic Matrices with a Negative Eigenvalue*, New York J. Math, **1**, pp 120–129.

[65] P. Cheeseman and J. Stutz (1996), *Bayesian Classification (Autoclass): Theory And Results In: Advances in Knowledge Discovery and Data Mining*, California: AAAIPress.

[66] P. Piela (2002), *Introduction to self-organizing maps modelling for imputation-techniques and technology*, Research in Official Statistics, **2**, pp 5–19.

[67] G. E. Batista and M. C. Monard (2003), *An analysis of four missing data treatment methods for supervised learning*, Applied Artificial Intelligence, **17(5-6)**, pp 519–533.

[68] P.J. Cameron (2013),*Notes on finite group theory.*

[69] P. J. Garca-Laencina, J. L. Sancho-Gmez, A. R. Figueiras-Vidal and M. Verleysen (2009), *K nearest neighbours with mutual information for simultaneous classification and missing data imputation*, Neurocomputing, **72(7)**, pp 1483–1493.

[70] P. K. Sharpe and R. J. Solly (1995), *Dealing with missing values in neural network-based diagnostic systems*, Neural Computing and Applications, **3(2)**, pp 73–77.

[71] P. Wilmott (2007), *Paul Wilmott introduces quantitative finance*, John Wiley and Sons.

[72] R. Lyons, R. Pemantle and Y. Peres (1996), *Random walks on the lamplighter group*, The Annals of Probability, **24**, pp 1993–2006.

[73] R. A. Jarrow, D. Lando and S. M. Turnbull (1997), *A Markov model for the term structure of credit risk spreads*, Review of Financial studies, **10(2)**, pp 481–523.

[74] R. A.Horn and C. R. Johnson (1991), *Topics in matrix analysis*, Cambridge UP, New York.

[75] R.B. Israel, J.S. Rosenthal and J.Z. Wei (2001), *Finding generators for Markov chains via empirical transition matrices, with applications to credit ratings*, Mathematical Finance, **11(2)**, pp 245–265.

[76] R. C. Barros, M. P. Basgalupp, A. C. De Carvalho and A. A. Freitas (2012), *A survey of evolutionary algorithms for decision-tree induction*, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), **42(3)**, pp 291–312.

[77] R. C. Merton (1971), *Optimum consumption and portfolio rules in a continuous-time model*, Journal of Economic Theory, **3**, pp 373–413.

[78] R. F. Engle and C. W. Granger (1987), *Co-integration and error correction: representation, estimation, and testing*, Econometrica: journal of the Econometric Society, pp 251–276.

[79] R. K. Niven and M. Grendar (2008), *Generalized Maxwell-Boltzmann, Bose-Einstein, Fermi-Dirac and Acharya-Swamy Statistics and the Polya Urn Model*, arXiv preprint arXiv:0808.2102.

[80] R. P. Feynman, A. R. Hibbs and D. F. Styer (2005), *Quantum mechanics and path integrals*, Courier Corporation.

[81] R.S. Karadeniz and S. Utev, *Modelling share prices via the random walk on the Lamplighter group*, (under revision).

[82] R.S. Karadeniz and S. Utev, *Embedding Problem for Financial Data*, (under revision).

[83] S. Boyd and L. Vandenberghe (2004), *Convex optimization*, Cambridge university press.

[84] S. I. Boyarchenko and S. Z. Levendorski (2008),*Pricing American options in regime-switching models*, FFT Realization.

[85] S. Johansen (1974), *Some results on the imbedding problem for finite Markov chains*, Journal of the London Mathematical Society, **2(2)**, pp 345–351.

[86] S. Lang (1993),*Algebra Addison*, Third edition, Addison-Wesley, New York.

[87] S. P. Whitten and R. G. Thomas (1999), *A non-linear stochastic asset model for actuarial use*, British Actuarial Journal, **5(05)**, pp 919–953.

[88] S. F. Messner (1992), *Exploring the Consequences of Erratic Data Reporting for Cross-National Research on Homicide*, Journal of Quantitative Criminology, **8 (2)**, pp 155-173.

[89] S. G. Kou (2002), *A Jump-Diffusion Model for Option Pricing*, Management Science, **48**, pp 1086–1101.

[90] T. S. Cubitt, J. Eisert and M. M. Wolf (2012), *Extracting dynamical equations from experimental data is NP hard*, Physical review letters, **108(12)**.

[91] T. Orchard, M. A. Woodbury (1972), *A missing information principle: theory and applications*. Sixth Berkeley Symp. on Math. Statist. and Prob.,**Vol.1**, Univ. of California Press, pp 697–715.

[92] T. Samad, S. A. Harp (1992), *Self-organization with partial data*, Network Computation in Neural Systems, **3(2)**, pp 205-12.

[93] T. E. Cooley and E. C. Prescott (1973), *The adaptive regression model*, International economic review, **15**, pp 35–371.

[94] W. Culver (1996), *On the existence and uniqueness of the real logarithm of a matrix*, Proc. Amer. Math. Soc, **17**, pp 1146–1151.

[95] W. Hao and S. Utev, *Quantum mechanics approach to Option pricing*, (working paper).

[96] W. Woess (2000), *Random walks on infinite graphs and groups*, Cambridge university press, **138**.

[97] W. J. Conover (1980), *Practical Nonparametric Statistics*, New York: Wiley and Sons.

[98] X. Ma (2012), *Stochastic modelling in finance*, PhD dissertation, University of Notingham.

[99] X. Ma and S. Utev (2012), *Modelling the share prices as a hidden random walk on the lamplighter group*, In Mathematical and Statistical Methods for Actuarial Sciences and Finance, pp 263–270.

[100] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, and Z. H. Zhou (2008), *Top 10 algorithms in data mining*, Knowledge and information systems, **14(1)**, pp 1–37.

[101] Y. Guedon (2003), *Estimating hidden semi-Markov chains from discrete sequences.* Journal of computational and graphical statistics, **12**, pp 604–639.

[102] Z. Chen (2004), *Quantum theory for the binomial model in finance theory,* Journal of Systems Science and Complexity, **17(4)**.

Appendix A

Dataset

In the following figures, we illustrate the original data in this research are used to be analysed. We choose the BP daily share price data for four different financial years. The main point of this research is analysing of the each financial year share price data.

Figure A.1: BP Share Prices Chart between April 2009 and April 2010

Figure A.2: BP Share Prices Chart between April 2010 and April 2011

Figure A.3: BP Share Prices Chart between April 2011 and April 2012

Appendix B

Tables and Figures: Embedding Problem

Cases	$\hat{P}$
BP(2009-2010)	$\begin{pmatrix} 0.1392 & 0.8009 & 0.0598 \\ 0.1499 & 0.8168 & 0.0333 \\ 0.0333 & 0.9333 & 0.0333 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.2860 & 0.6650 & 0.0491 \\ 0.2097 & 0.7569 & 0.0333 \\ 0.9333 & 0.0333 & 0.0333 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.0655 & 0.9012 & 0.0333 \\ 0.1378 & 0.8249 & 0.0374 \\ 0.0333 & 0.9333 & 0.0333 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1205 & 0.7882 & 0.0914 \\ 0.1413 & 0.8253 & 0.0333 \\ 0.0333 & 0.9333 & 0.0333 \end{pmatrix}$

Table B.1: 3-by-3 pertubated transition matrices for volatilities

Q_X matrices $(P_X viaEM)$	Q_X matrices $(P_X viaC4.5)$
$\begin{pmatrix} -4.7935 & 4.7935 \\ 0.8059 & -0.8059 \end{pmatrix}$	$\begin{pmatrix} -2.6998+2.7272i & 2.6998-2.7272i \\ 0.4103-0.4144i & 0.4103+0.4144i \end{pmatrix}$
$\begin{pmatrix} -1.8522 & 1.8522 \\ 0.5425 & -0.5425 \end{pmatrix}$	$\begin{pmatrix} -2.4439 & 2.4439 \\ 0.5680 & -0.5680 \end{pmatrix}$
$\begin{pmatrix} -2.3475+2.7629i & 2.3475-2.7629i \\ 0.3218-0.3787i & -0.3218+0.3787i \end{pmatrix}$	$\begin{pmatrix} -5.3095 & 5.3095 \\ 0.7654 & -0.7654 \end{pmatrix}$
$\begin{pmatrix} -2.9885+2.7127i & 2.9885-2.7127i \\ 0.472-0.4289i & -0.4725+0.4289i \end{pmatrix}$	$\begin{pmatrix} -2.5729 & 2.5729 \\ 0.1437 & -0.1437 \end{pmatrix}$

Table B.2: Q_X matrices via EM and C4.5 algorithm for volatilities

Cases	Q_Y
BP(2009-2010)	$\begin{pmatrix} -0.1067+0.0974i & 0.1067-0.0974i \\ 3.3353-3.0442i & -3.3353+3.0442i \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} -0.0648+0.0492i & 0.0648-0.0492i \\ 4.0766-3.0924i & -4.0766+3.0924i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} -0.0593 & 0.0593 \\ 2.0744 & -2.0744 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} -0.1248+0.1223i & 0.1248-0.1223i \\ 3.0816-3.0193i & -3.0816+3.0193i \end{pmatrix}$

Table B.3: Q_Y for volatilities

Cases	P_X	Q matrices
BP(2009-2010)	$\begin{pmatrix} 0.2844 & 0.7156 \\ 0.2357 & 0.7463 \end{pmatrix}$	$\begin{pmatrix} -2.5733 & 2.5733 \\ 0.9124 & -0.9124 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.5454 & 0.4546 \\ 0.5420 & 0.4580 \end{pmatrix}$	$\begin{pmatrix} -2.5977 & 2.5977 \\ 3.0971 & -3.0971 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.2766 & 0.7234 \\ 0.2443 & 0.7557 \end{pmatrix}$	$\begin{pmatrix} -2.5651 & 2.5651 \\ 0.8663 & -0.8663 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.3175 & 0.6825 \\ 0.2400 & 0.7600 \end{pmatrix}$	$\begin{pmatrix} -1.8920 & 1.8920 \\ 0.6652 & -0.6652 \end{pmatrix}$

Table B.4: Perturbated transition matrices P_X via EM and their Q matrices

Cases	P_X	Q matrices
BP(2009-2010)	$\begin{pmatrix} 0.2531 & 0.7469 \\ 0.3689 & 0.6311 \end{pmatrix}$	$\begin{pmatrix} -1.4431+2.1028i & 1.4431-2.1028i \\ 0.7129-1.0388i & -0.7129+1.0388i \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.3807 & 0.6193 \\ 0.6187 & 0.3813 \end{pmatrix}$	$\begin{pmatrix} -0.7181+1.5716i & 0.7181-1.5716i \\ 0.7173-1.5700i & -0.7173+1.5700i \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.1083 & 0.8917 \\ 0.3232 & 0.6768 \end{pmatrix}$	$\begin{pmatrix} -1.1286+2.3058i & 1.1286-2.3058i \\ 0.4091-0.8358i & -0.4091+0.8358i \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1865 & 0.8135 \\ 0.3165 & 0.6835 \end{pmatrix}$	$\begin{pmatrix} -1.4692+2.2617i & 1.4692-2.2617i \\ 0.5716-0.8799i & -0.5716+0.8799i \end{pmatrix}$

