ESTIMATION OF A JUMP DIFFUSION PROCESS USING A CLOSED-FORM APPROXIMATION LIKELIHOOD UNDER A STRONG DISCRETIZATION SCHEME OF ORDER ONE

FADONOUGBO RENAUD

MASTER OF SCIENCE

(Mathematics-Financial Option)

PAN AFRICAN UNIVERSITY

INSTITUTE FOR BASIC SCIENCES, TECHNOLOGY

AND INNOVATION

2016

ESTIMATION OF A JUMP DIFFUSION PROCESS USING A CLOSED-FORM APPROXIMATION LIKELIHOOD UNDER A STRONG DISCRETIZATION SCHEME OF ORDER ONE

FADONOUGBO Renaud

MF 300-0005/15

A Thesis submitted to Pan African University Institute for Basic Sciences, Technology and Innovation in partial fulfillment of the requirement for the award of the degree of Master in Mathematics - Financial option

2016

DECLARATION

This research thesis is my own work and has not been presented elsewhere for a degree award.

Signature .... Date............................

Renaud FADONOUGBO

Declaration by supervisors.

This research thesis has been submitted for examination with our approval as university supervisors.

Signature .... Date...

Prof. Eugēne Kouassi,

University of Botswana

Signature .... Date..

Dr George Orwa,

Jomo Kenyatta University

ACKNOWLEDGEMENT

The completion of this research thesis has been possible through the help of many individuals who gave me their support. I thank the almighty God for granting me the knowledge to carry out this study. I thank Pan African University for this opportunity that has been offered to me. I can't overlook the love and support from my family, my lovely mother, my father, my wife, my brothers, sisters and friends who encouraged me throughout the research.

Special thanks to my supervisor Prof. E. Kouassi for accepting to supervise this work despite of his tight schedule. I do also thank and Dr George O. ORWA for his scholarly and professional assistance and his presence to correct my work. I am personally grateful to Dr Takam Patrice Soh, who really assisted me from the beginning to the end of this thesis. I highly appreciate the moral and academic support they gave me.

DEDICATION

To my fiancée and my son.

ABSTRACT

The jump diffusion process plays a major role in financial modeling. The process was introduced for the first time by Merton in option pricing, to overcome the shortcoming of the Black Scholes formula. This study provides a closed-form expansion of the likelihood function to estimate the model parameters of the jump-diffusion process. We use the so called Jump adapted discretization to approximate the solution of the SDE under consideration. The approximation converges strongly with order one to the exact solution which is available in an explicit form for a few cases of model. That discretization presents some tractability that we used to derive the characteristic function of the process, and the method of saddlepoint is used thereon to get the transition probability. The process being a Markov process, the joint density is deduced as a product of the transition probabilities . Therefore, using a Monte-Carlo simulation, we carried out a maximization program on the closed likelihood function obtained for parameters estimation followed by a bootstrap to check the efficiency of the estimators.

RESUME

Dans cette étude nous avons souligné l'importance des processus de diffusion à saut et leur rôle en modélisation financière. Introduit pour la premiere fois par Merton pour la formulation des prix des options financières, le processus de

diffusion à saut a permi de surmonter les faiblesses de la célèbre formule de Black Scholes. Cette étude a permi de developper une approximation de la fonction de vraisemblance utilisée en methode de Maximum de Vraisemblance afin d'estimer les paramètres du modèle caractérisé par le processus de diffusion à saut. Nous avons utilisé un schéma de discrétisation appelé Jump adapted discretization pour approcher le processus exact, solution de l' équation différentielle stochastique decrivant le mouvement du processus. Nous avons prouvé que le schéma cvon-vergerge fortement à l'ordre un, vers la solution exacte. En utilisant le processus approché et en supposant la taille du saut comme étant une variable aléatoire dont la distribution appartient à la famille exponentielle naturelle, nous avons derivé la fonction charactéristique du processus. Cette dernière nous a permi de déduire par la suite la probabilité de transition à l'aide de la méthode de saddlepoint. Le processus considéré, étant un processus de Markov, nous avons exprimé la densité jointe comme un produit des probabilités de transition. Nous avons à l'aide de la methode de Monte-Carlo, simulé le processus et évalué le programme de maximisation pour l'estimation des paramètres et par Boostrap verifié les propriétés asymptotiques des estimateurs obtenus.

ABBREVIATIONS

- Eq=Equation
- NEF=Natural exponential family
- GBM=Geometric Brownian motion
- GMM= Generalized method of moment
- JAD= Jump adapted discretization
- SDE= stochastic differential equation
- SMM= Simulated method of moment
- SV= affine diffusion with stochastic volatility process
- SVJ= affine jump-diffusion with stochastic volatility and jump in the asset price process
- SVJJ= affine jump-diffusion with jumps both in the asset price and stochastic volatility process

Contents

List of Figures

List of Tables

Chapter 1

INTRODUCTION

1.1 Background and Motivation

In options or derivatives pricing, when it is assumed that the underlying security has continuous trajectory, the diffusion process as Geometric Brownian Motion(GBM) is the most used to desscribe its dynamics, for instance see (Black & Scholes, 1973) in pricing option and corporate liabilities, (Cox, Ingersoll, & Ross, 1985) and (Vasicek, 1977) in the modeling of rate and option of rate.
However, despite its popularity in option pricing and risk management, the diffusion process fails in capturing some stylised facts presented by financial data. Andersen and Lund (1997) have highlighted its failure to reproduce the fat-tailedness and non Gaussian features. In 2004, Johannes found out that none of the diffusion model can reflect the treasury rate dynamics.
The Jump diffusion process, having discontinuous movement at some discrete points in time rather than continuous has been found as alternative process that is used to describe the kurtosis and some other characteristics, (Glasserman, 2004, p. 134). Therefore Jump process has drawn much attention in modeling and became a major tool in representing dynamics of various financial asset such as

stock prices, bond prices, exchange rate, interest rate etc.

Merton (1976) was one of the first to introduce jump stochastic process combined with the diffusion process for option pricing. He argued that the assumption in the Black-Scholes derivation stating that trading holds continuously in time and hence the continuous sample path of price dynamics is not always valid. Merton pointed out that there exists two types of changes that can affect the stock price. There is the one due to a monetary disparity between supply and demand movement in capitalization rate, and this can be modeled by a GBM. The second type of change is caused by the arrival of new information at discrete points in time and this will be perfectly represented by a jump process.

Johannes (2004) studied the economic role of jumps and found out a link between the incidence of jumps and macroeconomics news arrivals, in the modeling of interest rate. Since we know that there exists a negative relation between the interest rate and the treasury bond price that can be influenced by the stability of the economy, he claimed that the jumps provide the instrument through which, information from the macroeconomic aspect get into the treasury market. It is the jump process that can capture the surprise events occurring in the market. The mixed diffusion jump processes can adequately model the time paths of exchange rate dynamics,(Akgiray & Booth, 1988). Glasserman and Kou (2003) found out that jump processes in the modeling of interest rate help to capture the skews and smiles of implied volatility.

Jump processes are also very useful in the modeling of default event in credit, models dealing with that consider a stochastic process to represent the dynamics of the risky asset while the jump helps to model the default event affecting the process.Some of those models are illustrated in Jarrow and Turnbull (1995) and Duffie and Singleton (1999).

The importance of Jump process as well diffusion process is therefore a fact and there is a wide literature on it, see for instance (Glasserman, 2004). Therefore an appropriate estimation is required of a parametric jump diffusion process will be useful.

1.2 Presentation of the model

The model under study is a multivariate parametric jump diffusion process defined on the probability space $(\Omega, \mathbb{F}, P)$ as:

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t + J(X_t)dN_t \tag{1.1}$$

with, X_t the vector of dimension n that represents the state, W_t a standard Brownian motion of dimension d, $\mu(X_t, \theta) : \mathbb{R}^n \to \mathbb{R}^n$ is the drift, $\sigma(X_t, \theta) : \mathbb{R}^n \to \mathbb{R}^{n\times d}$ is called diffusion and the variance matrix of the process can be defined as: $V(x, \theta) = \sigma(X_t, \theta)\sigma(X_t, \theta)^{T}$[1].

N_t the pure jump process whose intensity is $\lambda(X_t, \theta)$, J_t it is a random variable that denotes the jump size, whose distribution belongs to Natural Exponential Family(NEF) with parameter ω and $\theta \in \Theta$, a compact subset of $\mathbb{R}^k$. To be consistent we have to differentiate the values of the process before and after the occurrence of a jump. Therefore we will denote by X_{t-} and J_{t-} the value and the size respectively of the process and the jump before the occurrence of jump. Thus 1.1 can be rewritten as follows:

$$dX_t = \mu(X_{t-}, \theta)dt + \sigma(X_{t-}, \theta)dW_t + J(X_{t-})dN_t \tag{1.2}$$

where $X_{t-} = \lim_{u \overset{\rightarrow}{<} t} X_u$.

X has Markov property and is connected to the infinitesimal partial operator $\mathcal{A}$, which is defined as follows: for $h : \mathbb{R}^n \to \mathbb{R}$ a bounded function in C^2 we have:

$$\begin{aligned}\mathcal{A}h(x) &= \sum_{i=1}^{n} \mu(x_i, \theta)\frac{\partial}{\partial x_i}h(x) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} v_{ij}(x, \theta)\frac{\partial^2}{\partial x_i x_j}h(x) \\ &\quad + \lambda(x, \theta)\int_C [h(x + c) - h(x)]\,\nu(c, \theta)dc \end{aligned}\tag{1.3}$$

[1] $\sigma(X_t, \theta)^T$ is the transposition of $\sigma(X_t, \theta)$

Let us assume the following conditions on the coefficient functions.

- Lipschitz conditions

$$|\mu(x(t),\theta) - \mu(y_t,\theta)|^2 \leq C_1|x_t - y_t|^2, |\sigma(x_t,\theta) - \sigma(y_t,\theta)|^2 \leq C_2|x_t - y_t|^2, \\ |J(x_t) - J(y_t)|^2 \leq C_3|x_t - y_t|^2, \tag{1.4}$$

  for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$

- Linear growth conditions

$$|\mu(x_t,\theta)|^2 \leq K_1(1 + |x|^2), |\sigma(x_t,\theta)|^2 \leq K_2(1 + |x|^2), \\ |J(x_t,)|^2 \leq K_3(1 + |x|^2), \tag{1.5}$$

  for all $t \in [0,T]$ and $x, \in \mathbb{R}^d$.

Therefore under conditions(1.4) and (1.5) the stochastic differential equation(SDE) (1.1)admits a unique strong solution, see Bruti-Liberati and Platen (2007).

1.3 Statement of the problem

To estimate the model 1.1 and 1.2 defined above, when it is well specified, the method of maximum likelihood is feasible. However that is not easy to implement since it is difficult to get a closed-form of the likelihood function. An alternative approach is to use an approximation to the likelihood, see (Aït-Sahalia, 2008), (Aït-Sahalia, Fan, & Peng, 2009)and (Li, 2013). In spite of the wide range of research carried out on it, the methods used by researchers involve discrete observations yet the process is necessarily not, and also some are computationally challenging. So in order to address the challenges mentioned above, in this we have derived a closed likelihood function based on discretization.

1.4 Objectives

1.4.1 General objective

To provide a maximum likelihood estimator of the jump diffusion process.

1.4.2 Specific objectives

1. To approximate the Jump diffusion process using a discretization scheme
2. To test the convergence of the approximation
3. To derive an approximation of the likelihood function using the probability transition, based on the discretization.
4. To carry out some Monte Carlo's simulation to investigate the efficiency of the estimators obtained.

1.5 Justification of the study

The jump diffusion process is becoming more and more useful in financial economics, and its estimation remains a challenge for econometricians. This study will provide a tractable tool of estimation of jump diffusion processes, that will be applicable to a large class of jump diffusion process. The purpose is to get a method that, in spite of using discrete observation we can still get an accurate estimate. We use the discretization scheme to derive the probability transition and from the Markov property derive the likelihood function. The method addresses some of the shortcomings faced by practitioners when it comes to asset pricing, modeling exchange rate, analyzing currency crises and financial market crashes. The work can still be useful in extreme event modeling as the jump itself is a "rare event", that can have an important implication on the market return, as it incorporates the new information that arrives into the market.

1.6 Significance of the study

In this study we have used strong convergence scheme to derive a transition probability for jump diffusion process, an approach which was used so far only for diffusion process may be due to the challenge presenting by the presence of jump modeled by a Poisson process. This helps to better conciliate the discrete observations and the continuous feature of the process.

1.7 Outline of the thesis

The remaining of the work is organized as follows, chapter presents the definitions of some basic concepts and literature review. In chapter three we present the methodology used,while in chapter 4 it is about the derivation of the main results, its applications and results for evidence. Finally in the chapter 5 we give a conclusion of the study and some recommendations.

Chapter 2

LITERATURE REVIEW

The jump diffusion processes play a major role in financial modeling. It wast first introduced for the by Merton in option pricing, to overcome the shortcoming of the Black Scholes formula. It has gained a great attention from researchers in various fields especially in finance and has been used intensively as an important modeling tool by most financial institution, due to the fact that it is closed to the market realities. They differ basically from the diffusion processes from their discontinuity path and help to include the effect of occurrence of rare event. Both a diffusion and jump solve a stochastic differential equation(SDE) $dS = fdt+gdW$ but the later allows the solution to jump and take S to $S + J$ and we have the following SDE $dS = fdt + gdW + JdN$, where N is used to count the number of jumps that occurred, J is the random jump magnitude.

As a model it has to be estimated parametrically as well as nonparametrically or numerically. Our study focused on the parametric estimation and unfortunately there is no a wide literature does exist on the methods.

In the remaining pats of this chapter, we present in the first section an overview on jump process followed by sections that present a review of the literature that exists on the discretization and estimation approaches.

2.1 Poisson Jump Processes

This section is essentially about the definition of Poisson process, for more details refer to (Privault, 2012).

The standard Poisson process is the basic known Poisson Process. It is a continuous stochastic process $\{N_t, t \in \mathbb{R}_+\}$ with maximum jump size 1 and present a constant paths between two jumps. The value N_s of the process at time s is defined as

$$N_s = \sum_{k=1}^{\infty} 1_{[T_k,\infty)}(s) \tag{2.1}$$

with

$$1_{[T_k,\infty)}(s) = \begin{cases} 1 & \text{if} \quad s \geq T_k \\ 0 & \text{if} \quad 0 \leq s < T_k \end{cases} \tag{2.2}$$

where $k \geq 1$, $(T_k)_{k\geq 1}$ is a an increasing family of jump times of the Poisson process,such that $\lim_{k\to\infty} T_k = \infty$. The process satisfies two important conditions

Condition 1 The Poisson process $(N_t)_{t\in\mathbb{R}_+}$ has an independent increment, meaning that $\forall\, 0 \leq s_0 \leq s_1 \leq \ldots s_n$ with $n \geq 1$ the random variables $N_{s_1} - N_{s_0}, \ldots, N_{s_n} - N_{s_{n-1}}$ are independent.

Condition 2 The Process has stationary increments meaning for all $h > 0$, the random variables $N_{t+h} - N_{s+h}$ and $N_t - N_s$ have the same probability distribution.

The two conditions described above confer to the Poisson Process the property of Levy process. For addition the increment $N_t - N_s$ of the Poisson standard process is proved to follow a Poisson distribution $\lambda(t-s)$ where λ,the jump arrival rate is called the intensity of the Poisson process.

The standard Poisson presents a shortcoming with his constant size that can't go beyond 1, therefore it is limited in providing accurate underlying dynamics model. Then people have shown interest in considering the compound Poisson process that satisfies all the conditions of the standard Poisson, but its improvement is from the fact that it has a random jump's size. The compound Poisson process

$(Z_t)_{t\in\mathbb{R}_+}$ is a stochastic process, whose value Z_s at the time s is defined from a standard Poisson process with intensity λ and from $Y_{i,i=1\ldots}$, independent and identically distributed random variables with density probability $\nu(dy)$ as

$$Z_s = \sum_{k=1}^{N_s} Y_k \tag{2.3}$$

It is important to mention that the standard Poisson and compound Poisson processes can be transformed easily into martingales processes called Compensated Poisson process with respect to their own filtration just subtracting their expected value. Since we know how important this martingale property is in option pricing, see (Glasserman, 2004).

2.2 Discretization method

As the transition density has a closed form expression for a few cases, and in general not many approaches have been proposed to get an approximation. In fact most of the method have been proposed for diffusion process. Since the jump diffusion process in the absence of jump remains a diffusion process, it is better to have an overview of these different methods.

As first method, we have the Euler scheme approximation for which the diffusion process X_t solution of the following (SDE):

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t \tag{2.4}$$

is approximated by the following discrete time model:

$$X_{ih} = X_{(i-1)h} - \mu(X_{(i-1)h}, \theta) + \sigma(X_{(i-1)h}, \theta) + \sqrt{h}\varepsilon_i \tag{2.5}$$

where $\varepsilon_i \sim i.i.dN(0,1)$. Meanwhile of how easy it is to implement the closed form expression of the transition, the accuracy of this approximation depends on the interval observation h. When it is small, the approximation is accurate, but when it is large the approximation can be poor.

