Computational Option Pricing under Jump Diffusion and Lévy processes

Eleftheria Chatzipanagou

A thesis submitted in partial fulfillment of the
requirements of the University of Greenwich
for the Degree of Doctor of Philosophy

August 2015

DECLARATION

I certify that this work has not been accepted in substance for any degree, and is not concurrently being submitted for any degree other than that of Doctor of Philosophy being studied at the University of Greenwich. I also declare that this work is the result of my own investigations except where otherwise identified by references and that I have not plagiarized the work of others.

Student

……..

Eleftheria Chatzipanagou

1st Supervisor

……..

Professor Kevin Parrott

2nd Supervisor

……..

Professor Choi-Hong Lai

ACKNOWLEDGEMENTS

Firstly, I want to express my deepest gratitude to my supervisor Prof. Kevin A. Parrott for his patience, constant encouragement and unmeasurable help in completing this project. He and his wife Charlotte Cory were a constant source of strength, support and motivation. Without them I wouldn't have reached this point.

Secondly I want to thank Dr. Nicholas Christakis, my supervisor during my first degree. His encouragement and guidance was essential in pursuing a continuation of my studies in London. This has defined my personal and professional life and I will always be grateful to him for showing me this road.

I am forever indebted to Anees Saqui. He to this day, kept reminding me how important the completion of this project is. His encouragement remained unchanged throughout the years and his support during challenging moments kept me afloat. He, like Kevin and Charlotte never lost faith in my ability to complete this work.

I thank also all my friends for their support and encouragement. I am very thankful to Dr. Yasmine Rosunally, for showing the way of completing a PhD and being so supportive throughout the process, Katerina Tzigkounaki and Theoni Fotopoulou for their unconditional friendship and support. Also Maria-Eleni Vlachakou for her inspiration and reassurance.

Finally, I want to thank my family, my parents Kon/nos Chatzipanagos and Despoina Papastergiou and my brother Michail Chatzipanagos for their support, guidance and trust.

ABSTRACT

The shortcomings of diffusion models in representing the risk related to large market movements have led to the development of various option pricing models with jumps. These models allow for a more realistic representation of price dynamics and greater flexibility in modelling and have therefore been the focus of much recent work. In this thesis the development of a robust finite difference method for the option pricing under jump-diffusion and Lévy processes is presented and its effectiveness is demonstrated on a range of pricing models.

LIST OF PUBLICATIONS

The research work carried out during the course of this PhD programme has been published in the following:

- Parrott, A. K, Chatzipanagou, E. *Finite Difference Methods for European and American Options under Jump-Diffusion Processes*, GH workshop on Applied and Numerical Mathematics, January 2010.

- Parrott, A. K, Chatzipanagou, Chatzipanagou, *Finite Differences with Coordinate Transformations for American Options under Jump- Diffusion Processes* Mathematical Finance Conference, Bachelier Finance Society, July 2008

- Parrott, A. K, Chatzipanagou, Chatzipanagou, E. *A Finite Difference Approach to Pricing American Options under Jump-Diffusion Processes.* Computational Methods in Finance Conference, July 2007.

NOMENCLATURE

Term	Definition
SDE	Stochastic Differential Equation
PDE	Partial Differential Equation
PIDE	Partial Integro-Differential Equation
FFT	Fast Fourier Transformation
VG	Variance Gamma
CGMY	Carr-Geman-Madan-Yor
i.i.d.	Independent identically distributed
FD	Finite Difference
p.d.f	Probability Density Function
LCP	Linear Complementarity Problem
Pr.	Peer review
IT	Ikonen-Toivanen

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Introduction

In this introductory section, we discuss the motivation underpinning the work presented in this thesis, starting from the Black-Scholes model's (Section 1.1) shortcomings which led to the development of various models which try to capture more accurately the pricing dynamics of financial derivatives. One of those models, the jump-diffusion one, is the focus of this work. We summarize (Section 1.2) the work of our research peers and discuss the contribution of our work. Finally the structure of this document is presented (Section 1.3).

1.1 Jump-diffusion processes: Motivation

The Black-Scholes paper (Black, 1973) published in 1973, presented the derivation of a partial differential equation for option prices, where the dynamics of prices $X_t = (X_t^1, \dots, X_t^m)$ of the underlying asset were described by a diffusion process, driven by a Geometric Brownian Motion:

$$dX_t = X_t \sigma(\tau, X_\tau) dW_t + X_t \mu_t dt \qquad (1.1)$$

The pricing formula obtained in that paper, presented a major breakthrough in understanding the way financial derivatives behave within the capital markets and for that reason this methodology and its generalizations were immediately adapted by financial institutions and are still widely used for the modelling of derivative products. Despite the success of the Black-Scholes model, its main characteristics, as described below, were linked to its fundamental drawbacks in trying to capture asset price behavior. In what follows we discuss those inefficiencies.

In continuous path models (e.g. diffusion models) the price process behaves locally like a Brownian motion. This implies that the price is unlikely to move by large amount in a short period of time, unless the input volatility parameter over the same time period is set unrealistically high. For example, in the case of a short-term out-of-the money option, this would mean that there is small probability of the option expiring in-the-money. However, in the market, it is possible to observe a sudden sharp change in the price of the underlying asset that would result in the option expiring in-the-money (and the opposite is true for an in-the-money short-term option close to expiry).

From the hedging point of view, continuous models for describing asset price dynamics lead to complete markets or to markets that can be made complete by introducing additional instruments (Zhang, 2007).

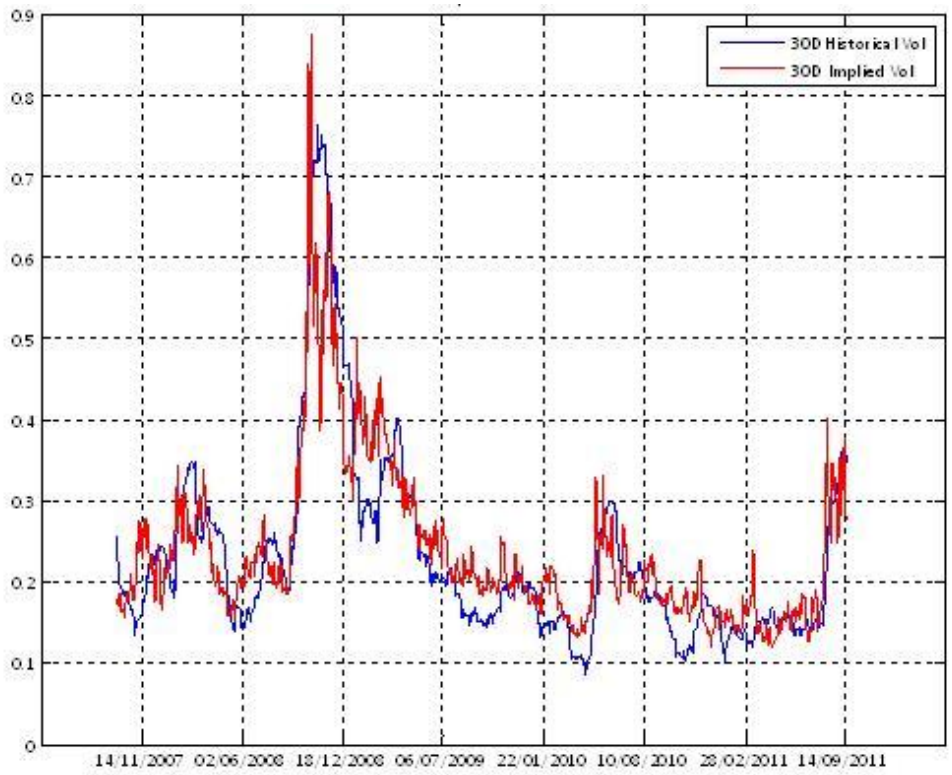
Under the assumption of stochastic volatility this property is lost, but even in this case the market can be made complete by the addition of vanilla options. Under the complete market assumption any derivative product can be dynamically replicated via cash and the underlying asset. From a probabilistic point of view, this means that in such models, every contingent claim C is attainable, in other words, under a unique equivalent martingale measure Q; there exists a self-financing trading strategy θ, whose value process:

$$V_T = V_0 + \sum_{i=0}^{N} \theta_T^i \left(S_T^i - S_0^i\right) \quad (1.2)$$

satisfies $V_T = C$. In this case any change in the portfolio value will be due to capital gain or losses within the portfolio and not due to the addition or withdrawal of funds.

However, real markets are incomplete and asset prices exhibit sudden sharp movements. The existence of jumps in asset price movements in the real market forces the market participants to hedge for risks that cannot be hedged only by using cash and the underlying asset and therefore make perfect hedging impossible. Even in portfolio valuation theory, the assumption of continuous price movements neglects the asymmetric correlation in the portfolio assets imposed by jumps or correlated signs across markets.

Figure 1-1 Evolution of Historical Prices and Implied Volatility-UKX Index (2007-2011)



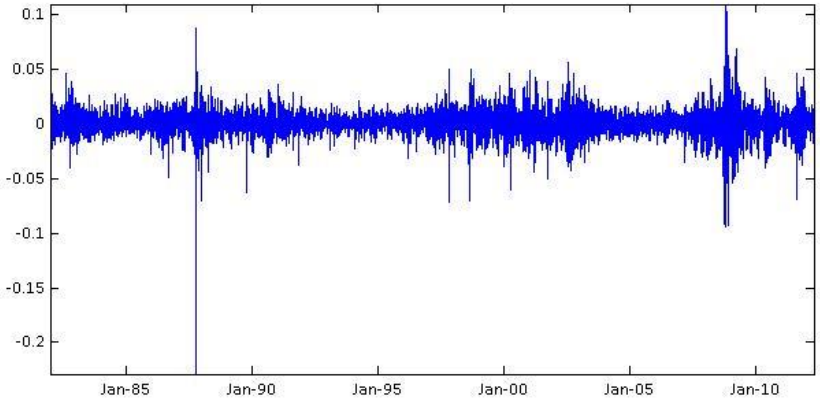
Famous market crashes along with the more recent financial crisis in 2008-2010, prove that the continuous models used by financial institutions are not able to predict the high volatility levels. Figure 1-1 shows the extreme market movements observed during periods of adverse macroeconomic conditions. It is therefore widely acknowledged that the behavior of speculative fluctuations deviated from the standard Geometric Brownian Motion (GBM) in several ways. That realization prompted the study of statistical properties of the financial times series and has revealed a range of interesting stylized facts which seem to be common to a wide range of instruments, time periods and markets.

Lack of log normality in the returns: An important property of the Brownian motion process is the continuity of its sample paths: a typical path $t \to B_t$ will be a continuous function of time. Figure 1-2 shows the daily log returns of the SPX index.

The daily log returns are defined as: $R_t = \{ln\frac{S(t)}{S(t-1)}\}$, at time t, where $S(t)$ denotes the asset price. It is obvious from the figure below that big spikes

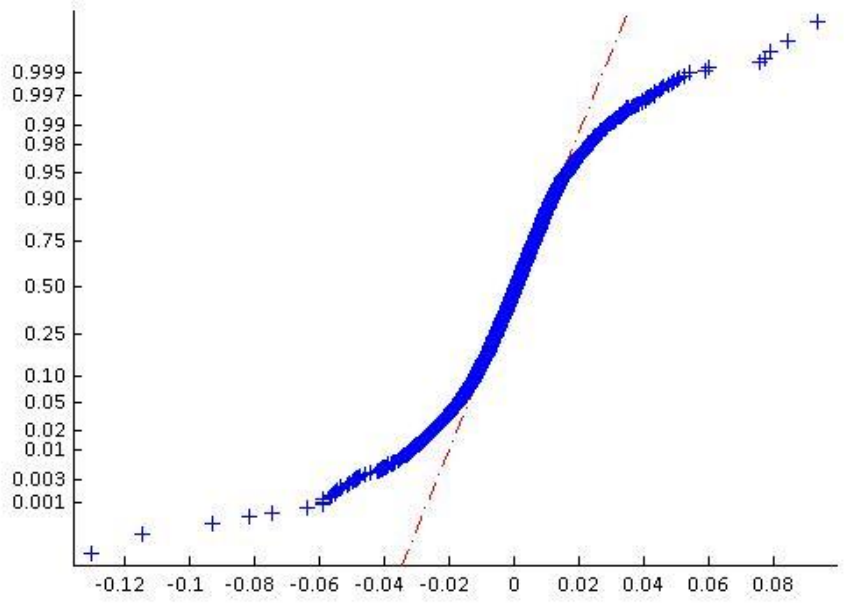
occurred during 1987 and 2008. That highlights the fact that extreme movements in asset price paths (in this work referred to as jumps) that have been observed in the markets are considered highly unlikely when modelled by a Gaussian distribution. In Figure 1-3 the normal probability $Q - Q$ plot for the FTSE 100 Index, further reveals the deviation of the data sample from the normal probability distribution.

Figure 1-2 Daily log returns of S&P 500 Index (1982-2012)



Scale invariance: Another important property of the Brownian motion is its scale invariance. That means that the statistical properties characterizing a Brownian motion do not change at all-time resolutions. In other words in any small time interval of the random walk the Brownian motion limit will still look like a Brownian motion. However, while Brownian motion does not distinguish between time scales, the observed market price behavior does. If stock price paths are analyzed on an intraday scale basis, it will be observed that they move essentially by jumps and still exhibit discontinuous behavior at the scale of months. Only under longer time horizons do the stock price paths resemble the paths of a Brownian motion processes.

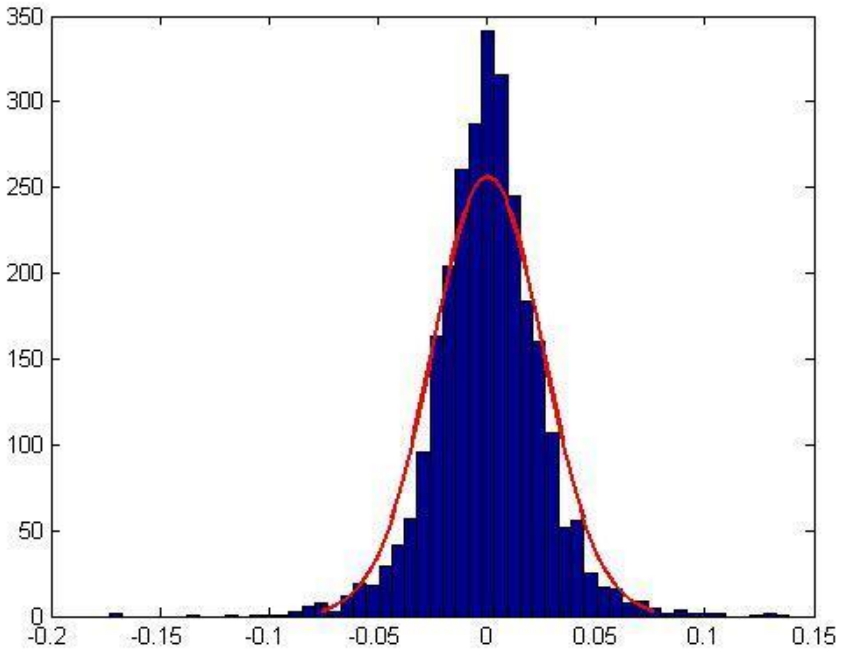
Figure 1-3 Normal Probability plot for FTSE 100 Index



Excess Volatility: It is a market accepted fact that it is difficult to justify the observed level of variability in asset returns by variations in fundamental economic factors driving the returns of the securities. In (Summers, 1989) it is argued that the occurrence of large negative or positive returns is not always attributable to the arrival of unexpected news. The volatility levels can be affected by the supply or demand for securities and can be positively correlated to trading volumes.

Volatility Clustering: In practice, it is often found that for financial time series (assuming log returns) the volatility fluctuation seems to change over time. More specifically, large changes tend to be followed by large changes, of either sign and small changes tend to be followed by small changes (Marcozzi, 2003). From Figure 1 2 it can be observed that long periods of high volatility are interspersed with periods or relative calm. This type of pattern is known as volatility clustering - while returns themselves are uncorrelated, absolute returns $|r_t|$ or their squares show positive, significant and slowly decaying autocorrelation function $corr(|r_t|, |r_{t+1}|)$, for t ranging from a few minutes to several weeks.

Figure 1-4 Histogram of daily log returns of FTSE 350 Electricity Index (1985-2012)



Heavy Tails, Excess Kurtosis and Leptokurtic characteristics: Figure 1-4 presents the histogram of FTSE 350 electricity index, along with the standard normal density function. It can be observed that the histogram displays a high peak (i.e. kurtosis, sometimes referred to as elongation) and asymmetric heavy tails. This is true, not only for the particular index, but is a commonly observed characteristic across the spectrum of financial asset prices (e.g. US and worldwide stock indices, individual stocks, foreign exchange rates, interest rates, commodities, etc.). More precisely, the kurtosis and skewness are defined as $Kurt = E(\frac{(X-\mu)^4}{\sigma^4})$, $\gamma_1 = E(\frac{(X-\mu)^3}{\sigma^3})$. For the standard normal density $Kurt = 3$. If $Kurt > 3$ then the distribution will be leptokurtic and will have a higher peak and two heavier tails than those of the normal distribution. The Black-Scholes Brownian motion model ignores this well observed feature.

The purpose behind the introduction of a continuous time stochastic process aiming to capture empirical properties of asset prices was ultimately the development of an option price model, which would be able to envisage the state of options market at a given time. For any model, this is achieved through its calibration with market data. More specifically, the parameters of the model are chosen to fit the observed market option prices and will produce modelled option prices not significantly different from the observed ones. The efficiency of the model is usually tested through backtesting methodologies developed and mostly utilized for the validation of Value at

Risk (VaR) models and exposure estimates like Potential Future Exposures (PFE) (Hull, 1989).

The Black-Scholes partial differential equation for a European or American call option on the underlying stock S with exercise price K is given by the following formula:

$$\frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S\frac{\partial^2 f}{\partial S^2} = rf \quad (1.3)$$

With $S_{min} < S < S_{max}$, $t_{min} < t < t_{max}$ where r is the risk free interest rate, q is the continuous dividend, T is the time to maturity of the option and σ is the volatility of the underlying stock price. The value of a European or American call option that pays zero dividends is defined as a contingent claim with payoff at maturity $max\ (0, S_T - K)^T$.