Table B.5: Perturbated transition matrices P_X via C4.5 and their Q matrices

Cases	P_Y	Q matrices
BP(2009-2010)	$\begin{pmatrix} 0.8821 & 0.1179 \\ 0.8854 & 0.1416 \end{pmatrix}$	$\begin{pmatrix} -0.4520 & 0.4520 \\ 3.2901 & -3.2901 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.8664 & 0.1356 \\ 0.6583 & 0.3417 \end{pmatrix}$	$\begin{pmatrix} -0.2697 & 0.2697 \\ 1.3095 & -1.3095 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.7230 & 0.2770 \\ 0.6418 & 0.3582 \end{pmatrix}$	$\begin{pmatrix} -0.7568 & 0.7568 \\ 1.7538 & -1.7538 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.8775 & 0.1225 \\ 0.8613 & 0.1387 \end{pmatrix}$	$\begin{pmatrix} -0.5139 & 0.5139 \\ 3.6124 & -3.6124 \end{pmatrix}$

Table B.6: Perturbated transition matrices P_Y and their Q matrices

Cases	P_X	Q matrices
BP(2009-2010)	$\begin{pmatrix} 0.1991 & 0.8009 \\ 0.1624 & 0.8376 \end{pmatrix}$	$\begin{pmatrix} -2.7487 & 2.7487 \\ 0.5573 & -0.5573 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.3341 & 0.6659 \\ 0.2186 & 0.7814 \end{pmatrix}$	$\begin{pmatrix} -1.6255 & 1.6255 \\ 0.5337 & -0.5337 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.1203 & 0.8797 \\ 0.1493 & 0.8507 \end{pmatrix}$	$\begin{pmatrix} -3.0256 + 2.6857i & 3.0256 - 2.6857i \\ 0.5136 - 0.4559i & -0.5136 + 0.4559i \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1651 & 0.8349 \\ 0.1601 & 0.8399 \end{pmatrix}$	$\begin{pmatrix} -4.4338 & 4.4338 \\ 0.8500 & -0.8500 \end{pmatrix}$

Table B.7: Perturbated transition matrices P_X via EM and their Q matrices for volatilities

Cases	P_X	Q matrices
BP(2009-2010)	$\begin{pmatrix} 0.1505 & 0.8495 \\ 0.1574 & 0.8426 \end{pmatrix}$	$\begin{pmatrix} -4.2100 + 2.6506i & 4.2100 - 2.6506i \\ 0.7799 - 0.4910i & -0.7799 + 0.4910i \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.2723 & 0.7277 \\ 0.1947 & 0.8053 \end{pmatrix}$	$\begin{pmatrix} -2.0165 & 2.0165 \\ 0.5395 & -0.5395 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.1819 & 0.8181 \\ 0.1465 & 0.8535 \end{pmatrix}$	$\begin{pmatrix} -2.8337 & 2.8337 \\ 0.5073 & -0.5073 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1706 & 0.8294 \\ 0.0778 & 0.9222 \end{pmatrix}$	$\begin{pmatrix} -2.1732 & 2.1732 \\ 0.2038 & -0.2038 \end{pmatrix}$

Table B.8: Perturbated transition matrices P_X via C4.5 and their Q matrices for volatilities

Cases	P_Y	Q matrices
BP(2009-2010)	$\begin{pmatrix} 0.9379 & 0.0621 \\ 0.9333 & 0.0667 \end{pmatrix}$	$\begin{pmatrix} -0.3368 & 0.3368 \\ 5.0595 & -5.0595 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.9524 & 0.0554 \\ 0.9333 & 0.0667 \end{pmatrix}$	$\begin{pmatrix} -0.1924 & 0.1924 \\ 3.7697 & -3.7697 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.9446 & 0.0554 \\ 0.8047 & 0.1953 \end{pmatrix}$	$\begin{pmatrix} -0.1266 & 0.1266 \\ 1.8402 & -1.8402 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.9302 & 0.0698 \\ 0.9333 & 0.0667 \end{pmatrix}$	$\begin{pmatrix} -0.4015+0.2185i & 0.4015-0.2186i \\ 5.3695-2.9230i & -5.3695+2.9230i \end{pmatrix}$

Table B.9: Perturbated transition matrices P_Y and their Q matrices for volatilities

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.1071 & 0.5714 & 0.3214 \\ 0.0957 & 0.4522 & 0.4522 \\ 0.1308 & 0.4486 & 0.4206 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.1500 & 0.3500 & 0.5000 \\ 0.1351 & 0.4775 & 0.3874 \\ 0.1919 & 0.4545 & 0.3535 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.0857 & 0.4286 & 0.4857 \\ 0.1810 & 0.4286 & 0.3905 \\ 0.1091 & 0.4091 & 0.4818 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.1220 & 0.3171 & 0.5610 \\ 0.1961 & 0.3824 & 0.4216 \\ 0.1619 & 0.4667 & 0.3714 \end{pmatrix}$

Table B.10: 3-by-3 transition matrices for all financial years of Barclays

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.3146 & 4.0294 & -1.7154 \\ -0.1590 & -1.6378 & 1.7970 \\ 0.7844 & 0.7290 & -1.5134 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.0881+1.3535i & 1.2544+1.2192i & 0.8337-2.5729i \\ 0.4331+0.2407i & -1.4201+0.2168i & 0.9871-0.4576i \\ 0.3557-0.8266i & 1.1119-0.7446i & -1.4678+1.5713i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -2.6648+3.1303i & 1.3120-0.9453i & 1.3527-2.1849i \\ 0.3807-1.8384i & -1.8342+0.5552i & 1.4538+1.2831i \\ 0.4568+0.7792i & 1.3327-0.2353i & -1.7896-0.5439i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.8272 & -2.9488 & 4.7765 \\ 0.9646 & -1.0017 & 0.0372 \\ -0.1975 & 2.1382 & -1.9408 \end{pmatrix}$

Table B.11: Q matrices of the 3-by-3 transition matrices for all financial years of Barclays

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.2222 & 0.4074 & 0.3704 \\ 0.2000 & 0.4667 & 0.3333 \\ 0.2308 & 0.3846 & 0.3846 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.1739 & 0.4130 & 0.4130 \\ 0.1560 & 0.4495 & 0.3945 \\ 0.2105 & 0.4316 & 0.3579 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2326 & 0.4186 & 0.3488 \\ 0.1000 & 0.4800 & 0.4200 \\ 0.2056 & 0.3178 & 0.4766 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.3019 & 0.3774 & 0.3208 \\ 0.1895 & 0.3368 & 0.4737 \\ 0.2000 & 0.4300 & 0.3700 \end{pmatrix}$

Table B.12: 3-by-3 transition matrices for all financial years of Burberry

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -7.1081+2.1916i & 2.5123-0.6083i & 4.5958-1.5834i \\ 0.5638-0.0002i & -1.5024+0.0001i & 0.9386+0.0001i \\ 3.6001-1.3148i & 0.2647+0.3649i & -3.8648+0.9499i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.6302+1.2341i & 1.5415+0.3314i & 1.0885-1.5654i \\ 0.7162+0.6859i & -1.9051+0.1842i & 1.1890-0.8700i \\ 0.4209-1.3586i & 1.4411-0.3648i & -1.8619+1.7233i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.8708 & 1.3414 & 0.5294 \\ -0.1078 & -1.0512 & 1.1589 \\ 0.8260 & 0.4502 & -1.2761 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.7636+0.0287i & 0.8790+0.5417i & 0.8848-0.5704i \\ 0.4899-0.0919i & -1.4406-1.7321i & 0.9506+1.8240i \\ 0.4945+0.0725i & 0.9016+1.3658i & -1.3962-1.4383i \end{pmatrix}$

Table B.13: Q matrices of the 3-by-3 transition matrices for all financial years of Burberry

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.1212 & 0.5152 & 0.3636 \\ 0.1204 & 0.4444 & 0.4352 \\ 0.1468 & 0.4037 & 0.4495 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2083 & 0.4167 & 0.3750 \\ 0.2062 & 0.3299 & 0.4639 \\ 0.1714 & 0.4286 & 0.4000 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.1591 & 0.3409 & 0.5000 \\ 0.2212 & 0.4231 & 0.3558 \\ 0.1373 & 0.4314 & 0.4314 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.3108 & 0.3649 & 0.3243 \\ 0.3103 & 0.3103 & 0.3793 \\ 0.2874 & 0.3678 & 0.3448 \end{pmatrix}$

Table B.14: 3-by-3 transition matrices for all financial years of Easyjet

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -3.2807 & 4.0439 & -0.7632 \\ 0.0713 & -1.6367 & 1.5654 \\ 0.9299 & 0.4177 & -1.3476 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.6493+0.2666i & 0.9786-1.0149i & 1.6707+0.7483i \\ 0.4840-0.5027i & -1.4398+1.9136i & 0.9558-1.4109i \\ 0.7639+0.3425i & 0.8829-1.3039i & -1.6468+0.9614i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -2.4070 & -0.1783 & 2.5852 \\ 1.2428 & -1.3195 & 0.0770 \\ -0.2163 & 1.3949 & -1.1785 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -2.5589-0.2926i & 1.0164+1.0213i & 1.5425-0.7286i \\ 0.9755+0.5839i & -1.8789-2.0383i & 0.9032+1.4542i \\ 1.2419-0.3255i & 0.9827+1.1363i & -2.2246-0.8107i \end{pmatrix}$

Table B.15: Q matrices of the 3-by-3 transition matrices for all financial years of Easyjet

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.1111 & 0.4722 & 0.4167 \\ 0.1513 & 0.4874 & 0.3613 \\ 0.1474 & 0.4737 & 0.3789 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2838 & 0.3649 & 0.3514 \\ 0.3118 & 0.3548 & 0.3333 \\ 0.2892 & 0.3855 & 0.3253 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2807 & 0.3860 & 0.3333 \\ 0.1705 & 0.2955 & 0.5341 \\ 0.2476 & 0.3714 & 0.3810 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.3289 & 0.3289 & 0.3421 \\ 0.2976 & 0.2619 & 0.4405 \\ 0.2955 & 0.4205 & 0.2841 \end{pmatrix}$

Table B.16: 3-by-3 transition matrices for all financial years of Hsbc

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.8022+2.8136i & 2.0901+0.0846i & 0.7121-2.8982i \\ 0.4379-0.5816i & -2.2407-0.0175i & 1.8028+0.5991i \\ 0.5146-0.3354i & 2.0606-0.0101i & -2.5752+0.3455i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.5177 & 0.0200 & 2.4980 \\ 2.5169 & -2.5789 & 0.0615 \\ -0.5401 & 2.8080 & -2.2676 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -3.4170+1.3078i & 0.7606+1.6531i & 2.6564-2.9611i \\ 1.5499-1.9338i & -1.6492-2.4444i & 0.0995+4.3785i \\ 0.5675+0.8856i & 0.9465+1.1194i & -1.5141-2.0050i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -2.3778+0.0002i & 1.1510-0.0201i & 1.2264+0.0198i \\ 1.0415-0.0181i & -1.4498+1.6249i & 0.4083-1.6067i \\ 1.0593+0.0171i & 0.3898-1.5337i & -1.4489+1.5165i \end{pmatrix}$

Table B.17: Q matrices of the 3-by-3 transition matrices for all financial years of Hsbc

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.1765 & 0.4412 & 0.3824 \\ 0.1311 & 0.5000 & 0.3689 \\ 0.1277 & 0.5000 & 0.3723 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2083 & 0.3333 & 0.4583 \\ 0.2308 & 0.3750 & 0.3942 \\ 0.1531 & 0.5000 & 0.3469 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2500 & 0.2778 & 0.4722 \\ 0.1250 & 0.4519 & 0.4231 \\ 0.1273 & 0.4273 & 0.4455 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.0952 & 0.5476 & 0.3571 \\ 0.1782 & 0.3861 & 0.4356 \\ 0.2000 & 0.3619 & 0.4381 \end{pmatrix}$