Merton (1980); Lo (1988) have found out that the approach is not consistent when h is fixed. A closely approach has been suggested by Gollnick, Houthakker, and Taylor (1968) and Bergstrom (1966), and it is defined on integrating the SDE and the use of the following trapezoidal rule approximation:

$$\int_{(i-1)h}^{ih} \mu(X_t, \theta)dt = \frac{h}{2}\left\{\mu(X_{ih}, \theta) + \mu(X_{(i-1)h}, \theta)\right\} \tag{2.6}$$

The two approximations are equivalent when h is small. Unfortunately the consistency of the approach is measured in term of h and some authors have showed the asymptotic bias in the estimate is $O(h^2)$.

A lot of work has been done to reduce the discretization bias on the Euler approximation, see (Elerian, 1998) and (Tse, Zhang, & Yu, 2004) which have improved the method of Milstein (1979). The first used a second order term in a stochastic Taylor series expansion to refine the approximation (2.5), while the later in a Bayesian context improved the Milstein's scheme.

Glasserman and Merener (2004) argued that the Euler scheme converges weakly with order one while the Milstein scheme with order two. They found that between the time of occurrence of jumps, the dynamics of process are purely diffusive and it is possible to be simulated using standard discretization methods. Under conditions on the coefficient functions of the jump-diffusion process, they showed that the method used conserves the same weak convergence order for both jump diffusion and pure diffusion processes and the construction of jumps does not degrade the convergence of the method. This argument leads then to one appropriate way of using diffusion discretization scheme for approximating jump diffusion process. This approximation scheme is the so called Jump adapted discretization, first introduced by ? (?). Even though This discretization scheme offers a good tractability as it is easier to manipulate and avoid double stochastic integrals involving Poisson measure and Wiener process, the proof of its convergence is not an easy task. There are two types of convergence namely weak and strong convergence. Bruti-Liberati and Platen (2007) proposed a general proof of the strong convergence theorem with a dependent mark for the jump. We find interesting to test the strong convergence of the non dependent Mark jump with a compound Poisson process, which is our model of interest.

2.3 Estimation approach

The difficulty of estimating jump diffusion has been pointed out about the fact that it is not easy to carry out an analytic form of the likelihood function, Sundaresan (2000), mentioned by Yu (2007). Rockinger and Semenova (2005) found that the estimation of the jump diffusion processes is subject to three main difficulties such as: requirement of analytic form of solution for the density function and sometimes its moment, only discrete observations are possible and also some state variables are non-observable. But the shortcomings of non-observable state variables are overcome by some methods especially, the simulated method of moment(SMM) of (Darrell Duffie, 1993) that is an extension to a considerable class of asset pricing models of generalized method of moment(GMM) estimator (Hansen, 1982). A. Ronald Gallant (1996) have proposed the method of moment that consist of using the score of density(i.e the first derivative of the density) that is analytically known to define the GMM principle used in (Hansen, 1982). But those methods fail in differentiating between the continuous path and the jump path dynamics which is discontinuous.

Rockinger and Semenova (2005) worked on three variant of the jump diffusion processes where the method of estimation is based on the characteristic function. It is about the affine diffusion with diffusive stochastic volatility(SV), affine jump-diffusion with stochastic volatility and jumps in the asset price(SVJ) and affine jump-diffusion with jumps both in the asset price and stochastic volatility processes(SVJJ). They argue that using the characteristic function does not imply any loss of information compared of using the density, cause of the existence of an injection relation between distribution and characteristic function.

They are few works that has used the maximum likelihood method which is possible cause of the Markov property of the process, some of those have used specific class of model to be able to get a closed-form of the likelihood function. For instance, Ball and Torous (1983) approximated the Poisson jump process by a Bernoulli jump process in the estimation of the jump diffusion process. Fher and Rosenfeld (1979) estimated the Poisson jump-diffusion process through maximum likelihood method and assumed that the jump size is a fixed known constant. But unfortunately , the results of parameter estimates have not been accurate with a large standard errors.

There is no wide research that has been done about the jump diffusion process in multivariate case. Aït-Sahalia (2002) has derived a closed-form approximation of likelihood function for univariate diffusion processes that has been proved well accurate and refined by Bakshi and Ju (2005). That approach is extended to a multivariate case by Aït-Sahalia (2008). Yu (2007) based on this closed-form, approximated the likelihood function of multivariate jump diffusion processes, solved using Kolmogorov equations. But this expansion faces a challenge on reducibility and explicitly of the Lamperti transform . Indeed the closed-form likelihood for jump diffusion proposed by Yu (2007) is deduced from the one of citeauthorNPsahalia2008 for the diffusion case, and that is applied differently according to the reducibility of the process. Li (2013) developed an expansion of transition for multivariate diffusion processes that overcomes those challenges cited above and is applicable to a large class of processes. But unfortunately all the method are still not able to overcome the problem of discretization. Aït-Sahalia and Yu (2006) proved the use of saddlepoint method introduced by ? (?) in the approximation of probability transition. The method essentially relies on the determination of characteristic function and the latter may not be tractable for a broad range of diffusion process and jump diffusion process in particular.

Based on that mean to approximate the probability transition, and the argument on the characteristic function from Rockinger and Semenova (2005), we develop a method that can allow us to get the characteristic function for a wide range of (JDP) and that can at the same time consider the discrete feature of the observation data.

Chapter 3

METHODOLOGY

This chapter presents an overall view of the approach we will use to achieve the objectives given in Chapter(3).

3.1 Method of estimation

The method of estimation is based on the maximum likelihood approach. But as stated, the problem is to find a closed-form expression of the likelihood function. In our study we use an approach of the transition probability density as used by Yu (2007) and Aït-Sahalia et al. (2009).

Definition 3.1.1 *The transition probability density $p(\Delta, z \mid y, \theta)$, if it exists, represents the probability that $X_{t+\Delta} = z \in \mathbb{R}^n$ knowing $X_t = y \in \mathbb{R}^n$*

The probability density function under a suitable conditions, satisfies the forward and backward Kolmogorov equations respectively described as follows:

$$\frac{\partial}{\partial \Delta} p = \mathcal{L}^F p + E\left[\lambda(z-j)p(\Delta, z-j \mid y) - \lambda(z)p(\Delta, z \mid y)\right] \quad (3.1)$$

$$\frac{\partial}{\partial \Delta} p = \mathcal{L}^B p + \lambda(y)E\left[p(\Delta, z-j \mid y+j) - p(\Delta, z \mid y)\right] \quad (3.2)$$

where the operators $\mathcal{L}^F$ and $\mathcal{L}^B$ are defined by:

$$\begin{aligned}
\mathcal{L}^F p &= -\sum_{i=1}^{n} \frac{\partial}{\partial z_i}[\mu_i(z)p] + \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\partial^2}{\partial z_i \partial z_k}[v_{ik}(z)p] \\
\mathcal{L}^B p &= \sum_{i=1}^{n} \frac{\partial}{\partial y_i}[\mu_i(z)p] + \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\partial^2}{\partial y_i \partial y_k}[v_{ik}(z)p] \qquad (3.3)
\end{aligned}$$

The infinitesimal generators for the forward and backward Kolmogorov equation are given by:

$$\begin{aligned}
\mathcal{A}^F p &= \mathcal{L}^F p + E\left[\lambda(z-j)p(\Delta, z-j \mid y) - \lambda(z)p(\Delta, z \mid y)\right] \\
\mathcal{A}^B p &= \mathcal{L}^B p + \lambda(y)E\left[p(\Delta, z-j \mid y+j) - p(\Delta, z \mid y)\right]
\end{aligned}$$

The expectations are computed with the distribution of the Jumps and are given as:

$$\begin{aligned}
E\left[\lambda(z-j)p(\Delta, z-j \mid y) - \lambda(z)p(\Delta, z \mid y)\right] &= \int_C [\lambda(z-j)p(\Delta, z-j \mid y) \\
&\quad -\lambda(z)p(\Delta, z \mid y)]\, \nu(c)dc \\
E\left[p(\Delta, z-j \mid y+j) - p(\Delta, z \mid y)\right] &= \int_C [p(\Delta, z-j \mid y+j) \\
&\quad -p(\Delta, z \mid y)]\, \nu(c)dc
\end{aligned}$$

The transition probability is an important tool as it plays a main role in financial statistical, it gives a proper description of the dynamics of the underlying, see Aït-Sahalia et al. (2009). In our study as in others we will focus on the transition probability density, obtain its expansion to approximate the likelihood function. Indeed, considering discrete observations of X in a time width T $\{\Delta, 2\Delta, \ldots, T\Delta\}$ which could be daily, weekly, etc,due to the Markovian property of the jump diffusion process the likelihood function can be expressed as :

$$L_T(\theta) = \Pi_{i=1}^{T} p_X(\Delta t, X(i\Delta) \mid X((i-1)\Delta), \theta) \qquad (3.4)$$

The parameters estimator of θ is obtained by maximizing the log-likelihood function i.e

$$\hat{\theta}_T = \arg\max_{\theta \in \Theta} l_T(\theta) \tag{3.5}$$

where

$$l_T(\theta) = \log L_T(\theta) = \sum_{i=1}^{T} \log p_X(\Delta t, X(i\Delta) \mid X((i-1)\Delta), \theta)$$

To get an appropriate approximation of the transition probability density, we look for an appropriate method of discretization as observations are obtained at discrete time. The method of discretization used is presented in the following section.

3.2 Discretization approach

3.2.1 Discretisation Scheme

In this section we derive the discrete approximation of order one of the strong solution of the SDE(1.2)and prove its convergence. The discretization approach used is called Jump-diffusion adapted approximation(JAD) that consists of separating the diffusion part and the jump one. An appropriate discretization is used for the diffusion part where the effect of the jump is added at due time. This method is based on Itô-Taylor expansion of the drift and diffusion term regardless of the jump part, see Bruti-Liberati and Platen (2007). Since we assumed that the jump occurs at discrete point time and between two discretization points the process has a diffusion dynamics. Therefore the solution of (1.1) can be expressed as follows:

$$X_t = X_{t-} + J(X_{t-})\Delta N_t \tag{3.6}$$

where X_{t-} is a diffusion process.

We consider a jump time discretization $0 = t_0 < t_1 < \ldots < t_N = T$; constructed by a superposition of the jump times $\{\tau_1, \ldots\}$ to a deterministic equidistant grid with maximum step size $\Delta > 0$. From the SDE(1.2) and for $t \in [t_i, t_{i+1}[$ we have :

$$\int_{t_i}^{t_{i+1}} dX_t = \int_{t_i}^{t_{i+1}} \mu(X_{t-}, \theta)dt + \int_{t_i}^{t_{i+1}} \sigma(X_{t-}, \theta)dW_t + \int_{t_i}^{t_{i+1}} J_{t-}dN_t \qquad (3.7)$$

As X_{t-} denotes the value of the process just before the occurrence of a potential jump, it may be then expressed as a solution of the following SDE:

$$dX_s = \mu(X_s, \theta)dt + \sigma(X_s, \theta)dW_s,$$

and applying the Itô formula, this leads to:

$$\begin{aligned} \mu(X_t, \theta) &= \mu(X_{t_i}, \theta) + \int_{t_i}^{t} \mu_x(X_s, \theta)\mu(X_s, \theta)ds + \frac{1}{2}\int_{t_i}^{t} \mu_{xx}(X_s, \theta)\sigma^2(X_s, \theta)ds \\ &\quad + \int_{t_i}^{t} \mu_x(X_s, \theta)\sigma(X_s, \theta)dW_s \end{aligned}$$

and

$$\begin{aligned} \sigma(X_t, \theta) &= \sigma(X_{t_i}, \theta) + \int_{t_i}^{t} \sigma_x(X_s, \theta)\mu(X_s, \theta)ds + \frac{1}{2}\int_{t_i}^{t} \sigma_{xx}(X_s, \theta)\sigma^2(X_s, \theta)ds \\ &\quad + \int_{t_i}^{t} \sigma_x(X_s, \theta)\sigma(X_s, \theta)dW_s \end{aligned}$$

Subtituting $\mu(X_t, \theta)$ and $\sigma(X_t, \theta)$ into (3.7) we have:

$$\begin{aligned} X_{t_{i+1}} &= X_{t_i} \\ &+ \int_{t_i}^{t_{i+1}} \mu(X_{t_i}, \theta)dt + \int_{t_i}^{t_{i+1}} \sigma(X_{t_i}, \theta)dW_t + \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \sigma_x(X_s, \theta)\sigma(X_s, \theta)dW_s dW_t \\ &+ \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \left[\mu_x(X_s, \theta)\mu(X_s, \theta) + \frac{1}{2}(\mu_{xx}(X_s, \theta)\sigma^2(X_s, \theta))\right] ds dt \\ &+ \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \left[\sigma_x(X_s, \theta)\mu(X_s, \theta) + \frac{1}{2}(\sigma_{xx}(X_s, \theta)\sigma^2(X_s, \theta))\right] ds dW_t \\ &+ \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \mu_x(X_s, \theta)\sigma(X_s, \theta)dW_s dt + \int_{t_i}^{t_{i+1}} J_{t-}dN_t \\ &= X_{t_{i+1}-} + \int_{t_i}^{t_{i+1}} J_{t-}dN_t \end{aligned} \qquad (3.8)$$

From (3.8) we deduce the following approximation $Y^\Delta = \{Y_{t_i}, i = 0, \dots, N\}$ of X, where:

$$Y_{t_{i+1}} = Y_{t_{i+1}-} + \int_{t_i}^{t_{i+1}} J(Y_{t_{i+1}-})dN_t \tag{3.9}$$

where

$$Y_{t_{i+1}-} = Y_{t_i} + \int_{t_i}^{t_{i+1}} \mu(Y_{t_i}, \theta)dt + \int_{t_i}^{t_{i+1}} \sigma(Y_{t_i}, \theta)dW_t + \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \sigma_x(Y_s, \theta)\sigma(Y_s, \theta)dW_s dW_t$$

Y^Δ is a jump adapted scheme and it converges strongly at time t with order one to the solution X of the SDE(1.2), seeBruti-Liberati and Platen (2007). In the following section we provide the proof of its convergence in the one dimension case which is a known result in the general form and for mark dependent jump, but in our case is adapted to a non mark dependent jump diffusion process, where the Poisson measure is changed.

The following section presents a complete proof of the convergence of the scheme described above to the exact solution of the model in consideration.

3.2.2 Convergence

Before stating the result of the convergence we have to define a certain compact notation that we use to make the proof explicit.