The Black-Scholes formula for the value of a European call option is:

$$C^{BS}(S_t, K, \tau, \sigma) = S_t N(d_1) - Ke^{-r\tau}N(d_2) \quad (1.4)$$

$$d_1 = \frac{-\ln(m)+\tau(r+\frac{\sigma^2}{2})}{\sigma\sqrt{\tau}}, \qquad d_2 = \frac{-\ln(m)+\tau(r-\frac{\sigma^2}{2})}{\sigma\sqrt{\tau}} \quad (1.5)$$

Where $m = K/S_t$ is the moneyness[1] of the option and:

$$N(x) = (2\pi)^{-\frac{1}{2}}\int_{-\infty}^{x} exp\,(-\frac{z^2}{2})dz \quad (1.6)$$

One attractive feature of the Black-Scholes model is that its parameters are unambiguously observable. That makes the model easily calibrated to market prices. A very important parameter in the Black-Scholes option-pricing model is the volatility of the underlying asset. The level of volatility used in the model has significant impact on the option price behavior and

[1] Moneyness is a measure of the degree to which a derivative is likely to have a positive monetary value at its expiration under the risk neutral measure. Here, moneyness is defined as the fraction of the option strike over the spot value of the underlying ($M = \frac{K}{S}$)

understanding the price-volatility dynamic (i.e. relationship between directional changes of the underlying and directional changes in the volatility) can help cushion against losses. Fortunately, this relationship in equity markets is easily observable. One can calculate the volatility implied by the option prices in the market, by solving the inverse Black-Scholes problem. The implied volatility $\sigma_{\underline{C}} = g(\underline{C}), \ .)$ of an option C will depend on the characteristics of the option. Those are the strike price K and the maturity T. However, the Black-Scholes model predicts a flat profile for the implied volatility surface:

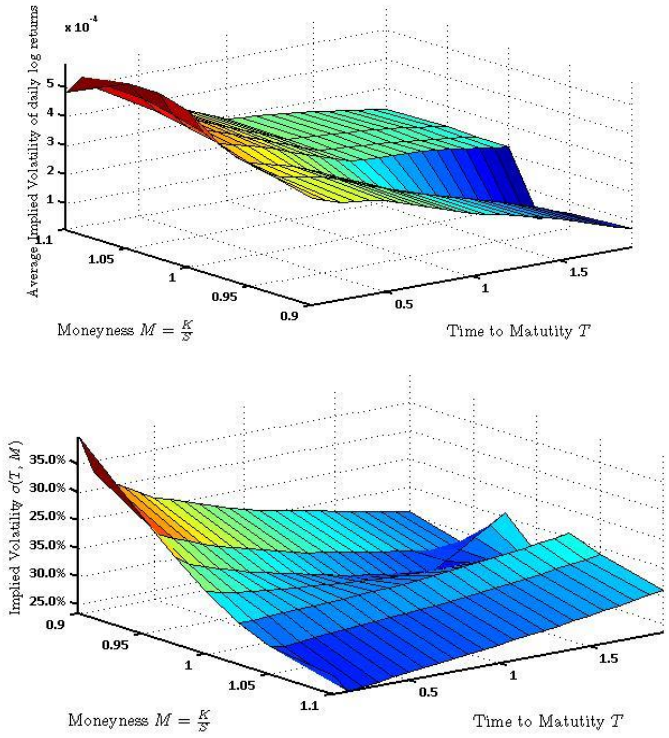
$$\sigma_{\underline{C}} = g\left(\underline{C}\right), .\Big) = \sigma \quad (1.7)$$

It is a well observed empirical fact that the volatility implied by options traded in the market is not constant as a function of strike nor as a function of time to maturity. This dependency is easily observed when looking at the shape of implied volatility surfaces, where the non-flat instantaneous profile of the surface points out the existence of term structure volatility and its deforming shape confirms the change of levels of implied volatilities over time, capturing the evolution of prices in the options market.

Figure 1-5 top presents the average implied volatilities for DJ EUROSTOXX index for the period of [01/2008-11/2011]. Maturities range from 1 to 24 months and moneyness levels range from 90% to 100%.

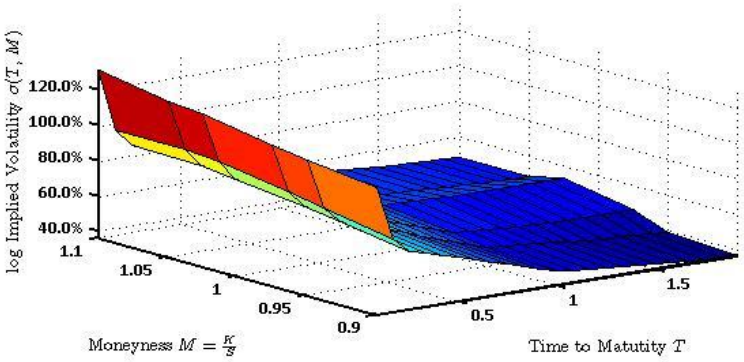
Figure 1-5 bottom shows the surface obtained for DJ EUROSTOXX when logarithmic changes of implied volatility are taken before averaging. The volatility skew and term structure characteristics persist. The sample standard deviation of implied volatilities shown in Figure 1-6 illustrates that the surface is not static and fluctuates around its average profile. It is observed that the daily standard deviation of the implied volatility can be as large as one half of its typical value for out-of-the money options, resulting in an important impact on option prices and prices of other financial derivatives (e.g. variance swaps). The above highlights the main weakness behind the assumption of constant volatility as implied by the Black-Scholes formulae which is further exacerbated in the case of short term options, where utilizing continuous models of diffusion type with constant or even stochastic volatility, cannot capture observed discontinuities in the underlying price path unless the volatility of volatility is set to be unrealistically high.

Figure 1-5 Average Implied (top) and log-implied (bottom) Volatility Surface DJ EUROSTOXX (2008-2011)



Those findings regarding the ability of the Black-Scholes model to capture the real market asset price dynamics, created the need for the development and integration with the existing pricing methodologies of a model that will be able to account for unexpected large price movements due to the arrival of information regarding unforeseen market events. An efficient option pricing model should be able to replicate the prices observed in the market, but should also assign values to complex derivative products without allowing for arbitrage opportunities. Several generalizations of the Black-Scholes model have been proposed in order to deal with the smile shaped implied volatilities problem. In (Dupire, 1994) a time dependent local volatility is inferred from the term structure of implied volatilities, essentially inducing a unique diffusion process. However for options with longer maturities the resulting local volatilities are roughly constant, predicting a future smile that is much flatter than smiles observed in the market. Proposed diffusion based stochastic volatility models (Heston, 1993a), (Papanicolaou, 2000), can reproduce the profile of implied volatilities at given maturities but fail to do so across maturities (Hull, 1989).

Figure 1-6 Standard Deviation of Implied Volatility DJ EUROSTOXX (2008-2011)



In contrast to the above, models with jumps can lead to a variety of smile and skew patterns since they account for discontinuities in the asset price paths and provide a way of capturing the fear of unexpected changes in the price of underlying assets, which is the main source of the skewness observed in diffusion models, even in the case of the short term options. Therefore, those models allow market participants to perform efficient risk management by quantifying and accounting for the risk of stock price shocks over a short interval of time, something not possible in a diffusion type model.

1.2 Relation to literature and contribution

Since the Geometric Brownian motion model was commonly accepted to be inconsistent with the market behavior, many studies have been conducted in order to modify the existent model to produce a pricing formula which will effectively capture the market behavior. A number of extensions have been proposed. Stochastic volatility and ARCH models (Hull, 1987); (Papanicolaou, 2000); normal jump models (Merton, 1975); affine stochastic-volatility and affine jump-diffusion models (Dupire); (Heston, 1993b); infinite activity Lévy models (Madan, 2002); (Madan, 1998). Jump-diffusion and Lévy models are the most attractive ones, since they can capture the jump patterns exhibited by stock prices and are more realistic when closer to maturity.

In particular, models with jumps are considered to be more representative of the actual market. Over the last decade research departments of major investment banks began to accept jump-diffusions as a valuable tool in their day-to-day modelling. The acceptance of jump-diffusion models compared

to the Black-Scholes is mostly supported by the following reasons: (1) in diffusion models and as discussed previously, the price process behaves locally like a Brownian motion and the probability that the stock price moves by a large amount over a short period of time is very small, (2) from the hedging point of view, jumps allow the practitioner to quantify and take into account the risk of strong price movements over short time intervals.

Starting with Merton's paper (Merton, 1975), and up to present date, various types of jump-diffusion models have been studied across the academic community. A number of finite difference methods for valuing options under jump-diffusion and Lévy processes have been published, (Cont, 2004), (d' Halluin, 2004), (Oosterlee, 2007), (Kou, 2004). These valuations require the numerical solution of a partial-integro-differential equation (PIDE), which involves, in addition to a possible degenerate second order differential operator, a non-local term that requires specific treatment at both a theoretical and numerical level. Various numerical methods for solving such parabolic integro-differential equations have been proposed.

Lévy models and in particular exponential Lévy models, where the market price of an asset is modelled as the exponential $S_t = S_0 \exp(rt + X_t)$ of a Lévy process X_t, provide a convenient framework to model the behavior of an asset price, as it appears both in the "real" and the "risk-neutral" world. This is because the sample paths can have jumps, the generating distributions can be fat-tailed and skewed and the volatilities implied from option prices can have "smile" shape, due to the flexibility of choice of the Lévy process X_t. Option pricing under exponential Lévy models has been discussed in (Dupire, 2000), (Voltchkova, 2005a).

In parallel with jump-diffusion and finite activity models another category of formulation was developed. Processes with infinite activity without a diffusion component, represent a family of models where all paths have infinitely many jumps along any time interval of finite length. However, the pricing equations in that case, are numerically more challenging and the market turns out to be incomplete in the sense that a hedging strategy leading to instantaneous risk free portfolio does not, in general, exist (Madan, 2002). The numerical approximation technique that is presented as part of this thesis is implemented in two models of this category: (i) Variance Gamma model (VG) (Madan, 1998) and (ii) the CGMY model (Madan, 2002) model.

In (Voltchkova, 2005b), the authors propose a finite difference scheme for solving PIDE's for the case of European and barrier options when the underlying is following a jump-diffusion model. Their numerical solution is

based on splitting the operator into a local and non-local part, treating the local term using an implicit step and the non-local term using an explicit term respectively.

For jump-diffusion models with finite intensity, Andersen and Andreasen in (Andreasen, 2000) propose an operator splitting method where the differential part is treated using a Crank-Nicholson step and the jump integral is computed using an explicit timestep.

Almendral and Oosterlee in (Oosterlee, 2007), solve for the value of a European contingent claim option, localizing and discretizing the PIDE in space using finite differences and finite elements and in time by using the second order backward differentiation formula, (BDF2). Then they use an iterative method based on the Fourier Transformation to solve the resulting system, based on the splitting of the matrix.

d' Halluin, Forsyth and Labahn (d' Halluin, 2004) discretize explicitly the jump-diffusion term and treat the usual PDE with a fully implicit method, proving unconditional stability for the resulting timestepping method. A fixed point iteration scheme is then used to solve the discretized algebraic equations. The correlation integral term is computed with FFT methods. They apply the proposed method to a variety of contingent claims.

Carr and Madan in (Madan, 1999) develop a simple analytic expression for the Fast Fourier Transformation (FFT) of the European option price or its time value for the Variance Gamma case.

Hirsa and Madan in (Madan, 2002) derive a PIDE for pricing American options when log-price dynamics of the underlying asset are given by the variance Gamma (VG) law. For the evaluation of the integral term, they expand the integrand near its singularity of $y = 0$ and treat this part implicitly, while the rest of the integral is treated explicitly.

Almendral and Oosterlee in (Oosterlee, 2006) propose a second order accurate finite-difference method for the computation of an American option price and its exercise boundary under the VG case. They formulate the problem as a linear complementarity problem and they solve it numerically using a splitting method. FFT is also used to accelerate the computations.

CGMY in (Madan, 2002) develop a continuous time model that allows for both diffusion and jumps of both finite/ infinite activity. In their article a pure jump process is defined that is of finite/ infinite activity if the number

of price jumps in any interval of time is finite/ infinite. The parameters of the process developed, allow the jump component to have either finite of infinite variation. A closed-form expression for the characteristic function of log-prices is generated but not for the return density. The Fourier inversion with the FFT is used as a technique to determine numerically the statistical and risk-neutral densities. On the empirical side, they conclude that one can dispense with the diffusion in describing the fine structure of asset returns as long as the jump process that is used is of infinite activity and finite variation.

Wang, Wan and Forsyth in (Wang, 2007) develop an implicit discretization method for pricing European and American options under the CGMY process. Taylor expansion approximation is used to treat the jump component in the neighborhood that the log-jump size equals zero and the drift term is dealt with using a semi-Langrangian scheme. The resulting PIDE is then solved using a preconditioned BiCGSTAB method coupled with FFT.

Almendral and Oosterlee in (Oosterlee, 2007) use a finite difference method for pricing European and American options under the CGMY process. The equations are discretized in space by the collocation method and in time by an explicit backward differentiation formula. Also the FFT transformation is used in the computation.

Ikonen and Toivanen (Toivanen, 2004), propose an operating splitting method for the timestepping and demonstrate it on the case of American options with stochastic volatility. This results to the decoupling of the early exercise constraint treatment and the solution of the resulting (from the discretization system) linear equations into separate fractional timesteps. They show that this enables the application of any efficient numerical method for the solution of the linear equations. Their method is proved to have same accuracy with Crank-Nicolson and its efficiency is tested on a multigrid period.

In (Toivanen, 2014) the authors define by a convex combination parameter, a set of implicit-explicit schemes for the solution of the PIDE resulting from the pricing of options following a jump-diffusion process. Those schemes lead to tridiagonal systems that are shown to be solved efficiently.

In the work presented here a finite difference scheme for solving the pre-mentioned PIDE is explored. An extension to jump-diffusion processes of the coordinate transformation approach (Parrott, 1999) is being developed along with a compact mesh based quadrature approach to the integral term.

The quadrature formulation is effective, has good localization properties and results to second-order accurate prices and Greeks for Merton's classical jump-diffusion pricing model. An implicit time discretization is used and early exercise is treated by classical iterative projection approach. The accuracy of the method presented, will be compared with results for Merton's model, VG, and CGMY processes published in (d' Halluin, 2005); (Madan, 2002) and (Wang, 2007). The effectiveness of the approach is also tested with the IT splitting method (Toivanen, 2004).

1.3 Thesis structure

This thesis contains of 4 main chapters excluding the introductory one, the conclusion and the appendix (Chapter 7). The contents of the chapters are described briefly in this section.

Chapter 2: This chapter summarizes the basic financial mathematics notions which will help us construct the building blocks of the research subject of this work[2]. Section 2.1 presents the basic definitions from stochastic calculus, the Brownian motion, the Markov property and their linkage to financial mathematics. The section then touches upon the properties of Martingale processes, a tool extremely important in the financial engineering field. Section 2.2 discuses Lévy processes, their main characteristic properties and several important results about them. We then turn our attention to their application in financial modelling and option pricing, with examples of Lévy processes that are widely used in the field discussing their activity and variation.

Chapter 3: In this chapter we present the PIDE whose efficient solution via finite difference approximation will be the focus of this work. In Section 3.1 we focus on the PIDE derived from the pricing of options following Merton's jump-diffusion process. We discuss the challenges arising in its numerical approximation and present our proposal for the efficient numerical solution for the case of European options. In Section 3.2 we present the changes in the approach to account for the early exercise in the American option case. Finally in Section 3.3 we present the results derived from the implementation of our proposed approach in the Merton case for a European and an American put option.

Chapter 4 In this section we discuss the changes in the PIDE when singularities arise for the Variance Gamma and CGMY process (Section 4.1) and derive the amended numerical approximation to treat those singularities. Results are presented and discussed (Section 4.2).

Chapter 5 In this chapter we present the numerical approximation described in Chapter 3, incorporating the IT splitting iterative method instead of PSOR. We re-calculate results for Merton's classical model, VG and CGMY and comment on the performance of IT splitting compared to PSOR.

[2] Important theorems will be presented, while the proofs of those theorems are not provided, as this is outside the scope of this work. The reader is referred to stochastic calculous literature instead.

2. Modelling the Stock price movements: Basic tools and Theorems

In this chapter we review the basic definitions and theorems, essential for the understanding of stochastic calculus tools and their application in financial engineering. Proofs of theorems widely available in the financial engineering literature will be omitted. Section 2.1 summarizes the basic stochastic calculus notions (σ-algebras, Brownian motion, Markov properties, martingales), Section 2.2 provides definitions and the main properties of *Lévy* processes. It then expands on their application in financial engineering and provides characteristic examples of *Lévy* processes that are of interest for the work presented here.

2.1 Stochastic processes-main notions and theorems

This work is aiming to numerically explore the dynamics of the arrival of random jumps in stock price paths. Those arrivals could be thought of as probabilistic events. In order to define mathematically the notion of a probabilistic event we need to define σ-algebras.

Definition 2.1 (σ-algebra) Let Ω be a set. A σ-algebra $\mathcal{F}$ on the set Ω is a family of subsets of Ω with the following properties.

(i) $\emptyset \in \mathcal{F}$
(ii) $F \in \mathcal{F} \Rightarrow F^C \equiv \Omega \backslash F \in \mathcal{F}$
(iii) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A \coloneqq \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

From the above definition we can derive the following property:

If $\mathcal{F}$ is a σ-algebra and $A_i \in \mathcal{F}$ then:

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcap_{i=1}^{\infty} A_i^C\right)^C \in \mathcal{F} \quad (2.1)$$

We can think of σ-algebras as information structures. Ω contains all possible outcomes of an experiment. Since by definition a σ-algebra is a collection of Ω's subsets, we can interpret it as a set containing all the possible questions one could pose for the experiment. The subsets of Ω are

called events. The properties of a σ-algebra are such that one could pose the questions:

(i) Could the event $A \subset \Omega$ occur, and:
(ii) Could event $B \subset \Omega$ occur, and:
(iii) Could both events A and B $(A \cap B)$ occur, or
(iv) Could only event A or event B $(A \cup B)$ occur, or
(v) Could event A not occur at all (A^C).

Definition 2.2 (Smallest σ-algebra) The smallest σ-algebra defined from a set A is the smallest σ-algebra that contains set A. This algebra is denoted as $\sigma(A)$.

For example, we can assume the algebra $\{\emptyset, \Omega, A, A^C\}$. This is the smallest σ-algebra that contains the set A, so this is the algebra $\sigma(A)$.

Definition 2.3 (Borel sigma Algebra) The smallest σ-algebra that contains the class C of all intervals $(-\infty, x)$ that can been seen as subsets of the real line is called Borel sigma Algebra and is denoted as $\mathcal{B}$. The elements of σ-algebra are called Borel sets.

Definition 2.4 The Borel σ-algebra, $\mathcal{B}(\mathbb{R}^d)$ is the smallest σ-algebra that contains all the parallelograms $(a, b]$, or in other words the σ-algebra that is created by parallelograms.

$\mathcal{B}(\mathbb{R}^d)$ contains all the subsets of $\mathbb{R}^d$ relevant to the work presented here.

Having defined the notion of a probabilistic event, the question that comes naturally is regarding the likelihood of this event occurring. This likelihood, is linked to a set belonging to an appropriately defined σ-algebra. The next step is to introduce the notion of a function that will link a set to a likelihood. Such functions are called measures. In what follows we will define the probability measure.