Table B.18: 3-by-3 transition matrices for all financial years of ITV

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.5653 & -0.1346 & 2.7003 \\ 0.4879 & -2.4152 & 1.9273 \\ 0.2927 & 3.2435 & -3.5362 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -0.8765 & -1.7673 & 2.6437 \\ 1.3140 & -2.0785 & 0.7646 \\ -0.9625 & 3.1155 & -2.1532 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.8054 & -0.5943 & 2.3996 \\ 0.2838 & -1.8498 & 1.5660 \\ 0.3226 & 1.9434 & -2.2658 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.5531 - 13.0049i & -1.9339 + 37.2338i & 3.4866 - 24.2266i \\ 0.6086 - 0.5316i & -2.0978 + 1.8432i & 1.4890 - 1.3114i \\ 0.0544 + 5.8271i & 2.7933 - 16.9899i & -2.8475 + 11.1617i \end{pmatrix}$

Table B.19: Q matrices of the 3-by-3 transition matrices for all financial years of ITV

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.2245 & 0.3265 & 0.4490 \\ 0.2396 & 0.3958 & 0.3646 \\ 0.1429 & 0.4095 & 0.4476 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2353 & 0.3922 & 0.3725 \\ 0.1373 & 0.4216 & 0.4412 \\ 0.2577 & 0.4124 & 0.3299 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2222 & 0.2963 & 0.4815 \\ 0.1954 & 0.2989 & 0.5057 \\ 0.2294 & 0.4128 & 0.3578 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.2656 & 0.3594 & 0.3750 \\ 0.2826 & 0.3587 & 0.3587 \\ 0.2283 & 0.3913 & 0.3804 \end{pmatrix}$

Table B.20: 3-by-3 transition matrices for all financial years of Marks and Spencer

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -1.9870 & 0.2340 & 1.7530 \\ 1.2440 & -1.7176 & 0.4736 \\ -0.2218 & 1.4922 & -1.2704 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.4567-0.1271i & 1.2794-0.0380i & 1.1772+0.1652i \\ 0.7721-1.7305i & -1.7787-0.5174i & 1.0068+2.2480i \\ 0.4731+1.9223i & 1.2293+0.5747i & -1.7024-2.4971i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -2.8686+0.2099i & 1.8293+0.8388i & 1.0393-1.0487i \\ 1.0051+0.3021i & -1.7696+1.2073i & 0.7644-1.5094i \\ 0.6190-0.3452i & 0.5061-1.3792i & -1.1251+1.7244i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -2.4322 & 0.7091 & 1.7230 \\ 2.6087 & -3.2734 & 0.6647 \\ -0.9168 & 2.7802 & -1.8635 \end{pmatrix}$

Table B.21: Q matrices of the 3-by-3 transition matrices for all financial years of Marks and Spencer

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.3735 & 0.2651 & 0.3614 \\ 0.3165 & 0.2405 & 0.4430 \\ 0.3256 & 0.4419 & 0.2326 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.3125 & 0.3500 & 0.3375 \\ 0.2857 & 0.3571 & 0.3571 \\ 0.3721 & 0.3023 & 0.3256 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2903 & 0.3548 & 0.3548 \\ 0.2308 & 0.3407 & 0.4286 \\ 0.2371 & 0.4021 & 0.3608 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.2632 & 0.4386 & 0.2982 \\ 0.1771 & 0.3542 & 0.4688 \\ 0.2577 & 0.3918 & 0.3505 \end{pmatrix}$

Table B.22: 3-by-3 transition matrices for all financial years of Next

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -1.9206+0.0152i & 1.1094+0.4271i & 0.8111-0.4423i \\ 1.0088+0.0487i & -1.4014+1.3727i & 0.3926-1.4213i \\ 0.9623-0.0601i & 0.2027-1.6938i & -1.1650+1.7538i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.9385 & 1.9280 & 1.0107 \\ -1.0607 & -1.0656 & 2.1261 \\ 3.8464 & -0.7823 & -3.0643 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -2.1712-0.0107i & 1.0528-0.1985i & 1.1181+0.2092i \\ 0.7197+0.0883i & -1.7532+1.6364i & 1.0338-1.7248i \\ 0.7124-0.0776i & 0.9987-1.4382i & -1.7111+1.5159i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -0.6405 & 2.5194 & -1.8791 \\ -1.0974 & -2.4896 & 3.5873 \\ 1.4864 & 1.0196 & -2.5061 \end{pmatrix}$

Table B.23: Q matrices of the 3-by-3 transition matrices for all financial years of Next

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.3594 & 0.2812 & 0.3594 \\ 0.1667 & 0.3111 & 0.5222 \\ 0.2812 & 0.4583 & 0.2604 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2344 & 0.3438 & 0.4219 \\ 0.2500 & 0.3696 & 0.3804 \\ 0.2766 & 0.3830 & 0.3404 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2537 & 0.4328 & 0.3134 \\ 0.2360 & 0.3258 & 0.4382 \\ 0.3085 & 0.3404 & 0.3511 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.2687 & 0.3433 & 0.3881 \\ 0.2857 & 0.3516 & 0.3626 \\ 0.2667 & 0.3889 & 0.3444 \end{pmatrix}$

Table B.24: 3-by-3 transition matrices for all financial years of Pearson

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -1.4071-0.0850i & 0.7154-0.2184i & 0.6917+0.3033i \\ 0.5579-0.4793i & -1.0771-1.2319i & 0.5192+1.7113i \\ 0.4365+0.5111i & 0.5277+1.3135i & -0.9643-1.8248i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.8163-1.2101i & 2.4422-0.9536i & 0.3744+2.1638i \\ 1.3724-0.1805i & -2.9908-0.1422i & 1.6183+0.3227i \\ 0.5743+1.0006i & 1.2647+0.7885i & -1.8390-1.7893i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.5070 & 2.3909 & -0.8841 \\ -0.3702 & -1.8630 & 2.2332 \\ 1.4413 & 0.0815 & -1.5227 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.8407 & -1.3628 & 3.2035 \\ 1.6675 & -2.2727 & 0.6048 \\ -0.2768 & 3.3012 & -3.0240 \end{pmatrix}$

Table B.25: Q matrices of the 3-by-3 transition matrices for all financial years of Pearson

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.1143 & 0.3143 & 0.5714 \\ 0.1238 & 0.4286 & 0.4476 \\ 0.1759 & 0.4444 & 0.3796 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.1034 & 0.4828 & 0.4138 \\ 0.1560 & 0.4037 & 0.4404 \\ 0.0885 & 0.4425 & 0.4690 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2286 & 0.4857 & 0.2857 \\ 0.1261 & 0.4775 & 0.3964 \\ 0.1250 & 0.4038 & 0.4712 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.1364 & 0.4091 & 0.4545 \\ 0.0667 & 0.5250 & 0.4083 \\ 0.1019 & 0.4537 & 0.4444 \end{pmatrix}$

Table B.26: 3-by-3 transition matrices for all financial years of Royal bank

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.2137+1.3011i & 1.7692+1.5626i & 0.4445-2.8639i \\ 0.5278+0.2675i & -1.6799+0.3212i & 1.1522-0.5888i \\ 0.2298-0.6902i & 1.0252-0.8289i & -1.2552+1.5193i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.4972+1.5278i & 1.0642-1.9363i & 1.4329+0.4087i \\ 0.2167-1.1305i & -1.5559+1.4328i & 1.3393-0.3024i \\ 0.4548+0.6765i & 1.2064-0.8574i & -1.6612+0.1810i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.9600 & 1.9774 & -0.0174 \\ 0.3251 & -1.5734 & 1.2483 \\ 0.3125 & 1.0409 & -1.3534 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -3.1964 & 0.3673 & 2.8291 \\ -0.1360 & -1.4506 & 1.5866 \\ 0.8106 & 1.5667 & -2.3774 \end{pmatrix}$

Table B.27: Q matrices of the 3-by-3 transition matrices for all financial years of Royal bank

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.3056 & 0.4306 & 0.2639 \\ 0.2386 & 0.2841 & 0.4773 \\ 0.3222 & 0.3667 & 0.3111 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.3188 & 0.3043 & 0.3768 \\ 0.2065 & 0.4022 & 0.3913 \\ 0.3034 & 0.3820 & 0.3146 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.3553 & 0.4079 & 0.2368 \\ 0.2738 & 0.2976 & 0.4286 \\ 0.2889 & 0.3111 & 0.4000 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.3750 & 0.3295 & 0.2955 \\ 0.3291 & 0.3671 & 0.3038 \\ 0.3704 & 0.2469 & 0.3827 \end{pmatrix}$

Table B.28: 3-by-3 transition matrices for all financial years of Royal mail

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} 0.0314 & 2.8395 & -2.8709 \\ -1.4386 & -2.9619 & 4.4007 \\ 1.4095 & 0.6637 & -2.0732 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -1.7140-0.5291i & 0.7704-0.4609i & 0.9434+0.9900i \\ 0.5252-0.6691i & -1.5033-0.5828i & 0.9782+1.2518i \\ 0.7577+1.0849i & 0.9568+0.9449i & -1.7145-2.0297i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.2691 & 4.5326 & -3.2634 \\ 0.2377 & -4.7518 & 4.5142 \\ 0.8500 & 0.6073 & -1.4572 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -2.1556 & 1.4231 & 0.7325 \\ 0.8724 & -1.7371 & 0.8647 \\ 1.5303 & 0.1073 & -1.6377 \end{pmatrix}$

Table B.29: Q matrices of the 3-by-3 transition matrices for all financial years of Royal mail

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.2933 & 0.4133 & 0.2933 \\ 0.2921 & 0.3034 & 0.4045 \\ 0.3140 & 0.3721 & 0.3140 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.3485 & 0.3333 & 0.3182 \\ 0.2647 & 0.3824 & 0.3529 \\ 0.1951 & 0.5000 & 0.3049 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.3284 & 0.3284 & 0.3433 \\ 0.3068 & 0.2955 & 0.3977 \\ 0.1895 & 0.4105 & 0.4000 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.4444 & 0.2667 & 0.2889 \\ 0.2754 & 0.3188 & 0.4058 \\ 0.3596 & 0.2472 & 0.3933 \end{pmatrix}$

Table B.30: 3-by-3 transition matrices for all financial years of Sainsbury

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} 1.6453 & 6.8228 & -8.4671 \\ -2.3286 & -6.1802 & 8.5077 \\ 1.0115 & 0.5201 & -1.5313 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -1.6648-0.2069i & 0.8479+0.6471i & 0.8169-0.4402i \\ 0.6806-0.5185i & -1.5909+1.6217i & 0.9103-1.1033i \\ 0.4933+0.8114i & 1.2966-2.5382i & -1.7899+1.7267i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.7569+0.1337i & 0.8524-0.2365i & 0.9048+0.1027i \\ 0.7559-1.2490i & -1.7616+2.2087i & 1.0057-0.9596i \\ 0.5381+1.0402i & 1.0047-1.8394i & -1.5429+0.7992i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.2159 & 0.8336 & 0.3823 \\ 0.1435 & -1.9072 & 1.7636 \\ 1.1381 & 0.6029 & -1.7408 \end{pmatrix}$

Table B.31: Q matrices of the 3-by-3 transition matrices for all financial years of Sainsbury

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.1707 & 0.3415 & 0.4878 \\ 0.1792 & 0.4057 & 0.4151 \\ 0.1456 & 0.4854 & 0.3689 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2295 & 0.4262 & 0.3443 \\ 0.2174 & 0.3587 & 0.4239 \\ 0.2887 & 0.3402 & 0.3711 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.1724 & 0.4483 & 0.3793 \\ 0.2366 & 0.3548 & 0.4086 \\ 0.2626 & 0.3434 & 0.3939 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.4752 & 0.2574 & 0.2673 \\ 0.3239 & 0.2817 & 0.3944 \\ 0.4079 & 0.3158 & 0.2763 \end{pmatrix}$