For $m \in \mathbb{N}$, we define the set of all multi-indices α as follows:
$\mathcal{M}_m = \{\alpha = (j_1, j_2, \dots, j_l); j_i \in \{0, 1, \dots, m\}; i \in \{1, \dots, l\}, \text{ for } l \in \mathbb{N}\} \cup \{v\}$ where v is the multi-index of length zero.
$n(\alpha)$ the number of component of α that equals to 0, $\alpha-$ denotes the multi index obtained by deleting the last component of α, $-\alpha$ is the one obtained deleting the first component of α.
Let $\bar{\mathcal{L}}^0$ be the set of functions $f(t, x) : [0, T] \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ from $\mathcal{C}^{1,2}$ and $\bar{\mathcal{L}}^k$ the set of $f(t, x)$ with partial derivatives $\frac{\partial}{\partial x^i} f(t, x), i \in \{1, \dots, d\}$, we also introduce

the following operator for a function $f(t,x) \in \bar{\mathcal{L}}^k$

$$\bar{L}^{(0)} f(t,x) := \frac{\partial}{\partial t} f(t,x) + \sum_{i=1}^{d} \mu^i(t,x) \frac{\partial}{\partial x^i} f(t,x) + \frac{1}{2} \sum_{i,r=1}^{d} \sum_{j=1}^{m} \sigma^{i,j} \sigma^{r,j} \frac{\partial^2}{\partial x^i \partial x^j} f(t,x) \tag{3.10}$$

and

$$\bar{L}^{(k)} f(t,x) := \sum_{i=1}^{d} \sigma^{i,k} \frac{\partial}{\partial x^i} f(t,x), \ \ \text{k} \ \in \{1,\dots,m\} \ \forall t \in [0,T] \text{ and } x \in \mathbb{R}^d \tag{3.11}$$

The Itô coefficient function $\bar{f}_\alpha$ for all $\alpha \in \mathcal{M}_m$ and for $f(t,x) : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is defined by:

$$\bar{f}_\alpha(t,x) := \begin{cases} f(t,x) & \text{for} \quad l(\alpha) = 0 \\ f(t,\mu(t,x)) & \text{for} \quad l(\alpha) = 1, j_1 = 0 \\ f(t,\sigma^{j_1}(t,x)) & \text{for} \quad l(\alpha) = 1, j_1 \in \{1,\dots,m\} \\ \bar{L}^{(j_1)} \bar{f}_{-\alpha}(t,x) & \text{for} \quad l(\alpha) \geq 2, j_1 \in \{0,\dots,m\} \end{cases} \tag{3.12}$$

We also require to define some compact notation to represent the multiple stochastic integral obtained from the Itô's Taylor expansion. So it is fundamental to define those integral on an appropriate set of integrable function. We define

$$\begin{aligned} \mathcal{H}_{(v)} &= \left\{ g : \sup_{t \in [0,T]} E(|g(t)|) < \infty \right\} \\ \mathcal{H}_{(0)} &= \left\{ g : E\left(\int_0^T |g(s)| ds \right) < \infty \right\} \\ \mathcal{H}_{(j)} &= \left\{ g : E\left(\int_0^T |g(s)|^2 ds \right) < \infty \right\}, \end{aligned} \tag{3.13}$$

for $j \in \{1, 2, \dots m\}$, and $\mathcal{H}_{(\alpha)}$ define the set of all predictable process $g()$ for which the following multiple stochastic integral is defined.

$$I_\alpha[g()]_{\rho,\tau} = \begin{cases} g(\rho) & \text{if} \ \ l = 0 \ \ \text{and} \ \ \alpha = v \\ \int_\rho^\tau I_{\alpha-}[g()]_{\rho,z} dz & \text{if} \ \ l \geq 1 \ \ \text{and} \ \ j_l = 0 \\ \int_\rho^\tau I_{\alpha-}[g()]_{\rho,z} dW_z^{j_l} & \text{if} \ \ l \geq 1 \ \ \text{and} \ \ j_l \in \{1,2,\dots m\} \end{cases} \tag{3.14}$$

where ρ and τ are two stopping times with $0 \leq \rho \leq \tau \leq T, a.s.$

In the more explicit case i.e in dimension one where $m = 1$ and $d = 1$, we have:

$(3.10) \Longrightarrow \bar{L}^{(0)} f(t,x) = \frac{\partial}{\partial t} f(t,x) + \mu(t,x) \frac{\partial}{\partial x} f(t,x) + \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2} f(t,x)$

and

$(3.11) \Longrightarrow \bar{L}^{(1)} f(t,x) := \sigma \frac{\partial}{\partial x} f(t,x), \quad \forall t \in [0,T] \text{and} x \in \mathbb{R}.$

Then considering the function $f(t,x) = x$ we have:

$$(3.12) \Longrightarrow \bar{f}_\alpha(t,x) = \begin{cases} x & \text{for} \quad l(\alpha) = 0 \\ \mu(t,x) & \text{for} \quad l(\alpha) = 1, j_1 = 0 \\ \sigma(t,x) & \text{for} \quad l(\alpha) = 1, j_1 = 1 \\ \bar{L}^{(j_1)} \bar{f}_{-\alpha}(t,x) & \text{for} \quad l(\alpha) \geq 2, j_1 \in \{0,1\} \end{cases} \tag{3.15}$$

and

$$(3.14) \Longrightarrow I_\alpha[g()]_{\rho,\tau} = \begin{cases} g(\rho) & \text{if } \ l = 0 \ \text{ and } \ \alpha = v \\ \int_\rho^\tau I_{\alpha-}[g()]_{\rho,z} dz & \text{if } \ l \geq 1 \ \text{ and } \ j_l = 0 \\ \int_\rho^\tau I_{\alpha-}[g()]_{\rho,z} dW_z & \text{if } \ l \geq 1 \ \text{ and } \ j_l = 1 \end{cases} \tag{3.16}$$

Let's define now the set for $\gamma = 1$, the hierachical set

$$\mathcal{A}_\gamma := \{\alpha \in \mathcal{M}_m : l(\alpha) + n(\alpha) \leq 2\gamma \ \ \},$$

we also define the remainder set $\mathcal{B}(\mathcal{A}_\gamma)$ of $\mathcal{A}_\gamma$ by

$$\mathcal{B}(\mathcal{A}_\gamma) = (\alpha \in \mathcal{M}_m - \mathcal{A}, -\alpha \in \mathcal{A})$$

For illustration we found that: $\mathcal{A}_\gamma = \{v; (0); (1); (1,1)\}$ and $\mathcal{B}(\mathcal{A}_\gamma) = \{(0,0); (0,1); (1,0)\}$.

Using (3.16) and (3.15) we can write $Y_{t_{i+1}-}$ and $X_{t_{i+1}-}$ as follows:

$$Y_{t_{i+1}-} = Y_{t_i} + \sum_{\alpha \in \mathcal{A}_\gamma \setminus v} I_\alpha[\bar{f}_\alpha(t_i, Y_{t_i})]_{t_i,t_{i+1}} = \sum_{\alpha \in \mathcal{A}_\gamma} I_\alpha[\bar{f}_\alpha(t, Y_{t_i})]_{t_i,t_{i+1}} \tag{3.17}$$

$$X_{t_{i+1}-} = X_{t_i} + \sum_{\alpha \in \mathcal{A}_\gamma \backslash v} I_\alpha[\bar{f}_\alpha(t_i, Y_{t_i})]_{t_i,t_{i+1}} + \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} I_\alpha[\bar{f}_\alpha(t_i, X_{t_i})]_{t_i,t_{i+1}} \tag{3.18}$$

(3.17) and (3.18) are called Wagner-Platen expansion see Bruti-Liberati and Platen (2007)

Definition 3.2.1 *The scheme Y^Δ is said to converge strongly with order $\gamma > 0$ at time T to the solution X of a given SDE if there exists a positive constant C such that:*

$$\varepsilon(\Delta) = \sqrt{\mathbb{E}(\mid X_T - Y_T \mid^2)} \leq C\Delta^\gamma$$

Assumption 3.2.1 *For $\gamma = 1$, let $Y^\Delta = \{Y_t, t \in [0,T]\}$ be the order γ jump adapted time discretization with step size Δ. Let us also consider the following assumptions:*

$$E(X_0) \leq \infty \quad and \quad E(|X_0 - Y_0|^2) \leq K\Delta^{2\gamma}. \tag{3.19}$$

For $\alpha \in \mathcal{A}_\gamma, t \in [0,T]$ and $x, y \in \mathbb{R}^d$, the Itô coefficients satisfy the following Lipschitz condition

$$\mid \bar{f}_\alpha(t,x) - \bar{f}_\alpha(t,y) \mid \leq K_1 \mid x - y \mid . \tag{3.20}$$

For all $\gamma \in \mathcal{A}_\gamma \cup \mathcal{B}(\mathcal{A}_\gamma)$

$$\bar{f}_{-\alpha} \in \mathcal{C}^{1,2} and \bar{f}_\alpha \in \mathcal{H}_\alpha, \tag{3.21}$$

for $t \in [0,T]$ and $x \in \mathbb{R}^d$

$$\mid \bar{f}_\alpha(t,x) \mid^2 \leq K_2(1 + |x|^2). \tag{3.22}$$

Theorem 3.2.1 *Under the Assumptions(3.2.1), the following inequality*

$$\sqrt{\mathbb{E}\left(\sup_{0\leq s\geq T} \mid X_s - Y_s \mid^2 / \mathcal{A}_0\right)} \leq K_3 \Delta^{\gamma} \tag{3.23}$$

holds where the constant K_3 does not depend on Δ

Before proving the above theorem we have to get some preliminary results. The following lemma is about the boundedness of the Itô multi-stochastic integral.

Lemma 3.2.1 *For $\alpha \in \mathcal{M}_m \setminus \{v\}, g \in \mathcal{H}_\alpha$ and for two stopping times τ_1 and τ_2, $\mathcal{F}_{\tau_1}$-measurable and satisfying the conditions $t_0 \leq \tau_1 \leq \tau_2 \leq \tau_1 + \Delta \leq T, a.s$ Then,*

$$\Phi_{\tau_2}^{\alpha} = E\left(\sup_{\tau_1\leq s\leq \tau_2} \mid I_\alpha[g(.)]_{\tau_1,s} \mid^2 |\mathcal{F}_{\tau_1}\right) \leq 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \int_{\tau_1}^{\tau_2} P_{\tau_1,z}, \tag{3.24}$$

where

$$P_{\tau_1,z} = E\left(\sup_{\tau_1\leq t\leq z} |g(t)|\,|\mathcal{F}_{\tau_1}\right) < \infty \tag{3.25}$$

for $z \in [\tau_1, \tau_2]$.

Proof: The proof of (3.24) is done by induction on the different cases of α, applying the Cauchy-Schwarz inequality we get: When $\alpha = (0)$ with $l(\alpha) = 1$ and $n(\alpha) = 1$

$$\left|\int_{\tau_1}^{s} g(z)dz\right|^2 \leq (s-\tau_1)\int_{\tau_1}^{s} |g(z)|^2 dz. \tag{3.26}$$

That leads to:

$$\begin{aligned}
\Phi_{\tau_2}^{\alpha} &= E\left(\sup_{\tau_1\leq s\leq \tau_2}\left|\int_{\tau_1}^{s} g(z)dz\right|^2 \middle| \mathcal{F}_{\tau_1}\right) \\
&\leq E\left(\sup_{\tau_1\leq s\leq \tau_2}(s-\tau_1)\int_{\tau_1}^{s}|g(z)|^2 dz \middle| \mathcal{F}_{\tau_1}\right) \\
&= E\left((\tau_2-\tau_1)\int_{\tau_1}^{\tau_2}|g(z)|^2 dz \middle| \mathcal{F}_{\tau_1}\right) \\
&\leq \Delta E\left(\int_{\tau_1}^{\tau_2}|g(z)|^2 dz \middle| \mathcal{F}_{\tau_1}\right)
\end{aligned} \tag{3.27}$$

$$
\begin{aligned}
\Phi^{\alpha}_{\tau_2} &\leq \Delta \int_{\tau_1}^{\tau_2} E(|g(z)|^2 \big| \mathcal{F}_{\tau_1}) dz \quad \text{since} \quad \tau_2 \quad is \quad \mathcal{F}_{\tau_1} - measurable \\
&= 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \int_{\tau_1}^{\tau_2} P_{\tau_1,z}, \qquad (3.28)
\end{aligned}
$$

When $\alpha = (1)$ with $l(\alpha) = 1$ and $n(\alpha) = 0$. The process

$$
\{I_{\alpha}[g(.)]_{\tau_1,t}, t \in [\tau_1, T]\} = \left\{ \int_{\tau_1}^{t} g(s) dW_s, t \in [\tau_1, T] \right\} \qquad (3.29)
$$

is a martingale. We have

$$
\begin{aligned}
\Phi^{\alpha}_{\tau_2} &= E\left(\sup_{\tau_1 \leq s \leq \tau_2} \left| \int_{\tau_1}^{s} g(z) dW_z \right|^2 \middle| \mathcal{F}_{\tau_1} \right) \\
&\leq 4E\left(\left| \int_{\tau_1}^{s} g(z) dW_z \right|^2 \middle| \mathcal{F}_{\tau_1} \right) \quad \text{from the application of Doob's inequality} \\
&= 4E\left(\int_{\tau_1}^{\tau_2} |g(z)|^2 dz \middle| \mathcal{F}_{\tau_1} \right) \quad \text{from the isometry formula}
\end{aligned}
$$

Since τ_2 is $\mathcal{F}_{\tau_1}$-measurable we have

$$
\begin{aligned}
\Phi^{\alpha}_{\tau_2} &\leq E\left(\int_{\tau_1}^{\tau_2} |g(z)|^2 dz \middle| \mathcal{F}_{\tau_1} \right) \\
&= 4 \int_{\tau_1}^{\tau_2} E\left(|g(t)|^2 \big| \mathcal{F}_{\tau_1} \right) dz \\
&\leq 4 \int_{\tau_1}^{\tau_2} E\left(\sup_{\tau_1 \leq t \leq z} |g(t)|^2 \middle| \mathcal{F}_{\tau_1} \right) dz \\
&= 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \int_{\tau_1}^{\tau_2} P_{\tau_1,z}, \qquad (3.30)
\end{aligned}
$$

When $\alpha = (j_1, j_2)$ with $j_2 = 0$ then $\forall j_1$ we have by applying the Cauchy-

Schwarz inequality we get:

$$\begin{aligned}
\Phi^{\alpha}_{\tau_2} &= E\left(\sup_{\tau_1 \le s \le \tau_2} \left| \int_{\tau_1}^{s} I_{\alpha-}[g(.)]_{\tau_1,z} dz \right|^2 \middle| \mathcal{F}_{\tau_1} \right) \\
&\le E\left(\sup_{\tau_1 \le s \le \tau_2} (s-\tau_1) \int_{\tau_1}^{s} |I_{\alpha-}[g(.)]_{\tau_1,z}|^2 dz \middle| \mathcal{F}_{\tau_1} \right) \\
&= E\left((\tau_2-\tau_1) \int_{\tau_1}^{\tau_2} |I_{\alpha-}[g(.)]_{\tau_1,z}|^2 dz \middle| \mathcal{F}_{\tau_1} \right) \\
&\le \Delta E\left(\int_{\tau_1}^{\tau_2} |I_{\alpha-}[g(.)]_{\tau_1,z}|^2 dz \middle| \mathcal{F}_{\tau_1} \right) \\
&\le \Delta E\left(\int_{\tau_1}^{\tau_2} \sup_{\tau_1 \le s \le \tau_2} |I_{\alpha-}[g(.)]_{\tau_1,s}|^2 dz \middle| \mathcal{F}_{\tau_1} \right) \\
&= \Delta E\left(\int_{\tau_1}^{\tau_2} dz \times \sup_{\tau_1 \le s \le \tau_2} |I_{\alpha-}[g(.)]_{\tau_1,s}|^2 \middle| \mathcal{F}_{\tau_1} \right) \\
&\le \Delta^2 E\left(\sup_{\tau_1 \le s \le \tau_2} |I_{\alpha-}[g(.)]_{\tau_1,s}|^2 \middle| \mathcal{F}_{\tau_1} \right) \\
&= \Delta^2 \Phi^{\alpha-}_{\tau_2}
\end{aligned} \tag{3.31}$$

Considering the previous result we conclude that,

$$\begin{aligned}
\Phi^{\alpha}_{\tau_2} &\le \Delta^2 4^{l(\alpha-)-n(\alpha-)} \Delta^{l(\alpha-)+n(\alpha-)-1} \int_{\tau_1}^{\tau_2} P_{\tau_1,z} \\
&= 4^{l(\alpha)-n(\alpha)} \Delta^{l(\alpha)+n(\alpha)-1} \int_{\tau_1}^{\tau_2} P_{\tau_1,z}
\end{aligned} \tag{3.32}$$

Since $l(\alpha) = l(\alpha-) + 1$ and $n(\alpha) = n(\alpha-) + 1$

When $\alpha = (j_1, j_2)$ with $j_2 = 1$ then $\forall j_1$, using the fact that the process

$$\{I_{\alpha}[g(.)]_{\tau_1,t}, t \in [\tau_1, T]\} \tag{3.33}$$

is a martingale we obtain by the application of Doob's inequality and Itô's isometry we get the expected result as follows

$$\begin{aligned}
\Phi^{\alpha}_{\tau_2} &= E\left(\sup_{\tau_1\le s\le\tau_2}\left|\int_{\tau_1}^{s} I_{\alpha-}[g(.)]_{\tau_1,z}dW_z\right|^2\middle|\mathcal{F}_{\tau_1}\right)\\
&\le 4E\left(\left|\int_{\tau_1}^{\tau_2} I_{\alpha-}[g(.)]_{\tau_1,z}dW_z\right|^2\middle|\mathcal{F}_{\tau_1}\right) \text{ Doob's inequality}\\
&= 4E\left(\int_{\tau_1}^{\tau_2} |I_{\alpha-}[g(.)]_{\tau_1,z}|^2dz\middle|\mathcal{F}_{\tau_1}\right) \text{ Itô's isometry}\\
&\le 4\left(\int_{\tau_1}^{\tau_2}\sup_{\tau_1\le s\le\tau_2} |I_{\alpha-}[g(.)]_{\tau_1,s}|^2dz\middle|\mathcal{F}_{\tau_1}\right)\\
&= 4\left(\int_{\tau_1}^{\tau_2}dz\times\sup_{\tau_1\le s\le\tau_2} |I_{\alpha-}[g(.)]_{\tau_1,s}|^2\middle|\mathcal{F}_{\tau_1}\right)\\
&= 4\left((\tau_2-\tau_1)\sup_{\tau_1\le s\le\tau_2} |I_{\alpha-}[g(.)]_{\tau_1,s}|^2\middle|\mathcal{F}_{\tau_1}\right)\\
&\le 4\Delta\left(\sup_{\tau_1\le s\le\tau_2} |I_{\alpha-}[g(.)]_{\tau_1,s}|^2\middle|\mathcal{F}_{\tau_1}\right)\\
&= 4^{l(\alpha-)+1-n(\alpha-)}\Delta^{l(\alpha-)+1+n(\alpha-)-1}\int_{\tau_1}^{\tau_2}P_{\tau_1,z}\\
&= 4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{\tau_1}^{\tau_2}P_{\tau_1,z} \qquad (3.34)
\end{aligned}$$

Since $l(\alpha) = l(\alpha-) + 1$ and $n(\alpha) = n(\alpha-)$ and this completes the proof of the lemma(3.2.1).