Definition 2.5 (Probability measure) A probability measure P on a measurable space $(\Omega, \mathcal{F})$ is a mapping $P: \mathcal{F} \rightarrow [0,1]$ with the properties:

(i) $P(\emptyset) = 0, P(\Omega) = 1$
(ii) If $A_1, A_2, \ldots \in \mathcal{F}$ and $\{A_i\}$ are pairwise disjoined, then

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \quad (2.2)$$

Definition 2.6 (Counting measure) Let $(\Omega, \mathcal{F})$ be a measurable space. For every set $A \in \mathcal{F}$ we can define the function $\delta: \mathcal{F} \to \mathbb{N} \subset \mathbb{R}$ in such a way so $\mu(A)$ will denote the number of the elements contained in set A. Function μ is a measure on space $(\Omega, \mathcal{F})$, called a counting measure. If we amend slightly the definition and define function $P: \mathcal{F} \to \mathbb{N} \subset \mathbb{R}$ such that $P(A) = \mu(A)/\#\Omega$, where $\#\Omega$ is the number of elements in Ω. Function P is then a probability measure.

For example, let $\Omega = \mathbb{R}$ and $\mathcal{F} = \mathcal{B}$. We can define measure $P: \mathcal{B} \to [0,1]$ such that the image of the space $I = (-\infty, x]$ is given by:

$$P(I) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-\frac{-y^2}{2}) dy \quad (2.3)$$

We can easily see that measure P is a probability measure on $\mathbb{R}$.

Definition 2.7* (*Probability space) Assume a σ-algebra on set Ω and a probability measure P. The triplet $(\Omega, \mathcal{F}, P)$ is called a probability space.

For example, let's assume $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}$ and P probability measure as defined in (2.3) above. The triplet $(\mathbb{R}, \mathcal{B}, P)$ is a probability space.

Definition 2.8* (*Measurable space) Subsets F of set Ω in the σ-algebra $\mathcal{F}$ are called $\mathcal{F}$-measurable. Measurable sets are always defined in relation to a specific σ-algebra.

For example, assume $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{B}$. Every interval $[a, b]$ where $a, b \in \mathbb{R}$ is $\mathcal{B}$-measurable.

Before moving on with the definition of stochastic processes, we will attempt to summarize the definitions provided above and provide their link to probability theory notions.

An experiment that contains a random element could be described as a probability space $(\Omega, \mathcal{F}, P)$.

More specifically:

- Ω is the ***sample space***. This set contains all possible outcomes of the experiment.
- An element $\omega \in \Omega$ is called a ***sample point*** and is a specific outcome of the experiment.
- The σ-algebra $\mathcal{F}$ is called a **family of events**. This algebra contains all possible questions one could pose in relation to the experiment.
- An ***even***t is an element of $\mathcal{F}$, in other words it is an $\mathcal{F}$-measurable subset of Ω. Those events could be more complex than a simple outcome of the experiment.
- ***Probability space*** P contains information on how likely is the occurrence of an event. In particular $P(F)$ contains information on how easy is for event $F \in \mathcal{F}$ to occur. For example if $P(F_1) > P(F_2)$ for two sets $F_1, F_2 \in \mathcal{F}$ then we can say that event $F_1$is more likely to occur than event F_2.

Definition 2.9 ($\mathcal{F}$-measurable function) For a set Ω, the function $Y: \Omega \to \mathbb{R}^d$ is called $\mathcal{F}$-measurable if for any measurable set $U \in \mathcal{F}$ the following holds:

$$Y^{-1}(U) = \{\omega \in \Omega; Y(\omega) \in U \quad (2.4)$$

for every open set $U \in \mathbb{R}^d$.

In order to answer the question whether the function Y takes a value in U we need to know the information contained in $\mathcal{F}$.

We can now define random variables.

Definition 2.10 (Random variable) A real random variable is an $\mathcal{F}$-measurable function $X: \Omega \to \mathbb{R}^d$ where $(\Omega, \mathcal{F}, P)$ is a probability space.

We can think of random variable X as a variable whose value depends on the outcome of a random experiment.

Every random variable is linked to a measure $\mu_X(B) = P(X^{-1}(B))$ for a Borel set $B \in \mathcal{B}(\mathbb{R}^d)$. This measure is called the ***distribution*** or ***law*** of the random variable. Setting $d = 1$ and $\mathcal{F} = \mathcal{B}$ we can define the distribution function of a random variable X as:

$$P(X \leq x) = F_X(x)$$
$$P(a < X \leq b) = F_X(b) - F_X(a) \quad (2.5)$$

F_X defines the probability measure P. It also satisfies the following:

- F_X is an increasing function
- F_X is right continuous: ($\lim_{x \to -\infty} F_X(x) = 0$) and $\lim_{x \to \infty} F_X(x) = 1$
- It holds:

$$F_X(x) = \int_{-\infty}^{+\infty} f(x)dx = 1 \quad (2.6)$$

for $f(x)$ the probability density function.

We can now move on to defining stochastic processes.

Definition 2.11 (Stochastic process) A stochastic process is a parametric set of random variables $\{X_t\}_{t \epsilon T}$ which are defined in a probability space $(\Omega, \mathcal{F}, P)$ and take values at $\mathbb{R}^d$. A stochastic process has two variables, t and ω.

1. For any given and constant $t \in T$ there is a random variable ω that satisfies:
$$\omega \to X_t(\omega); \omega \in \Omega \quad (2.7)$$

2. For a constant $\omega \in \Omega$ the function:

$$t \to X_t(\omega);\ t \in T \quad (2.8)$$

presents the path of X_t.

In order to understand the notion of the stochastic process we need to consider a set of particles which are observed in time. t presents time and can be either continuous or discrete. Let ω represent a specific particle. A specific ω is called a realization of the stochastic process. Then $X_t(\omega)$ is the position of the particle ω at time t. We can identify each ω with the function $\tau \to X_t(\omega)$ that reflects $T \to \mathbb{R}^d$. Then the σ-algebra $\mathcal{F}$ will include the σ-algebra $\mathcal{B}$ that is derived from sets of the following format:

$$\{\omega, \omega(t_1) \in F_1; \dots; \omega(t_k) \in F_k\} \; F_i \in \mathbb{R}^d \text{ Borel sets} \quad (2.9)$$

Stochastic processes that are commonly used by researchers and market practitioners in order to model stock price behavior are the continuous-variable, continuous-time stochastic processes. While in the real market it is not observed that stock prices follow continuous-variable, continuous-time processes, an understanding of these types of processes is the first step towards the understanding of option pricing models and of pricing models for more complicated derivative products.

2.1.1 Brownian motion in financial modelling

The Brownian motion is undoubtedly the most commonly used tool amongst the stochastic processes for modelling price fluctuations. It is the most widely studied stochastic process and the cornerstone of the modern stochastic analysis. It is the simplest continuous-time stochastic process and it is a limit of both simpler and more complicated stochastic processes.

Definition 2.12 (Brownian motion) A Brownian motion is a stochastic process B_t that takes values in $\mathbb{R}$ and satisfies the following properties:

1. The process has independent increments, i.e:

For all times $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$, the increments:

$$B(t_n) - B(t_{n-1}), B(t_{n-1}) - B(t_{n-2}), \dots, B(t_2) - B(t_1) \quad (2.10)$$

are independent random variables.

2. If $s, t \geq 0$, then

$$P\{(B_{s+t} - B_s) \in A\} = \int_A \frac{1}{2\pi t^{1/2}} \exp(-\frac{|x|^2}{2t})\, dx \qquad (2.11)$$

with A denoting a Borel set, so changes in the Brownian motion follow the Normal Distribution (Gaussian distribution).

3. The paths followed by the Brownian motion are continuous with probability 1. So, $t \to B(t)$ is a continuous function.

An important feature of the Brownian motion is that is satisfies the Markov property and the strong Markov property[3], both important properties in the calculations involving Brownian motion, such us estimation of conditional expectations with respect to filtrations defining various types of stochastic processes that appear in the modelling of financial derivatives.

2.1.2 The Markov property

The Markov process is a particular type of stochastic process where only the present value of a variable is relevant to its future path. Predictions of the future movements of the variable are uncertain and must be expressed in terms of probability distributions. The Markov property implies that the probability distribution of the underlying variable at any particular future point is not dependent on the path followed by the variable in the past. This property is closely related to the weak form of *market efficiency*[4], which implies that the present price of a stock contains all information of the price path in the past.

Theorem 2.1 (Markov Process) Let $\mathcal{F}_s = \sigma(B_u, u \leq s)$, the σ-algebra created by the Brownian motion up to time s. This is the smallest σ-algebra where the random variable B_r, $r \leq s$ is countable. $\mathcal{F}_s$ contains all the information of the Brownian motion path up to s. If f is a bounded function, then for every $x \in \mathbb{R}^d$:

$$E_x[f(B_{t+s} - B_s)|\mathcal{F}_s] = E_x[f(B_{t+s} - B_s)] = \int_{-\infty}^{\infty} f(y) \frac{1}{\sqrt{2\pi t}} \exp(-\frac{-y^2}{2t}) dy \quad (2.12)$$

due to the independence of the Brownian motion. We can therefore write:

$$E_x[f(B_{t+s})|\mathcal{F}_s] = E_x[f(B_{t+s} - B_s + B_s)|\mathcal{F}_s] \quad (2.13)$$

[3] Here, only the Markov property will be discussed as this section serves as a primer of the basic stochastic tools.

[4] In the weak form of market efficiency, prices reflect all information contained in the market trading data. Historical pricing data are efficiently digested by millions of competing market participants, making the information useless in identifying mispriced securities. It is therefore impossible for a single investor to consistently outperform the market.

Since $B_{t+s} - B_s$ is independent of the $\mathcal{F}_s$ algebra and B_s fully known until time s, we can assume that $B_s = z$. Then:

$$E_x[f(B_{t+s} - B_s + B_s)|\mathcal{F}_s] = E_x[f(B_{t+s} - B_s + z)] \quad (2.14)$$

However, $B_{t+s} - B_s$ is also a Brownian motion that starts at 0 and *runs* for time $t = s - s = t$. Therefore, the process $B_{t+s} - B_s + z$ has the same distribution with a Brownian motion that starts at time z and *runs* for time t, so that:

$$\begin{aligned} E_x[f(B_{t+s} - B_s) + z] \\ &= \int_{-\infty}^{\infty} f(y+z) \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{-y^2}{2t}\right) dy \\ &= \int_{-\infty}^{\infty} f(y) \frac{1}{\sqrt{2\pi t}} \exp(-\frac{(y-z)^2}{2t}) dy = E_z[f(B_t)] \\ &= E_{B_s}[f(B_t)] \end{aligned} \quad (2.15)$$

We therefore conclude that:

$$E_x[f(B_{t+s})|\mathcal{F}_s] = E_{B_s}[f(B_t)] = \int_{-\infty}^{\infty} f(y) \frac{1}{\sqrt{2t}} \exp(-\frac{(y-B_s)^2}{2t}) dy \quad (2.16)$$

This property is an expression of the Markov Property for the Brownian motion. An equivalent expression for the Markov Property is:

$$E_x[f(B_t)|\mathcal{F}_s] = E_{B_s}(f(B_{t-s})), s \leq t \quad (2.17)$$

The above implies that in order to calculate the expectation at time t of a Brownian motion, conditional on its *history* up to time s, we can simply calculate the expectation of a *new* Brownian motion that starts at the position that the initial Brownian motion reached at time s , i.e. B_s, and runs for time $t - s$. In other words, the entire history of the initial Brownian motion prior to time s does not contain any useful information.

2.1.3 Martingales and arbitrage Theory

Martingales are a special class of stochastic processes which play an important role in probability theory, stochastic calculus and financial engineering. Here we present the basic notions of martingale theory.

Definition 2.13 (Martingale) Let $(\Omega, \mathcal{F}, P)$ be a probability space with $\mathcal{F}_t$ the filtration of $\mathcal{F}(\mathcal{F}_t \subset \mathcal{F})$ and X_t a group of real, integrable $(E[|X_t|] < \infty])$ random variables of the $\mathcal{F}_t$ filtration. Then:

1. X_t is a martingale if:

$$E[X_t|\mathcal{F}_s] = X_s \, , s \leq t$$

2. X_t is a supermartingale if:

$$E[X_t|\mathcal{F}_s] \leq X_s, s \leq t$$

3. X_t is a submartingale if:

$$E[X_t|\mathcal{F}_s] \geq X_s, s \leq t$$

Where t can be a continuous indicator $t \in \mathbb{R}$ or a discrete indicator $\in \mathcal{N}$.

If for a Martingale the information contained in $\mathcal{F}_s$ is known, then the best prediction that can be made for the value of X_t is X_s. If X_t is a supermartingale then the best prediction that can be made about the value of X_t based on the information contained in the filtration $\mathcal{F}_s$, will be greater than the value of X_s. Accordingly if X_t is a submartingale then the best prediction will be less than X_s.

Definition 2.14 If X_t is a Martingale then:

1. $E[X_t] = E[X_0]$

2. $E[X_t - X_s] = 0$

Let X_t a stochastic process with t denoting time. $\mathcal{F}_t$ can be any filtration. A possible choice that can be made is the normal filtration $\mathcal{F}_t = \sigma(X_u, u \leq t)$, which is the filtration produced from the trajectories of the random process.

In this case $\mathcal{F}_t$ can be viewed as the information derived from the behavior of the stochastic process X_t from time $t = 0$ up to time t.

Theorem 2.2 (Arbitrage) A portfolio can be an arbitrage portfolio if it is self-financing[5] and if the following holds:

$$V_0(a) = 0, V_t(a) \geq 0 \forall\, t \text{ and } E[V_T(a)] > 0$$

Arbitrage implies that an investor can have sure profit from a portfolio without imposing him/ herself to any risk. The absence of arbitrage opportunities in the market is equivalent to the existence of an equivalent probability measure under which the discounted price process is a martingale. The existence of a martingale measure allows for the fair pricing of the contingent claims.

The theorem below presents the first fundamental theorem in asset pricing.

Theorem 2.3 There are no arbitrage opportunities in the market if and only if there exists an equivalent martingale measure.

The value of a self-financing portfolio is given from:

$$V_t = V_0 + \sum_{u=1}^{t} a_u \Delta X_u \quad (2.18)$$

where

$$a_u \Delta X_u = \sum_{n=0}^{N} a_{n,u} X_u^n \quad (2.19)$$

We assume that under the Q probability measure the discounted stochastic process X describes the price change under filtration $\mathcal{F}$. V is a martingale transformation of the process X. Consequently V is also a martingale under the measure Q. The martingale property implies that:

$$E_Q[V_t] = E_Q[V_0] \quad (2.20)$$

with $V_0 = 0$ and $V_T \geq 0$ under Q. Using the martingale property $E_Q[V_T]$ combined with the fact that $V_T \geq 0$ under Q, we can conlclude that $V_T = 0$ under Q.

[5] Please see Section 7.1 for a definition of a self-financing portfolio.

Remark 2.1 The P measure is the probability measure derived from the stochastic process X_t. The above theorem also holds for continuous time models. In that case there are specific stochastic calculus tools that can be used to test the existence of equivalent probability measures (Girsanov's theorem). Under the equivalent martingale measure all asset prices discounted by the current interest rate process are martingales. This result is used extensively in derivative's pricing.

The theorem below plays an important role in defining derivative hedging strategies. In its simplified form it states that for a one asset model where the evolution of its price is modelled by a Browning motion, the existence of a hedging strategy depends on the following:

Theorem 2.4 (Martingale representation, one dimension) Let B_t, $0 \leq t \leq T$, be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and let $B(t)$ be a filtration generated by this Brownian motion. Let $M(t), 0 \leq t \leq T$, be a martingale with respect to this filtration. That is for every $t, M(t)$ is $B(t)$-measurable and for $0 \leq s \leq t \leq T$, $(\mathbb{E}[M(t)|B(s)] = M(s))$. Then there is an adapted process $\Gamma(u), 0 \leq u \leq T$, such that:

$$M(t) = M(0) + \int_0^t \Gamma(u) dW(u), 0 \leq \tau \leq T \quad (2.21)$$

Defining this adapted process $\Gamma(u)$ poses a challenge in the solution of any Brownian motion based model.

Below we present another major breakthrough in the stochastic calculus field relating to stochastic integration.

Definition 2.15 (Itô processes) An Itô process is a stochastic process X_t of the following form:

$$X_t = X_0 + \int_0^t u(s, \omega) ds + \int_0^t v(s, \omega) dB_s \quad (2.22)$$

Where u and v satisfy the conditions:

$$\int_0^t v^2(s, \omega) ds < \infty, \int_0^t u(s, \omega) ds < \infty, \quad (2.23)$$

This process can be also written in the differential form:

$$dX_T = u dt + v dB_t \quad (2.24)$$

From the above we can see that an Itô process can break into two parts:

The $M_t \coloneqq \int_0^t v dB_t$ part, which is a martingale and the $A_t \coloneqq \int_0^t u ds$ part, which is a process of finite variation.

It is important to explore whether a function of an Itô process is going to be an Itô process itself. If this it true, then the question that arises is which is the exact form of this function, as a summation of a Riemann integral and an Itô integral. Itô's Lemma provides an answer to that question suggesting a change of variable formula that also holds for stochastic integrals.

Definition 2.16 (Itô-Doeblin Lemma) Let X_t represent an Itô process that can also be expressed as:

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dB_s \quad (2.25)$$

Then, any function of the process X_t, of the form $g(t,x) \in C^{1,2}$, can be expressed as a stochastic integral of the following form:

$$g(t,X_t) = g(0,X_0) + \int_0^t \left(\frac{\partial g}{\partial s} + u\frac{\partial g}{\partial x} + \frac{1}{2}v^2\frac{\partial^2 g}{\partial x^2}\right)ds + \int_0^t v\frac{\partial g}{\partial x}dB_s \quad (2.26)$$

The above result can be written as:

$$dg(t,X_t) = \left(\frac{\partial g}{\partial t} + u\frac{\partial g}{\partial x} + \frac{1}{2}v^2\frac{\partial^2 g}{\partial x^2}\right)dt + v\frac{\partial g}{\partial x}dB_t \quad (2.27)$$

With $C^{1,2}$ denoting the space of the $g(t,x)$ functions that have continuous first derivative of the first variable and continuous second derivative of the second variable.

Remark 2.2 In the above the following notation has been used:

$$\frac{\partial g}{\partial s} \coloneqq \frac{\partial g}{\partial t}(s,X_s),\ u\frac{\partial g}{\partial x} \coloneqq u(s,X_s)\frac{\partial g}{\partial x}(s,X_s), \text{ etc} \quad (2.28)$$

A more tractable form of the Itô Lemma is the following:

Let $Y_t = g(t,X_t)$ be a function of an Itô process. Then:

$$dY_t = \frac{\partial g}{\partial t}dt + \frac{\partial g}{\partial x}dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(dX_t)^2 \quad (2.29)$$

In what follows we provide a summary on Girsanov's theorem and build up to the first option pricing PDE. Girsanov's theorem, also known as the measure change theorem, is essential in understanding the connection of two stochastic processes defined under different probability measures.

Before presenting the theorem, the equivalent martingale measure is presented.

Definition 2.17 (Equivalent probability measure) Let Q and P two probability measures. These measures are called equivalent probability martingale measures if for every event A, $P(A) = 0$ if and only if $Q(A) = 0$. In this case there always exists a random variable ξ which is called Radon-Nikodym derivative of Q in respect to P. ξ satisfies the following property:

Let a stochastic process Z_t. If Z satisfies $E_Q[|Z|] < \infty$, then $E_Q[Z] = E_P[\xi Z]$. The above assumption is established by the Radon-Nikodym theorem.