Table B.32: 3-by-3 transition matrices for all financial years of Shell

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -1.0916 & -3.2671 & 4.3591 \\ 0.9717 & -1.9729 & 1.0013 \\ -0.5794 & 3.3807 & -2.8016 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.0712 & 2.3539 & -0.2828 \\ -0.3142 & -1.5826 & 1.8968 \\ 1.6381 & -0.0009 & -1.6373 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -9.2765 & 15.5368 & -6.2609 \\ -0.3218 & -0.2394 & 0.5612 \\ 5.7372 & -8.8780 & 3.1410 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.3302-0.0487i & 0.6509-0.0795i & 0.6792+0.1282i \\ 0.8334+0.8280i & -1.7987+1.3523i & 0.9653-2.1802i \\ 1.0225-0.6981i & 0.7824-1.1400i & -1.8048+1.8380i \end{pmatrix}$

Table B.33: Q matrices of the 3-by-3 transition matrices for all financial years of Shell

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.3125 & 0.4688 & 0.2188 \\ 0.2323 & 0.3636 & 0.4040 \\ 0.2414 & 0.3908 & 0.3678 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2969 & 0.3125 & 0.3906 \\ 0.2283 & 0.4022 & 0.3696 \\ 0.2447 & 0.3723 & 0.3830 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.3291 & 0.2785 & 0.3924 \\ 0.2625 & 0.3500 & 0.3875 \\ 0.3407 & 0.3297 & 0.3297 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.3750 & 0.2596 & 0.3654 \\ 0.3788 & 0.3333 & 0.2879 \\ 0.5256 & 0.2051 & 0.2692 \end{pmatrix}$

Table B.34: 3-by-3 transition matrices for all financial years of Smith

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -1.9183-0.3596i & 2.4554-3.5820i & -0.5367+3.9413i \\ 0.5997+0.2834i & -2.7218+2.8231i & 2.1218-3.1063i \\ 0.7300-0.0619i & 1.3365-0.6163i & -2.0664+0.6781i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.3011 & -0.0062 & 2.3073 \\ 0.4913 & -2.0842 & 1.5932 \\ 1.0488 & 2.0240 & -3.0730 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.8652+0.7422i & 0.8333+0.4678i & 1.0318-1.2099i \\ 0.8028+0.6698i & -1.8297+0.4222i & 1.0268-1.0919i \\ 0.8835-1.2129i & 0.8861-0.7646i & -1.7695+1.9772i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.2508-1.5218i & 0.5916+0.3365i & 0.6592+1.1853i \\ 0.9976-0.0602i & -1.7332+0.0133i & 0.7356+0.0469i \\ 0.8527+2.0967i & 0.6453-0.4637i & -1.4982-1.6331i \end{pmatrix}$

Table B.35: Q matrices of the 3-by-3 transition matrices for all financial years of Smith

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.0741 & 0.4815 & 0.4444 \\ 0.1120 & 0.5440 & 0.3440 \\ 0.1122 & 0.4592 & 0.4286 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.1538 & 0.4615 & 0.3846 \\ 0.0952 & 0.3810 & 0.5238 \\ 0.1008 & 0.4538 & 0.4454 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.1860 & 0.5116 & 0.3023 \\ 0.1546 & 0.3814 & 0.4639 \\ 0.1818 & 0.3455 & 0.4727 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.2750 & 0.3500 & 0.3750 \\ 0.0816 & 0.3980 & 0.5204 \\ 0.2000 & 0.4000 & 0.4000 \end{pmatrix}$

Table B.36: 3-by-3 transition matrices for all financial years of Taylor

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.9151+2.8006i & 1.3743-0.5083i & 1.5408-2.2923i \\ 0.3523-0.3368i & -1.2387+0.0611i & 0.8864+0.2757i \\ 0.3536-0.3419i & 1.2289+0.0621i & -1.5826+0.2799i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.6447-0.0674i & 1.0941-1.5091i & 1.5503+1.5765i \\ 0.3143+0.0819i & -1.5100+1.8355i & 1.1957-1.9175i \\ 0.3000-0.0587i & 1.1140-1.3147i & -1.4140+1.3734i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -2.2724 & 3.6020 & -1.3295 \\ 0.1905 & -2.0100 & 1.8193 \\ 0.7202 & 0.3646 & -1.0847 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.7155-0.1214i & 0.8575-0.1276i & 0.8580+0.2490i \\ 0.0109+1.1724i & -1.6342+1.2332i & 1.6233-2.4056i \\ 0.6341-0.9892i & 1.1206-1.0405i & -1.7547+2.0298i \end{pmatrix}$

Table B.37: Q matrices of the 3-by-3 transition matrices for all financial years of Taylor

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.2239 & 0.4179 & 0.3582 \\ 0.3061 & 0.3571 & 0.3367 \\ 0.2588 & 0.4118 & 0.3294 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.3846 & 0.3077 & 0.3077 \\ 0.3171 & 0.3415 & 0.3415 \\ 0.2556 & 0.3222 & 0.4222 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.2603 & 0.3288 & 0.4110 \\ 0.2674 & 0.3140 & 0.4186 \\ 0.3407 & 0.3736 & 0.2857 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.3711 & 0.2680 & 0.3608 \\ 0.4167 & 0.2500 & 0.3333 \\ 0.3924 & 0.3544 & 0.2532 \end{pmatrix}$

Table B.38: 3-by-3 transition matrices for all financial years of Unilever

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.0807+1.8093i & 1.0598-1.3328i & 1.0210-0.4766i \\ 0.2351-4.1759i & -1.7705+1.2215i & 1.5352+2.9545i \\ 1.3690+3.3886i & 1.2059-0.3578i & -2.5749-3.0308i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -1.6675 & 1.0454 & 0.6220 \\ 1.2822 & -2.4369 & 1.1550 \\ 0.3119 & 1.2743 & -1.5863 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -2.6595-0.9472i & 1.8290+1.9020i & 0.8309-0.9547i \\ 1.8790+2.4434i & -2.5532-1.0771i & 0.6741-1.3666i \\ 0.3755-1.5087i & 0.9092-0.5154i & -1.2847+2.0243i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 1.0389 & -5.1947 & 4.1558 \\ 1.2105 & -1.2997 & 0.0892 \\ -2.3789 & 7.5628 & -5.1841 \end{pmatrix}$

Table B.39: Q matrices of the 3-by-3 transition matrices for all financial years of Unilever

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.2456 & 0.3509 & 0.4035 \\ 0.1798 & 0.3596 & 0.4607 \\ 0.2596 & 0.3558 & 0.3846 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.3043 & 0.3188 & 0.3768 \\ 0.2069 & 0.3908 & 0.4023 \\ 0.3191 & 0.3298 & 0.3511 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.3239 & 0.4085 & 0.2676 \\ 0.2471 & 0.2941 & 0.4588 \\ 0.2766 & 0.3404 & 0.3830 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.3293 & 0.2927 & 0.3780 \\ 0.3256 & 0.3721 & 0.3023 \\ 0.3375 & 0.3750 & 0.2875 \end{pmatrix}$

Table B.40: 3-by-3 transition matrices for all financial years of Vodafone

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -3.2193-0.3495i & 1.4513-0.0560i & 1.7679+0.4056i \\ 1.4281+3.2195i & -2.5615+0.5162i & 1.1337-3.7358i \\ 0.5422-2.5639i & 1.3969-0.4110i & -1.9392+2.9750i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.1184-0.7839i & 0.8731-0.0961i & 1.2451+0.8799i \\ 0.2637-1.4741i & -1.7465-0.1806i & 1.4828+1.6545i \\ 1.3107+1.9396i & 0.9755+0.2376i & -2.2861-2.1771i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -2.9824-3.9836i & 1.6598-8.1379i & 1.3226+12.1216i \\ 1.1155+3.3470i & -2.9656+6.8374i & 1.8501-10.1844i \\ 1.1969-0.0946i & 1.4719-0.1932i & -2.6688+0.2878i \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -2.3404+0.9337i & 1.5241+2.2564i & 0.8163-3.1901i \\ 1.1888-0.1745i & -2.4356-0.4216i & 1.2468+0.5960i \\ 1.1211-0.7696i & 1.0561-1.8598i & -2.1772+2.6294i \end{pmatrix}$

Table B.41: Q matrices of the 3-by-3 transition matrices for all financial years of Vodafone

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.3247 & 0.3377 & 0.3377 \\ 0.3011 & 0.3871 & 0.3118 \\ 0.3077 & 0.3974 & 0.2949 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2364 & 0.3818 & 0.3818 \\ 0.2653 & 0.3980 & 0.3367 \\ 0.1753 & 0.3918 & 0.4330 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.3182 & 0.3485 & 0.3333 \\ 0.2449 & 0.4490 & 0.3061 \\ 0.2442 & 0.3721 & 0.3837 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.2075 & 0.3396 & 0.4528 \\ 0.2212 & 0.4327 & 0.3462 \\ 0.2043 & 0.4516 & 0.3441 \end{pmatrix}$

Table B.42: 3-by-3 transition matrices for all financial years of Whitbread

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.5071-0.3086i & 1.1923-1.7637i & 1.3150+2.0723i \\ 1.1116-0.0574i & -2.2769-0.3279i & 1.1653+0.3853i \\ 1.1497+0.3730i & 1.5377+2.1321i & -2.6875-2.5051i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.9835 & 1.3921 & 1.5912 \\ 2.1329 & -2.4377 & 0.3048 \\ -0.4349 & 1.6764 & -1.2412 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.9189 & 0.7862 & 1.1327 \\ 0.6926 & -1.4651 & 0.7725 \\ 0.6834 & 1.0987 & -1.7820 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -2.4150 & -2.1666 & 4.5815 \\ 0.9793 & -1.1052 & 0.1262 \\ 0.2742 & 2.5095 & -2.7839 \end{pmatrix}$

Table B.43: Q matrices of the 3-by-3 transition matrices for all financial years of Whitbread

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.2794 & 0.3824 & 0.3382 \\ 0.2903 & 0.3441 & 0.3656 \\ 0.2529 & 0.4023 & 0.3448 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.4177 & 0.3544 & 0.2278 \\ 0.2892 & 0.2410 & 0.4699 \\ 0.2500 & 0.3977 & 0.3523 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.4189 & 0.2568 & 0.3243 \\ 0.2442 & 0.3953 & 0.3605 \\ 0.2444 & 0.3667 & 0.3889 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.2041 & 0.3469 & 0.4490 \\ 0.1376 & 0.4771 & 0.3853 \\ 0.2609 & 0.4348 & 0.3043 \end{pmatrix}$

Table B.44: 3-by-3 transition matrices for all financial years of WhSmith

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.8648+0.3297i & 1.3138-0.6588i & 1.5510+0.3292i \\ 0.7836-1.0011i & -1.8820+2.0006i & 1.0983-0.9995i \\ 1.4015+0.8125i & 0.9850-1.6238i & -2.3865+0.8113i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -1.2454+0.0938i & 0.6204-0.7644i & 0.6247+0.6706i \\ 0.5799-0.2452i & -1.2587+1.9972i & 0.6789-1.7522i \\ 0.5711+0.1470i & 0.6302-1.1975i & -1.2012+1.0506i \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -1.2290 & 0.2636 & 0.9654 \\ 0.5154 & -1.9301 & 1.4147 \\ 0.5179 & 1.6276 & -2.1455 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -1.7460+1.0732i & 0.9932+0.5723i & 0.7529-1.6455i \\ 0.4565+0.5072i & -1.2441+0.2705i & 0.7877-0.7777i \\ 0.3891-1.1726i & 0.9452-0.6253i & -1.3343+1.7979i \end{pmatrix}$

Table B.45: Q matrices of the 3-by-3 transition matrices for all financial years of WhSmith