Lemma 3.2.2 *For α a given multi-index in $\mathcal{M}\{v\}$, a time discretization $\{t_i, i = 0\ldots N\}$ with step size $\Delta \in (0,1)$, and $g \in \mathcal{H}_\alpha$*

$$P_{t_0,u} = E\left(\sup_{t_0\le z\le u}|g(z)|^2\middle|\mathcal{F}_{t_0}\right) < \infty \qquad (3.35)$$

and

$$\Phi^{\alpha}_{t} = E\left(\sup_{t_0\le z\le t}\left|\sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}} + I_\alpha[g(.)]_{t_{i_z},z}\right|^2\middle|\mathcal{F}_{t_0}\right) \qquad (3.36)$$

where

$$i_t = \max\{i \in \{0,1\ldots\} : t_i \le t\} \qquad (3.37)$$

Then

$$\Phi_t^\alpha = \begin{cases} (t-t_0)\Delta^{2(l(\alpha)-1)} \int_{t_0}^t P_{t_0,u} du & when: \quad l(\alpha) = n(\alpha) \\ 4^{l(\alpha)-n(\alpha)+2} \Delta^{(l(\alpha)+n(\alpha)-1} \int_{t_0}^t P_{t_0,u} du & when: \quad l(\alpha) \neq n(\alpha) \end{cases} \tag{3.38}$$

Proof:

1. From the definition of i_z, for $z \in [t_i, t_{i+1})$, the relation $t_{i_z} = t_i$ holds. Then, for a multi index $\alpha = (j_1, j_2)$ with $j_2 = 0$, we have

$$\begin{aligned} & \sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}} + I_\alpha[g(.)]_{t_{i_z},z} \\ = & \sum_{i=0}^{i_z-1} \int_{t_i}^{t_{i+1}} I_{\alpha-}[g(.)]_{t_i,s} ds + \int_{t_{i_z}}^{z} I_{\alpha-}[g(.)]_{t_{i_z},s} ds \\ = & \sum_{i=0}^{i_z-1} \int_{t_i}^{t_{i+1}} I_{\alpha-}[g(.)]_{t_{i_s},s} ds + \int_{t_{i_z}}^{z} I_{\alpha-}[g(.)]_{t_{i_s},s} ds \\ = & \int_{t_0}^{z} I_{\alpha-}[g(.)]_{t_{i_s},s} ds \end{aligned} \tag{3.39}$$

For $\alpha = (j_1, j_2)$ with $j_2 = 1$, we have

$$\begin{aligned} & \sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}} + I_\alpha[g(.)]_{t_{i_z},z} \\ = & \sum_{i=0}^{i_z-1} \int_{t_i}^{t_{i+1}} I_{\alpha-}[g(.)]_{t_i,s} dW_s + \int_{t_{i_z}}^{z} I_{\alpha-}[g(.)]_{t_{i_z},s} dW_s \\ = & \sum_{i=0}^{i_z-1} \int_{t_i}^{t_{i+1}} I_{\alpha-}[g(.)]_{t_{i_s},s} dW_s + \int_{t_{i_z}}^{z} I_{\alpha-}[g(.)]_{t_{i_s},s} dW_s \\ = & \int_{t_0}^{z} I_{\alpha-}[g(.)]_{t_{i_s},s} dW_s \end{aligned} \tag{3.40}$$

2. Let us consider the case $l(\alpha) = n(\alpha)$

$$
\begin{aligned}
\Phi_t^\alpha &= E\left(\sup_{t_0\le z\le t}\left|\int_{t_0}^{z} I_{\alpha-}[g(.)]_{t_{i_s},s}ds\right|^2 \middle| \mathcal{F}_{t_0}\right) \\
&\le E\left(\sup_{t_0\le z\le t}(z-t_0)\int_{t_0}^{z}|I_{\alpha-}[g(.)]_{t_{i_s},s}|^2 ds\middle| \mathcal{F}_{t_0}\right) \quad \text{Cauchy-Schwarz inequality} \\
&\le (t-t_0)E\left(\int_{t_0}^{t}|I_{\alpha-}[g(.)]_{t_{i_s},s}|^2 ds\middle| \mathcal{F}_{t_0}\right) \\
&= (t-t_0)\int_{t_0}^{t}E\left(|I_{\alpha-}[g(.)]_{t_{i_s},s}|^2\middle| \mathcal{F}_{t_0}\right)ds \\
&\le (t-t_0)\int_{t_0}^{t}E\left(\sup_{t_{i_s}\le z\le s}|I_{\alpha-}[g(.)]_{t_{i_s},z}|^2\middle| \mathcal{F}_{t_0}\right)ds \\
&= (t-t_0)\int_{t_0}^{t}E\left[\underbrace{E(\sup_{t_{i_s}\le z\le s}|I_{\alpha-}[g(.)]_{t_{i_s},z}|^2|\mathcal{F}_{t_{i_s}})}_{=\Phi_s^{\alpha-}}\middle| \mathcal{F}_{t_0}\right]ds
\end{aligned} \tag{3.41}
$$

Since $t_0 \le t_{i_s}$ and then $\mathcal{F}_{t_0} \subseteq \mathcal{F}_{t_{i_s}}$ for $s \in [t_0, t]$.

From Lemma(3.2.1) we subtitute the value of $\Phi_s^{\alpha-}$ into (3.41) and yields

$$
\begin{aligned}
\Phi_t^\alpha &\le (t-t_0)4^{l(\alpha-)-n(\alpha-)} \\
&\quad \times \int_{t_0}^{t}E\left(\Delta^{l(\alpha-)+n(\alpha-)-1}\int_{t_{i_s}}^{s}P_{t_{i_s},z}dz\middle| \mathcal{F}_{t_0}\right)ds \\
&\le (t-t_0)4^{l(\alpha-)-n(\alpha-)}\int_{t_0}^{t}E\left(\Delta^{l(\alpha-)+n(\alpha-)-1}(s-t_{i_s})P_{t_{i_s},s}\middle| \mathcal{F}_{t_0}\right)ds \\
&\le (t-t_0)4^{l(\alpha-)-n(\alpha-)}\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^{t}E\left(P_{t_{i_s},s}\middle| \mathcal{F}_{t_0}\right)ds
\end{aligned} \tag{3.42}
$$

since $(s - t_{i_s}) \le \Delta$. Given that $\mathcal{F}_{t_0} \subseteq \mathcal{F}_{t_{i_s}}$ for $s \in [t_0, t]$

$$
\begin{aligned}
E\left(P_{t_{i_s},s}\middle| \mathcal{F}_{t_0}\right) &= E\left(E(\sup_{t_{i_s}\le u\le s}|g(u)|^2|\mathcal{F}_{t_{i_s}})\middle| \mathcal{F}_{t_0}\right) \\
&= E\left(\sup_{t_{i_s}\le u\le s}|g(u)|^2\middle| \mathcal{F}_{t_0}\right) \\
&\le E\left(\sup_{t_0\le u\le s}|g(u)|^2\middle| \mathcal{F}_{t_0}\right) \\
&= P_{t_0,s}
\end{aligned} \tag{3.43}
$$

It then follows that

$$\begin{aligned}
\Phi_t^\alpha &\le (t-t_0)4^{l(\alpha-)-n(\alpha-)}\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^t P_{t_0,s}ds \\
&= (t-t_0)\Delta^{2(l(\alpha)-1)}
\end{aligned} \tag{3.44}$$

This completes the proof of the case $l(\alpha)=n(\alpha)$.

3. Let us now consider the case for the multi-index $\alpha=(j_1,j_2)$ such that $l(\alpha)\neq n(\alpha)$ and $j_2=1$. For that we know that the multiple stochastic integral $\left\{\int_{t_0}^z I_{\alpha-}[g(.)]_{t_{i_s},s}dW_s\right\}$ is a martingale. We then have

$$\begin{aligned}
\Phi_t^\alpha &= E\left(\sup_{t_0\le z\le t}|\int_{t_0}^z I_{\alpha-}[g(.)]_{t_{i_s},s}dW_s|^2\middle|\mathcal{F}_{t_0}\right) \\
&\le 4E\left(|\int_{t_0}^z I_{\alpha-}[g(.)]_{t_{i_s},s}dW_s|^2\middle|\mathcal{F}_{t_0}\right) \text{ Doob's inequality} \\
&= 4E\left(\int_{t_0}^z |I_{\alpha-}[g(.)]_{t_{i_s},s}|^2ds\middle|\mathcal{F}_{t_0}\right) \text{ Itô's isometry} \\
&\le \int_{t_0}^z E\left(|I_{\alpha-}[g(.)]_{t_{i_s},s}|^2\middle|\mathcal{F}_{t_0}\right)ds
\end{aligned} \tag{3.45}$$

$$\begin{aligned}
\Phi_t^\alpha &\le \int_{t_0}^z E\left(E(|I_{\alpha-}[g(.)]_{t_{i_s},s}|^2|\mathcal{F}_{t_{i_s}})\middle|\mathcal{F}_{t_0}\right)ds \\
&\le 4E\left(\underbrace{E(\sup_{t_{i_s}\le z\le s}|I_{\alpha-}[g(.)]_{t_{i_s},z}|^2|\mathcal{F}_{t_{i_s}})}_{=\Phi_s^{\alpha-}}\middle|\mathcal{F}_{t_0}\right)ds \\
&\le 44^{l(\alpha-)-n(\alpha-)} \\
&\quad\times\int_{t_0}^t E\left(\Delta^{l(\alpha-)+n(\alpha-)-1}\int_{t_{i_s}}^s P_{t_{i_s},z}dz\middle|\mathcal{F}_{t_0}\right)ds \\
&\le 4^{l(\alpha-)-n(\alpha-)+1}\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^t E\left(P_{t_{i_s},s}\middle|\mathcal{F}_{t_0}\right)ds
\end{aligned} \tag{3.46}$$

And as the previous proof this leads to

$$\begin{aligned}
\Phi_t^\alpha &\le 4^{l(\alpha-)-n(\alpha-)+1}\Delta^{l(\alpha-)+n(\alpha-)}\int_{t_0}^t P_{t_0,s}ds \\
&= 4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^t P_{t_0,s}ds
\end{aligned} \tag{3.47}$$

since $l(\alpha-) = l(\alpha) - 1$ and $n(\alpha-) = n(\alpha)$.

4 To complete the proof we finally consider the case of $\alpha = (j_1, j_2)$ such that $l(\alpha) \neq n(\alpha)$ and $j_2 = 0$. Applying Cauchy-Schwarz inequality on Φ_t^α, we get

$$\begin{aligned}\Phi_t^\alpha &= E\left(\sup_{t_0\le z\le t}\left|\sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}} + I_\alpha[g(.)]_{t_{i_z},z}\right|^2\middle|\mathcal{F}_{t_0}\right)\\ &\le 2E\left(\sup_{t_0\le z\le t}\left|\sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}}\right|^2\middle|\mathcal{F}_{t_0}\right)\\ &\quad +2E\left(\sup_{t_0\le z\le t}\left|I_\alpha[g(.)]_{t_{i_z},z}\right|^2\middle|\mathcal{F}_{t_0}\right) \qquad (3.48)\end{aligned}$$

$\sum_{i=0}^{n} I_\alpha[g(.)]_{t_i,t_{i+1}} + I_\alpha[g(.)]_{t_{i_z},z}$ is a discrete time martingale, see Kloeden (1992). Then considering the first term of (3.48) we have,

$$\begin{aligned}& E\left(\sup_{t_0\le z\le t}\left|\sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}}\right|^2\middle|\mathcal{F}_{t_0}\right)\\ \le\ & 4E\left(\left|\sum_{i=0}^{i_t-1} I_\alpha[g(.)]_{t_i,t_{i+1}}\right|^2\middle|\mathcal{F}_{t_0}\right) \text{ Doob's Inequality}\\ =\ & 4E\left(\left|\sum_{i=0}^{i_t-2} I_\alpha[g(.)]_{t_i,t_{i+1}} + I_\alpha[g(.)]_{t_{i_t-1},t_{i_t}}\right|^2\middle|\mathcal{F}_{t_0}\right)\\ \le\ & 4E\Bigg(\Bigg[|\sum_{i=0}^{i_t-2} I_\alpha[g(.)]_{t_i,t_{i+1}}|^2 + 2|\sum_{i=0}^{i_t-2} I_\alpha[g(.)]_{t_i,t_{i+1}}|E(|I_\alpha[g(.)]_{t_{i_t-1},t_{i_t}}|\big|\mathcal{F}_{t_{i_t}-1})\\ & + E(|I_\alpha[g(.)]_{t_{i_t-1},t_{i_t}}|^2\big|\mathcal{F}_{t_{i_t}-1}\big|\mathcal{F}_{t_0}\Big)\\ \le\ & 4E\Bigg(\Bigg[|\sum_{i=0}^{i_t-2} I_\alpha[g(.)]_{t_i,t_{i+1}}|^2\\ & + E(|I_\alpha[g(.)]_{t_{i_t-1},t_{i_t}}|^2\big|\mathcal{F}_{t_{i_t}-1})\big|\mathcal{F}_{t_0}\Big)\end{aligned} \qquad (3.49)$$

since from the discrete martingale property of the stochastic integral $I_\alpha[g(.)]_{t_{i_t-1},t_{i_t}}$ we have $E(I_\alpha[g(.)]_{t_{i_t-1},t_{i_t}}\big|\mathcal{F}_{t_{i_t}-1} = 0$.

By applying Lemma(3.2.1) on $E(|I_\alpha[g(.)]_{t_{i_t-1},t_{i_t}}|^2 \big| \mathcal{F}_{t_{i_t}-1})$ we get

$$\begin{aligned} & E\left(\sup_{t_0\le z\le t}\left|\sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}}\right|^2 \middle| \mathcal{F}_{t_0}\right) \\ \le\ & 4E\left(\left[|\sum_{i=0}^{i_t-2} I_\alpha[g(.)]_{t_i,t_{i+1}}|^2\right.\right. \\ & \left.\left. +\ 4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{i_t-1}}^{t_{i_t}} P_{t_{i_t-1},u}du\right] \middle| \mathcal{F}_{t_0}\right) \end{aligned}$$

Expanding $\sum_{i=0}^{i_t-2} I_\alpha[g(.)]_{t_i,t_{i+1}}$ as $\sum_{i=0}^{i_t-1} I_\alpha[g(.)]_{t_i,t_{i+1}}$ we obtain

$$\begin{aligned} & E\left(\sup_{t_0\le z\le t}\left|\sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}}\right|^2 \middle| \mathcal{F}_{t_0}\right) \\ \le\ & 4E\left(\left[|\sum_{i=0}^{i_t-3} I_\alpha[g(.)]_{t_i,t_{i+1}}|^2\right.\right. \\ & +4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{i_t-2}}^{t_{i_t}-1} P_{t_{i_t-2},u}du \\ & \left.\left. +\ 4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{i_t-1}}^{t_{i_t}} P_{t_{i_t-1},u}du\right] \middle| \mathcal{F}_{t_0}\right) \\ \le\ & 4E\left(\left[|\sum_{i=0}^{i_t-3} I_\alpha[g(.)]_{t_i,t_{i+1}}|^2\right.\right. \\ & +4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{i_t-2}}^{t_{i_t}-1} P_{t_{i_t-2},u}du \\ & \left.\left. +\ 4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{i_t-1}}^{t_{i_t}} P_{t_{i_t-2},u}du\right] \middle| \mathcal{F}_{t_0}\right) \text{ since } P_{t_{i_t-1},u} \le P_{t_{i_t-2},u} \\ \le\ & 4E\left(\left[|\sum_{i=0}^{i_t-3} I_\alpha[g(.)]_{t_i,t_{i+1}}|^2\right.\right. \\ & \left.\left. +\ 4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_{i_t-2}}^{t_{i_t}} P_{t_{i_t-2},u}du\right] \middle| \mathcal{F}_{t_0}\right) \end{aligned} \tag{3.50}$$

and iteratively this leads to

$$\begin{aligned} & E\left(\sup_{t_0\le z\le t}\left|\sum_{i=0}^{i_z-1} I_\alpha[g(.)]_{t_i,t_{i+1}}\right|^2 \middle| \mathcal{F}_{t_0}\right) \\ \le\ & 4^{l(\alpha)-n(\alpha)+1}\Delta^{l(\alpha)+n(\alpha)-1}E\left(\int_{t_0}^{t} P_{t_0,u}du \middle| \mathcal{F}_{t_0}\right) \\ =\ & 4^{l(\alpha)-n(\alpha)+1}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t} P_{t_0,u}du \end{aligned} \tag{3.51}$$

Considering now the second term of equation (3.48) using Cauchy Schwarz inequality we have

$$\begin{aligned} E\left(\sup_{t_0\le z\le t}\left|I_\alpha[g(.)]_{t_{i_z},z}\right|^2\middle|\mathcal{F}_{t_0}\right) &= E\left(\sup_{t_0\le z\le t}\left|\int_{t_{i_z}}^{z} I_\alpha[g(.)]_{t_{i_z},u}du\right|^2\middle|\mathcal{F}_{t_0}\right) \\ &\le E\left(\sup_{t_0\le z\le t}(z-t_{i_z})\int_{t_{i_z}}^{z}|I_\alpha[g(.)]_{t_{i_z},u}|^2du\middle|\mathcal{F}_{t_0}\right) \end{aligned}$$

Similarly as in the proof of Lemma(3.2.1) we obtain

$$\begin{aligned} E\left(\sup_{t_0\le z\le t}\left|I_\alpha[g(.)]_{t_{i_z},z}\right|^2\middle|\mathcal{F}_{t_0}\right) &\le 4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1} \\ &\quad\times\int_{t_0}^{t}P_{t_0,u}du \end{aligned} \tag{3.52}$$

Therefore, from the equations (3.51) and (3.52) we finally get

$$\begin{aligned} \Phi_t^\alpha &\le 2\left(4^{l(\alpha)-n(\alpha)+1}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t}P_{t_0,u}du\right. \\ &\quad\left.+\,4^{l(\alpha)-n(\alpha)}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t}P_{t_0,u}du\right) \\ &\le 4^{l(\alpha)-n(\alpha)+2}\Delta^{l(\alpha)+n(\alpha)-1}\int_{t_0}^{t}P_{t_0,u}du, \end{aligned} \tag{3.53}$$

and that completes the proof of Lemma(3.2.2).