Theorem 2.5 (Radon-Nikodym) Let a probability space $(\Omega, \mathcal{F}, P)$ and Q and P finite probability measures defined in σ-algebra $\mathcal{F}$. The following two statements are equivalent:

1. For every $A \in \mathcal{F}$, $P(A) = 0 \Rightarrow Q(A) = 0$
2. There exists a $k \in L^1(P), k \geq 0$ such that $Q(A) = \int_A k dP$

The density factor k is called Random-Nikodym derivative and is defined as $k = \frac{dQ}{dP}$.

Theorem 2.6 (Girsanov's) Let $u_t = (u_{1_t}, \dots, u_{n_t})$ a vector that consists of (square integrable) stochastic processes which satisfy the Novikov condition:

$$E[\exp(\frac{1}{2}\int_0^T u_s . u_s ds)] < \infty \quad (2.30)$$

Let B_t be a typical Brownian motion under the measure P. We define the stochastic process:

$$M_t^u = \exp\left[-\int_0^t u_s dB_s - \frac{1}{2}\int_0^t u_s . u_s ds\right], t \in [0, T] \qquad (2.31)$$

This process is a Martingale under the measure P.

If we assume that Q^u is the equivalent probability measure defined as:

$$\frac{dQ^u}{dP} = M_t^u \quad (2.32)$$

then the stochastic process B_t^u, defined by:

$$B_t^u = B_t + \int_0^t u_s ds, \qquad 0 \leq t \leq T \qquad (2.33)$$

is a typical Brownian motion and a Martingale under the measure Q^u. Moreover, B_t^u has the following property:

for every local $Q^u - martingale$ there is a φ stochastic process which is square integrable and satisfies:

$$L_t = L_0 + \int_0^t \varphi_s dB_s^u,\ t \leq T \qquad (2.34)$$

Collaterally the following theorem is extracted:

Theorem 2.7 Let X an Itô process defined in $\mathbb{R}^n$:

$$X_t = x + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s \qquad (2.35)$$

Let v_t a vector of integrable stochastic processes and let's also assume that there is a vector that consists of stochastic processes u_t which are square integrable and satisfy:

$$u_t \sigma_t = \mu_t - v_t \quad (2.36)$$

Then if X_t is a martingale under measure P (for which measure the Novikov property holds), the process X_t is also and Itô process with:

$$X_t = x + \int_0^t v_t ds + \int_0^t \sigma_s dB_s^u \quad (2.37)$$

where B_s^u is the typical Brownian motion under the measure Q^u which is defined as:

$$B_t^u = B_t + \int_0^t u_s ds \quad (2.38)$$

The proof of this theorem is an implementation of Girsanov's theorem presented above. In an effort to interpret Girsanov's theorem, let's assume that we keep the trajectories of the stochastic process B_t as they are. We actually keep a collection of functions of the form $B_t(\omega)$ and we change the frequency of the occurrence of different realisations of ω. This equals to the change of measure P to the equivalent measure Q. Girsanov's theorem allows us to change the velocity of a stochastic process. In a way it combines the solutions of the stochastic processes. Furthermore, by using the Feynman-Kac representation we can combine the solutions of the stochastic processes with their partial derivatives.

Lastly, we will present the Feynman-Kac format which offers a useful tool towards the solution of parabolic problems.

Let the operator:

$$A = \sum_i b_i(x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{ij} (\sigma\sigma)^T{}_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \quad (2.39)$$

The Feynman-Kac format implies that the solution of the Cauchy problem:

$$\frac{\partial u}{\partial t} = AU + cU \quad (2.40)$$

$$u(t,x) \in C^{1,2}([0, \infty \times \mathbb{R}^n]) \quad (2.41)$$

$$u(x,0) = f(x) \quad (2.42)$$

for continuous c and bounded function $f \in C_0^2$, can be written as the mean value over the trajectories of the Itô process X_t which has the operator A as a generator operator.

Theorem 2.8 (Feynman-Kac representation) The solution of the Cauchy problem can be written as:

$$u(t,x) = E_x[f(X_t)\exp(\int_0^t c(X_s)ds)] \qquad (2.43)$$

Let's assume that we want to estimate the solution of the Cauchy problem at the point (x,t). We then take a number of $X_t(\omega_i)$ diffusions, which have A as an operator and share the same starting point $(x,0)$. We allow the diffusion to run for time t and for each one of the trajectories we estimate the quantity $f(X_t(\omega_i))\exp(\int_0^t c(X_s(\omega_i))ds)$. We then take the mean value over all the trajectories (for all the ω_i). This mean is the solution to the equation.

2.2 Lévy processes

This section will present key examples of continuous-time stochastic processes-*Lévy processes*-an important tool in describing the asset price behavior observed in financial markets. We will summarize basic theorems and transition to their application in financial engineering. For a comprehensive read on Lévy processes and their properties the reader is referred to (Cont, 2004) and (Papantoleon, 2000).

Let $(\Omega, \mathcal{F}, F, P)$ be a filtered probability space, where $\mathcal{F} = F$ and the filtration $F = (\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions. Let $T \in [0,\infty]$ denote the time horizon which in general can be infinite.

Definition 2.18 (Lévy process) A cadlag, adapted, real valued stochastic process $L = (L_t)_{t\geq 0}$ with $L_0 = 0$ is a Lévy process if the following conditions are satisfied:

1. L has independent increments, i.e. $L_t - L_s$ is independent of $\mathcal{F}_s$ for any $0 \leq s < t \leq T$
2. L has stationary increments, i.e.for any $0 \leq s, t \leq T$ the distribution of $L_{t+s} - L_t$ does not depend on t.
3. L is stochastically continuous, i.e. for every $t \geq 0$ and $\varepsilon > 0$: $lim_{s\to t} P(|L_t - L_s| > \varepsilon) = 0$

There is a strong interplay between Lévy processes and random walk models. By sampling a Lévy process at regular time intervals $0, \Delta t, 2\Delta t, \dots, n\Delta t$ we obtain a random walk.

For any sampling interval Δ, and by defining $S_n(\Delta) = X_{n\Delta}$ we can write $S_n(\Delta) = \sum_{k=0}^{n-1} Y_k$ where $Y_k = X_{(k+1)\Delta} - X_{k\Delta}$ are i.i.d random variables whose distribution is the same as the distribution of X_Δ. We can therefore conclude that for any interval Δ, $n > 1$ and time $t > 0$ the distribution $S_n(\Delta)$ can be broken into n i.i.d parts. In other words, by sampling a Lévy process at regular time intervals $0, \Delta t, 2\Delta t, \dots, n\Delta t$ we obtain a random walk.

This property is called infinite divisibility:

Let X be a real valued random variable with characteristic function:

$$\varphi_X(u) = \int_{\mathbb{R}} e^{iux} P_X(dx) \quad (2.44)$$

where P_X denotes the distribution law. Let $\mu * v$ denote the convolution of the measures μ and v, given by:

$$(\mu * v)(A) = \int_{\mathbb{R}} v(A - x)\mu(dx) \quad (2.45)$$

Definition 2.19 (Infinite Divisibility) The law P_X of random variables is infinitely divisible if for all $n \in \mathbb{N}$ there exist i.i.d. random variables $X_1^{(1/n)}, \dots, X_n^{(1/n)}$ such that:

$$X = X_1^{(1/n)} + \cdots + X_n^{(1/n)} \quad (2.46)$$

Equivalently, we can say that the law P_X of a random variable X is infinitely divisible is for $n \in \mathbb{N}$ there exists another law $P_{X^{1/n}}$ of a random variable $X^{(1/n)}$ such that:

$$P_X = \underbrace{P_{X^{(1/n)}} * \dots * P_{X^{(1/n)}}}_{n\ times} \quad (2.47)$$

Another way to characterize an infinitely divisible random variable is via its characteristic function.

Definition 2.20 The law of a random variable X is infinitely divisible, if for all $n \in \mathbb{N}$, there exists a random variable $X^{(1/n)}$, such that:

$$\varphi_X(u) = (\varphi_{X^{(1/n)}}(u))^n \quad (2.48)$$

The simplest example of an infinite divisible process is the linear drift, which is a deterministic process. Brownian motion is the only (non-deterministic) Lévy process with continuous sample paths. Other famous examples of infinitely divisible processes are the Poisson and Compound Poisson process. The sum of a linear drift, a Brownian motion and a compound Poisson process is again a Lévy process; this summation presents a so called "jump-diffusion" model or Lévy jump-diffusion model which will be the focus of this work.

Next, we provide a complete characterization of random variables with infinitely divisible distributions via their characteristic functions. This is the *Lévy-Khintchine* formula.

Theorem 2.9 The law P_X of a random variable X is infinitely divisible if and only if there exists a triplet (c, v, b), with $b \in \mathbb{R}$, $c \in \mathbb{R}$, $c \geq 0$ and a measure satisfying $v(\{0\}) = 0$, and $\int_{\mathbb{R}}(1 \wedge |x|^2 v(dx)) < \infty$, such that:

$$\mathbb{E}\left[e^{iuX}\right] = \exp[ibu - \frac{u^2 c}{2} \int_{\mathbb{R}} (e^{iux} - 1 - iux1_{\{|x|<1\}})v(dx)] \quad (2.49)$$

The above is based on the following Lemma presented in (Sato, 1999).

Lemma 2.1 If $(P_k)_{k\geq 0}$ is a sequence of infinitely divisible laws and $P_k \to P$, then P is also infinitely divisible.

The full proof of the theorem can be found in (Sato, 1999).

The exponent in equation (2.49):

$$ibu - \frac{u^2 c}{2} \int_{\mathbb{R}} (e^{iux} - 1 - iux1_{\{|x|<1\}})v(dx) \quad (2.50)$$

is called the Lévy or characteristic exponent, with the triplet (c, v, b) called the Lévy or characteristic triplet. Moreover, $b \in \mathbb{R}$ is called the drift term, $c \in \mathbb{R}, c \geq 0$ the Gaussian or diffusion coefficient and v the Lévy measure.

This leads us to the following theorem:

Theorem 2.10 For every Lévy process $L = (L_t)_{0 \le t \le T}$ we have that:

$$\mathbb{E}\left[e^{iuL_t}\right] = e^{t\psi(u)}$$
$$= \exp[t(ibu - \frac{u^2 c}{2} + \int_{\mathbb{R}} (e^{iux} - 1 - iux1_{\{|x|<1\}}\, v(dx))] \quad (2.51)$$

with $\psi(u)$, the characteristic exponent of $L_i, i \in \mathbb{N}$-a random variable with an infinitely divisible distribution.

The quantity v presents the Lévy measure, which gives the expected number per unit of time of jumps that are of size x. If the measure v presents a density with respect to the Lebesque measure[6], then it can be called the Lévy density of X devoted as $v(x)$.

We have previously seen that every Lévy process can be associated with the law of an infinitely divisible distribution. The opposite, i.e. that given any random variable X, whose law is infinitely divisible, we can construct a Lévy process $L = (L_t)_{0 \le t \le T}$ such that $\mathcal{L}aw(L_i) \coloneqq \mathcal{L}aw(X)$, is also true.

2.2.1 Jumps and poisson processes

The two basic tools for building a jump-diffusion process are the Brownian motion, which represents the diffusion part and the Poisson process, which represents the jump part. In the following section some primer information about the Poisson process is going to be presented.

Definition 2.21 (The Poisson Process) A sequence $\{\tau_\iota\}_{\iota \ge 1}$ of independent exponential random variables with parameter λ and the process $(N_t, t \ge 0)$ defined by:

$$N_t = \sum_{n \ge 1} 1_{t \ge T_n}$$

is called Poisson process with intensity λ.

[6] Given a set A of real numbers, μ(A) will denote its Lebesque measure if it's defined.

The trajectories of a Poisson process are piecewise constant, with jump size 1 only. The $(N_t, t \geq 0)$ process can also be described as a counting process.

Proposition 2.1 Let $(N_t, t \geq 0)$ be a Poisson process. Then:

1. For any $t > 0,\ N_t$ is almost surely finite.

2. For any ω, the sample path $t \to N_t$ is piecewise constant and increases by jumps of size 1.

3. The sample paths $t \to N_t$ are right continuous with left limit (cadlag process).

4. For any $t > 0$, $N_{t-} = N_t$ with probability 1.

5. (N_t) is continuous in probability:

$$\forall t > 0, N_s \underset{s \to t}{\longrightarrow} N_t$$

6. For any $t > 0$, N_t follows a Poisson distribution with parameter λt:

$$\forall n \in N, P(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

7. The characteristic function of N_t is given by:

$$E\left[e^{iuN_t}\right] = \exp\left\{\lambda t\left(e^{iu} - 1\right)\right\},\ \text{for every } u \in \mathbb{R}.$$

8. N_t has independent increments: for any $t_1 < \cdots < t_n$, $N_{t_n} - N_{t_{n-1}}, \ldots, N_{t_2} - N_{t_1}, N_{t_1}$ are independent random variables.

9. The increments of N are homogeneous: for any $t > s$, $N_t - N_s$ has the same distribution as N_{t-s}.

10. (N_t) satisfies the Markov property:

$$\forall t > s, E[f(N_t)|N_u, u \leq s] = E[f(N_t)|N_s]$$

Properties 1 and 2 actually imply that any path of Poisson process is almost surely discontinuous and moves only by jumps. On the other hand property 4 implies that for every given point in time t the sample function of a Poisson process is continuous with probability 1. This paradox is due to the

fact that as for all jump processes the points of discontinuity of the process form a set of zero numbers.

Moreover the cadlag property of the Poisson process is actually chosen to be like that, as then the jumps in the process are interpreted as sudden events, a feature that we will use in this work (i.e. unpredictable jumps in option pricing dynamics).

When it comes to the study of jump-diffusion processes, the notion of the characteristic function plays an essential role.

It is possible not to know the distribution function of a jump-diffusion process in a closed form, while knowing explicitly its characteristic function. The characteristic function of a random variable X is defined by:

$$\varphi_X(u) \equiv E\left[e^{iuX}\right] \quad \forall u \in \mathbb{R}^d \quad (2.52)$$

For the Poisson process, this gives:

$$E\left[e^{iuN_t}\right] = \exp\left\{\lambda t\left(e^{iu-1}\right)\right\} \ \forall u \in \mathbb{R} \qquad (2.53)$$

Let $\Delta L = \Delta L_{t\,0\le t\le T}$ present the jump process associated to a Lévy process L which is defined for each $0 \le t \le T$, via:

$$\Delta L_t = L_t - L_{t-1} \quad (2.54)$$

where $L_{t-1} = \lim_{s\uparrow t} L_s$. Given the condition of stochastic continuity of a Lévy process, we can immediately derive that for any Lévy process L and any fixed $t > 0$, then $\Delta L_t = 0$, i.e. a Lévy process has no fixed times of discontinuity.

Generally, the sum of the jumps of a Lévy process does not converge. It is therefore possible that:

$$\sum_{s\le t} |\Delta L_s| = \infty$$

But the below equality will always hold:

$$\sum_{s\le t} |\Delta L_s|^2 < \infty$$

This property allows us to handle Lévy process using martingale techniques.

In order to analyse the jumps of a Lévy process we need to understand the nature of the measure of those jumps.

We first define a process that counts the jumps of the process L.

Definition 2.22 (Counting process) A counting process counts the number of random times T_n which occur between $(0, T)$, where $(T_n - T_{n-1})_{n\geq 1}$ is a sequence of i.i.d. exponential variables. Given an increasing sequence of random times T_n, $n \geq 1$ with $P(T_n \to \infty) = 1$, we can define the associated counting process $(X_t)_{t\geq 0}$ as:

$$X_t = \sum_{n\geq 1} 1_{t\geq T_n} = \#\{\, n \geq 1,\ T_n \geq t\} \quad (2.55)$$

X_t is actually the number of random times $T_n,\ n \geq 1$ that occur in the time interval $[0, t]$. By imposing the condition $P(T_n \to \infty) = 1$ we guarantee that with probability 1 the process X_t is finite for every $t \geq 0$ and like the *Poisson* process it is a cadlag process with piecewise constant trajectories, with sample paths moving also by jumps of size $+1$.

Definition 2.23 (Compensated Poisson process) Let $\tilde{N}_t = N_t - \lambda t$ be a centred version of the Poisson process N_t. N_t follows this centred version of Poisson law with the characteristic function:

$$\Phi_{\tilde{N}_t}(z) = \exp[\lambda t(e^{iz} - 1 - iz)] \quad (2.56)$$

Like the Poisson process also $\tilde{N}_t$ has independent increments and it follows that:

$$E[N_t | N_s, s \leq t] = E[N_t - N_s + N_s | N_s] \quad (2.57)$$

$$= E[N_t - N_s] + N_s = \lambda(t - s) + N_s$$

So the process $(\tilde{N}_t)$ has the martingale property:

$$\forall t > s, E[\tilde{N}_t | \tilde{N}_s] = \tilde{N}_s \quad (2.58)$$

The process $(\tilde{N}_t)_{t\geq 0}$ is a compensated Poisson process and $(\lambda t)_{t\geq 0}$ is the compensator of $(N\_t)_{t\geq 0}$ and is actually the quantity that needs to be subtracted from N_t to get a martingale.

Definition 2.24 (Compound Poisson Process) Let N_t, a Poisson process with parameter λ and $Y_i \geq 1$ a sequence of i.i.d random variables with distribution$f$. Then, the process:

$$X_t = \sum_{i=1}^{N_t} Y_i \quad (2.59)$$

is called a compound Poisson process.

Its trajectories are right continuous with left limits and piecewise constant and the jump sizes are now random with distribution f. The compound Poisson process has independent and stationary increments. Its characteristic function has the following form:

$$E\left[e^{iuX_t}\right] = \exp\left\{t\lambda \int_{\mathbb{R}} \left(e^{iux} - 1\right) f(dx)\right\} \; \forall\, u \in \mathbb{R}^d \quad (2.60)$$

where λ denotes the jump intensity and f the jump size distribution.

By comparing the characteristic function of the compound Poisson process to the characteristic function of the Poisson process we conclude that a compound Poisson random variable can be understood as a superposition of independent Poisson processes with different jump sizes. The total intensity of a Poisson process with jump sizes in the interval $[x, x + dx]$ is determined by density $\lambda f(dx)$.

Proposition 2.2 $(X_t)_{t\geq 0}$ is a compound Poisson process if and only if it is a Lévy process and its sample paths are piecewise constant functions[7].

Definition 2.25 (Jump measure of a compound Poisson process) Let $(X_t)_{t\geq 0}$ be a compound Poisson process with λ intensity and jump size distribution f. The jump measure J_X is a Poisson random measure counting the number of jumps occurring per unit of time with jump sizes in $A \subset \mathbb{R}^d \times [0, \infty)$:

[7] Proof in (Cont, 2004)

$$J_x(A) = \#\{(t, X_t - X_{t\_} \in A\} \quad (2.61)$$

This leads to the following proposition:

Proposition 2.3 Let $(X_t)_{\geq 0}$ be a compound Poisson process with intensity λ and jump size distribution f. Its jump measure J_X is a Poisson random measure on $\mathbb{R}^d \times [0, \infty)$ with intensity measure $\mu(dx \times dt) = v(dx)dt = \lambda f(dx)dt$.