Cases	Transition matrices
(2009-2010)	$\begin{pmatrix} 0.0938 & 0.5312 & 0.3750 \\ 0.1481 & 0.4352 & 0.4167 \\ 0.1182 & 0.4091 & 0.4727 \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} 0.2586 & 0.3103 & 0.4310 \\ 0.2188 & 0.4479 & 0.3333 \\ 0.2292 & 0.3646 & 0.4062 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} 0.1707 & 0.4390 & 0.3902 \\ 0.1553 & 0.4272 & 0.4175 \\ 0.1589 & 0.3832 & 0.4579 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} 0.2453 & 0.3962 & 0.3585 \\ 0.2347 & 0.3776 & 0.3878 \\ 0.1753 & 0.4124 & 0.4124 \end{pmatrix}$

Table B.46: 3-by-3 transition matrices for all financial years of Wolseley

Cases	Q matrices
(2009-2010)	$\begin{pmatrix} -2.3837+2.2104i & 1.2078-2.9745i & 1.1759+0.7641i \\ 0.3532-0.6833i & -1.5374+0.9194i & 1.1842-0.2362i \\ 0.3473+0.0340i & 1.1836-0.0457i & -1.5309+0.0117i \end{pmatrix}$
(2010-2011)	$\begin{pmatrix} -2.7590 & 0.1130 & 2.6457 \\ 0.7245 & -1.3198 & 0.5953 \\ 0.9427 & 1.2516 & -2.1942 \end{pmatrix}$
(2011-2012)	$\begin{pmatrix} -3.4843 & 2.9667 & 0.5172 \\ 0.5925 & -2.1119 & 1.5194 \\ 0.7248 & 0.9140 & -1.6388 \end{pmatrix}$
(2012-2013)	$\begin{pmatrix} -2.3764-0.5760i & 1.3717+0.9181i & 1.0045-0.3422i \\ 0.9560+1.2546i & -2.2675-1.9997i & 1.3117+0.7452i \\ 0.3327-0.9527i & 1.5414+1.5186i & -1.8739-0.5659i \end{pmatrix}$

Table B.47: Q matrices of the 3-by-3 transition matrices for all financial years of Wolseley

Appendix C

Tables and Figures: Lamplighter Group

Cayley graph with 24 elements:

e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12	e13	e14	e15	e16	e17	e18	e19	e20	e21	e22	e23	e24
e2	e1	e5	e6	e3	e4	e8	e7	e10	e9	e13	e14	e11	e12	e16	e15	e18	e17	e21	e22	e19	e20	e24	e23
e3	e5	e1	e7	e2	e8	e4	e6	e11	e13	e9	e15	e10	e16	e12	e14	e19	e21	e17	e23	e18	e24	e20	e22
e4	e6	e7	e1	e8	e2	e3	e5	e12	e14	e15	e9	e16	e10	e11	e13	e20	e22	e23	e17	e24	e18	e19	e21
e5	e3	e2	e8	e1	e7	e6	e4	e13	e11	e10	e16	e9	e15	e14	e12	e21	e19	e18	e24	e17	e23	e22	e20
e6	e4	e8	e2	e7	e1	e5	e3	e14	e12	e15	e10	e16	e9	e13	e11	e22	e20	e24	e18	e23	e17	e21	e19
e7	e8	e4	e3	e6	e5	e1	e2	e15	e16	e12	e11	e14	e13	e9	e10	e23	e24	e20	e19	e22	e21	e17	e18
e8	e7	e6	e5	e4	e3	e2	e1	e16	e15	e14	e13	e12	e11	e10	e9	e24	e23	e22	e21	e20	e19	e18	e17
e9	e12	e10	e11	e14	e15	e13	e16	e17	e20	e18	e19	e22	e23	e21	e24	e1	e4	e2	e3	e6	e7	e5	e8
e10	e14	e9	e13	e12	e16	e11	e15	e18	e22	e17	e21	e20	e24	e19	e23	e2	e6	e1	e5	e4	e8	e3	e7
e11	e15	e13	e9	e16	e12	e10	e14	e19	e23	e21	e17	e24	e20	e18	e22	e19	e7	e5	e1	e8	e4	e2	e6
e12	e9	e14	e15	e10	e11	e16	e13	e20	e17	e22	e23	e18	e19	e24	e21	e20	e1	e6	e7	e2	e3	e8	e5
e13	e16	e11	e10	e15	e14	e9	e12	e21	e24	e19	e18	e23	e22	e17	e20	e5	e8	e3	e2	e7	e6	e1	e4
e14	e10	e12	e16	e9	e13	e15	e11	e22	e18	e20	e24	e17	e21	e23	e19	e6	e2	e4	e8	e1	e5	e7	e3
e15	e11	e16	e12	e13	e9	e14	e10	e23	e19	e24	e20	e21	e17	e22	e18	e7	e3	e8	e4	e5	e1	e6	e2
e16	e13	e15	e14	e11	e10	e12	e9	e24	e21	e23	e22	e19	e18	e20	e17	e8	e5	e7	e6	e3	e2	e4	e1
e17	e19	e20	e18	e23	e21	e22	e24	e1	e3	e4	e2	e7	e5	e6	e8	e9	e11	e12	e10	e15	e13	e14	e16
e18	e21	e22	e17	e24	e19	e20	e23	e2	e5	e6	e1	e8	e3	e4	e7	e10	e13	e14	e9	e16	e11	e12	e15
e19	e17	e23	e21	e20	e18	e24	e22	e3	e1	e7	e5	e4	e2	e8	e6	e11	e9	e15	e13	e12	e10	e16	e14
e20	e23	e17	e22	e19	e24	e18	e21	e4	e7	e1	e6	e3	e8	e2	e5	e12	e15	e9	e14	e11	e16	e10	e13
e21	e18	e24	e19	e22	e17	e23	e20	e5	e2	e8	e3	e6	e1	e7	e4	e13	e10	e16	e11	e14	e9	e15	e12
e22	e24	e18	e20	e21	e23	e17	e19	e6	e8	e2	e4	e5	e7	e1	e3	e14	e16	e10	e12	e13	e15	e9	e11
e23	e20	e19	e24	e17	e22	e21	e18	e7	e4	e3	e8	e1	e6	e5	e2	e15	e12	e11	e16	e9	e14	e13	e10
e24	e22	e21	e23	e18	e20	e19	e17	e8	e6	e5	e7	e2	e4	e3	e1	e16	e14	e13	e15	e10	e12	e11	e9

Figure C.1: Cayley graph with 24 elements.

Transition matrices for the first generator set,

$$S_1 = \{e4 = (0, (1, 0, 0)), e11 = (1, (0, 1, 0))\}$$

and second generator set

$$S_2 = \{e10 = (1,(0,0,1)), e20 = (2,(1,0,0))\}$$

0	0	0	1/2	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1/2	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1/2	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1/2	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0
0	0	1/2	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1/2	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	1/2	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	1/2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	1/2	0	0
0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	1/2	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	1/2
0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	1/2	0
0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0
0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0
0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0
1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0
0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0
0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0
0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2
0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0

Figure C.2: Transition matrix of the simple random walk for the first generator (S_1)

$$
\begin{matrix}
\alpha & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 \\
0 & \alpha & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 \\
0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) \\
0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & \frac{1}{2}(1-\alpha) & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & \alpha & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{matrix}
$$

$$
\begin{matrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 \\
\alpha & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 \\
0 & \alpha & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) \\
0 & \frac{1}{2}(1-\alpha) & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 \\
0 & 0 & 0 & 0 & \alpha & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 1/2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & \alpha & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & \frac{1}{2}(1-\alpha) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & \alpha
\end{matrix}
$$

Figure C.3: Transition matrix of the α- biased random walk for the first generator (S_1)

$$
\begin{matrix}
0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\lambda+1} & 0 & \frac{\lambda}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\lambda+1} & 0 & \frac{\lambda}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{\lambda}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{\lambda}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\lambda}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 \\
0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 \\
1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\lambda}{\lambda+1} & 0 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\
0 & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0
\end{matrix}
$$

Figure C.4: Transition matrix of the λ-biased random walk for the first generator (S_1)

α	0	0	$(1-\alpha)1/2$	0	0	0	0	0	0	$(1-\alpha)1/2$	0
0	α	0	0	0	$(1-\alpha)1/2$	0	0	0	0	0	0
0	0	α	0	0	0	$(1-\alpha)\frac{1}{\lambda+1}$	0	$(1-\alpha)\frac{\lambda}{\lambda+1}$	0	0	0
$(1-\alpha)1/2$	0	0	α	0	0	0	0	0	0	0	0
0	0	0	0	α	0	0	$(1-\alpha)\frac{1}{\lambda+1}$	0	$(1-\alpha)\frac{1}{\lambda+1}$	0	0
0	$(1-\alpha)1/2$	0	0	0	α	0	0	0	0	0	0
0	0	$(1-\alpha)\frac{\lambda}{\lambda+1}$	0	0	0	α	0	0	0	0	$(1-\alpha)\frac{1}{\lambda+1}$
0	0	0	0	$(1-\alpha)\frac{1}{\lambda+1}$	0	0	α	0	0	0	0
0	0	0	0	0	0	0	0	α	0	$(1-\alpha)\frac{1}{\lambda+1}$	0
0	0	0	0	0	0	0	0	0	α	0	0
0	0	0	0	0	0	0	0	$(1-\alpha)1/2$	0	α	0
0	0	0	0	0	0	0	0	0	0	0	α

0	0	0	0	0	0	0	0	0	0	0	0
$(1-\alpha)1/2$	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	$(1-\alpha)1/2$	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	0	$(1-\alpha)1/2$	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	$(1-\alpha)\frac{\lambda}{\lambda+1}$	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	$(1-\alpha)\frac{\lambda}{\lambda+1}$	0	0	0	0	0	0
1/2	0	0	0	$(1-\alpha)1/2$	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	$(1-\alpha)1/2$	0	0	0
0	0	$(1-\alpha)1/2$	0	0	0	0	0	0	$(1-\alpha)1/2$	0	0

$$\begin{matrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 \\
(1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 \\
0 & (1-\alpha)\frac{1}{\lambda+1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{matrix}$$

$$\begin{matrix}
\alpha & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & \alpha & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 \\
0 & (1-\alpha)1/2 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 \\
0 & 0 & 0 & 0 & \alpha & (1-\alpha)1/2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & (1-\alpha)1/2 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & (1-\alpha)1/2 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & (1-\alpha)1/2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & 0 & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)\frac{\lambda}{\lambda+1} & 0 & \alpha & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & (1-\alpha)1/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\alpha)1/2 & \alpha
\end{matrix}$$

Figure C.5: Transition matrix of the $\alpha-\lambda$- biased random walk for the first generator (S_1)