The following Lemma is an intermediary result required to handle the moment of the solution of the SDE(1.2).

Lemma 3.2.3 *Assuming conditions (1.4) and (1.5) are satisfied and let*

$$E(|X_{t_0}|^2) < \infty. \tag{3.54}$$

Then the solution X_t of (1.2) satisfies

$$E\left(\sup_{t_0\le s\le T}|X_s|^2\middle|\mathcal{F}_{t_0}\right) \le C(1+E(|X_{t_0}|^2)) \tag{3.55}$$

for every $t\in[t_0,T]$, with C a positive constant that depends only on $(T-t_0)$ and the linear growth bound.

Proof of Theorem(3.2.1)

Considering the Wagner-Platen expansion for the diffusion process as described in equations(3.17) and (3.18) the solution of the SDE(1.2) at time $t \in [0, T]$ can be written as follows:

$$\begin{aligned} X_t &= X_0 + \sum_{\alpha \in \mathcal{A}_\gamma \backslash v} \left\{ \sum_{i=o}^{i_t - 1} I_\alpha[\bar{f}_\alpha(t_i, X_{t_i})]_{t_i, t_{i+1}} + I_\alpha[\bar{f}_\alpha(t_{i_t}, X_{t_i})]_{t_{i_t}, t} \right\} \\ &+ \sum_{\alpha \in \mathcal{B}(\mathcal{A}_\gamma)} \left\{ \sum_{i=o}^{i_t - 1} I_\alpha[\bar{f}_\alpha(t_i, X_{t_i})]_{t_i, t_{i+1}} + I_\alpha[\bar{f}_\alpha(t_{i_t}, X_{t_i})]_{t_{i_t}, t} \right\} \\ &+ \int_0^t J(X_{t_{i_s}-}) dN_s, \end{aligned} \tag{3.56}$$

and the jump adapted scheme is

$$\begin{aligned} Y_t = Y_0 + \sum_{\alpha \in \mathcal{A}_\gamma \backslash v} \left\{ \sum_{i=o}^{i_t - 1} I_\alpha[\bar{f}_\alpha(t_i, Y_{t_i})]_{t_i, t_{i+1}} + I_\alpha[\bar{f}_\alpha(t_{i_t}, Y_{t_i})]_{t_{i_t}, t} \right\} \\ + \int_0^t J(Y_{t_{i_s}-}) dN_s, \end{aligned} \tag{3.57}$$

for $t \in [0, T]$. From Lemma(3.2.3) we have

$$E\left(\sup_{t_0 \le s \le T} |X_s|^2 \middle| \mathcal{F}_{t_0} \right) \le C(1 + E(|X_{t_0}|^2))$$

Now let us prove the same result for the jump adapted scheme.

$$\begin{aligned} E\left(\sup_{0 \le s \le T} |Y_s|^2 \middle| \mathcal{F}_{t_0} \right) &\le E\left(\sup_{0 \le s \le T} (1 + |Y_s|^2) \middle| \mathcal{F}_0 \right) \\ &\le E\Bigg(\sup_{0 \le s \le T} (1 + |Y_0 + \sum_{\alpha \in \mathcal{A}_\gamma \backslash v} \{ \sum_{i=o}^{i_t - 1} I_\alpha[\bar{f}_\alpha(t_i, Y_{t_i})]_{t_i, t_{i+1}} \\ &\quad + I_\alpha[\bar{f}_\alpha(t_{i_t}, Y_{t_i})]_{t_{i_t}, t} + \int_0^s J(X_{t_{i_u}}-) dN_u|^2) \Bigg| \mathcal{F}_{t_0} \Bigg) \\ &\le E\Bigg(\sup_{0 \le s \le T} (1 + 2|Y_0|^2 + 4 \Bigg| \sum_{\alpha \in \mathcal{A}_\gamma \backslash v} \{ \sum_{i=o}^{i_t - 1} I_\alpha[\bar{f}_\alpha(t_i, Y_{t_i})]_{t_i, t_{i+1}} \\ &\quad + I_\alpha[\bar{f}_\alpha(t_{i_t}, Y_{t_i})]_{t_{i_t}, t} \Big|^2 + \\ &\quad 4 \left| \int_0^s J(X_{t_{i_u}-}) dN_u \right|^2 \Bigg| \mathcal{F}_0 \Bigg) \end{aligned} \tag{3.58}$$

$$\leq E\left(\sup_{0\leq s\leq T}(1+|Y_0|^2)\right)+4\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}E\left(\sup_{0\leq s\leq T}\left|\sum_{i=o}^{i_t-1}I_\alpha[\bar{f}_\alpha(t_i,Y_{t_i})]_{t_i,t_{i+1}}\right.\right.$$

$$\left.\left.+\ I_\alpha[\bar{f}_\alpha(t_{i_t},Y_{t_i})]_{t_{i_t},t}\right|^2\right|\mathcal{F}_0\Bigg)+4E\left(\sup_{0\leq s\leq T}\left|\int_0^s J(X_{t_{i_u}-})dN_u\right|^2\middle|\mathcal{F}_0\right)$$

Then from the Lema(3.2.2) and the Linear growth condition(3.22) we have

$$\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}E\left(\sup_{0\leq s\leq T}\left|\sum_{i=o}^{i_t-1}I_\alpha[\bar{f}_\alpha(t_i,Y_{t_i})]_{t_i,t_{i+1}}+I_\alpha[\bar{f}_\alpha(t_{i_t},Y_{t_i})]_{t_{i_t},t}\right|^2\middle|\mathcal{F}_0\right)$$

$$\leq K_2\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\int_0^T E\left(\sup_{t_0\leq s\leq u}|\bar{f}_\alpha(s,Y_s)|^2\middle|\mathcal{F}_0\right)du$$

$$\leq K_2\sum_{\alpha\in\mathcal{A}_\gamma\setminus\{v\}}\int_0^T E\left(\sup_{t_0\leq s\leq u}(1+|Y_s|^2)\middle|\mathcal{F}_0\right)du$$

$$\leq C_2^{'}\int_0^T E\left(\sup_{t_0\leq s\leq u}(1+|Y_s|^2)\middle|\mathcal{F}_0\right)du \tag{3.59}$$

$$E\left(\sup_{0\leq s\leq T}\left|\int_0^s J(X_{t_{i_u}-})dN_u\right|^2\middle|\mathcal{F}_0\right)$$

$$\leq E\left(\sup_{t_0\leq s\leq T}\left|\int_0^s J(X_{t_{i_u}-})d\tilde{N}_u+\lambda\int_0^s J(X_{t_{i_u}-})du\right|^2\middle|\mathcal{F}_0\right),$$ where $\tilde{N}$ is the compensated poisson measure

$$\leq 2E\left(\sup_{t_0\leq s\leq T}\left|\int_0^s J(X_{t_{i_u}-})d\tilde{N}_u\right|^2\middle|\mathcal{F}_0\right)+2E\left(\sup_{t_0\leq s\leq T}\left|\int_0^s J(X_{t_{i_u}-})du\right|^2\middle|\mathcal{F}_0\right)$$

$$\leq 8E\left(\left|\int_0^T J(X_{t_{i_u}-}d\tilde{N}_u\right|^2\middle|\mathcal{F}_0\right)+2E\left(\sup_{t_0\leq s\leq T}\left|\int_0^s J(X_{t_{i_u}-})du\right|^2\middle|\mathcal{F}_0\right)$$

Using the Itô's isometry for jump and the linear growth condition (1.5) we get

$$E\left(\sup_{0\leq s\leq T}\left|\int_0^s J(X_{t_{i_u}-})dN_u\right|^2\middle|\mathcal{F}_0\right)\ \leq\ C_3^{'}\int_0^T E\left(\sup_{t_0\leq s\leq u}(1+|Y_s|^2)\middle|\mathcal{F}_0\right)du \tag{3.60}$$

Substituting (3.59) and (3.60) into (3.58) we have

$$E\left(\sup_{0\leq s\leq T}|Y_s|^2\middle|\mathcal{F}_0\right)\ \leq\ C_1^{'}(1+E(|Y_0|^2))+C_4^{'}\int_0^T E\left(\sup_{t_0\leq s\leq u}(1+|Y_s|^2)\middle|\mathcal{F}_0\right)du \tag{3.61}$$

Then by applying the Gronwall inequality, we obtain

$$E\left(\sup_{0\le s\le T}|Y_s|^2\bigg|\mathcal{F}_0\right) \;\le\; C^{'}(1+E(|Y_0|^2)), \tag{3.62}$$

where $C^{'}$ is a positive constant. The final step of proof consists of analyzing the mean square given by

$$\begin{aligned}
Z(t) \;&:=\; E\left(\sup_{0\le s\le t}|X_s-Y_s|^2\bigg|\mathcal{F}_0\right)\\
&\le\; E\left(\sup_{0\le s\le t}\left|X_0-Y_0+\sum_{\alpha\in\mathcal{A}_\gamma\backslash\{v\}}\{\sum_{i=0}^{i_s-1}I_\alpha[\bar f_\alpha(t_i,X_{t_i})-\bar f_\alpha(t_i,Y_{t_i})]_{t_i,t_{i+1}}\right.\right.\\
&\quad +I_\alpha[\bar f_\alpha(t_{i_s},X_{t_{i_s}})-\bar f_\alpha(t_{i_s},Y_{t_{i_s}})]_{t_{i_s},s}\\
&\quad +\sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)}\left\{\sum_{i=o}^{i_t-1}I_\alpha[\bar f_\alpha(t_i,X_{t_i})]_{t_i,t_{i+1}}+I_\alpha[\bar f_\alpha(t_{i_s},X_{t_{i_s}})]_{t_{i_s},s}\right\}\\
&\quad \left.\left.+\int_0^s\{J(X_{t_{i_u}-})-J(Y_{t_{i_u}-})\}dN_u\right|^2\Bigg|\mathcal{F}_0\right)\\
&\le\; Cste\left\{|x_0-Y_0|^2+\sum_{\alpha\in\mathcal{A}_\gamma\backslash\{v\}}A_t^\alpha+\sum_{\alpha\in\mathcal{B}(\mathcal{A}_\gamma)}B_t^\alpha+D_t\right\}
\end{aligned} \tag{3.63}$$

$\forall t\in[0,T]$ where A_t^α, B_t^α and D_t are defined as follows

$$\begin{aligned}
A_t^\alpha \;&=\; E\left(\sup_{0\le s\le t}\left|\sum_{i=0}^{i_s-1}I_\alpha[\bar f_\alpha(t_i,X_{t_i})-\bar f_\alpha(t_i,Y_{t_i})]_{t_i,t_{i+1}}\right.\right.\\
&\quad \left.\left.+\;I_\alpha[\bar f_\alpha(t_{i_s},X_{t_{i_s}})-\bar f_\alpha(t_{i_s},Y_{t_{i_s}})]_{t_{i_s},s}\right|^2\bigg|\mathcal{F}_0\right),
\end{aligned} \tag{3.64}$$

$$B_t^\alpha \;=\; E\left(\sup_{0\le s\le t}\left|\sum_{i=o}^{i_t-1}I_\alpha[\bar f_\alpha(t_i,X_{t_i})]_{t_i,t_{i+1}}+I_\alpha[\bar f_\alpha(t_{i_s},X_{t_{i_s}})]_{t_{i_s},s}\right|^2\Bigg|\mathcal{F}_0\right), \tag{3.65}$$

and

$$D_t \;=\; E\left(\sup_{0\le s\le t}\left|\int_0^s\{J(X_{t_{i_u}-})-J(Y_{t_{i_u}-})\}dN_u\right|^2\bigg|\mathcal{F}_0\right) \tag{3.66}$$

Now by using the Lemma(3.2.2)and Lipschitz condition(3.20) we get

$$
\begin{aligned}
A_t^\alpha &:= E\left(\sup_{0\le s\le t}\left|\sum_{i=0}^{i_s-1} I_\alpha[\bar{f}_\alpha(t_i,X_{t_i})-\bar{f}_\alpha(t_i,Y_{t_i})]_{t_i,t_{i+1}}\right.\right.\\
&+ \left.\left. I_\alpha[\bar{f}_\alpha(t_{i_s},X_{t_{i_s}})-\bar{f}_\alpha(t_{i_s},Y_{t_{i_s}})]_{t_{i_s},s}\right|^2\middle|\mathcal{F}_0\right)\\
&\le cste\int_0^t E\left(\sup_{0\le s\le u}|X_{t_{i_u}}-Y_{t_{i_u}}|^2\middle|\mathcal{F}_0\right)du\\
&\le cste\int_0^t Z(u)du \qquad (3.67)
\end{aligned}
$$

In the same way using the Lemma(3.2.2)and Linear growth condition(3.22) we get

$$
\begin{aligned}
B_t^\alpha &:= E\left(\sup_{0\le s\le t}\left|\sum_{i=o}^{i_t-1} I_\alpha[\bar{f}_\alpha(t_i,X_{t_i})]_{t_i,t_{i+1}}+I_\alpha[\bar{f}_\alpha(t_{i_s},X_{t_{i_s}})]_{t_{i_s},s}\right|^2\middle|\mathcal{F}_0\right)\\
&\le cste\int_0^t E\left(\sup_{0\le s\le u}|\bar{f}_\alpha(s,X_s)|^2\middle|\mathcal{F}_0\right)du\\
&\le cste\Delta^{\psi(\alpha)}\int_0^t E\left(\sup_{0\le s\le u}(1+|X_s|^2\middle|\mathcal{F}_0\right)du\\
&\le cste\Delta^{\psi(\alpha)}\left(t+\int_0^t E\left(\sup_{0\le s\le u}|X_s|^2\middle|\mathcal{F}_0\right)du\right) \qquad (3.68)
\end{aligned}
$$

where $\psi(\alpha) = \begin{cases} 2l(\alpha)-2 & : l(\alpha)=n(\alpha) \\ l(\alpha)+n(\alpha)-1 & : l(\alpha)\neq n(\alpha) \end{cases}$ So for $\alpha \in \mathcal{B}(\mathcal{A}_\gamma)$, we have $l(\alpha) \ge \gamma+1$ if $l(\alpha)=n(\alpha)$ and $l(\alpha)+n(\alpha) \ge 2\gamma$ if $l(\alpha)\neq n(\alpha)$ so that $\psi(\alpha)\ge 2\gamma$. Therefore applying Lemma(3.2.3) together with the above result we have

$$
B_t^\alpha \;\le\; Cste\Delta^{2\gamma}(1+|X_0|^2) \qquad (3.69)
$$

Now let us analyze the last term

$$
\begin{aligned}
D_t &= E\left(\sup_{0\leq s\leq t}\left|\int_0^s\{J(X_{t_{iu}-})-J(Y_{t_{iu}-})\}dN_u\right|^2\middle|\mathcal{F}_0\right)\\
&= E\left(\sup_{0\leq s\leq t}\left|\int_0^s\{J(X_{t_{iu}-})-J(Y_{t_{iu}-})\}d\tilde{N}_u\right.\right.\\
&\quad+\left.\left.\int_0^s\{J(X_{t_{iu}-})-J(Y_{t_{iu}-})\}du\right|^2\middle|\mathcal{F}_0\right)\\
&\leq 2E\left(\sup_{0\leq s\leq t}\left|\int_0^s\{J(X_{t_{iu}-})-J(Y_{t_{iu}-})\}d\tilde{N}_u\right|^2\middle|\mathcal{F}_0\right)\\
&\quad+2E\left(\sup_{0\leq s\leq t}\left|\int_0^s\{J(X_{t_{iu}-})-J(Y_{t_{iu}-})\}du\right|^2\middle|\mathcal{F}_0\right)\\
&\leq 8E\left(\int_0^t\left|J(X_{t_{iu}-})-J(Y_{t_{iu}-})\right|^2du\middle|\mathcal{F}_0\right) \quad \text{applying Doob's inequality and Itô's isometry for jump}\\
&\quad+2\lambda E\left(\int_0^t\left|J(X_{t_{iu}-})-J(Y_{t_{iu}-})\right|^2du\middle|\mathcal{F}_0\right)\\
&\leq Cste\int_0^t Z(u)du. \qquad (3.70)
\end{aligned}
$$

Then from the results (3.67), (3.69) and (3.70) we have for $t\in[0,T]$

$$Z(t)\leq cste\{E(|x_0-Y_0|^2)+\Delta^{2\gamma}(1+|X_0|^2)+\int_0^t Z(u)du \qquad (3.71)$$

Therefore for $t=T$ we have

$$Z(T)\leq cste\{E(|x_0-Y_0|^2)+\Delta^{2\gamma}(1+|X_0|^2)+\int_0^T Z(u)du \qquad (3.72)$$

Using the Assumptions(3.19) and applying the Gronwall inequality we obtain

$$
\begin{aligned}
Z(T) &\leq K\Delta^{2\gamma}+\int_0^T Z(u)du \qquad (3.73)\\
&\leq K\Delta^{2\gamma} \quad \text{where } K \text{ is a positive constant} \qquad (3.74)
\end{aligned}
$$

Finally we get the desired result that concludes the proof as

$$\sqrt{E\left(\sup_{0\leq s\leq T}|X_s-Y_s|^2\middle|\mathcal{F}_0\right)}=\sqrt{Z(T)}\leq K\Delta^{\gamma} \qquad (3.75)$$

3.3 Saddlepoint

This method originated from (?, ?) and is used in the context of Markov process by Aït-Sahalia and Yu (2006).