This proposition actually suggests another way of interpreting the Lévy measure of a compound Poisson process as the average number of jumps per unit of time. It can be used to define the Lévy measure for all Lévy processes.

Definition 2.26 (Lévy measure) Measure v of a $(X_t)_{\geq 0}$ Lévy process on $\mathbb{R}^d$ defined as:

$$v(A) = E[\#\{t \in [0,1]: \Delta X_t \neq 0, \Delta X_t \in A\}], A \in \boldsymbol{\mathcal{B}}(\mathbb{R}^d) \quad (2.62)$$

is called the Lévy measure of X and denotes the expected number per unit of time of jump sizes in A.

Generally v is not a probability measure and $\int v(dx)$ need not to be finite. In the case where $\lambda = \int_{\mathbb{R}^d} v(dx) < +\infty$, the measure v can be normalized to define a probability measure μ which can now be interpreted as the distribution of jump size x:

$$\mu(dx) = \frac{v(dx)}{\lambda} \quad (2.63)$$

The jumps of X are then described by a compound Poisson process with λ as a jump intensity (average number of jumps per unit of time) and jump distribution $\mu(.)$. Generally, if $\int_{|x|\leq 1} |x| v(dx) < \infty$, the sum of jumps is absolutely convergent with probability 1 and X_t can be represented as a path-wise sum of a Brownian motion plus jumps:

$$X_t = \sigma W_t + \gamma_0 t + \sum_{0<s\leq t} \Delta X_t \quad (2.64)$$

where $\gamma_0 = \gamma - \int_{|x|\leq 1} xv(dx)$. In this case the compensation of small jumps is not needed and the Lévy-Khintchine representation reduces to:

$$E\left[e^{izx}\right] = \exp\{t(-\frac{1}{2}az^2 + i\gamma z + \int_{-\infty}^{\infty}(e^{izx} - 1)v(dx))\} \quad (2.65)$$

In the case where $\int_{|x|\leq 1}|x|v(dx) = \infty$ the jumps have infinite variation and small jumps need to be compensated.

As noted in the previous section, to every compound process $(X_t)_{t\geq 0}$ we can associate a random measure on $\mathbb{R}^d \times [0, \infty)$, which describes the jumps of process X.

Theorem 2.11 (Lévy-Itô decomposition)

Consider a triple (c, v, b) where $b \in \mathbb{R}, c \in \mathbb{R}, c \geq 0$, and v is a measure satisfying $v(\{0\}) = 0$ and $\int_{\mathbb{R}}(1 \wedge |x|^2 v(dx) < \infty)$. Then, there exists a probability space $(\Omega, \mathcal{F}, P)$ on which four independent Lévy processes exist: $L^{(1)}, L^{(2)}, L^{(3)}, L^{(4)}$, where $L^{(1)}$ is a constant drift, $L^{(2)}$ is a Brownian motion, $L^{(3)}$ is a compound Poisson process and $L^{(4)}$, is a square integrable (pure jump) martingale with a countable number of jumps of magnitude less than 1 on each finite time interval. Taking $L^{(1)} + L^{(2)} + L^{(3)} + L^{(4)}$ we have that there exists a probability space on which a Lévy process $L = (L_t)_{0\leq t\leq T}$ with characteristic exponent:

$$\psi(u) = iub - \frac{u^2 c}{2} + \int_{\mathbb{R}}(e^{iux} - 1 - iux1_{\{|x|<1\}})v(dx) \quad (2.66)$$

for all $u \in \mathbb{R}$, is defined.

We can therefore decompose any Lévy process into four independent Lévy processes $L^{(1)} + L^{(2)} + L^{(3)} + L^{(4)}$ as follows:

$$L_t = bt + \sqrt{c}W_t + X_t^L + \lim_{\varepsilon\to 0}\tilde{L}_t^{\varepsilon} \quad (2.67)$$

or,

$$
\begin{aligned}
L_t = bt + \sqrt{c}W_t & \\
&+ \int_0^t \int_{|x|\geq 1} x J_L(ds \times dx) \\
&+ \int_0^t \int_{|x|<1} x\{J_L(ds \times dx) - v(dx)ds\} \\
&\equiv \int_0^t \int_{|x|<1} x\tilde{J}_L(ds \times dx)
\end{aligned}
\tag{2.68}
$$

The terms of (2.67) are independent and convergence in the last term is almost sure and uniform in t on [0, T].

The Lévy-Itô decomposition entails that for every Lévy process there exists a vector b, a positive definite matrix c and a positive measure v that uniquely determines its distribution. This triplet (c, v, b) is called characteristic triplet or Lévy triplet of the process X_t.

The Lévy-Itô decomposition therefore suggests that it is possible that every Lévy process is a combination of a Brownian motion with drift and a possibly infinite sum of independent Poisson processes. This also means that every Lévy process can be approximated with arbitrary precision by a jump-diffusion process-that is by the sum of a Brownian motion with drift and a compound Poisson process.

2.2.2 Activity and variation of Lévy processes

Let L be a Lévy process with triplet (c, v, b).

1. If $v(\mathbb{R}) < \infty$ then almost all paths of L have a finite number of jumps on every compact interval. In this case, the Lévy process is said to have finite activity.

2. If $v(\mathbb{R}) = \infty$ then almost all paths of L have an infinite number of jumps on every compact interval. In this case, the Lévy process is said to have infinite activity.

Proposition 2.4 Let L be a Lévy process with triplet (c, v, b).

1. If $c = 0$ and $\int_{|x|\leq 1}|x|v(dx) < \infty$ then almost all paths of L have finite variation.

2. If $c \neq 0$ or $\int_{|x|\leq 1}|x|v(dx) = \infty$ then almost all paths of L have infinite variation.

A Lévy process is a strong Markov process: The associated semigroup is a convolution semigroup and its infinitesimal generator $L^X: \rightarrow L^X f$ is an integro-differential operator given by:

$$L^X f(x) = lim_{t\to 0}\frac{E[f(x+X_t)] - f(x)}{t} = \frac{\sigma^2}{2}\frac{\partial^2 f}{\partial x} + \gamma\frac{\partial f}{\partial x} + \int v(dy)[f(x+y) - f(x) - y1_{\{|y|\leq 1\}}\frac{\partial f}{\partial x}(x)] \quad (2.69)$$

which is well defined for $f \in C^2(\mathbb{R})$ with compact support.

2.2.3 Exponential Lévy models

It is more convenient to model the asset price processes as exponential Lévy processes, where log returns are independent and stationary increments are distributed according to an infinitely divisible law, estimated from the data.

Let $(S_t)_{t\in[0,T]}$ be the price of a financial asset, modelled as a stochastic process on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. $\mathcal{F}_t$ contains the price history up to t. Under the hypothesis of absence of arbitrage opportunities, there exists a measure Q equivalent to P under which the discounted prices of all traded financial assets are Q-martingales. In particular the discounted underlying $(e^{-rt}S_t)_{t\in[0,T]}$ is a Q martingale.

In exponential Lévy models, the risk-neutral dynamics of S_t under Q are presented as the exponential of a Lévy process:

$$S_t = S_0 e^{rt+X_t} \quad (2.70)$$

In that case X_t is a Lévy process (under Q) with the characteristic triplet (σ, v, γ). The absence of arbitrage imposes that $\hat{S}_t = S_t e^{-rt} = exp(X_t)$ is a martingale, which is equivalent to the following conditions on the characteristic triplet (σ, v, γ):

$$\int_{|y|>1} v(dy)e^y < \infty,$$
$$\gamma = \gamma(\sigma, v) = -\frac{\sigma^2}{2} \int (e^y - 1 - y1_{|y|\leq 1})v(dy) \quad (2.71)$$

Then the infinitesimal generator L becomes:

$$Lf(x) =$$
$$\frac{\sigma^2}{2}\left[\frac{\partial^2 f}{\partial x^2} - \frac{\partial f}{\partial x}\right](x) + \int_{-\infty}^{\infty} v(dy)[f(x+y) - f(x) - (e^y - 1)\frac{\partial f}{\partial x}(x)] \quad (2.72)$$

The notation $Y_t = rt + X_t$ is used, where Y_t is a strong Markov process with infinitesimal generator:

$$Lf = L^X f + r\frac{\partial f}{\partial x} \quad (2.73)$$

2.2.4 Examples of Lévy processes in finance

While in principle one can have both a non-zero diffusion component ($\sigma \neq 0$) and an infinite activity jump component, in practice the models encountered in the financial literature fall into two categories.

In the first category, the jump-diffusion one, the *normal* evolution of prices is given by a diffusion process, punctuated by jumps at random intervals. In this case the total change in the stock price is due to both *normal* and *abnormal* vibrations. The *normal* component is modelled by a standard geometric Brownian motion with a constant variance per unit of time and has continuous sample paths. In this case the jumps present rare events, crashes and large drawdowns and reflect the non-marginal impact of information. This means that there can be *active* and *quiet* times which are totally random-discrete points in time. Examples of those models are the Merton jump-diffusion models with Gaussian jumps (Merton, 1975) and the Kou et. Al. model with double exponential jumps (Kou, 2004).

The second category consists of models with an infinite number of jumps in every interval, which are called infinite activity models. In those models there is no need to introduce a Brownian motion component since the dynamics of the jumps are such so they can generate nontrivial small diffusion behaviour (Voltchkova, 2005b) and it has been argued (Madan, 2002), (Oosterlee, 2006), (Wang, 2007) that these models give a more realistic description of the price process at various time scales. There are also models (Oosterlee, 2007) that allow for both diffusions and for jumps of both finite and infinite activity.

Generally finite activity processes are said to be more useful if one is aiming to group assets by their activity levels. On the other hand, infinitely activity models are more suitable when one deals with highly liquid markets with large activity. However, it is important to keep in mind that since the piece-wise process is observed on a discrete grid, it is difficult if not impossible to see empirically to which category the price belongs. The choice is more a question of modelling convenience than an empirical one. In practice, different exponential Lévy models proposed in the financial modelling literature simply correspond to different choices for the Lévy measure v:

Compound Poisson jumps: $\sigma > 0$ and $\int_{|x|\leq 1}|x|v(dx) < \infty$.

- Merton Model (Merton, 1975): $v(x) = \frac{\lambda}{\delta\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\delta^2}}$
- Kou Model (Kou, 2004): $v(x) = pa_1 e^{-a_1 x} 1_{x>0} + (1-p)a_2 e^{a_2 x} 1_{x<0}$

Compound Poisson jumps: $\sigma = 0$ and $\int_{|x|\leq 1}|x|v(dx) = \infty$.

- Variance Gamma (Madan, 1998): $v(x) = \frac{1}{\kappa|u|} e^{Au - B|u|}$, $A = \frac{\theta}{\sigma^2}$, $B = \frac{\sqrt{\theta^2 + 2\sigma^2/\kappa}}{\sigma^2}$
- Tempered stable processes (Cont, 2004)
- Normal inverse Gaussian process (Cont, 2004)
- Hyperbolic and generalised hyperbolic processes (Cont, 2004)
- Meixner process (Schoutens, 2002)

2.2.4.1 Jump-diffusion models

Let's assume a general class of exponential Lévy models $S_t = S_0 e^{X_t}$ where X_t is a Lévy process consisting of a drift term, a Brownian motion term and a superposition of Poisson processes with various jump sizes:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_t \quad (2.74)$$

where $(N_t)_{t\geq 0}$ is a Poisson process that counts the jumps of X_t and Y_t denotes the jump sizes which are i.i.d. variables.

The characteristic function of X_t is:

$$E\left[e^{iuX_t}\right] = \exp\{t(i\gamma u - \frac{\sigma^2 u^2}{2} + \lambda \int_{\mathbb{R}} (e^{iux} - 1) f(dx))\} \quad (2.75)$$

In what follows we will present the three main Lévy processes that will be used in this work as test cases for the implementation of the numerical methodology proposed.

Test Case 2.1 (Merton's Model) Merton's model consists of a jump term added to a Geometric Brownian motion, that is:

$$\frac{dS}{S} = (\mu - \lambda\kappa)dt + \sigma dW_t + dJ_t \quad (2.76)$$

where J_t is a compound Poisson process with rate λ, that is:

$J_t = \sum_{j=1}^{N_t}(Y_t - 1)$ where $E[N_t] = \lambda t$ and $P\{N_t = n\} = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$.

If the p^{th} jump occurs at time t and S_{t^-} is the asset price immediately before the jump, then $S_t = Y_p S_{t^-}$. The compensation factor $\kappa = E[Y - 1]$ reflects the expected relative jump size. $Y - 1$ is an impulse function producing a jump from S_{t-} to S_t.

Merton (Merton, 1975) derives an analytical solution for a European call option when the Y are log normally distributed: $lnY \sim N(a, b^2)$, that is:

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma}\exp\{(-\frac{(lny - \mu)^2}{2\sigma^2})\} \quad (2.77)$$

$$\Rightarrow \kappa = e^{(a+0.5b^2)} - 1$$

The solution to the jump diffusion SDE (2.76) is:

$$S_t = S_0((\mu - \lambda\kappa - 0.5\sigma^2)t + \sigma W_t)\prod_{p=1}^{N_t} Y_p \quad (2.78)$$

So, multiplicative jumps can be seen to be a natural extension of GBM.

Theorem 2.12 The price of a European call option under Merton's jump-diffusion model using the Black-Scholes formula is:

$$C_{Merton}(S_0, K, T, \sigma, r) = e^{-rT}\sum_{n=0}^{\infty} e^{-\lambda T}\frac{(\lambda T)^n}{n!} e^{r_n T} C_{BS}(S_0 e^{\frac{n\sigma^2}{T}}, T, \sigma_n, r_n) \quad (2.79)$$

Test Case 2.2 (Variance Gamma) The model introduced in (Madan, 1998) proposes a three parameter generalization of the Brownian motion as a model for the dynamics of the logarithm of the stock price. It is obtained by evaluating the Brownian motion with constant drift and volatility at random time change given by a Gamma process. Under a VG process the unit period continuously compounded return is normally distributed, conditional on the realization of a random time. This random time has a Gamma density. In this model comparing to the Black-Scholes there are two more parameters apart from the volatility σ that provide control for the kurtosis and the skewness.

The Lévy measure of the VG process is:

$$v(u) = \frac{1}{\kappa|u|} e^{Au - B|u|} \quad (2.80)$$

where:

$$A = \frac{\theta}{\sigma^2} \text{ and } B = \frac{\sqrt{\theta^2 + 2\sigma^2/\kappa}}{\sigma^2} \quad (2.81)$$

with σ and θ the volatility and drift of Brownian motion and κ denoting the variance rate of the gamma time change.

The characteristic function of the process X_t is:

$$\Phi_t(u) = (1 + \frac{u^2\sigma^2\kappa}{2} - i\theta\kappa u)^{-\frac{t}{\kappa}} \quad (2.82)$$

Theorem 2.13 The European call option price on a stock, when risk neural dynamics of the stock price are given by the VG process for the risk neutral parameters,σ, v, θ is given by:

$$c(S(0);K,t) = S(0)\Psi(d\sqrt{\frac{1-c_1}{v}}, (a+s)\sqrt{\frac{v}{1-c_1}}, \frac{t}{v}) \quad (2.83)$$

$$-Kexp(-rt)\Psi(d\sqrt{\frac{1-c_2}{v}}, a\sqrt{\frac{v}{1-c_2}}, \frac{t}{v})$$

where:

$$d = \frac{1}{s}[\ln\left(\frac{S(0)}{K}\right) + rt + \frac{t}{v}\ln(\frac{1-c_1}{1-c_2})]$$

with:

$$a = \zeta s, \text{ where } \zeta = -\frac{\theta}{\sigma^2} \text{ and } \zeta = \frac{\sigma}{\sqrt{1+(\frac{\theta}{\sigma})^2\frac{v}{2}}}$$

and

$$c_1 = \frac{v(a+s)^2}{2}, c_2 = \frac{va^2}{2}$$

The Ψ function is defined in terms of modified Bessel function of the second kind and the degenerate hypergeometric function of two variables (Cont, 2004).

Test Case 2.3 (CGMY process) Under the CGMY process the Lévy measure is given by:

$$v(u) = \frac{Ce^{-Mu}}{|u|^{1+Y}}1_{u>0} + \frac{Ce^{-G|u|}}{|u|^{1+Y}}1_{u<0} \quad (2.84)$$

where $1_{u>0}$ and $1_{u<0}$ are the indicator variables. The parameter $C > 0$ is the measure of the overall level of activity. $G \geq 0$ and $M \geq 0$ provide control over the rate of the exponential decay on the left and the right of the Lévy density and $Y < 2$ describes the behaviour of the Lévy density in the area around zero where the density tends to infinity. If $Y < 0$ then the measure v

integrates to a finite value that yields a process of finite activity. If $Y \in [0,1]$, the process displays infinite activity but finite variation since $\int_{|u|<1} uv(du) < \infty$.

This chapter provided a summary of the main stochastic calculus definitions and theorems that will be necessary for the understanding of the financial mathematics notions underpinning the work presented here. Lévy processes, their characteristic properties and examples of interest were also presented.

3. Partial Integro-Differential Equation for Option pricing

In this section we present the proposed partial integro-differential formulation for the price of options whose underlying asset follows a jump diffusion process (Section 3.1). We will consider a one factor Merton's model (i.e. finite activity jumps) with deterministic volatility and describe the approximation of European (Section 3.1.3) and American option prices. (Section 3.2). Results for both cases will be presented and discussed in Section 3.3. In Section 3.4 we touch upon the calibration of jump-diffusion models.

3.1 Pricing PIDE

We saw in the previous section that the solution to the jump-diffusion s.d.e in Merton's case is:

$$S_t = S_0 e^{((\mu - \lambda\kappa - 0.5\sigma^2)t + \sigma W_t)} \prod_{j=1}^{N_t} Y_j$$

Assuming that the Brownian motion and the jump process are independent and based on the above s.d.e we can derive the PIDE (Cont, 2004) for a contingent claim $V(S,t)$ that depends on S by:

$$-\frac{\partial V}{\partial \tau} = \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \int_{-\infty}^{+\infty} \left(V(Se^u) - V(S) - S(e^u - 1)\frac{\partial V}{\partial S} \right) v(u)du \quad (3.1)$$

or in a more compact form:

$$-\frac{\partial V}{\partial t} = LV + I_c V \quad (3.2)$$

where L is a differential operator and I_c is the convolution-like integral operator.

Inequality holds for the case of the American option prices:

$$-\frac{\partial V}{\partial t} \geq LV + I_c V \quad (3.3)$$

which results to a linear complementarity formulation (LCP), discussed in (Section 3.2).

This PIDE should be augmented by defining the usual payoff terminal conditions and the boundary conditions when S is zero and when it tends to infinity. One can choose the linearity boundary conditions at $S \to \infty$, namely the second order derivatives with respect to S to be zero.