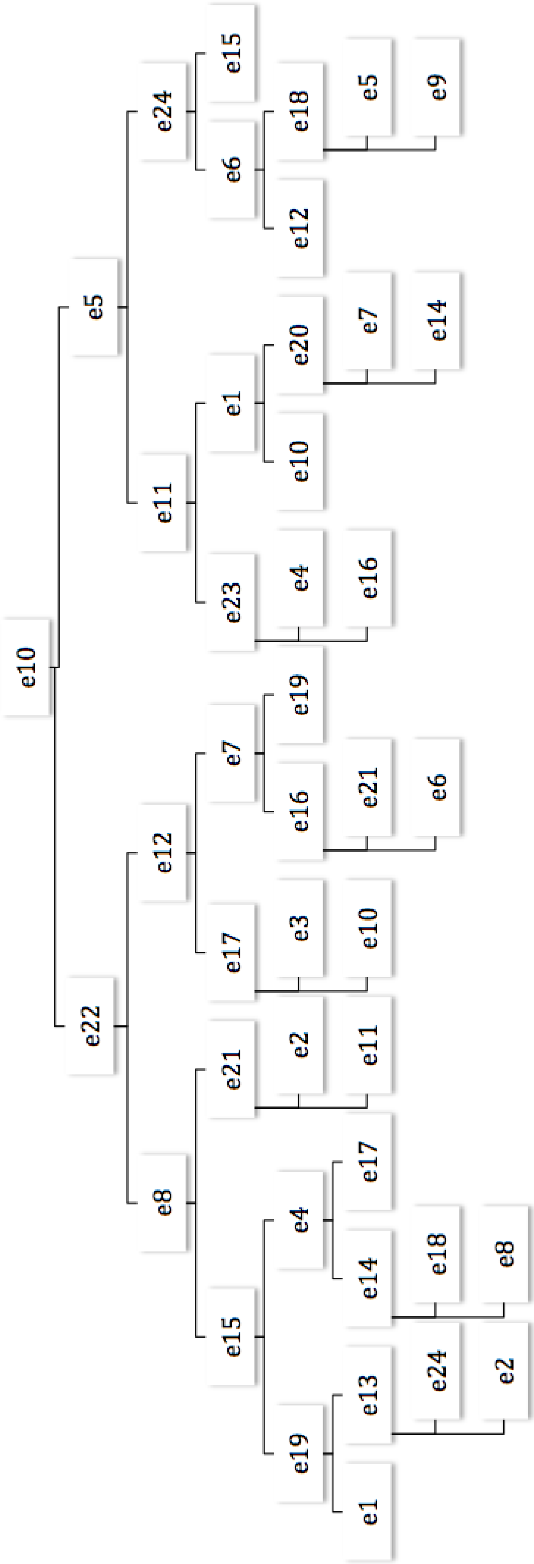


Figure C.6: Branching tree for the second generators (S_2) via method 1, initial point is e10

0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0
0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	1/2	0
0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	1/2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	1/2
0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	1/2	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	1/2	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	1/2	0	0	0
0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0
0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0
1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0
0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0
0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2
0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0
0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0
0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0
0	0	1/2	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1/2	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1/2	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1/2	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0
0	1/2	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1/2	0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1/2	0	0	0	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0
0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	1/2	0	0	0	0	0	0	0	0	0

Figure C.7: Transition matrix of the simple random walk for the second generator (S_2)

$$
\begin{matrix}
\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 \\
0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 \\
0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 \\
0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & \alpha & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & \alpha
\end{matrix}
$$

$$
\begin{matrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 \\
0 & \frac{1}{2}(1-\alpha) & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{matrix}
$$

$$
\begin{matrix}
0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) \\
0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0
\end{matrix}
$$

$$
\begin{matrix}
\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) \\
0 & \alpha & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \alpha & 0 & 0 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 & 0 \\
0 & 0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & \alpha & 0 \\
0 & 0 & \frac{1}{2}(1-\alpha) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha
\end{matrix}
$$

Figure C.8: Transition matrix of the α- biased random walk for the second generator (S_2)

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.4845 & 0.5155 & 0 \\ 0 & 0.4908 & 0.5092 \\ 0.4716 & 0 & 0.5284 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5271 & 0.4729 \\ 0.5638 & 0 & 0.4362 \\ 0.5128 & 0.4872 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.4702 & 0.5298 & 0 \\ 0 & 0.4847 & 0.5153 \\ 0.4732 & 0 & 0.5268 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.4738 & 0.5262 \\ 0.5385 & 0 & 0.4615 \\ 0.5106 & 0.4894 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.4947 & 0.5093 & 0 \\ 0 & 0.5015 & 0.4985 \\ 0.4670 & 0 & 0.5330 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5233 & 0.4767 \\ 0.5444 & 0 & 0.4556 \\ 0.5031 & 0.4969 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.4654 & 0.4346 & 0 \\ 0 & 0.5157 & 0.4843 \\ 0.5136 & 0 & 0.4864 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5701 & 0.4299 \\ 0.4766 & 0 & 0.5234 \\ 0.5294 & 0.4706 & 0 \end{pmatrix}$

Table C.1: Branching-type random walk on the lamplighter group.

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.4088 & 0.5912 & 0 \\ 0 & 0.5955 & 0.4045 \\ 0.4909 & 0 & 0.5091 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5836 & 0.4164 \\ 0.3623 & 0.1836 & 0.4541 \\ 0.4295 & 0.5705 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.4718 & 0.5282 & 0 \\ 0 & 0.6084 & 0.3916 \\ 0.4474 & 0 & 0.5526 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5604 & 0.4396 \\ 0.4144 & 0.1663 & 0.4194 \\ 0.4348 & 0.5652 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.3723 & 0.6277 & 0 \\ 0 & 0.5438 & 0.4562 \\ 0.4928 & 0 & 0.5072 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5749 & 0.4251 \\ 0.3687 & 0.2048 & 0.4265 \\ 0.4463 & 0.5537 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.4418 & 0.5582 & 0 \\ 0 & 0.5888 & 0.4112 \\ 0.5159 & 0 & 0.4841 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5019 & 0.4981 \\ 0.3131 & 0.2015 & 0.4854 \\ 0.3921 & 0.6079 & 0 \end{pmatrix}$

Table C.2: λ biased random walk on the lamplighter group.

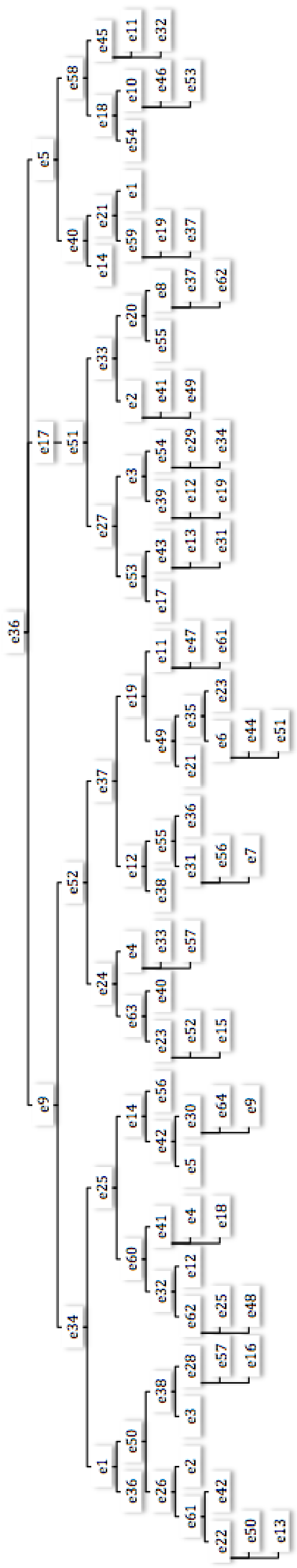


Figure C.9: Branching Tree for the chosen generator e36 as a initial point via method 1

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.6860 & 0.3140 & 0 \\ 0 & 0.7927 & 0.2073 \\ 0.2247 & 0 & 0.7753 \end{pmatrix}$	$\begin{pmatrix} 0.6167 & 0.3833 & 0 \\ 0.3174 & 0.6807 & 0.0018 \\ 0 & 1.0000 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.6070 & 0.3930 & 0 \\ 0 & 0.6990 & 0.3010 \\ 0.1886 & 0 & 0.8114 \end{pmatrix}$	$\begin{pmatrix} 0.5482 & 0.2193 & 0.2326 \\ 0.2295 & 0.5184 & 0.2521 \\ 0.1561 & 0.3035 & 0.5405 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.6808 & 0.3192 & 0 \\ 0 & 0.7629 & 0.2371 \\ 0.2103 & 0 & 0.7897 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.4000 & 0.6000 \\ 0.0061 & 0.3347 & 0.6592 \\ 0.0020 & 0.6514 & 0.3466 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.2857 & 0.7143 & 0 \\ 0 & 0.8000 & 0.2000 \\ 0.0041 & 0 & 0.9959 \end{pmatrix}$	$\begin{pmatrix} 0.3438 & 0.2188 & 0.4375 \\ 0.0181 & 0.5221 & 0.4598 \\ 0.0234 & 0.4936 & 0.4830 \end{pmatrix}$

Table C.3: α biased random walk on the lamplighter group.

Cases	$\hat{P}$
BP(2009-2010)	$\begin{pmatrix} 0 & 0.4323 & 0.5677 & 0 \\ 0 & 0 & 0.5973 & 0.4027 \\ 0.5018 & 0 & 0 & 0.4982 \\ 0.5391 & 0.4609 & 0 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0 & 0.4568 & 0.5432 & 0 \\ 0 & 0 & 0.5261 & 0.4739 \\ 0.5919 & 0 & 0 & 0.4081 \\ 0.5273 & 0.4727 & 0 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0 & 0.3827 & 0.6173 & 0 \\ 0 & 0 & 0.5577 & 0.4423 \\ 0.5261 & 0 & 0 & 0.4739 \\ 0.5526 & 0.4474 & 0 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0 & 0.4167 & 0.5833 & 0 \\ 0 & 0 & 0.5804 & 0.4196 \\ 0.5246 & 0 & 0 & 0.4754 \\ 0.5000 & 0.5000 & 0 & 0 \end{pmatrix}$

Table C.4: λ biased random walk on the tensor product of the lamplighter group with S_3.

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.4702 & 0.5298 & 0 \\ 0 & 0.4847 & 0.5153 \\ 0.4732 & 0 & 0.5268 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.4684 & 0.5316 \\ 0.5314 & 0 & 0.4686 \\ 0.5359 & 0.4641 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.4907 & 0.5093 & 0 \\ 0 & 0.5015 & 0.4985 \\ 0.4670 & 0 & 0.5330 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5015 & 0.4985 \\ 0.5015 & 0 & 0.4985 \\ 0.5301 & 0.4699 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.4654 & 0.5346 & 0 \\ 0 & 0.5157 & 0.4843 \\ 0.5136 & 0 & 0.4864 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.4832 & 0.5168 \\ 0.5045 & 0 & 0.4955 \\ 0.4643 & 0.5357 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.5248 & 0.4752 & 0 \\ 0 & 0.5090 & 0.4910 \\ 0.4985 & 0 & 0.5015 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5138 & 0.4862 \\ 0.5272 & 0 & 0.4728 \\ 0.4414 & 0.5586 & 0 \end{pmatrix}$

Table C.5: Branching-type random walk on the lamplighter group for volatility.

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.4071 & 0.5929 & 0 \\ 0 & 0.5833 & 0.4167 \\ 0.5093 & 0 & 0.4907 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5431 & 0.4569 \\ 0.3187 & 0.1946 & 0.4866 \\ 0.4224 & 0.5776 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.5103 & 0.4897 & 0 \\ 0 & 0.5812 & 0.4188 \\ 0.6180 & 0 & 0.3820 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5879 & 0.4121 \\ 0.3605 & 0.1684 & 0.4711 \\ 0.5700 & 0.4300 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.3827 & 0.6173 & 0 \\ 0 & 0.5535 & 0.4465 \\ 0.5000 & 0 & 0.5000 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5654 & 0.4346 \\ 0.3873 & 0.1569 & 0.4559 \\ 0.4013 & 0.5987 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.3893 & 0.6107 & 0 \\ 0 & 0.5604 & 0.4396 \\ 0.5136 & 0 & 0.4864 \end{pmatrix}$	$\begin{pmatrix} 0 & 0.5208 & 0.4792 \\ 0.2961 & 0.2257 & 0.4782 \\ 0.4396 & 0.5604 & 0 \end{pmatrix}$

Table C.6: λ biased random walk on the lamplighter group for volatility.