This method exploits the relation between the characteristic function and the probability distribution through the so called cumulative distribution function. In the following we bring down that relationship and explain how to move from the characteristic function to the probability density function. By assuming that for each state $(\Delta, x) \in \mathbb{R}_+ \times \mathbb{R}^m$ the probability density function exists, we have:

The conditional Laplace transform of the process X is defined for $u \in \mathbb{R}^m$ as

$$\phi(\Delta, u|x) = E(\exp(u^T X_\Delta)|X_0 = x) \tag{3.76}$$

where ^{T} denotes the transposition, for all values (Δ, u, x). And the domain of convergence of ϕ is denoted Λ. The cumulant generating function of X is the function defined as

$$K(\Delta, u|x) = \ln(\phi(\Delta, u|x)) \tag{3.77}$$

For given (Δ, x), K is a closed-convex function of u in $\mathbb{R}^m$ and if the variance matrix $Var(X)$ is positive definite, K is strictly convex on Λ. Also $K(\Delta, 0|x) = 0$ Derivatives of all order of j with respect to u exist and are given by

$$\frac{\partial^{r_1+\ldots+r_m}\phi(\Delta, u|x)}{\partial u^{r_1}\ldots\partial u^{r_m}} = E\left(X_{1_\Delta}^{r_1}\ldots X_{m_\Delta}^{r_m}\exp(u.X_\Delta)\middle| X_0 = x\right) \tag{3.78}$$

The characteristic function of X is the function $\phi(\Delta, iu|x)$ where $i^2 = -1$. The density function and characteristic functions are linked by the Fourier inversion formula which is presented as follows:

$$\begin{aligned} P(\Delta, y|x) &= (2\pi)^{-m}\int_{-\infty}^{\infty}\exp(-iu.y)\phi(\Delta, iu|x)du \\ &= (2\pi)^{-m}\int_{\hat{u}-i\infty}^{\hat{u}+i\infty}\exp(-u.y)\phi(\Delta, u|x)du \end{aligned} \tag{3.79}$$

The key point of the saddlepoint method is to choose the path of integration, i.e.,$\hat{u}$ in (3.79),well. Considering $\hat{u} = \hat{u}(\Delta, u|x)$ as solution in u of the equation

$$\frac{\partial K(\Delta, u|x)}{\partial u} = y \tag{3.80}$$

is found to be the optimal choice, see Aït-Sahalia and Yu (2006). The solution does exist and is unique because of the convex property of the cumulant generating function and it is only computed numerically due either to the form of k that may not be known explicitly or that Eq(3.80) is too involved, see Aït-Sahalia and Yu (2006).

The equation(3.79) can be written in the following way

$$\begin{aligned} P(\Delta, y|x) &= (2\pi)^{-m} \int_{\hat{u}-i\infty}^{\hat{u}+i\infty} \exp(-u.y) \exp(\ln(\phi(\Delta, u|x)))du \\ &= (2\pi)^{-m} \int_{\hat{u}-i\infty}^{\hat{u}+i\infty} \exp(K(\Delta, u|x) - u.y)du \end{aligned} \tag{3.81}$$

Where with a Taylor expansion of the function $u \longrightarrow K(\Delta, u|x) - u.y$ arround its minimum $\hat{u}$:

$$\begin{aligned} K(\Delta, u|x) - u.y &= K(\Delta, \hat{u}|x) - \hat{u}.y \\ &- \frac{1}{2}(u - \hat{u})^T \frac{\partial^2 K(\Delta, u|x)}{\partial u \partial u^T}(u - \hat{u})^T \\ &+ O(\|u - \hat{u}\|^3) \end{aligned} \tag{3.82}$$

On the path of integration relevant for (3.79), we have $u = \hat{u} + iv$ with $v \in \mathbb{R}^m$ hence $(u - \hat{u})$ is a purely imaginary complex vector. Thus,

$$K(\Delta, u|x) - u.y = K(\Delta, \hat{u}|x) - \hat{u}.y - \frac{1}{2}v^T \frac{\partial^2 K(\Delta, u|x)}{\partial u \partial u^T} v + O(\|v\|^3) \tag{3.83}$$

From this it follows that the leading term of an approximation to; $P(\Delta, y|x)$can be taken to be:

$$P(\Delta, y|x) = (2\pi)^{-m} \exp(K(\Delta, \hat{u}|x) - \hat{u}.y) \int_{\hat{u}-i\infty}^{\hat{u}+i\infty} \exp(-\frac{1}{2}v^T \frac{\partial^2 K(\Delta, u|x)}{\partial u \partial u^T} v)dv \tag{3.84}$$

Then we have

$$\int_{\hat{u}-i\infty}^{\hat{u}+i\infty} \exp(-\frac{1}{2}v^T \frac{\partial^2 K(\Delta, u|x)}{\partial u \partial u^T} v)dv = (2\pi)^{(-m/2)} \det(\frac{\partial^2 K(\Delta, u|x)}{\partial u \partial u^T})^{-1/2} \tag{3.85}$$

From which it follows

$$P(\Delta, y|x) = (2\pi)^{-m/2} \exp(K(\Delta, \hat{u}|x) - \hat{u}.y) \det(\frac{\partial^2 K(\Delta, u|x)}{\partial u \partial u^T})^{-1/2} \tag{3.86}$$

Chapter 4

Results And Discussion

In this chapter we first apply the discretization result to some known jump diffusion process to illustrate the approximation scheme. Further we use that discretization approach to get the characteristic function of the process considered together with the assumption made on the jump size distribution, and the probability transition is derived thereafter using the saddlepoint method described above. At the end a numerical application is conducted where the process is simulated through Monte-Carlo's approach and an estimation result is obtained.

4.1 Discretization Results

Let's recall the discretization scheme,

$$Y_{t_{i+1}} = Y_{t_{i+1}-} + \int_{t_i}^{t_{i+1}} J(Y_{t_{i+1}-}) dN_t \tag{4.1}$$

where

$$\begin{aligned}
Y_{t_{i+1}-} &= Y_{t_i} + \int_{t_i}^{t_{i+1}} \mu(Y_{t_i},\theta)dt + \int_{t_i}^{t_{i+1}} \sigma(Y_{t_i},\theta)dW_t \\
&\quad + \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} \sigma_x(Y_s,\theta)\sigma(Y_s,\theta)dW_s dW_t \\
&= Y_{t_i} + \mu(Y_{t_i},\theta)(t_{i+1}-t_i) + \sigma(Y_{t_i},\theta)(W_{t_{i+1}} - W_{t_i}) \\
&\quad +\sigma_x(Y_{t_i},\theta)\sigma(Y_{t_i},\theta) \int_{t_i}^{t_{i+1}} \int_{t_i}^{t} dW_s dW_t
\end{aligned}$$

$$\begin{aligned}
\int_{t_i}^{t_{i+1}} \int_{t_i}^{t} dW_s dW_t &= \int_{t_i}^{t_{i+1}} (W_t - W_{t_i})dW_t \\
&= \int_{t_i}^{t_{i+1}} W_t dW_t - W_{t_i} \int_{t_i}^{t_{i+1}} dW_t \\
&= \frac{1}{2}\left\{(W_{t_{i+1}}^2 - W_{t_i}^2) - (t_{i+1}-t_i)\right\} - W_{t_i}(W_{t_{i+1}} - W_{t_i}) \\
&= \frac{1}{2}\left\{(W_{t_{i+1}}^2 - W_{t_i}^2 - 2W_{t_i}W_{t_{i+1}} + 2W_{t_i}^2) - (t_{i+1}-t_i)\right\} \\
&= \frac{1}{2}\left\{(W_{t_{i+1}} - W_{t_i})^2 - (t_{i+1}-t_i)\right\}
\end{aligned}$$

Considering a uniform scheme size $h = \frac{T-0}{N}$ we have $t_{i+1} - t_i = h$ and $(W_{t_{i+1}} - W_{t_i}) = \sqrt{h}\varepsilon_i$ with $\varepsilon_i \rightsquigarrow \mathcal{N}(0,1)$ the standard normal distribution. Therefore

$$Y_{t_{i+1}-} = Y_{t_i} + \mu(Y_{t_i},\theta)h + \sigma(Y_{t_i},\theta)\sqrt{h}\varepsilon_i + \frac{1}{2}\sigma_x(Y_{t_i},\theta)\sigma(Y_{t_i},\theta)h(\varepsilon_i^2 - 1) \quad (4.2)$$

Therefore the jump adapted scheme is

$$\begin{aligned}
Y_{t_{i+1}} &= Y_{t_i} + \mu(Y_{t_i},\theta)h + \sigma(Y_{t_i},\theta)\sqrt{h}\varepsilon_i + \frac{1}{2}\sigma_x(Y_{t_i},\theta)\sigma(Y_{t_i},\theta)h(\varepsilon_i^2 - 1) \\
&\quad +J(Y_{t_{i+1}-})\Delta N_{t_i} \qquad (4.3)
\end{aligned}$$

The following is the Merton jump diffusion process defined by (Merton, 1976),described as follows:

$$dS_t = \mu S_{t-}dt + \sigma S_{t-}dW_t + S_{t-}dJ_t \quad (4.4)$$

with $J_t = \sum_{j=1}^{N_t}(Y_j - 1)$, where $Y_1, Y_2 \ldots$ are i.i.d. Log-Normal random variables. The solution of the SDE(4.4)is known and it is $S_t = S_0 e^{(\mu - \frac{1}{2})t + \sigma W(t)} \Pi_{j=1}^{N_t} Y_j$ see

Merton (1976). We use a Monte-Carlo approach, we simulate the using the JAD scheme, defined in Eq(4.1), and compare it with the exact path. Considering the discretization time $0 = t_0 < t_1 < \ldots < t_N = T$, combined with the jumps time $\tau_i, i = 1 \ldots$, the jump adapted scheme $Y^\Delta = \{Y_{t_i}, i = 0, \ldots, N\}$ of the process X is :

$$Y_{t_{i+1}} = Y_{t_{i+1}-} + \int_{t_i}^{t_{i+1}} J(Y_{t_{i+1}-}) dN_t \tag{4.5}$$

where

$$Y_{t_{i+1}-} = Y_{t_i} + \mu Y_{t_i} h + \sigma Y_{t_i} \sqrt{h} \varepsilon_i + \frac{1}{2} \sigma^2 Y_{t_i} h (\varepsilon_i^2 - 1) \tag{4.6}$$

For the following set of parameters values $\mu = 2$, $\sigma = 1$; $\lambda = 1$, $\mu_j = 0$, $\sigma_j = 2$, $T = 1$; we obtained:

- Number of iterations N=100 and step size h=1/100

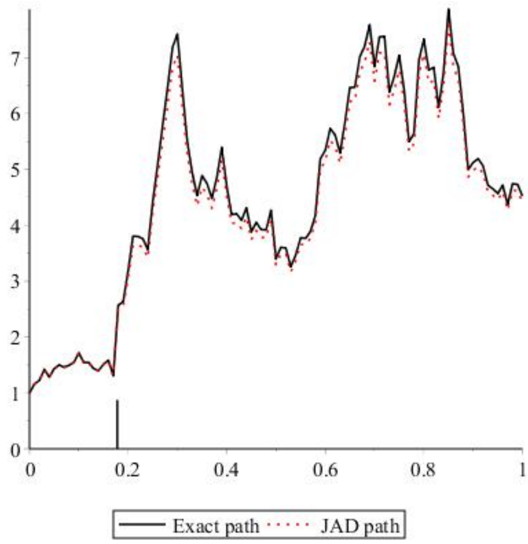

Figure 4.1: Paths of exact and approximated Merton diffusion process for $h = 1/100$

- Number of iterations N=100 and step size h=1/1000

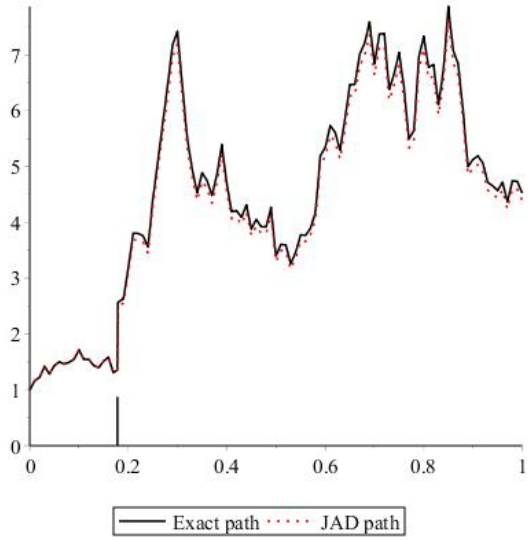

Figure 4.2: Paths of exact and approximated Merton diffusion process for $h = 1/1000$

4.2 Transition Probability

This section is about the derivation of the transition probability which is based on the strong Taylor jump adapted of order one approximation $Y^{\Delta} = \{Y_{t_i}, i = 0, \ldots, N\}$ obtained and explained in the previous section. Let's recall the definition of the transition probability density.

Definition 4.2.1 *The transition probability density $p(\Delta, z \mid y, \theta)$, if it exists, represents the probability that $X_{t_{i+1}} = z \in \mathbb{R}^n$ knowing $X_{t_i} = y \in \mathbb{R}^n$*

To establish it, we require to find some distribution especially the one of the variable $Y_{t_{i+1}-}$.

Considering the equation(4.2), let us analyze the distribution of the variable $Y = aX^2 + bX + c$ where $X \sim \mathcal{N}(0,1)$ and $a \neq 0$. Let F_Y be the distribution function of Y, we have

$$F_Y(y) = Prob(Y < y) = Prob(aX^2 + bX + c < y) \tag{4.7}$$

Solving the inequality $aX^2 + bX + c < y$ we handle two cases.