Boundary Conditions

As $S \to 0$ the PIDE reduces to

$$V_\tau = -rV$$

As $S \to \infty$ then the following assumption can be made:

$$V_{SS} \cong 0; S \to \infty$$

This naturally implies that:

$$V \cong A(\tau)S + B(\tau);\ S \to \infty$$

Assuming that the previous equation holds then V_τ reduces to:

$$V_\tau = \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV; S \to \infty$$

This means that at both $S = 0, S \to \infty$, the PIDE reduces to the Black-Scholes PDE and the usual boundary conditions can be imposed.

Smoothness with respect to the underlying process

In the case where the log price X_t has a non-degenerate diffusion component it is proved (Garroni, 2001) that the fundamental solution of the pricing PIDE, which corresponds to the density, X_t is in fact a smooth C^∞ solution. Consequently the option price $V(t,x)$ depends smoothly on the underlying process and based on the Feynman-Kac representation results discussed in (Section 2.1.1) we can use the solution of the PIDE for the computation of the option price.

3.1.1 Numerical challenges

The solution of the PIDE above needs to be approximated[8]. This is most commonly done using discretization methods and solving the system of resulting algebraic equations. The PIDE consists of two parts: The Black-Scholes part representing the convection-diffusion equation and the integral part reflecting jumps. The presence of the integral increases the complexity of the PIDE numerical solution compared to the Black-Scholes PDE solution. The main numerical challenges faced are described below:

1. ***Non-local character of the integral term*** The integral term of the PIDE is on a semi-infinite interval. This poses difficulties in the approximation of the integral term. The integrand may also have singularities.

2. The unknown solution appears in both the PDE and the integral.

[8] Exact solutions to the PIDE do not always exist. Merton's exact solution exists for the European case and will be used to assess the accuracy of the proposed approximation.

3. ***The PDE part of the PIDE is defined in an infinite interval*** The numerical solution of the PIDE requires the localization of the variables and of the integral term to a bounded domain $x \in (-B, B)$. This happens by defining boundary conditions (at $x = -B$ and $x = B$) in such a way as not to destroy the accuracy of the method. Similarly for the numerical integration of the jump component we need to reduce the integration region to a bounded domain. This corresponds to the truncation of large jumps.

4. The matrix resulting from the discrete system of equations while approximating the integral term on a bounded interval can be dense.

5. Finite difference schemes that are the scope of this work are widely used in combination with numerical integration techniques for the approximation of the PIDE. However, issues arise regarding the stability and accuracy that these methods can provide. A typical issue is the spurious oscillation problem that the Crank-Nicholson scheme is known for.

6. Smooth pasting may not always apply (e.g. singular cases) and this can necessitate an increased number of mesh points.

3.1.2 Peer review on numerical approximation

The integral term can be approximated via implicit or explicit treatment in the time dimension. In the American case the resulting LCP can be solved directly or iteratively. Choices regarding the localization of the integral term can affect the approximation accuracy. In what follows we present a brief survey on the topic of efficient discretization and algorithmic solutions developed by research peers[9].

In (d' Halluin, 2005) the resulting PIDE is solved using a fully implicit method for the PDE and a weighted timestepping method for the integral term.

$$V_\tau = \frac{1}{2}\sigma^2 V_{SS} + (r - \lambda\kappa)SV_S - (r + \lambda)V + \int_0^{+\infty} V(S\eta)g(\eta)d\eta \quad \text{(Pr 3.1)}$$

Note that in the above formulation the linear terms appear in the PDE part of the equation. In (Section 3.3) we discuss how this affects the shape of the convolution integral.

[9] Notation used here is as seen in the referenced papers.

The computational domain is bounded using $V_{SS} = 0$ and common Dirichlet conditions are applied at $0, S_{max}$.

Background to numerical approximation

The jump integral term can be transformed into a correlation integral as follows. Setting $x = logS$ and introducing the change of variable $y = \log(\eta)$ the resulting integral formulation is:

$$I = \int_{-\infty}^{+\infty} \bar{V}(x+y)\bar{f}(y)dy \quad \text{(Pr. 3.2)}$$

with $\bar{f}(y) = g(e^y)e^y$ denoting the probability density of a jump of size $y = \log(\eta)$ and $\bar{V}(y) = V(e^y)$. The discretized form of the integral is then:

$$I_i = \sum_{j=-\frac{N}{2}+1}^{j=\frac{N}{2}} \bar{V}_{i+j}\,\bar{f}_J \Delta y + O((\Delta y)^2) \quad \text{(Pr. 3.3)}$$

where $I_i = I(i\Delta x)$ and $\bar{V}_j = \bar{V}(j\Delta x)$ and:

$$\bar{f}_J = \frac{1}{\Delta x}\int_{x_j-\frac{\Delta x}{2}}^{x_j+\frac{\Delta x}{2}} \bar{f}_x dx \quad \text{(Pr. 3.4)}$$

with $x_j = j\Delta x$. The discrete form of the correlation integral (Pr 3.3) is using an equally spaced grid in $logS$. The equally spaced grid works well for the FFT evaluation of the correlation integral but is not suitable for the PDE discretization. The authors modify this approach to allow for non-uniform mesh coordinates in $x = lnS$. Their approach is to obtain values for $\bar{V}_j$ using second order accurate interpolation in S. However, FFT requires regular spacing, so I_i must be regularly spaced. I_i is spaced on grid points that do not coincide with the non-uniform discretization grid in S. Another interpolation is then performed to remediate that.

$\bar{V}_{\frac{N}{2}+j}, j > 0$ is approximated by an asymptotic boundary condition which does not affect accuracy for large values of N.

The resulting fully discretized PIDE is:

$$V_i^{n+1}[1 + (a_i + b_i + r + \lambda)\Delta\tau - \Delta\tau b_i V_{i+1}^{n+1} - \Delta\tau a_i V_{i-1}^{n+1}]$$

$$= V_i^n + (1-\theta_J)\Delta\tau\lambda \sum_{j=-\frac{N}{2}+1}^{j=\frac{N}{2}} x(V^{n+1}, i, j)\bar{f}_j \Delta y + \theta_J \Delta\tau\lambda \sum_{j=-\frac{N}{2}+1}^{j=\frac{N}{2}} x(V^n, i, j)\,\bar{f}_j \Delta y \quad \text{(Pr. 3.5)}$$

with $x(V, i, j)$ denoting the weighted timestep discretization and a_i, b_i denoting the central differences of the space derivatives of the PDE. The explicit discretization of the integral term is proven to be unconditionally stable, while the accuracy of the time discretization is improved using Crank-Nicholson's scheme. A fixed point iteration scheme is also proposed for the efficient solution of the linear algebraic system resulting from the implicit discretization of the integral term, with the method shown to exhibit global convergence. It is then proved that this discretized solution converges to a viscosity solution. The early exercise for the American option case is treated via a penalty method (d' Halluin, 2004). The cost of the proposed method is essentially the cost of performing two FFTs since the cost of computing the correlation integral by Fourier methods at each timestep can be significant (Carr, 2007).

In (Toivanen, 2014) the pricing PIDE is split into two parts:

$$Lv = Dv + \lambda(Jv - v)$$

With D denoting the differential operator and J the integral operator:

$$Dv := \frac{1}{2}\sigma^2 x^2 V_{xx} + (r - \lambda\kappa)xv_x - rv \quad \text{(Pr. 3.6)}$$

$$Jv := \int_0^{+\infty} v(xy, \tau) f(y) dy \quad \text{(Pr. 3.7)}$$

The authors study the performance of three 2nd order IMEX schemes in 3 time dimensions. Those schemes are derived by an implicit method combined with an extrapolation procedure for the explicit part. An extra parameter $c \in [0,1]$ is introduced, controlling the convex combination of the zeroth-order term λv between the implicit and explicit part. Their stability is studied via Fourier analysis and proved to hold true, conditional on $\lambda\Delta\tau$ with the mid-point scheme being the less restrictive in respect to choice of c.

IMEX mid-point:

$$\frac{v_{m+1} - v_{m-1}}{2} = \Delta\tau\lambda(J - cI) + \Delta\tau(D - \lambda(1-c)I)\frac{v_{m+1} + v_{m-1}}{2} \quad \text{(Pr. 3.8)}$$

where $v_m = v(m\Delta\tau)$, I the identity matrix, D, J matrices resulting from the discretization of the PIDE. For J strictly positive and real eigenvalues $v_D \leq 0$ the scheme is conditionally stable for $\lambda\Delta\tau < 1$ and $c = 0$.

IMEX mid-CNAB:

$$v_{m+1} - v_m = \Delta\tau\lambda(J - cI)\frac{3v_m - v_{m-1}}{2} + \Delta\tau(D - \lambda(1-c)I)\frac{v_{m+1} + v_m}{2} \quad \text{(Pr. 3.9)}$$

For the same conditions as above, the methods is stable for $\lambda\Delta\tau < \frac{1}{2}$ and $c \in [0,1]$.

IMEX mid-BDF2:

$$\frac{3}{2}v_{m+1} - 2v_m + \frac{1}{2}v_{m-1} = \Delta\tau\lambda(J - cI)(2V_m - V_{m-1} + \Delta\tau(D - \lambda(1-c)I))v_{m+1} \quad \text{(Pr. 3.10)}$$

With conditions as for the previous methods, the scheme is stable for $\lambda\Delta\tau < \frac{2}{3}$ and $c \in [0,1]$.

3.1.3 Numerical approximation

Generally for the numerical solution of the PIDE the following steps need to be followed:

Localization A bounded computational domain should be chosen and artificial boundary conditions should be imposed.

Truncation of large jumps This corresponds to truncating the integration domain of the jump component.

Discretization The derivatives of the solution are replaced by the usual finite differences and the integral term is approximated using common

methods (e.g. Trapezoidal, Simpson rule). The problem is then solved using implicit/ explicit schemes.

The PIDE (3.1) is integrated between two time levels at

$S = S_j$, i.e.

$$\int_{t^{m-1}}^{t^m} -\frac{\partial V}{\partial t}\, dt = V(S_j, t^{m-1}) - V(S_j, t^m)$$
$$= \int_{t^{m-1}}^{t^m} \left[L\, V(S_j, t) + I_j(V(S,t)) \right] dt \qquad (3.4)$$

where

$$I_j(V(S,t)) = \int_{-\infty}^{\infty} \left[V(S_j e^u, t) - V(S_j, t) - S_j(e^u - 1) \left(\frac{\partial V}{\partial S}\right)|_{S=S_j} \right] \nu(u) du \qquad (3.5)$$

is the option price impact from the probability of the asset price jumping from $S = S_j$.

We will use $L_{\Delta S}$ as the usual spatial second-order discrete approximation to the Black-Scholes operator however approximating the convolution integral requires some further discussion.

Using a uniform spatial mesh in S or in $u = lnS$ does not distribute the error efficiently and a better choice is either to use a mesh refined around the strike region ((Papanicolaou, 2000), (Voltchkova, 2006)), or via a coordinate stretching transformation (this work).

Coordinate Stretching For most options the region of interest is determined by the payoff. For vanilla call or put options, prices are normally quoted around the strike price ($S = S_K$). A suitable stretching transformation should enlarge this region of asset price.

If we consider a general transformation:

$$S = \text{asinh}(x - L) + c \text{ for } a, c, L \in R \quad (3.6)$$

and in order to determine the parameters a, c we first constrain the curve to pass through $(0,0)$, that is:

$$c = asinhL \quad (3.7)$$

While the maximum stretch (at $x = L$) must correspond to $S = S_K$, thus $c = S_K$, that is:

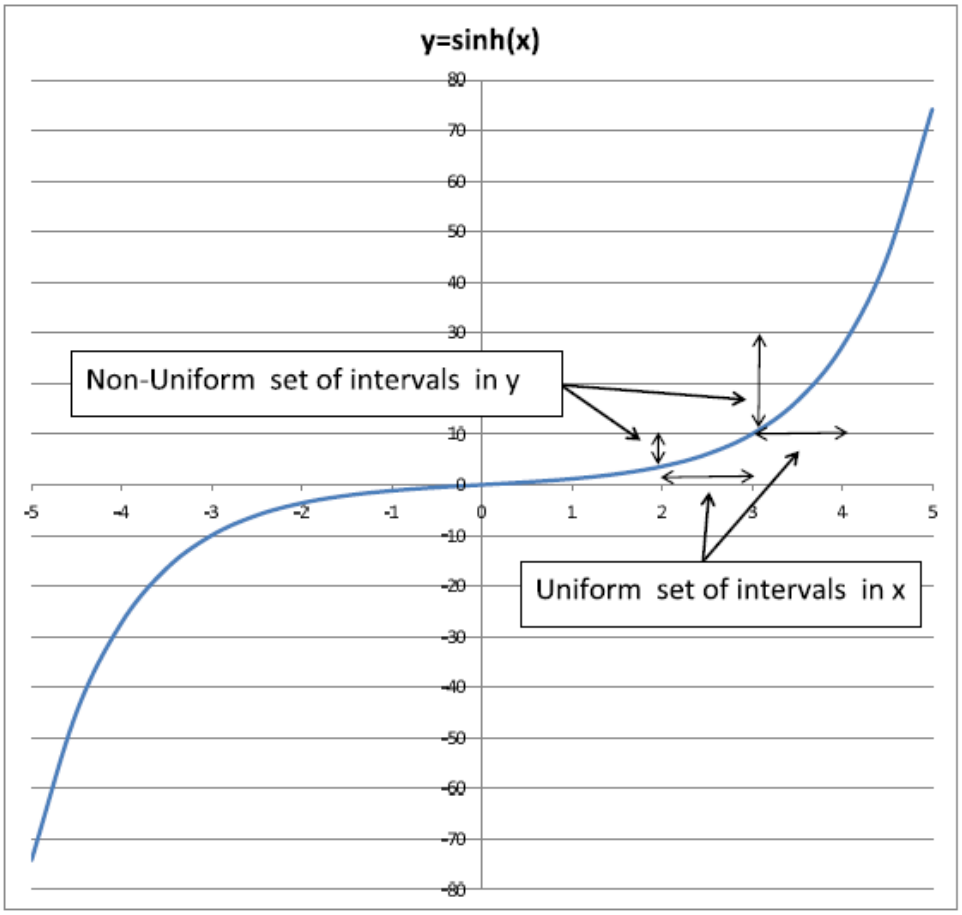
$$S = \frac{S_K}{\sinh(L)} \sinh(x - L) + S_K \quad (3.8)$$

Note that $x = 2L$ corresponds to $S = 2S_K$.

The maximum enlargement effect of the transformation is determined by the ratio of the gradient at the strike to that at $S = 0$. That is (the stretch):

$$\tau = \left(\frac{dS}{dx}\right)_{x=0} / \left(\frac{dS}{dx}\right)_{x=L} = \cosh(L) \quad (3.9)$$

Figure 3-1 Coordinate transformation



For a given stretch (e.g. $\tau = 5$) the transformation is determined since:

$$L = \log[\tau + \sqrt{\tau^2 - 1}] \text{ and } \sqrt{(\tau^2 - 1)} = \lambda = sinhL$$

This results to:

$$S = \frac{S_K}{\lambda}\sinh(x - L) + S_K, \text{ for } \tau > 1 \quad (3.10)$$

and inversely:

$$x = sinh^{-1}\left(\frac{\lambda}{S_K}(S - S_K)\right) + L \quad (3.11)$$

As discussed previously, S_κ is the focus for the refinement and $\lambda = sinhL$ is a parameter controlling the degree of the stretch. Substituting these expressions into the operator LV we get:

$$LV(x,t) = \frac{1}{2}\sigma^2 T(x)^2\left(\frac{\partial^2 V}{\partial x^2} - \tanh(x-L)\frac{\partial V}{\partial x}\right) + rT(x)\frac{\partial V}{\partial x} - rV \quad (3.12)$$

where

$$T(x) = \frac{\sinh(x-L)+\lambda}{\cosh(x-L)} \quad (3.13)$$

(3.4) with $V(S_j, t^m) \xrightarrow{s\to x} V(x_j, t^m)$ can then be written as:

$$\begin{aligned} V(x_j, t^{m-1}) - V(x_j - t^m) \\ = \Delta t[\theta\left(L_{\Delta x}V(x_j, t^{m-1}) + \lambda I_c^j(V(x, t^{m-1}))\right) \\ + (1-\theta)\left(L_{\Delta x}V(x, t^m) + \lambda I_c^j(V(x, t^m))\right)] \end{aligned} \quad (3.14)$$

The boundary condition at $S = S_{max}$ will also now change since:

$$S^2\frac{\partial^2 V}{\partial S^2} = T(x)^2\left(\frac{\partial^2 V}{\partial x^2} - \tanh(x-L)\frac{\partial V}{\partial x}\right) = 0 \quad (3.15)$$

Evaluation of the discrete convolution Integral

Comments on the convolution integral

- The integrand is localised by the exponential tapering of the Lévy density and moreover is identically zero when the prices are locally linear.
- The main accuracy requirement is consequently around the strike region (or between the strike and the optimal exercise boundary). Local mesh refinement of this region will tend to equidistribute the errors both in the Black-Scholes operator and convolution quadrature.
- Using $V_{SS} = 0\ at\ S = S_{max}$ as a mesh truncation condition enforces linearity in the prices beyond S_{max}. The coordinate stretching is also

an efficient way to provide a locally linear region to the right of the strike to reduce any localisation errors.

Writing the integral in terms of the jump variable y the first term in the integrand for $S = S_j$ and $t = t^m$ can be denoted:

$$F_j^m(y) = V(S_j y, t^m) - V(S_j, t^m) - S_j(y-1)\left(\frac{\partial V}{\partial S}\right)_{S=S_j}^{t=t^m} \quad (3.16)$$

The linear extrapolation of the option price from $S = S_j$ to $S = S_j y$ using price and gradient data at $S = S_j$ is:

$$V^{LE}(S_j y, t^m) = V(S_j, t^m) + S_j(y-1)(\frac{\partial V}{\partial S})_{S=S_j}^{t=t^m} \quad (3.17)$$

so that:

$$F_j^m(y) = V(S_j y, t^m) - V^{LE}(S_j y, t^m) \quad (3.18)$$

simply captures the discrepancy between this linear extrapolation and the actual option price. This is largest when the extrapolation passes through a region of high curvature, i.e. typically around the strike price. The integral term is then just the expected value of this discrepancy:

$$I_j(V(S, t^m)) = \int_0^\infty F_j^m(y)\tilde{\nu}(y)dy, \quad for\ 1 \le j \le N_S \quad (3.19)$$

where $\tilde{\nu}(y)$ is the density in terms of the jump variable y. For the Merton model this equates to a lognormal density:

$$\tilde{\nu}(y) = \frac{\lambda}{y\sqrt{2\pi}b} e^{-\frac{(\ln y - a)^2}{2b^2}} \quad (3.20)$$

Hence, I_j will only be significant for asset prices S_j from which asset jumps through the strike region are probable. If the price jumps are localised by the distribution to a region that is relatively linear in asset price, e.g. substantially to the left or the right of the strike, then the integral will have a relatively small effect on the option price.