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.8821 & 0.1179 & 0 \\ 0 & 0.9480 & 0.0520 \\ 0.3882 & 0 & 0.6118 \end{pmatrix}$	$\begin{pmatrix} 0.6099 & 0.2967 & 0.0934 \\ 0.2074 & 0.6878 & 0.1048 \\ 0.2584 & 0.2022 & 0.5393 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.3333 & 0.6667 & 0 \\ 0 & 0.8182 & 0.1818 \\ 0.0010 & 0 & 0.9990 \end{pmatrix}$	$\begin{pmatrix} 0.6030 & 0.3712 & 0.0258 \\ 0.3259 & 0.6437 & 0.0304 \\ 0.5750 & 0.0750 & 0.3500 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.6032 & 0.3968 & 0 \\ 0 & 0.7788 & 0.2212 \\ 0.1589 & 0 & 0.8411 \end{pmatrix}$	$\begin{pmatrix} 0.2941 & 0.2941 & 0.4118 \\ 0.0120 & 0.3406 & 0.6474 \\ 0.0104 & 0.6798 & 0.3098 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.3611 & 0.6389 & 0 \\ 0 & 0.5615 & 0.4385 \\ 0.2036 & 0 & 0.7964 \end{pmatrix}$	$\begin{pmatrix} 0.5534 & 0.2362 & 0.2104 \\ 0.1965 & 0.5869 & 0.2166 \\ 0.2041 & 0.3095 & 0.4864 \end{pmatrix}$

Table C.7: α biased random walk on the lamplighter group for volatility.

Cases	Generator S_1	Generator S_2
BP(2009-2010)	$\begin{pmatrix} 0.4884 & 0.5116 & 0 \\ 0 & 0.6218 & 0.3782 \\ 0.3359 & 0 & 0.6641 \end{pmatrix}$	$\begin{pmatrix} 0.2500 & 0.7500 & 0 \\ 0.0020 & 0.9970 & 0.0010 \\ 0 & 1.0000 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.9103 & 0.0897 & 0 \\ 0 & 0.9636 & 0.0364 \\ 0.0922 & 0 & 0.9078 \end{pmatrix}$	$\begin{pmatrix} 0.2857 & 0.5714 & 0.1429 \\ 0.0030 & 0.9970 & 0 \\ 1.0000 & 0 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.4672 & 0.5328 & 0 \\ 0 & 0.6254 & 0.3746 \\ 0.2701 & 0 & 0.7299 \end{pmatrix}$	$\begin{pmatrix} 0 & 1.0000 & 0 \\ 0 & 0.9990 & 0.0010 \\ 0 & 1.0000 & 0 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.7647 & 0.2353 & 0 \\ 0 & 0.8792 & 0.1208 \\ 0.0656 & 0 & 0.9344 \end{pmatrix}$	$\begin{pmatrix} 0 & 1.0000 & 0 \\ 0 & 0.9990 & 0.0010 \\ 0 & 1.0000 & 0 \end{pmatrix}$

Table C.8: $\alpha - \lambda$ biased random walk on the lamplighter group for volatility.

Cases	P_X	P_Y	$P_X \oplus P_Y$
BP(2009-2010)	$\begin{pmatrix} 0.2419 & 0.7581 \\ 0.2449 & 0.7551 \end{pmatrix}$	$\begin{pmatrix} 0.9060 & 0.0940 \\ 0.9167 & 0.0833 \end{pmatrix}$	$\begin{pmatrix} 0.2192 & 0.0227 & 0.6868 & 0.0713 \\ 0.2218 & 0.0202 & 0.6949 & 0.0632 \\ 0.2219 & 0.0230 & 0.6841 & 0.0710 \\ 0.2245 & 0.0204 & 0.6922 & 0.0629 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.5319 & 0.4681 \\ 0.5652 & 0.4348 \end{pmatrix}$	$\begin{pmatrix} 0.8864 & 0.1136 \\ 0.6944 & 0.3056 \end{pmatrix}$	$\begin{pmatrix} 0.4715 & 0.0604 & 0.4149 & 0.0532 \\ 0.3694 & 0.1625 & 0.3251 & 0.1430 \\ 0.5010 & 0.0642 & 0.3854 & 0.0494 \\ 0.3925 & 0.1727 & 0.3019 & 0.1329 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.2333 & 0.7667 \\ 0.2344 & 0.7656 \end{pmatrix}$	$\begin{pmatrix} 0.7293 & 0.2707 \\ 0.6761 & 0.3239 \end{pmatrix}$	$\begin{pmatrix} 0.3298 & 0.1224 & 0.3995 & 0.1483 \\ 0.3057 & 0.1465 & 0.3704 & 0.1775 \\ 0.3354 & 0.1245 & 0.3939 & 0.1462 \\ 0.3109 & 0.1490 & 0.3652 & 0.1750 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.2787 & 0.7213 \\ 0.2296 & 0.7704 \end{pmatrix}$	$\begin{pmatrix} 0.9009 & 0.0991 \\ 0.9200 & 0.0800 \end{pmatrix}$	$\begin{pmatrix} 0.2511 & 0.0276 & 0.6498 & 0.0715 \\ 0.2564 & 0.0223 & 0.6636 & 0.0577 \\ 0.2068 & 0.0228 & 0.6940 & 0.0764 \\ 0.2112 & 0.0184 & 0.7088 & 0.0616 \end{pmatrix}$

Table C.9: Transition matrices via EM Algorithm.

Cases	P_X	P_Y	$P_X \oplus P_Y$
BP(2009-2010)	$\begin{pmatrix} 0.2072 & 0.7928 \\ 0.3729 & 0.6271 \end{pmatrix}$	$\begin{pmatrix} 0.9060 & 0.0940 \\ 0.9167 & 0.0833 \end{pmatrix}$	$\begin{pmatrix} 0.1877 & 0.0195 & 0.7183 & 0.745 \\ 0.1899 & 0.0173 & 0.7268 & 0.0660 \\ 0.3378 & 0.351 & 0.5682 & 0.0589 \\ 0.3418 & 0.0311 & 0.5749 & 0.0522 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.3489 & 0.6511 \\ 0.6504 & 0.3496 \end{pmatrix}$	$\begin{pmatrix} 0.8864 & 0.1136 \\ 0.6944 & 0.3056 \end{pmatrix}$	$\begin{pmatrix} 0.3093 & 0.0396 & 0.5771 & 0.0740 \\ 0.2423 & 0.1066 & 0.4521 & 0.1990 \\ 0.5765 & 0.0739 & 0.3099 & 0.0397 \\ 0.4516 & 0.1988 & 0.2428 & 0.1068 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.0463 & 0.9537 \\ 0.3221 & 0.6779 \end{pmatrix}$	$\begin{pmatrix} 0.7293 & 0.2707 \\ 0.6761 & 0.3239 \end{pmatrix}$	$\begin{pmatrix} 0.0338 & 0.0125 & 0.6955 & 0.2582 \\ 0.0313 & 0.0150 & 0.6448 & 0.3089 \\ 0.2349 & 0.0872 & 0.4944 & 0.1835 \\ 0.2178 & 0.1043 & 0.4583 & 0.2196 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1332 & 0.8668 \\ 0.3146 & 0.6854 \end{pmatrix}$	$\begin{pmatrix} 0.9009 & 0.0991 \\ 0.9200 & 0.0800 \end{pmatrix}$	$\begin{pmatrix} 0.1200 & 0.0132 & 0.7809 & 0.0859 \\ 0.1225 & 0.0107 & 0.7975 & 0.0693 \\ 0.2834 & 0.0312 & 0.6175 & 0.0679 \\ 0.2894 & 0.0252 & 0.6306 & 0.0548 \end{pmatrix}$

Table C.10: Transition matrices via Machine Learning.

Cases	P_X	P_Y	$P_X \oplus P_Y$
BP(2009-2010)	$\begin{pmatrix} 0.1471 & 0.8529 \\ 0.1434 & 0.8566 \end{pmatrix}$	$\begin{pmatrix} 0.9680 & 0.0320 \\ 1.0000 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.1424 & 0.0047 & 0.8256 & 0.0273 \\ 0.1471 & 0 & 0.8529 & 0 \\ 0.1388 & 0.0046 & 0.8292 & 0.0274 \\ 0.1434 & 0 & 0.8566 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.2954 & 0.7046 \\ 0.2054 & 0.7946 \end{pmatrix}$	$\begin{pmatrix} 0.9841 & 0.0159 \\ 1.0000 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.2907 & 0.0047 & 0.6934 & 0.0112 \\ 0.2954 & 0 & 0.7046 & 0 \\ 0.2022 & 0.0033 & 0.7820 & 0.0126 \\ 0.2054 & 0 & 0.7946 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.0596 & 0.9404 \\ 0.1289 & 0.8711 \end{pmatrix}$	$\begin{pmatrix} 0.9755 & 0.0245 \\ 0.8571 & 0.1429 \end{pmatrix}$	$\begin{pmatrix} 0.0581 & 0.0015 & 0.9174 & 0.0230 \\ 0.0511 & 0.0085 & 0.8061 & 0.1343 \\ 0.1257 & 0.0032 & 0.8498 & 0.0213 \\ 0.1105 & 0.0184 & 0.7467 & 0.1244 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1094 & 0.8906 \\ 0.1408 & 0.8592 \end{pmatrix}$	$\begin{pmatrix} 0.9595 & 0.0405 \\ 1.0000 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.1049 & 0.0044 & 0.8546 & 0.0361 \\ 0.1094 & 0 & 0.8906 & 0 \\ 0.1351 & 0.0057 & 0.8245 & 0.03348 \\ 0.1408 & 0 & 0.8592 & 0 \end{pmatrix}$

Table C.11: Transition matrices via EM Algorithm for volatility.

Cases	P_X	P_Y	$P_X \oplus P_Y$
BP(2009-2010)	$\begin{pmatrix} 0.0932 & 0.9068 \\ 0.1378 & 0.8622 \end{pmatrix}$	$\begin{pmatrix} 0.9680 & 0.0320 \\ 1.0000 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.0902 & 0.0030 & 0.8778 & 0.0290 \\ 0.0932 & 0 & 0.9068 & 0 \\ 0.1334 & 0.0044 & 0.8346 & 0.0276 \\ 0.1378 & 0 & 0.8622 & 0 \end{pmatrix}$
BP(2010-2011)	$\begin{pmatrix} 0.2285 & 0.7715 \\ 0.1793 & 0.8207 \end{pmatrix}$	$\begin{pmatrix} 0.9841 & 0.0159 \\ 1.0000 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.2249 & 0.0036 & 0.7592 & 0.0123 \\ 0.2285 & 0 & 0.8077 & 0.0130 \\ 0.1793 & 0 & 0.8207 & 0 \end{pmatrix}$
BP(2011-2012)	$\begin{pmatrix} 0.1280 & 0.8720 \\ 0.1257 & 0.8743 \end{pmatrix}$	$\begin{pmatrix} 0.9755 & 0.0245 \\ 0.8571 & 0.1429 \end{pmatrix}$	$\begin{pmatrix} 0.1249 & 0.0031 & 0.8506 & 0.0214 \\ 0.1097 & 0.0183 & 0.7474 & 0.1246 \\ 0.1226 & 0.0031 & 0.8529 & 0.0214 \\ 0.1077 & 0.0180 & 0.7494 & 0.1249 \end{pmatrix}$
BP(2012-2013)	$\begin{pmatrix} 0.1155 & 0.8845 \\ 0.0494 & 0.9506 \end{pmatrix}$	$\begin{pmatrix} 0.9595 & 0.0405 \\ 1.0000 & 0 \end{pmatrix}$	$\begin{pmatrix} 0.1127 & 0.0028 & 0.8628 & 0.0217 \\ 0.0990 & 0.0165 & 0.7581 & 0.1264 \\ 0.0482 & 0.0012 & 0.9273 & 0.0233 \\ 0.0423 & 0.0071 & 0.8148 & 0.1358 \end{pmatrix}$

Table C.12: Transition matrices via Machine Learning for volatility.