1. $a < 0$, either $y = c - \frac{b^2}{4a}$ and $aX^2 + bX + c < y$ for $X \in \mathbb{R}$ or $y > c - \frac{b^2}{4a}$ and $aX^2 + bX + c < y$ for $X \in (-\infty, \frac{-b-\sqrt{b^2-4a(c-y)}}{2a}) \cup (\frac{-b+\sqrt{b^2-4a(c-y)}}{2a}, +\infty)$
2. $a < 0$ either $y = c - \frac{b^2}{4a}$ and $aX^2 + bX + c < y$ doesn't admit a solution, or $y > c - \frac{b^2}{4a}$ and $aX^2 + bX + c < y$ for $X \in (\frac{-b-\sqrt{b^2+4a(c-y)}}{2a}, \frac{-b+\sqrt{b^2+4a(c-y)}}{2a})$

Therefore for

- $a < 0$ and $y > c - \frac{b^2}{4a}$

$$\begin{aligned}
F_Y(y) &= Prob(Y < y) = Prob(aX^2 + bX + c < y) \\
&= Prob\left[X \in (-\infty, \frac{-b-\sqrt{b^2-4a(c-y)}}{2a}) \right. \\
&\quad \left. \cup(\frac{-b+\sqrt{b^2-4a(c-y)}}{2a}, +\infty)\right] \\
&= 1 + Prob\left[X < \frac{-b-\sqrt{b^2-4a(c-y)}}{2a}\right] \\
&\quad -Prob\left[X < \frac{-b+\sqrt{b^2-4a(c-y)}}{2a}\right] \\
F_Y(y) &= 1 + \Phi(\frac{-b-\sqrt{b^2-4a(c-y)}}{2a}) \\
&\quad -\Phi(\frac{-b+\sqrt{b^2-4a(c-y)}}{2a})
\end{aligned} \tag{4.8}$$

where Φ is the cumulative distribution function of the standard normal distribution. and it follows that the density function f_y of Y is :

$$\begin{aligned}
f_Y(y) &= \frac{1}{\sqrt{b^2-4a(c-y)}}\left[\phi(\frac{-b+\sqrt{b^2-4a(c-y)}}{2a})\right. \\
&\quad \left. - \phi(\frac{-b-\sqrt{b^2-4a(c-y)}}{2a})\right]
\end{aligned} \tag{4.9}$$

- $a > 0$ and $y = c - \frac{b^2}{4a}$ $F_Y(y) = 0$
- $a > 0$ and $y > c - \frac{b^2}{4a}$

$$\begin{aligned}
F_Y(y) &= Prob(Y < y) = Prob(aX^2 + bX + c < y) \\
&= Prob\left[X \in (\frac{-b - \sqrt{b^2 - 4a(c-y)}}{2a}, \frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a})\right] \\
&= Prob\left[X < \frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a}\right] \\
&\quad - Prob\left[X < \frac{-b - \sqrt{b^2 - 4a(c-y)}}{2a}\right] \\
F_Y(y) &= \Phi(\frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a}) - \Phi(\frac{-b - \sqrt{b^2 - 4a(c-y)}}{2a}) \qquad (4.10)
\end{aligned}$$

and the density function is given by

$$\begin{aligned}
f_Y(y) = \frac{1}{\sqrt{b^2 - 4a(c-y)}} &\left[\phi(\frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a})\right. \\
&\left. -\phi(\frac{-b - \sqrt{b^2 - 4a(c-y)}}{2a})\right] \qquad (4.11)
\end{aligned}$$

The Laplace transform $\phi_Y(u)$ of Y is given by

$$\begin{aligned}
\phi_Y(u) &= E\left(\exp(uY)\right) \\
&= \int_{c-\frac{b^2}{4a}}^{\infty} \frac{1}{\sqrt{b^2 - 4a(c-y)}} \exp(uy)\phi(\frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a})dy \\
&\quad - \\
&\quad \int_{c-\frac{b^2}{4a}}^{\infty} \frac{1}{\sqrt{b^2 - 4a(c-y)}} \exp(uy)\phi(\frac{-b - \sqrt{b^2 - 4a(c-y)}}{2a})dy \\
&\qquad (4.12)
\end{aligned}$$

Considering the first integral in (4.12), we have,

$$\begin{aligned}
&\int_{c-\frac{b^2}{4a}}^{\infty} \frac{1}{\sqrt{b^2 - 4a(c-y)}} \exp(uy)\phi(\frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a}) \\
&= \frac{1}{\sqrt{2\pi}} \int_{c-\frac{b^2}{4a}}^{\infty} \frac{1}{\sqrt{b^2 - 4a(c-y)}} \exp\{\frac{-1}{2}(\frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a})^2\} \exp(uy)dy \\
&= \frac{1}{\sqrt{2\pi}} \int_{c-\frac{b^2}{4a}}^{\infty} \frac{1}{\sqrt{b^2 - 4a(c-y)}} \exp\left\{\frac{-1}{2}\left[(\frac{-b + \sqrt{b^2 - 4a(c-y)}}{2a})^2 - 2uy\right]\right\} dy
\end{aligned}$$

Let's change $\frac{-b+\sqrt{b^2-4a(c-y)}}{2a}$ to x. That implies that $y = \frac{(2ax+b)^2-b^2}{4a} + c$ and $dy = \sqrt{b^2-4a(c-y)}dx$. Therefore

$$\begin{aligned}
& \int_{c-\frac{b^2}{4a}}^{\infty} \frac{1}{\sqrt{b^2-4a(c-y)}} \exp(uy)\phi(\frac{-b+\sqrt{b^2-4a(c-y)}}{2a})dy \\
= & \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \exp\left\{ \frac{-1}{2} \left[(x^2 - 2uc - 2u\frac{(2ax+b)^2-b^2}{4a} \right] \right\} dx \\
= & \frac{\exp(uc)}{\sqrt{2\pi}} \int_0^{\infty} \exp\left\{ \frac{-1}{2} [(1-2ua)x^2 - 2ubx] \right\} dx \\
= & \frac{\exp(uc)}{(1-2ua)^{\frac{1}{2}}} \exp[-\frac{u^2b^2}{2(1-2ua)}] \qquad (4.13)
\end{aligned}$$

In the same way, taking the integral and doing the same change of variable we have

$$\begin{aligned}
& \int_{c-\frac{b^2}{4a}}^{\infty} \frac{1}{\sqrt{b^2-4a(c-y)}} \exp(uy)\phi(\frac{-b-\sqrt{b^2-4a(c-y)}}{2a}) \\
= & -\frac{\exp(uc)}{(1-2ua)^{\frac{1}{2}}} \exp[-\frac{u^2b^2}{2(1-2ua)}] \qquad (4.14)
\end{aligned}$$

Therefore substituting (4.13) and (4.13) into (4.12) we lead to a conditional Laplace transform of $Y_{t_{i+1}-}$ as

$$\phi_{Y_{t_{i+1}-}}(u|Y_{t_i}) = \frac{\exp(uc)}{(1-2ua)^{\frac{1}{2}}} \exp[-\frac{u^2b^2}{2(1-2ua)}] \qquad (4.15)$$

The following lemma states the result of the characteristic function.

Lemma 4.2.1 *The conditional characteristic function* $\Phi_{Y_{t_{i+1}-}}$ *is given by:*

$$\Phi_{Y_{t_{i+1}-}}(u|Y_{t_i}) = \phi_{Y_{t_{i+1}-}}(iu|Y_{t_i}) = \frac{\exp(iuc)}{(1-2iua)^{\frac{1}{2}}} \exp[-\frac{u^2b^2}{2(1-2iua)}] \qquad (4.16)$$

where $a = \frac{1}{2}\sigma_x(Y_{t_i},\theta)\sigma(Y_{t_i},\theta)h$, $b = \sqrt{h}\sigma(Y_{t_i},\theta)$ *and* $c = Y_{t_i} + \mu(Y_{t_i},\theta)h - \frac{1}{2}\sigma_x(Y_{t_i},\theta)\sigma(Y_{t_i},\theta)h$.

Let's consider now the jump part. This is a compound Poisson process, for which the characteristic function is derived within an interval time of length h.

We have,

$$J_{t_i}\Delta N_{t_i} = \sum_{N_{t_i}+1}^{N_{t_{i+1}}} J_t i = \sum_{t=1}^{N_{t_{i+1}}-N_{t_i}} J_t$$

We know that the Poisson process has the Poisson distribution with parameter λt and its increment $N_{t_{i+1}} - N_{t_i}$ has also a Poisson distribution with parameter $\lambda(t_{i+1} - t_i)$. Since the random variables J_i are assumed identically and independently distributed and are also independent from the process N_t we have,

Lemma 4.2.2 *The characteristic function $\Phi_{X_{t_i}}(u)$ of the process $X_{t_i} = \sum_{t=1}^{N_{t_{i+1}}-N_{t_i}}$ is:*

$$\Phi_{X_{t_i}}(u) = e^{\lambda(t_{i+1}-t_i)[\Phi_J(u)-1]} \tag{4.17}$$

where $\Phi_J(u)$ represents the characteristic function of the random variable J.

Proof.

$$\begin{aligned}
\Phi_{X_{t_i}}(u) = E(e^{iuX_{t_i}}) &= = E\left(e^{iu \sum_{t=1}^{N_{t_{i+1}}-N_{t_i}} J_t}\right) \\
&= E\left(\prod_{t=1}^{N_{t_{i+1}}-N_{t_i}} e^{iuJ_t}\right) \qquad (4.18)
\end{aligned}$$

Using now the conditional expectation and the independence between the processes N_t and J_t, we have:

$$\begin{aligned}
E\left(\prod_{t=1}^{N_{t_{i+1}}-N_{t_i}} e^{iuJ_t}\right) &= E\left[E\left(\prod_{t=1}^{N_{t_{i+1}}-N_{t_i}} e^{iuJ_t}\middle| N_{t_{i+1}} - N_{t_i}\right)\right] \\
&= E\left[E\left(e^{iuJ_1}\right)^{N_{t_{i+1}}-N_{t_i}}\right] \\
&= \sum_{n=0}^{\infty} E\left(e^{iuJ_1}\right)^n P\left(N_{t_{i+1}} - N_{t_i} = n\right) \\
&= \sum_{n=0}^{\infty} e^{-\lambda(t_{i+1}-t_i)} \frac{E\left(e^{iuJ_1}\right)^n [\lambda(t_{i+1}-t_i)]^n}{n!} \\
&= e^{-\lambda(t_{i+1}-t_i)} e^{\lambda(t_{i+1}-t_i)E\left(e^{iuJ_1}\right)} \\
&= e^{\lambda(t_{i+1}-t_i)\left[E\left(e^{iuJ_1}\right)-1\right]} \qquad (4.19)
\end{aligned}$$

Then substituting (4.19) into (4.18) we get the expectected result.

As we assume the distribution of the jump size to belong to the NEF with parameter ω, the characteristic function of the jump's size is:

$$\Phi_J(u) \quad = \quad \exp(A(i.u + \omega) - A(\omega)) \tag{4.20}$$

with A a known function, see Appendix for details on the characteristics function. We can therefore from the Lemmas (4.16)and (4.2.2) deduce that

Lemma 4.2.3 *The conditional characteristic function of the approximated process*

$\{Y_{t_{i+1}}, i = 0 \ldots N\}$ *is :*

$$\Phi_{Y_{t_{i+1}}}(u|Y_{t_i}) = \Phi_{Y_{t_{i+1}-}}(u|Y_{t_i})\Phi_{X_{t_i}}(u) = \frac{\exp(iuc)}{(1-2iua)^{\frac{1}{2}}}\exp[-\frac{u^2b^2}{2(1-2iua)}]e^{h\lambda[\Phi_J(u)-1]} \tag{4.21}$$

and the cumulant generating function is

$$\begin{aligned} K(u|Y_{t_i}) &= \ln[\Phi_{Y_{t_{i+1}}}(-iu|Y_{t_i})] \\ &= \frac{2uc + (b^2 - 4ac)u^2}{2(1-2ua)} + \frac{1}{2}\ln(1-2ua) \\ &\quad + h\lambda\left(\exp(A(u+\omega) - A(\omega)) - 1\right) \end{aligned} \tag{4.22}$$

with $u < 1/2a$, $a = \frac{1}{2}\sigma_x(Y_{t_i},\theta)\sigma(Y_{t_i},\theta)h$, $b = \sqrt{h}\sigma(Y_{t_i},\theta)$ *and* $c = Y_{t_i} + \mu(Y_{t_i},\theta)h - \frac{1}{2}\sigma_x(Y_{t_i},\theta)\sigma(Y_{t_i},\theta)h$

Therefore from the saddlepoint method, (Aït-Sahalia & Yu, 2006) we have

Theorem 4.2.1 *A leading term of the probability transition as*

$$P(Y_{t_{i+1}} = y_i|Y_{t_i} = x_i) = (2\pi)^{\frac{-1}{2}}\frac{\partial^2 K(\hat{u}_i|x_i)}{\partial u^2}\exp(K(\hat{u}_i|x_i) - \hat{u}_i.y_i) \tag{4.23}$$

The Log-likelihood function is

$$l(\theta) = \sum_{i=0}^{N-1} \left[\frac{-1}{2} \log(\frac{\partial^2 K(\hat{u}_i|x_i)}{\partial u^2}) + K(\hat{u}_i|x_i) - \hat{u}_i.y_i \right] \tag{4.24}$$

The estimator is obtained as

$$\hat{\theta} = \arg\max l(\theta) \tag{4.25}$$

where $\hat{u}_i$ *is the solution of the equation* $\frac{\partial K(\hat{u}|x_i)}{\partial u} = y_i$

4.3 Aplication

In this section we verify the effectiveness of the theoretical results obtain above. As application, we use the Cox-Ingersoll-Rox(CIR) model with Jump. It is a famous model, frequently used in modeling interest rate or default in credit risk, where the jump part helps to capture the effect of unpredictable event that occur. Several authors used it in finance and related areas. But there is no many work that consider the Jump. In the following section we present the CIR diffusion process and the Jump diffusion CIR process used.

4.3.1 CIR Process

The CIR process as developed by (Cox et al., 1985) is the process X_t solution of the following equation::

$$dX_t = \kappa(\mu - X_t)dt + \sigma\sqrt{X_t}dW_t \tag{4.26}$$

where we have $\mu(X_t, \theta) = \kappa(\mu - X_t)$ and $\sigma(X_t, \theta) = \sigma\sqrt{X_t}$, therefore the parameter θ is the vector (κ, μ, σ). The CIR process with jump is presented as:

$$dX_t = \kappa(\mu - X_t)dt + \sigma\sqrt{X_t}dW_t + J_t dN_t \tag{4.27}$$

here the vector of parameter is $\theta = (\kappa, \mu, \sigma, \lambda'\omega)$. Considering a deterministic discretization time $0 = t_0 < t_1 < \ldots < t_N = T$ with an equidistant grid with step size h, the jump adapted convergent scheme $\hat{X}_t$ with order one,of the process X_t solution of the equation (4.27) is

$$\hat{X}_{t_{i+1}} = \hat{X}_{t_{i+1}-} + J_{t_{i+1}} \Delta N_{t_i} \tag{4.28}$$

with

$$\hat{X}_{t_{i+1}-} = \hat{X}_{t_i} - \frac{h\sigma^2}{4} + \kappa(\mu - \hat{X}_{t_i})h + \sigma\sqrt{h\hat{X}_{t_i}}\varepsilon_i + \frac{h\sigma^2}{4}\varepsilon_i^2 \tag{4.29}$$

where ε_i is a standard normal random variable.

Let's consider the distribution of J_t for all t being a normal distribution with unknown mean μ_j and known variance σ_j^2, the NEF parameter is $\omega = \frac{\mu_j}{\sigma_j}$ and the function $A(\omega) = 1/2\omega^2$, therefore the cumulant generating function is:

$$\begin{aligned} K(u|\hat{X}_{t_i} = x_i) &= uc + \frac{u^2b^2}{2(1-2ua)} - \frac{1}{2}\ln(1-2ua) \\ &\quad + h\lambda\left(\exp(1/2u^2 + u.\omega) - 1\right) \end{aligned} \tag{4.30}$$

with $u < 1/2a$, $a = \frac{h\sigma^2}{4}$, $b = \sigma\sqrt{hx_i}$ and $c = x_i - \frac{h\sigma^2}{4} + \kappa(\mu - x_i)h$

For numerical evidence we consider weekly observations for which $h = 1/52$ and the theoretical values as follows, $\kappa = 0.1$, $\mu = 0.4$, $\sigma = 0.5$, $\lambda = 3$, $\omega = 2$. From a Bootstrap method we deduce some statistics on the estimation such as the bias, the mean square error(MSE) and the standard deviation(sd) for each estimation. The results are presented in the Table(4.1) below.

Table 4.1: Estimation Results

Step's size	Parameters				
	κ	μ	σ	λ	ω
h=1/52	0.647	0.919	0.724	1.142	2.534
Bias	(0.326)	(-0.035)	(-0.122)	(1.465)	(0.438)
MSE	(1.878)	(2.343)	(0.414)	(5.93)	(1.1475)
sd	(1.332)	(1.532)	(0.632)	(1.94)	(0.9783)

The estimation results show a good point estimator of the parameters, even though biased with minor mean square error. The bias obtained in the estimation is proper to the maximum likelihood method itself, and as this work is about an approximation of the likelihood function, it is quiet normal that the method performs similarly to the exact. Despite the ML being biased it is argued of being the preferred choice of estimation method in continuous time finance due to its asymptotic properties of consistency and efficiency, ? (?). The result in our work is consistent with the argument presented above as illustrated by the figures below.