Note that the non-convolution terms can be integrated exactly so that:

$$I_j\big(V(S,t^m)\big) = \int_0^{\infty} V\big(S_j y, t^m\big)\tilde{\nu}(y)dy - \lambda V\big(S_j, t^m\big) - S_j\lambda\kappa\left(\frac{\partial V}{\partial S}\right)_{S=S_j}^{t=t^m} \quad (3.21)$$

for the Merton density since:

$$\int_0^{\infty} \tilde{\nu}(y)dy = \lambda \text{ and } \lambda\kappa = \int_0^{\infty} \tilde{\nu}(y)(y-1)dy = \lambda E[Y-1] \quad (3.22)$$

Localization of the integral term

The integral term can be efficiently localised to a finite domain $[0, y^*]$, such that:

$$\int_{y^*}^{\infty} F_j^m(y)\tilde{\nu}(y)dy \approx 0 \quad (3.23)$$

since $lim_{y\to\infty}\tilde{\nu}(y) = 0$ for any sensible density. We shall choose:

$$y^* = S_{max}/S_j$$

The PDE domain truncation price S_{max} has to be chosen so that there is no curvature in the option price for $S > S_{max}$. This equates to imposing the truncation conditions $V_{SS} = 0$ at $S = S_{max}$. This is sufficiently far from the strike price and the effect of asset price jumps outside the mesh is ignored.

The now localized integral:

$$I_j\big(V(S,t^m)\big) = \int_0^{y^*} F_j^m(y)\tilde{\nu}(y)dy, \quad for\ 1 \le j \le N_S \quad (3.24)$$

needs to be calculated using a quadrature approach. A common choice is the composite trapezoidal rule.

Quadrature approach

We define N_Q sub-intervals, $[y_0, y_1], \ldots [y_{N_Q-1}, y_{N_Q}]$ where $y_0 = 0, y_{N_Q} = y^*$ and apply the composite trapezoidal quadrature as below[10]:

$$\begin{aligned} I_j\big(V(S,t^m)\big) \approx\ & I_j^m \\ &= \sum_{p=0}^{p=N_Q-1} \frac{1}{2}(y_{p+1}-y_p)\big[F_j^m(y_{p+1})\tilde{v}(y_{p+1}) \\ &+ F_j^m(y_p)\tilde{v}(y_p)\big] = \sum_{p=0}^{p=N_Q-1} \omega_p F_j^m(y_p) \end{aligned} \tag{3.25}$$

where the discrete densities are given by:

$$\begin{aligned} \omega_p &= \frac{1}{2}(y_{p+1}-y_{p-1})\tilde{v}(y_p), \qquad p = 2, N_Q - 1 \\ \omega_0 &= \frac{1}{2}(y_1 - y_0)\tilde{v}(y_0), \omega_{N_Q} = \frac{1}{2}(y_{N_Q} - y_{N_Q-1})\tilde{v}\left(y_{N_Q}\right) \end{aligned} \tag{3.26}$$

Note that:

$$F_j^m(y_p) = V(S_p, t^m) - V(S_j, t^m) - S_j(y-1)\left(\frac{\partial V}{\partial S}\right)_{S=S_j}^{t=t^m} \tag{3.27}$$

where $S_p = S_j\, y_p$. We need to evaluate $V(S_p, t^m)$ using the FD mesh values $\{V_j^m\}$. This would normally require interpolation since $S_p = S_j y_p \notin \{S_0, S_1, \ldots S_{N_s}\}$ i.e. does not coincide with another mesh point.

Our approach is to construct the quadrature points y_p based on the requirement for $S_p = S_j y_p$ to be another mesh point i.e. for the j^{th} finite difference equation we define a set of (non-uniform) quadrature points:

$$y_{jp} = S_p/S_j\,, p = 0, \ldots N_S \tag{3.28}$$

Note that for the j_{th} equation the quadrature points correspond to jumps from the asset price S_j to every other finite difference mesh point.

[10] See (Section 7.2) for derivation

Hence we now have that the number of quadrature intervals is the same as the number of mesh points $N_Q = N_S$; note each equation has specific quadrature weights ω_{jp} as below:

$$\omega_{jp} = \frac{1}{2}(y_{jp+1} - y_{jp-1})\tilde{v}(y_{jp}) \quad (3.29)$$

Hence,

$$I_j^m = \sum_{p=0}^{p=N_S-1} \omega_{jp} F_j^m(y_{jp}) \quad (3.30)$$

and the function $F_j^m(y_{jp})$ can be estimated from a pair of FD mesh point option prices V_p^m and V_j^m, and a second order approximation to the option delta:

$$F_j^m(y_{jp}) = V_p^m - V_j^m - S_j(y_{jp} - 1)\frac{(V_{j+1}^m - V_{j-1}^m)}{2\Delta S} \quad (3.31)$$

The disadvantage of this approach is that the quadrature points will change for each S_j. On the advantages, this method is simple to implement (i.e. no interpolation needed) for a non-uniform or coordinate transformed mesh.

Note that for a coordinate transformed mesh:

$$\bar{F}_j^m(y_{jp}) = V_p^m - V_j^m - T(x_j)(y_{jp} - 1)\frac{(V_{j+1}^m - V_{j-1}^m)}{2\Delta x} \quad (3.32)$$

We can now write the final fully discretized form of the PIDE (3.1)

$$\begin{aligned} V_j^{m-1} - V_j^m = \Delta t[\theta(L_{\Delta x} V_j^{m-1} + \sum_{p=0}^{N_{Q-1}} \bar{F}_j^{m-1}(y_j p)\tilde{v}(y_{jp})) \\ + (1-\theta)(L_{\Delta x} V_j^m + \sum_{p=0}^{N_Q-1} \bar{F}_j^m(y_j p)\tilde{v}(y_{jp}))] \end{aligned} \quad (3.33)$$

Comments on quadrature with coordinate transformation

- The non-uniform mesh creates a non-uniform set of quadrature intervals in $y = {}^{S}/_{S_j}$.
- The quadrature sampling is most accurate in the region of the greatest transformation stretch (e.g. at the strike price $S = K$) where the integrand has its maximum variation.

Simpson's Rule

An alternative to the trapezoidal rule is the higher order Simpson's rule. This can be applied in composite form in a similar way.

$$I_j^m = \sum_{p=0}^{p=N_Q-1} \frac{1}{6}(y_{p+1}-y_p)\big[F_j^m(y_{p+1})\tilde{v}(y_{p+1}) + 2\big[F_j^m(y_{p+1})+F_j^m(y_p)\big]\tilde{v}(y_{p+1/2}) + F_j^m(y_p)\tilde{v}(y_p)\big] \quad (3.34)$$

$$= \sum_{p=0}^{p=N_Q-1} \frac{1}{6}(y_{p+1}-y_p)\left[F_j^m(y_{p+1})\left(\tilde{v}(y_{p+1}) + 2\tilde{v}(y_{p+1/2})\right) + F_j^m(y_p)\left(\tilde{v}(y_{p+1}) + 2\tilde{v}(y_{p+1/2})\right)\right]$$

$$= \sum_{p=0}^{p=N_Q-1} \bar{\omega}_p F_j^m(y_p)$$

where we again use mesh intervals as the quadrature intervals to avoid interpolation and $F_j^m(y_{p+1/2})$ has been estimated using:

$$F_j^m(y_{p+1/2}) = \frac{1}{2}\big[F_j^m(y_{p+1}) + F_j^m(y_p)\big] \quad (3.35)$$

and

$$y_{p+1/2} = \frac{1}{2}(y_{p+1}+y_p)$$

with the discrete densities given by:

$$\bar{\omega}_p = \frac{1}{6}(y_{p+1}-y_{p-1})\tilde{v}(y_p) + \frac{1}{3}(y_p-y_{p-1})\tilde{v}(y_{p-1/2}) + \frac{1}{3}(y_{p+1}-y_p)\tilde{v}(y_{p+1/2}), \qquad p = 2,..,N_Q-1$$

$$\bar{\omega}_0 = \frac{1}{6}(y_1-y_0)\left(\tilde{v}(y_0) + 2\tilde{v}(y_{1/2})\right) \quad (3.36)$$

$$\bar{\omega}_{N_Q} = \frac{1}{6}\left(y_{N_Q} - y_{N_Q-1}\right)\left(\tilde{v}\left(y_{N_Q}\right) + 2\tilde{v}\left(y_{N_Q-1/2}\right)\right)$$

Simpson's rule improves the discrete quadrature measures but computational tests show little effect on option pricing.

Both quadrature approaches are tested on the solution of a European put option using the parameters chosen in (d' Halluin, 2004). The results for a sequence of mesh refinements are shown in (Section 3.3) on a uniform mesh. Since the exact solution is known, the errors can be calculated exactly and the accuracy of the quadrature and the impact of the localization can be assessed precisely.

Discretization

For the discretization of the PIDE we are not constrained by any requirements to use the same θ value for the PDE and integral term discretization (θ_C) Along the same lines of (d' Halluin, 2004) we define separate values and use the following discretization:

$$V_j^{m-1} - V_j^m = \theta \Delta t L_{\Delta S} V_j^{m-1} + (1-\theta)\Delta t L_{\Delta S} V_j^m + \theta_C \Delta t I_j^m + (1-\theta_C)\Delta t I_j^m \quad (3.37)$$

As previously noted, the discrete integral (on a uniform mesh):

$$I_j^m = \sum_{p=0}^{p=N_S} \omega_{jp} F_j^m(y_{jp}) = \sum_{p=0}^{p=N_S} \omega_{jp}\left(V_p^m - V_j^m - S_j(y_{jp}-1)\frac{\left(V_{j+1}^m - V_{j-1}^m\right)}{2\Delta S}\right) \quad (3.38)$$

can be split defining:

$$\kappa_j = \sum_{p=0}^{p=N_S} \omega_{jp}\left(y_{jp} - 1\right), (\approx \lambda\kappa \text{ for Merton}) \quad (3.39)$$

and

$$\Omega_j = \sum_{p=0}^{p=N_S} \omega_{jp} \ , (\approx \lambda \text{ for Merton}) \quad (3.40)$$

so that:

$$I_j^m = \sum_{p=0}^{p=N_S} \omega_{jp} V_p^m - \Omega_j V_j^m - \kappa_j S_j \frac{\left(V_{j+1}^m - V_{j-1}^m\right)}{2\Delta S} \quad (3.41)$$

Then (3.38) can be written:

$$I_j^m = \sum_{p=0}^{p=N_S} \omega_{jp} V_p^m - \Omega_j V_j^m - \kappa_j S_j \frac{\left(V_{j+1}^m - V_{j-1}^m\right)}{2\Delta S} \quad (3.42)$$

Algebraic Structure

Substituting:

$$L_{\Delta S} V\left(S_j, t^m\right) = \frac{1}{2}\sigma^2 S_j^2 \left[\frac{V_{j+1}^m - 2V_j^m + V_{j-1}^m}{\Delta S^2}\right] + rS_j \left[\frac{V_{j+1}^m - V_{j-1}^m}{2\Delta S}\right] - rV_j^m \quad (3.43)$$

and equivalently for $L_{\Delta S} V\left(S_j, t^{m-1}\right)$ in (3.37) the θ-method finite differences approximation can be written out as:

$$\begin{aligned} a_j V_{j-1}^{m-1} + b_j V_j^{m-1} + c_j V_{j+1}^{m-1} - \theta_C \Delta t \sum_{p=0}^{p=N_S} \omega_{jp} V_p^{m-1} \\ = A_j V_{j-1}^m + B V_j^m + C_j V_{j+1}^{m-1} + (1-\theta_C)\Delta t \sum_{p=0}^{p=N_S} \omega_{jp} V_p^m \end{aligned} \quad (3.44)$$

The coefficients are given by:

$$a_j = -\theta \Delta t \left[\frac{\sigma^2 S_j^2}{2\Delta S^2}\right] + \theta \Delta t \frac{rS_j}{2\Delta S} - \theta_C \Delta t \frac{\kappa_j S_j}{2\Delta S} \quad (3.45)$$

$$b_j = 1 + \theta \Delta t \left[\frac{\sigma^2 S_j^2}{\Delta S^2} + r\right] + \theta_C \Omega_j \Delta t \quad (3.46)$$

$$c_j = -\theta \Delta t \left[\frac{\sigma^2 S_j^2}{2\Delta S^2}\right] - \theta \Delta t \frac{rS_j}{2\Delta S} + \theta_C \Delta t \frac{\kappa_j S_j}{2\Delta S} \quad (3.47)$$

and

$$A_j = (1-\theta)\Delta t \left[\frac{\sigma^2 S_j^2}{2\Delta S^2}\right] - (1-\theta)\Delta t \frac{rS_j}{2\Delta S} + (1-\theta_C)\Delta t \frac{\kappa_j S_j}{2\Delta S} \quad (3.48)$$

$$B_j = 1 - (1-\theta)\Delta t\left[\frac{\sigma^2 S_j^2}{\Delta S^2} + r\right] - (1-\theta_C)\Omega_j \Delta t \quad (3.49)$$

$$C_j = (1-\theta)\Delta t\left[\frac{\sigma^2 S_j^2}{2\Delta S^2}\right] + (1-\theta)\Delta t\frac{rS_j}{2\Delta S} - (1-\theta_C)\Delta t\frac{\kappa_j S_j}{2\Delta S} \quad (3.50)$$

If $a_j < 0, b_j \geq |a_j| + |c_j|,\ c_j < 0$ and for at least one $j,\ b_j > |a_j| + |c_j|$, then the system is termed positive definite and is non-singular so can be solved either by factorization or relaxation.

In the case of applying coordinate transformation, where for a stretch of $\tau = coshL$ focused at $S = S^*$, we use the mesh transformation described earlier:

$$S = S^*\left(1 + \frac{\sinh(x-L)}{\cosh L}\right) \quad (3.51)$$

We can re-write the coefficients (3.45-3.50) as:

$$a_j = -\theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] + \theta\Delta t\frac{\tilde{r}T(x_j)}{2\Delta x} - \theta_C\Delta t\frac{\kappa_j T(x_j)}{2\Delta x} \quad (3.52)$$

$$b_j = 1 + \theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{\Delta x^2} + r\right] + \theta_C\Omega_j\Delta t \quad (3.53)$$

$$c_j = -\theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] - \theta\Delta t\frac{(\tilde{r}-D)T(x_j)}{2\Delta x} + \theta_C\Delta t\frac{\kappa_j T(x_j)}{2\Delta x} \quad (3.54)$$

and

$$A_j = (1-\theta)\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] - (1-\theta)\Delta t\frac{(\tilde{r}-D)T(x_j)}{2\Delta x} + (1-\theta_C)\Delta t\frac{\kappa_j T(x_j)}{2\Delta x} \quad (3.55)$$

$$B_j = 1 - (1-\theta)\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2} + r\right] - (1-\theta_C)\Omega_j\Delta t \quad (3.56)$$

$$C_j = (1-\theta)\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] + (1-\theta)\Delta t\,\frac{(\tilde{r}-D)T(x_j)}{2\Delta x} - (1-\theta_C)\Delta t\,\frac{\kappa_j T(x_j)}{2\Delta x} \quad (3.57)$$

where:

$$\tilde{r} = r - \tfrac{1}{2}\sigma^2 T(x)\tanh(x-L)$$

$$T(x) = \frac{\sinh(x-L)+\sinh L}{\cosh(x-L)}$$

Notes on scheme (3.44)

1. V_j^m denotes the solution at node j and time level m.
2. $\theta_C = 0$ corresponds to an implicit handling of the jump integral, whereas $\theta_C = 1$ indicates an explicit treatment of this term and Crank-Nicholson when $\theta_C = \frac{1}{2}$
3. If a_j, b_j, c_j, do not satisfy the conditions for a positive definite non-singular system, oscillation may appear in the numerical solution. This can be avoided by using adaptive upwind differencing.
4. For typical parameter values and grid spacing, forward or backward differencing is rarely required for single factor options. In practice, since the errors occur at only a small number of nodes, in an area remote from the region of interest, the limited use of a low order scheme does not result in pure convergence as the mesh is refined.
5. Algorithmically we decide between central or forward discretization at each node of the scheme.

Adaptive Upwinding

We saw from the algebraic structure of scheme (3.44) that the system's positivity property requires:

$$a_j, c_J < 0, \qquad b_j > 0$$

That is,

$$a_j = -\theta\left[\frac{\sigma^2 T^2(x_j)}{\Delta x}\right] + \theta(\tilde{r} - D)T(x_j) - \theta_C \kappa_j T(x_j) < 0 \quad (3.58)$$

$$-\theta\left[\frac{\sigma^2 T(x_j)}{\Delta x}\right] + \theta(\tilde{r} - D) - \theta_C \kappa_j < 0 \quad (3.59)$$

and similarly for c_j. Hence:

$$\left|\theta(\tilde{r} - D) - \theta_C \kappa_j\right| < \theta\left[\frac{\sigma^2 T(x_j)}{\Delta x}\right] \quad (3.60)$$

If this condition is not satisfied for a given j then an upwind scheme is selectively used to generate the first derivative approximation.

Assuming that:

$$\theta(\tilde{r} - D) - \theta_C \kappa_j > 0 \quad (3.61)$$

then:

$$a_j = -\theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] \quad (3.62)$$

$$b_j = 1 + \theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{\Delta x^2} + r\right] + \theta\Delta t\frac{(\tilde{r} - D)T(x_j)}{\Delta x} - \theta_C\Delta t\frac{\kappa_j T(x_j)}{\Delta x} + \theta_C\Omega_j\Delta t \quad (3.63)$$

$$c_j = -\theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] - \theta\Delta t\frac{(\tilde{r} - D)T(x_j)}{\Delta x} + \theta_C\Delta t\frac{\kappa_j T(x_j)}{\Delta x} \quad (3.64)$$

If in the less likely case:

$$\theta(\tilde{r} - D) - \theta_C \kappa_j < 0 \quad (3.65)$$

then:

$$a_j = -\theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] + \theta\Delta t\frac{(\tilde{r} - D)T(x_j)}{\Delta x} - \theta_C\Delta t\frac{\kappa_j T(x_j)}{\Delta x} \quad (3.66)$$

$$b_j = 1 + \theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{\Delta x^2} + r\right] - \theta\Delta t\frac{(\tilde{r} - D)T(x_j)}{\Delta x} + \theta_C\Delta t\frac{\kappa_j T(x_j)}{\Delta x} + \theta_C\Omega_j\Delta t \quad (3.67)$$

$$c_j = -\theta\Delta t\left[\frac{\sigma^2 T^2(x_j)}{2\Delta x^2}\right] \quad (3.68)$$

The explicit terms are changed in a similar manner.

Stability and monotonicity

As presented in (Souganidis, 1991) stability and monotonicity are important properties for a numerical scheme in order to ensure convergence to a viscosity solution. In what follows we will present the stability assessment of the proposed scheme (3.44).