Appendix D

Quantum Binomial Model

D.1 Derivation of Single-Step Classical Binomial Model

A binomial market (B, S) consist of a risk-free bank account B and stock of price S. An arbitrage-free portfolio is:

$$B_1 = B_0(1 + r), S_1 = S_0(1 + R)$$

where the interest rate r is constant and the volatility rate R takes two values such as:

$$-1 \leq u < r < d$$

S_1 has two outcomes. Therefore, S_1 is described by using Bernoulli's model of the binomial market. $\Omega = \{1 + d, 1 + u\}$ with probability distribution P is the probabilistic model for S_1. Let's assume that:

$$q_u = P\{S_1 = S_0(1 + u)\} = P\{R = u\} > 0.$$

The variable q_u is interpreted as the probability of upward stock movement, and $q_d = 1 - q_u$ is interpreted as the probability of downward movement. There is a unique risk neutral measure M on Ω:

$$E_M[R] = uM\{R = u\} + dM\{R = d\} = r,$$

that is,

$$M\{R=u\}=\frac{r-d}{b-d}, M\{R=d\}=\frac{u-r}{b-d}.$$

Hence, the risk-neutral world of the classical model for the binomial market has only one element M.

$C_u = [S_0(1+u)-K]^+$ is the price of the call option if there is an upward movement u in the stock price and $C_d = [S_0(1+d)-K]^+$ is the price of the call option if there is an downward movement u in the stock price. Therefore, the formula of the current price of an option C is

$$C=\frac{1}{1+r}[\frac{r-d}{u-d}C_u+\frac{u-r}{u-d}C_d].$$

D.2 Derivation of Single-Step Quantum Binomial Model

In this section we present the quantum based no-arbitrage stock market in order to derive the single-step quantum model is presented for the binomial market (B, S). We consider the Hilbert space $\mathbf{C}^2$ with its canonical basis

$$|0>=\begin{pmatrix}1\\0\end{pmatrix}, |1>=\begin{pmatrix}0\\1\end{pmatrix}.$$

Assume a stock is in a quantum state:

$$\rho=\frac{1}{2}(wI_2+x\sigma_x+y\sigma_y+z\sigma_z)$$

where density matrix ρ is an arbitrary 2×2 Hermitian matrix. A Hermitian matrix is a complex square matrix that is equal to its own conjugate transpose. In quantum mechanics, a density matrix is a Hermitian matrix of trace one. Density matrices describe the statistical state of a group of systems or a single system where the pure quantum state the system is in. In addition, 2×2 Hermitian matrix can be written as a linear combination of the Pauli spin matrices

$$I_2=\begin{pmatrix}1&0\\0&1\end{pmatrix}, \sigma_x=\begin{pmatrix}0&1\\1&0\end{pmatrix}, \sigma_y=\begin{pmatrix}0&-i\\i&0\end{pmatrix}, \sigma_z=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

They form the basis for the Hilbert space of 2×2 complex Hermitian matrices ([48]). Let's consider a matrix A is a quantum operator in quantum mechanics. The stock is transformed from one state to the next by using this matrix where A is an arbitrary Hermitian matrix.

$$A = (x_0 I_2 + x_1\sigma_x + x_2\sigma_y + x_3\sigma_z)$$

Let a and b be the eigenvalues of A. The eigenvalues are represent the possible values of A when it is measured. And they expressed as follows:

$$a = x_0 - \sqrt{x_1^2 + x_2^2 + x_3^2}, b = x_0 + \sqrt{x_1^2 + x_2^2 + x_3^2} \tag{D.2.1}$$

where all x_j are real numbers, $x_1^2 + x_2^2 + x_3^2 \neq 0$ and $a, b > -1$. The following result is obtained after solving each equation for x_0 and taking them equal to each other:

$$a + \sqrt{x_1^2 + x_2^2 + x_3^2} = b - \sqrt{x_1^2 + x_2^2 + x_3^2}$$

which can be written as:

$$\frac{(b-a)^2}{4} = x_1^2 + x_2^2 + x_3^2$$

By substituting $(b-a)/2$ into equation D.2.1 in place of $\sqrt{x_1^2 + x_2^2 + x_3^2}$ it follows that:

$$x_0 = a + \frac{b-a}{2} = \frac{a+b}{2}$$

By the risk-neutral valuation, all individuals are indifferent to risk in a risk-neutral world, and the return earned on the stock must equal the risk-free interest rate. Considering A is used to transform the stock from one state to the next, one should expect that it evolves the stock at the risk-free rate. Thus, the expected value of measuring A should be the risk-free interest rate. In quantum mechanics, the expected value of a quantum operator is computed by using trace (see Chapter 2 for details of trace) and density matrix:

$$< A >_p = \mathrm{tr}(\rho A) = r$$

where r is the risk-free rate. As we mentioned before, a density matrix must have a trace of one:

$$\mathrm{tr}(\rho) = \frac{1}{2}(w+z) + \frac{1}{2}(w-z) = 1$$

This concludes that $w = 1$, which means that:

$$r = \mathrm{tr}\left(\frac{1}{2}\begin{pmatrix} w+z & x-iy \\ x+iy & w-z \end{pmatrix}\right)\begin{pmatrix} x_0+x_3 & x_1-x_2i \\ x_1+x_2i & x_0-x_3 \end{pmatrix}$$

which reduces to:

$$r = wx_0 + zx_3 + xx_1 + x_2y$$

The risk-neutral states are shown to be

$$x_1x + x_2y + x_3z = r - \frac{(a+b)}{2}$$

where $x_0 = (a+b)/2$ and $w = 1$. The eigenvalues of any density matrix are

$$\lambda_1 = \frac{1}{2}(w - \sqrt{x^2+y}^2 + z^2), \lambda_2 = \frac{1}{2}(w + \sqrt{x^2+y}^2 + z^2)$$

where $0 \le \lambda i \le 1$, $w = 1$ and all (x, y, z) satisfy

$$x^2 + y^2 + z^2 < 1$$

which is a disk of radius $\sqrt{1 - \frac{(2r-a-b)^2}{(b-a)^2}}$ in the unit ball of $\mathbf{R}^3$. The quantum binomial model replaces the single random variable R in the classical model with a complex Hermitian matrix A.

D.3 Derivation of Multi-Step Quantum Binomial Model

In this part, the single-step quantum binomial model is considered in order to derive $N-$period multi-step model. In the multi-step model, each step is taken tensor product (see Chapter 2 for details of the tensor product) with the previous step to build a composite quantum system. Let $\mathbf{H}_n = (C^2)^{\otimes n}$ and

$$|\epsilon_1 \dots \epsilon_n >= |\epsilon_1 > \otimes \dots \otimes |\epsilon_n >, \epsilon_1 \dots \epsilon_n = 0, 1.$$

$\{|\epsilon_1 \dots \epsilon_n > \epsilon_1 \dots \epsilon_n = 0, 1\}$ is the canonical basis of $\mathbf{H}_n$.

Then the N-period quantum binomial market (B, S) sets up with a risk-free bank account $B = (B_0, B_1, \dots, B_N)$ and a stock $S = (S_0, S_1, \dots, S_N)$ as follows:

$$B_n = B_0(1+r)^n, S_n = S_0 \otimes_{j=1}^{n} (I_2 + A) \otimes I_{N-n}$$

where I_{N-n} is the identity on $\mathbf{H}_{N-n}$ and stock price movement is represented by the quantum operator A which is a Hermitian matrix

$$A = (x_0 I_2 + x_{1j}\sigma_x + x_{2j}\sigma_y + x_{3j}\sigma_z) \tag{D.3.2}$$

where $\sigma_x, \sigma_y, \sigma_z$ are the Pauli matrices of quantum mechanics for all $j = 1, \ldots, N$.

S_N can be written as follows, assuming the Maxwell-Boltzmann classical statistics:

$$S_N = S_0 \sum_{n=0}^{N} (1+b)^n (1+a)^{N-n} [\sum_{|\sigma|} \otimes_{j=1}^{N} |w_{j\sigma} >< w_{j\sigma}|]$$

where all σ are subsets of $\{1, \ldots, N\}$, $w_{j\sigma} = u_\sigma$ for $j \in \sigma$ or $w_{j\sigma} = v_\sigma$ otherwise and form an orthonormal basis in the Hilbert space.

In here, $w_{j\sigma} = |1>= \binom{1}{0}$ meaning to choose $b = u - 1$ at time j in the tensor product. And, $w_{j\sigma} = |0>= \binom{0}{1}$ meaning to choose $a = d - 1$

$$f(S_N) = \sum_{n=0}^{N} f(u^n d^{N-n}) \sum_{|\sigma|} |w_{j\sigma} >< w_{j\sigma}|$$

The sum means that we fist fix the number of chosen u's $=n$, then automatically the number of chosen d's $= N - n$. The second sums is over all permutation with fixed number of chosen u's $= n$.

Now, let's consider call options in the $N-$ period quantum binomial market (B, S). Its payoff is

$$H_N = [S_N - K]^+$$

where K is the strike price. And,

$$H_N = [S_N - K]^+ = \sum_{n=0}^{N} [S_0(1+b)^n(1+a)^{N-n} - K][\sum_{|\sigma|} \otimes_{j=1}^{N} |w_{j\sigma} >< w_{j\sigma}|]$$

The stocks quantum states are represented by the density matrix ρ is:

$$\otimes_{j=1}^{N} \rho_j = \frac{1}{2^N} \otimes_{j=1}^{N} (I_2 + x_j\sigma_x + y_j\sigma_y + z_j\sigma_z)$$

All the states are faithful risk-neutral states of the N-period quantum binomial market (B, S) where (x_j, y_j, z_j) satisfies

$$\begin{cases} x_j^2 + y_j^2 + z_j^2 < 1, \\ x_1 j x_j + y_1 j y_j + z_1 j z_j = r - \frac{a+b}{2}, \end{cases}$$

for every $j = 1, \dots, N$. Also, by using the Maxwell-Boltzmann statistics, one has that

$$\mathrm{tr}[(\otimes_{j=1}^{N}\rho_j)\sum_{|\sigma|=n}\otimes_{j=1}^{N}|w_{j\sigma}><w_{j\sigma}|] = \frac{N!}{n!(N-n)!}q^n(1-q)^{N-n}$$

for $n = 0, 1, \dots, N$, where $q = \frac{ra}{ba}$.

Therefore, the price for the call option in this multi-step quantum binomial pricing can be written as follow:

$$C_0^N = \mathrm{tr}[(\otimes_{j=1}^{N}\rho_j)[S_N - K]^+]$$

Then this equation is taken and derived the equivalent of the Cox-Ross-Rubinstein option pricing formula in [102] as follows:

$$C_0^N = (1+r)^{-N}\sum_{n=0}^{N}\frac{N!}{n!(N-n)!}q^n(1-q)^{N-n}[S_0(1+b)^n(1+a)^{N-n} - K]^+].$$

Moreover, option price formula is via Bose-Einstein statistics instead of the classical Maxwell-Boltzmann statistics as follows ([102]):

$$C_0^N = \mathrm{tr}[\rho^{(\otimes N)})[S_N - K]^+]$$

Also this equation is taken and derived a new quantum option pricing formula in [102] as follows:

$$C_0^N = (1+r)^{-N}\sum_{n=0}^{N}(\frac{q^n(1-q)^{N-n}}{\sum_{k=0}^{N}q^k(1-q)^{N-k}}[S_0(1+b)^n(1+a)^{N-n} - K]^+].$$