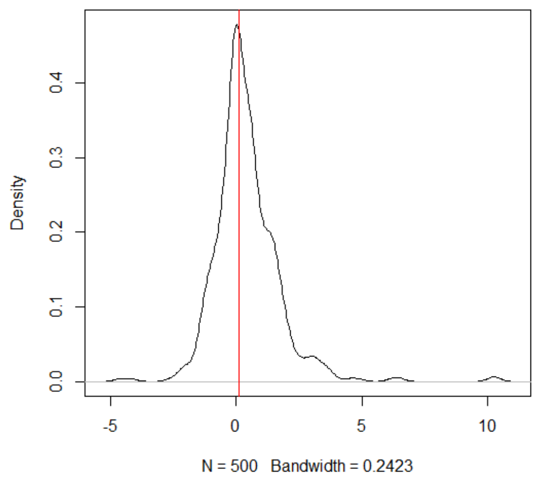

Figure 4.3: Density plot of parameter κ estiamtor

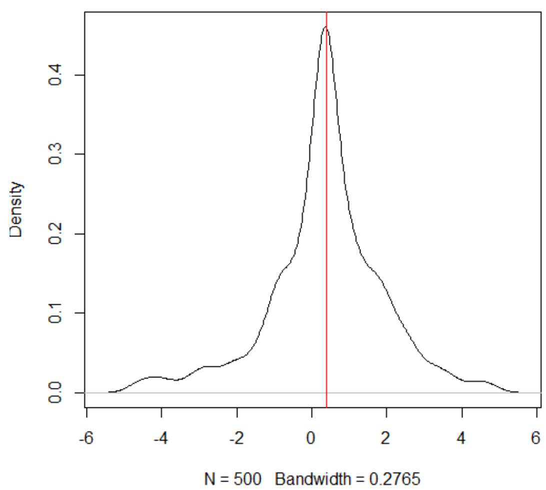

Figure 4.4: Density plot of parameter μ estiamtor

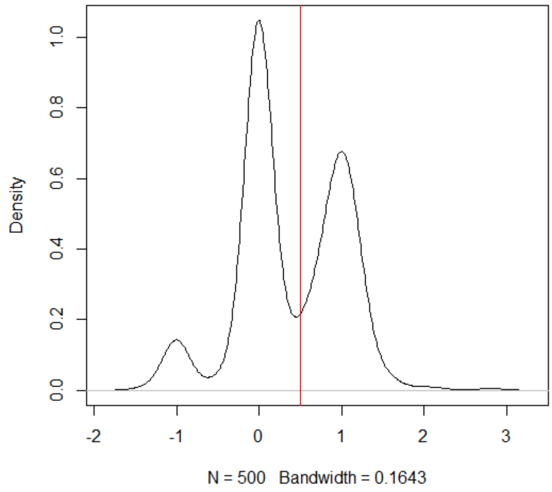

Figure 4.5: Density plot of parameter σ estiamtor

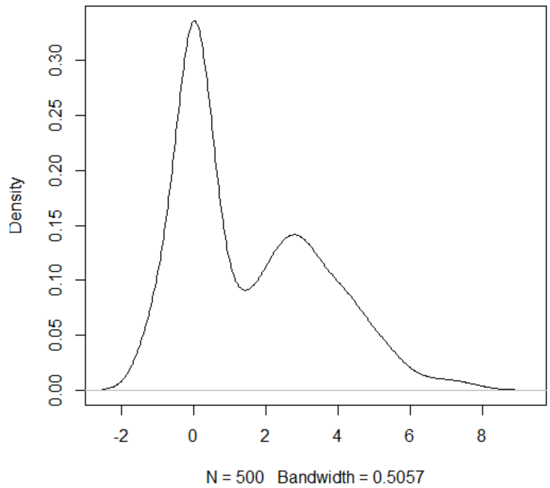

Figure 4.6: Density plot of parameter λ estiamtor

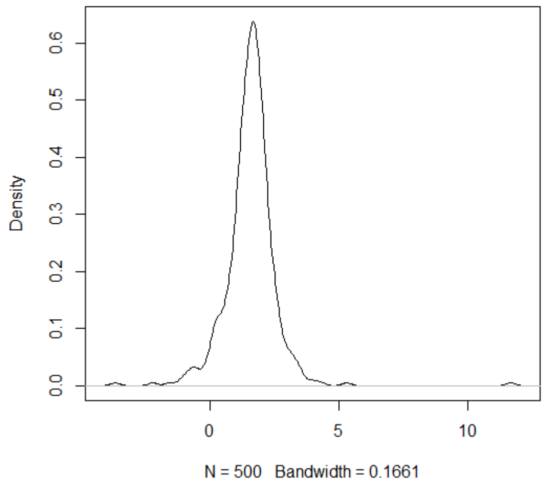

Figure 4.7: Density plot of parameter ω estimator

So for a large number of observations we find the normality asymptotic is reached as the density plot shows a curve apparent to a normal distribution density plot with mean the theoretical values chosen.

Chapter 5

Conclusion and Recommendation

5.1 Conclusion and summary

Through this study, we get a closed form expansion of the likelihood function that is applicable to a large class of jump diffusion processes due also to the large family of natural exponential family that has been considered. We use the so called jump adapted discretization scheme to approximate the solution of the considered model. Though the solution does exist under some regular conditions, it may only be known explicitly in few cases. Therefore in this study, the approximation scheme used is proved to converge converges strongly with order under some assumed regularities conditions. The discretization scheme used is such appropriate as it makes possible to derive the characteristic function of the diffusion part and from the assumption of the jump's size being a random variable whose distribution is in the Natural Exponential Family, we get the characteristic function of the compound Poisson process as well. Due to the independence between both processes, we deduce then the characteristic function of the jump diffusion process. Finally by applying the saddlepoint method which is based on the link between probability density function and characteristic function, we have been able get an approximation of the probability transition and the likelihood function thereafter. Therefore with appropriate method of simulation we show how accurate are the estimators obtained which support our choice of choosing the distribution of the jump's size from a family of distribution that has been proved to lead to efficient estimator from an asymptotic point of view for all the distribution in the family.

5.2 Recommendation

Through this work we rely on the Natural exponential family for deriving the characteristic function of the jump diffusion process, it will be interesting to explore some other distributions for instance the α-stable distributions, but in this context as the objective is to derive a likelihood function approximation based on the discretization scheme, the NEF is the more adaptable. Also this family of distribution contains a wide enough number of distribution that can describe the jumps magnitude observed from financial market. We can later check the bias induced by the approximation but recommend the use of some bias reduction techniques for instance the Jacknife estimation as a solution to the finite sample bias. It will also very important to prove theoretically the asymptotic property of the parameter estimator. A possible extension of this work will be to construct a test specification on the parameters estimated.

References

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica*, *70*(1), 223-262. doi: 10.1111/1468-0262.00274

Aït-Sahalia, Y. (2008). Closed-form likelihood expansions for multivariate diffusions. *Ann. Statist.*, *36*(2), 906-937. doi: 10.1214/009053607000000622

Aït-Sahalia, Y., Fan, J., & Peng, H. (2009). Nonparametric transition-based tests for jump diffusions. *Journal of the American Statistical Association*, *104*(487), 1102-1116. doi: 10.1198/jasa.2009.tm08198

Aït-Sahalia, Y., & Yu, J. (2006). Saddlepoint approximations for continuous-time markov processes. *Journal of Econometrics*, *134*(2), 507-551. doi: 10.1016/j.jeconom.2005.07.004

Akgiray, V., & Booth, G. G. (1988). Mixed diffusion-jump process modeling of exchange rate movements. *The Review of Economics and Statistics*, *70*(4), 631. doi: 10.2307/1935826

Andersen, T. G., & Lund, J. (1997). Estimating continuous-time stochastic volatility models of the short-term interest rate. *Journal of Econometrics*, *77*(2), 343-377. doi: 10.1016/s0304-4076(96)01819-2

A. Ronald Gallant, G. T. (1996). Which moments to match? *Econometric Theory*, *12*(4), 657-681. Retrieved from `http://www.jstor.org/stable/3532789`

Bakshi, G., & Ju, N. (2005). A refinement to aït-sahalia's (2002) "maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach"*. *The Journal of Business*, *78*(5), 2037-2052. doi: 10.1086/431451

Ball, C. A., & Torous, W. N. (1983). A simplified jump process for common stock returns. *The Journal of Financial and Quantitative Analysis*, *18*(1), 53. doi: 10.2307/2330804

Bergstrom, A. R. (1966). *Survey of continuous time econometrics.* [s.n.].

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637. doi: 10.1086/260062

Bruti-Liberati, N., & Platen, E. (2007). Strong approximations of stochastic differential equations with jumps. *Journal of Computational and Applied Mathematics*, *205*(2), 982-1001. doi: 10.1016/j.cam.2006.03.040

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*(2), 385. doi: 10.2307/1911242

Darrell Duffie, K. J. S. (1993). Simulated moments estimation of markov models of asset prices. *Econometrica*, *61*(4), 929-952. Retrieved from `http://www.jstor.org/stable/2951768`

Duffie, D., & Singleton, K. J. (1999). Modeling term structures of defaultable bonds. *Review of Financial Studies*, *12*(4), 687-720. doi: 10.1093/rfs/

12.4.687

Elerian, O. (1998). A note on the existence of a closed-form conditional transition density for the milstein scheme. In *Economics discussion paper 1998- w18,*.

Fher, W., D, & Rosenfeld, E. (1979). *Maximum likelihood parameter estimation for stochastic processes* (Tech. Rep.). Harvard University.

Glasserman, P. (2004). *Monte carlo methods in financial engineering.* Springer.

Glasserman, P., & Kou, S. G. (2003). The term structure of simple forward rates with jump risk. *Mathematical Finance*, *13*(3), 383-410. doi: 10.1111/1467-9965.00021

Glasserman, P., & Merener, N. (2004). Convergence of a discretization scheme for jump-diffusion processes with state-dependent intensities. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *460*(2041), 111-127. doi: 10.1098/rspa.2003.1237

Gollnick, H., Houthakker, H. S., & Taylor, L. D. (1968). Consumer demand in the united states, 1929-1970. analyses and projections. *Econometrica*, *36*(1), 203. doi: 10.2307/1909620

Gourieroux, C., Monfort, A., & Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica*, *52*(3), 681. doi: 10.2307/1913471

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*(4), 1029. doi: 10.2307/1912775

Jarrow, R. A., & Turnbull, S. M. (1995). Pricing derivatives on financial securities subject to credit risk. *The Journal of Finance*, *50*(1), 53. doi: 10.2307/2329239

Johannes, M. (2004). The statistical and economic role of jumps in continuous-time interest rate models. *The Journal of Finance*, *59*(1), 227-260. doi: 10.1111/j.1540-6321.2004.00632.x

Kloeden, E., Peter EPlaten. (1992). *Numerical solution of stochastic differential equations*. Springer.

Li, C. (2013). Maximum-likelihood estimation for diffusion processes via closed-form density expansions. *Ann. Statist.*, *41*(3), 1350-1380. doi: 10.1214/13-aos1118

Lo, A. W. (1988). Maximum likelihood estimation of generalized it processes with discretely sampled data. *Econ Theory*, *4*(02), 231-247. doi: 10.1017/s0266466600012044

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1-2), 125-144. doi: 10.1016/0304-405x(76)90022-2

Merton, R. C. (1980). On estimating the expected return on the market. *Journal of Financial Economics*, *8*(4), 323-361. doi: 10.1016/0304-405x(80)90007-0

Milstein, G. N. (1979). A method of second-order accuracy integration of stochastic differential equations. *Theory of Probability ;& Its Applications*, *23*(2), 396-401. doi: 10.1137/1123045

Privault, N. (2012). *An elementary introduction to stochastic interest rate modeling*. World Scientific.

Rockinger, M., & Semenova, M. (2005). Estimation of jump-diffusion processes via empirical characteristic functions. *SSRN Electronic Journal*. doi: 10.2139/ssrn.675202

Sundaresan, S. M. (2000). Continuous-time methods in finance: A review and an assessment. *J Finance*, *55*(4), 1569-1622. doi: 10.1111/0022-1082.00261

Tse, Y. K., Zhang, X., & Yu, J. (2004). Estimation of hyperbolic diffusion using the markov chain monte carlo method. *Quantitative Finance*, *4*(2), 158-169. doi: 10.1080/14697680400000020

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, *5*(2), 177-188. doi: 10.1016/0304-405x(77)90016-2

Yu, J. (2007). Closed-form likelihood approximation and estimation of jump-diffusions with an application to the realignment risk of the chinese yuan. *Journal of Econometrics*, *141*(2), 1245-1280. doi: 10.1016/j.jeconom.2007.02.003

Appendix

Definition .0.1 *A Brownian motion or Wiener process is a stochastic process denoted W that satisfied the following conditions:*

- $W(0) = 0$*;*
- W *has independent increment;*
- *For $s < t$ the stochastic variable $W(t)-W(s)$ has the Gaussian distribution $N(0, \sqrt{t-s})$;*
- W *has continuous trajectory with probaility one.*

Definition .0.2 *A diffusion process X is a stochastic process whose local dynamics can be approximated by sthochastic differential equation of the form:*

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t,$$

where W_t is a Wiener process

Definition .0.3 *An option is a derivative whose value depends on the price stock. It gives the holder the right to buy(call option) or to sell(put option) the underlying security at a given price by a given period of time.*

Definition .0.4 *The SVJJ model is presented as follows:*

$$\frac{dS_t}{S_t} = (\mu - \lambda m)dt + \sqrt{V_t}dW_t^{(1)} + (\exp^{Z_s} - 1)dq_t; \ where \ dV_t = \beta(\alpha - V_t)dt + \sigma + \sqrt{V_t}dW_t^{(2)} + Z_v dq_t$$

$W_t^{(1)}$ and $W_t^{(2)}$ are standard Brownian motion such that $corr(dW_t^{(1)}, dW_t^{(2)}) = \rho dt$; SV and SVJ models are restriction of SVJJ

Definition .0.5 *For a Markov process X, its joint distribution may be expressed as follows:*

$$f_{X_{t_0}, X_{t_1}, \ldots, X_{t_n}}(x_{t_0}, x_{t_1}, \ldots x_{t_n}) = f_{X_{t_0}}(x_{t_0}) \Pi_{k=1}^{n} f_{X_{t_k}}(x_{t_k} \mid x_{t_{(k-1)}})$$

Definition .0.6 *The diffusion process X is said to be reducible to unit diffusion if there exists an injective mapping of X into a diffusion Y whose diffusion matrix is the idendity.*

Definition .0.7 *A process $l_T(\theta)$ is said to converge in probability uniformly to a function $l(\theta)$ for $\theta \in \Theta$ if:*

$$\lim_{T \to \infty} Pr\left[\sup_{\theta \in \Theta} \mid l_T(\theta) - l(\theta) < \varepsilon \mid\right] = 1, \forall \varepsilon > 0$$

Definition .0.8 *A process X_t is said to converge in probability to X if:*

$$\lim_{T \to \infty} Pr[\mid X_T - X \mid < \varepsilon] = 1; \forall \varepsilon > 0$$

Proposition .0.1 (Doob's inequality) *Let's pbeapositiveinteger, $p \geq 1$ and (M_n) a martingale, then*

$$E\left(\max_{0 \leq k \leq n} |M_k|^p\right) \leq \left(\frac{p}{p-1}\right)^p \sup_{n \geq 0} E(|M_n|^p) \quad (1)$$

Proposition .0.2 (Gronwall Inequality) *Let $h(t)$ and $g(t)$ be nonnegative continuous functions on $I = [0, \infty)$ for which the inequality*

$$h(t) \leq c + \int_a^t g(s)h(t), t \in I \tag{2}$$

holds where c is a nonnegative constant. Then

$$h(t) \leq c\exp(\int_a^t g(s)ds) \tag{3}$$

Definition .0.9 *The natural exponential family(NEF) is a subfamily of the exponential family, see (Gourieroux, Monfort, & Trognon, 1984) for details on the exponential family. A distribution is said to belong to the NEF with parameter ω if it can be written with the probability distribution(pdf)*

$$f_X(x|\omega) = B(x)\exp(\omega.x - A(\omega)) \tag{4}$$

with B and A two known functions. The characteristic function is as follows:

$$\begin{aligned}
\phi_X(x) = E(e^{iuX}) &= B(x)\int \exp\left((iu+\omega)x - A(\omega)\right)dx \\
&= \exp\left(A(iu+\omega) - A(\omega)\right)\underbrace{\int B(x)\exp\left((iu+\omega)x - A(iu+\omega)\right)}_{=1} \\
&= \exp\left(A(iu+\omega) - A(\omega)\right)
\end{aligned}$$

Examples

- *Suppose the random variable J has a normal distribution with an unknown mean m and variance 1. Its pdf is*

$$\begin{aligned}
f_J(j,m) &= \frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}(j-m)^2} \\
&= \frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}j^2}e^{mj-\frac{1}{2}m^2}
\end{aligned}$$

Setting $\omega = m$; $B(j) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2} j^2}$ *and* $A(\omega) = \frac{1}{2}\omega^2$, *we obtain the form defined in Eq(4).*

- *Suppose the random variable* J *has a Gamma distribution with a known shape parameter* α. *Its pdf is*

$$\begin{aligned} f_J(j, \beta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} j^{\alpha-1} e^{(-\beta j)} \\ &= \frac{j^{\alpha-1}}{\Gamma(\alpha)} e^{[-\beta j - (-\alpha \ln(\beta)]} \end{aligned}$$

Setting $\omega = -\beta$; $B(j) = \frac{j^{\alpha-1}}{\Gamma(\alpha)}$ *and* $A(\omega) = -\alpha \ln(-\omega)$ *we obtain the form defined in equation(4). For* $\alpha = 1$ *we have the exponential distribution with parameter* β