Theorem 3.1 (Stability of scheme (3.44)) The discretization method (3.44) is unconditionally stable for any choice of θ_C, where $0 \leq \theta_C \leq 1$, provided that:

- $a_j, c_j < 0$ and $b_j > 0$;
- $r, \lambda \geq 0$

Proof Similarly to (d' Halluin, 2004) we define $V^m = [V_0^m, V_1^m, .., V_{NS}^m]'$ to be the discrete solution vector to equation (3.44) and we suppose a perturbation in the initial solution:

$$\hat{V}^0 = V^0 + E^0 \quad (3.69)$$

with $E^m = [E_0^m, E_1^m, .., E_{NS}^m]'$ denoting the perturbation vector. Note that $E_p^m = 0$ assuming Dirichlet boundary conditions are imposed at this node ($V_{SS} = 0$). We can then write the equation of the propagation of the perturbation as:

$$a_j E_{j-1}^{m-1} + b_j E_j^{m-1} + c_j E_{j+1}^{m-1} - \theta_C \Delta t \sum_{p=0}^{p=N_S} \omega_{jp} E_p^{m-1} = A_j E_{j-1}^{m} + B_j E_j^{m} + C_j E_{j+1}^{m} + (1-\theta_C)\Delta t \sum_{p=0}^{p=N_S} \omega_{jp} E_p^{m} \quad (3.70)$$

We now define:

$$||E||^m = \max_j |E_j|^m$$

It follows from $\sum_{p=0}^{p=N_S} \omega_{jp} = \Omega$,[11] with $\Omega \geq 0$ that:

$$b_j|E_j|^{m-1} \leq A_j\|E\|^m + B_j\|E\|^m + C_j\|E\|^m + (1-\theta_C)\Delta t\Omega_j\|E\|^m - a_j|E|_{j-1}^{m-1} - c_j|E|_{j+1}^{m-1} - \theta_C\Delta t\Omega_j\|E\|^{m-1} \quad (3.71)$$

This implies:

$$b_j|E_j|^{m-1} \leq A_j\|E\|^m + B_j\|E\|^m + C_j\|E\|^m + (1-\theta_C)\Delta t\Omega_j\|E\|^m - a_j\|E\|^{m-1} - c_j\|E\|^{m-1} - \theta_C\Delta t\Omega_j\|E\|^{m-1} \quad (3.72)$$

Equation (3.72) holds for all $i < p$. In particular equality holds for $i = i^*$ where:

$$||E||^{m-1} = \max_j |E_j|^{m-1} \quad (3.73)$$

Re-writing equation (3.71) for $i = i^*$ we get:

$$\|E\|^{m-1}\left(b_j + a_j + c_j - \theta_C \Delta t\Omega_j\right) = \|E\|^m(A_j + B_j + C_j + (1-\theta_C)\Delta t\Omega_j) \quad (3.74)$$

Substituting the coefficients[12] (3.45-3.50) into (3.74) we get:

$$\|E\|^{m-1} = \frac{(1-\Omega_j \Delta t)}{1 + r\Delta t}\|E\|^m \quad (3.75)$$

Therefore:

[11] $\sum_{p=0}^{p=N_S} \omega_{jp} \approx \lambda$ for Merton's case

[12] Please note that coefficients (3.45-3.50) contain the $\Delta t\Omega_j$ terms. In the above proof and for presentation purposes we assumed the $\Delta t\Omega_j$ do not appear in the coefficients.

$$\|E\|^{m-1} \leq \|E\|^m \quad (3.76)$$

■

The above results indicates that the stability of the scheme is unaffected by the discretization choice (implicit/ explicit) of the integral term.

Remark 3.1 Following (Briani, 2004) is easy to show the discretization scheme (3.44) is monotone and consistent and since it is proven to be unconditionally stable, the discretized solution converges in a viscosity sense (Barles, 1997).

3.2 American options

Early exercise options allow the holder to exercise before maturity. The pricing PIDE is now replaced by an inequality since the hedge portfolio value can only have a return bounded above the risk-free return:

$$-\frac{\partial V}{\partial \tau} \geq \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV + \int_{-\infty}^{+\infty} \left[V\left(Se^u - V(S) - S(e^u - 1)\frac{\partial V}{\partial S}\right)\right] V(u)du \quad (3.77)$$

or in the compact form:

$$-\frac{\partial V}{\partial \tau} \geq \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + rS - rV + I_C(V) \quad (3.78)$$

Additionally an American option cannot fall beneath its immediate payoff $g(S,t)$ so:

$$V \geq g \quad (3.79)$$

must also hold. The above two conditions combine into a linear complementarity problem:

$$(V - g)\left(\frac{\partial V}{\partial t} + LV + I_C(V)\right) = 0 \quad (3.80)$$

Discrete complementarity The finite difference equations now also become inequalities:

$$V_j^{m-1} \geq V_j^m + \theta \Delta t L_{\Delta S} V_j^{m-1} + (1-\theta)\Delta t L_{\Delta S} V_j^m + \theta_C \Delta t I_j^m + (1-\theta_C)\Delta t I_j^m \quad (3.81)$$

$$V_j^{m-1} \geq g(S_j, t^{m-1}) \quad (3.82)$$

Leading to a discrete linear complementarity problem:

$$\begin{aligned}\left(V_j^{m-1} - g(S_j, t^{m-1})\right) & \\ \times \left(V_j^{m-1} - V_j^m - \theta \Delta t L_{\Delta S} V_j^{m-1}\right. & \\ \left. - (1-\theta)\Delta t L_{\Delta S} V_j^m - \theta_C \Delta t I_j^m - (1-\theta_C)\Delta t I_j^m\right) & \\ = 0 & \end{aligned} \quad (3.83)$$

Early Exercise

Projected SOR The discrete LCP (3.83) can be solved for $\theta > 0$ using a relaxation combined with a simple projection step. First the finite difference inequality is relaxed as if it was an equation:

$$V^{k+1_*} = V^k = \omega(D+L)^{-1} r^k \quad (3.84)$$

Then the price vector V^{k+1_*} is corrected by adjusting any value greater than the payoff to be equal to the payoff.

$$V^{k+1_*} = \max(V_i^{k+1_*}, g(S_j, t^m)) \quad (3.85)$$

This is done at the end of each relaxation step. If $\theta = 0$, there are no equations to relax and this adjustment to V^m is done just once after each timestep.

3.3 Numerical experiments on Merton's model

This section presents results computed with the proposed finite difference scheme on European and American options where the underlying is following Merton's jump-diffusion process. We will assess the performance of our proposed scheme through comparison with published results of other researchers in the field, and we will draw conclusions on the impact specific numerical approximation choices we are proposing have on the consistency and accuracy of the scheme.

We start our experiment using a European put and the parameters used in (d' Halluin, 2005):

$S = 100, K = 100, T = 0.25, r = 0.05, \sigma = 0.15, \lambda = 0{,}1, , a = -0{,}9, b = 0.45$, with the exact value equal to 3.1490.

Where a and b are the parameters of the Merton jump amplitude distribution.

A common way to assess the accuracy of a finite difference option pricing approximation is to measure the error at the strike for $t = 0$ that is at the start of the option duration.

$$E_K = |V(K, 0) - V_{j_K}^0| \quad (3.86)$$

Mesh point j_K will be assumed to be located at the strike asset price, unless otherwise stated.

Quadrature Error

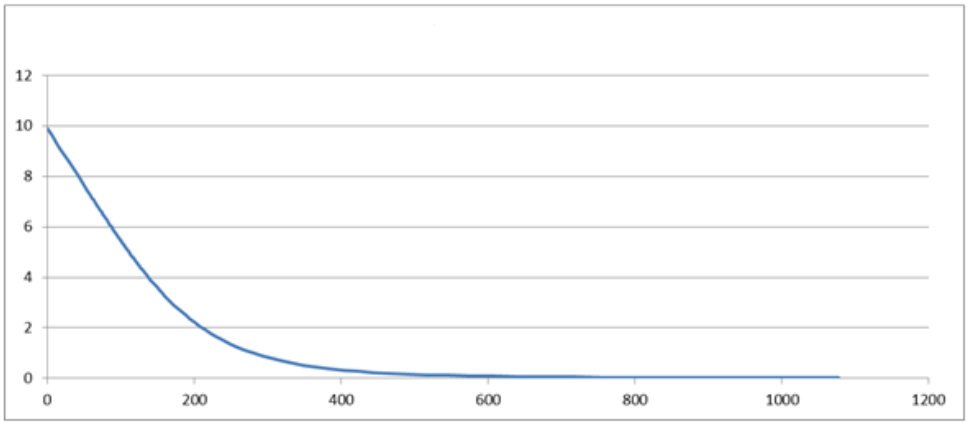
With $S_{max} = 1000$, a stretched mesh and variable timesteps we will assess the quadrature error. This will be assessed by calculating the error in the jump integral I_j, that is:

$$E_I(S_j) = \max_j \left| I_j\big(V(S, 0)\big) - \hat{I}_j^0 \right| \quad (3.87)$$

We discussed in (Section 3.1.2) that in (d' Halluin, 2004) formulation (Pr 3.1) the actual value of the price before the jump is removed from the integral term and appears exactly in the PDE part of the equation.

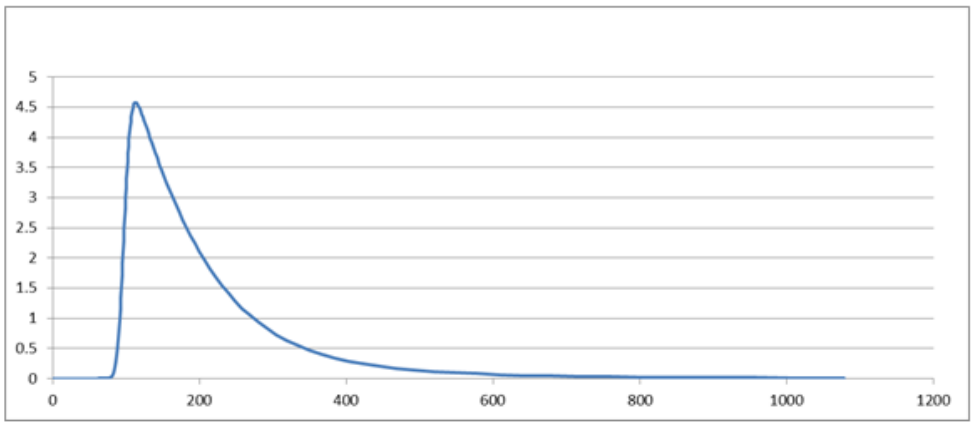
The exact values of the integral term $I_j(V(S,t^m))$ are calculated and shown in Figure 3-2 (without linear terms) and Figure 3-3 (with linear terms) as a function of S_j.

Figure 3-2 Exact Integral w/o linear terms



The exact integral without the linear terms for $a = -0.9$ is exhibiting the expected behavior, following the shape of the option price. Those two terms inside the integral can be removed with exact integration, however the shape of the integral term (with the same jump parameters) is then changed (Figure 3-3). Retaining the term helps the localization of the convolution integral, focusing that way the numerical quadrature on the region of interest that is, around the asset strike price.

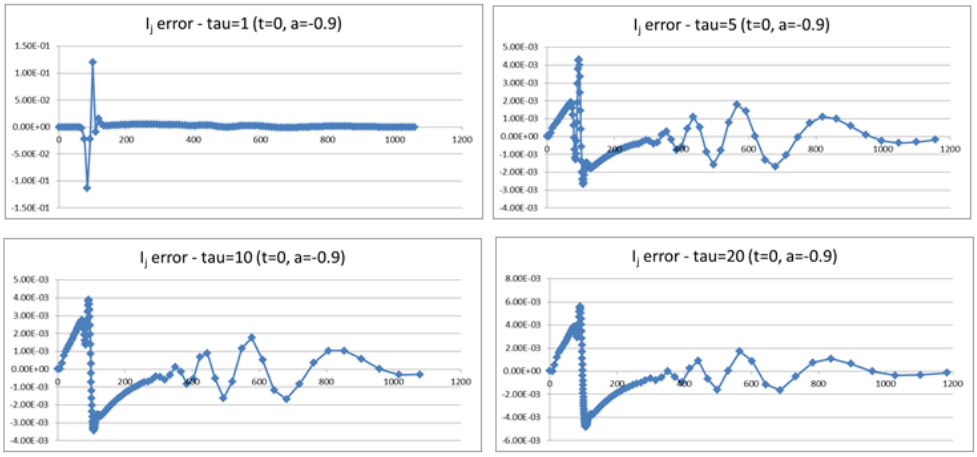
Figure 3-3 Exact Integral with linear terms



The quadrature error, when keeping the linear terms inside the integral and using the mesh quadrature described in (Section 3.1.1) with $t = 0$ for $S_j \in [0, S_{max}]$ is shown below for $N_s = 127$ (Figure 3-4) and $N_s = 254$ (Figure 3-5) with different degrees of stretch. The errors shown are total

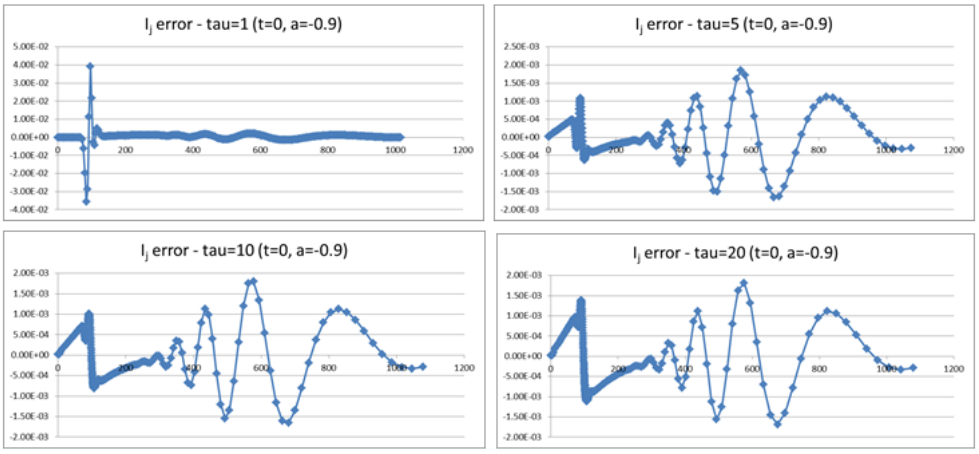
errors, i.e. both finite difference mesh price errors and quadrature errors, hence the behavior around the strike.

Figure 3-4 Quadrature error in evaluating I_j^0 ($N_s = 127$)



It can be seen that for a mesh of $N_s = 127$, a coordinate stretch of $\tau = 5$ is sufficient to remove the large quadrature errors at the strike, incurred on a uniform mesh ($tau = 1$).

Figure 3-5 Quadrature error in evaluating I_j^0 ($N_s = 254$)

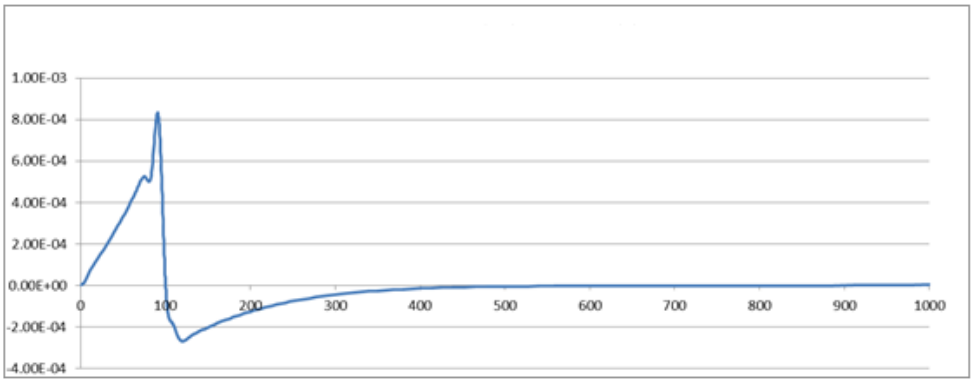


For a mesh of $N_s = 254$ a stretching of $tau = 5$ is sufficient to remove the large quadrature errors at the strike.

From Figure 3-4 and Figure 3-5 for $tau = 1$, we can see the expected reduction of 4 in the maximum quadrature error after the doubling of the mesh points. Using a stretch of $tau = 5$ should show an error reduction by a factor of 25 if the refinement is in the region of maximum quadrature error. This is approximately observed.

The use of Merton's distribution with bias towards downwards jumps will extend the computational region more than is required for the approximation of the PDE terms alone. The choice of $S_{max} = 1000$ is required for the 6 figure accuracy needed for the comparison with the results in (d' Halluin, 2005).

Figure 3-6 Price error ($N_s = 254$, tau $= 20$)



Price Accuracy and Price Error comparison

Table 3-1 shows the price accuracy at the strike with varying degrees of coordinate stretching with an explicit treatment ($\theta_C = 0$) of the integral term. The results are for a uniform mesh in the transformed variable. We are comparing against the results in (d' Halluin, 2005). The comparison results are for an unspecified variable mesh with additional nodes used for the extension to large asset values. The authors use a form of adaptive timestepping that ensures that very small timesteps are used for the initial evolution of the solution.

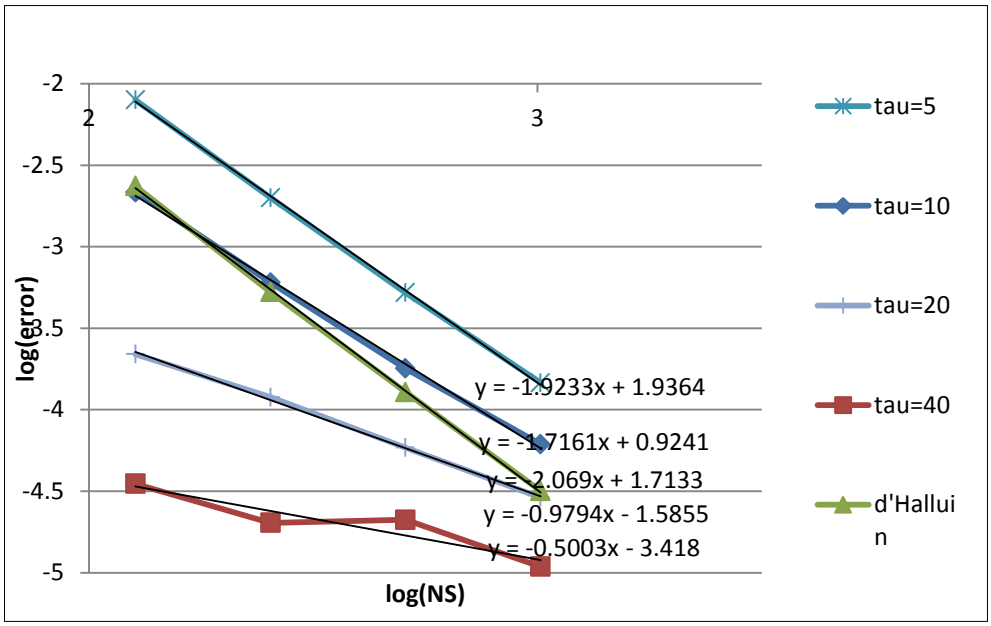
Table 3-1 Price accuracy for explicit integral treatment

Discretization		(d' Halluin, 2004)		$\theta_C = 0$			
NS	M	Price	error	tau=5	tau=10	tau=20	tau=40
127	40	3.146666	2.36E-03	7.97E-03	2.16E-03	2.19E-04	3.51E-05
254	80	3.148498	5.28E-04	2.00E-03	6.03E-04	1.19E-04	2.02E-05
508	160	3.148897	1.28E-04	5.24E-04	1.79E-04	5.85E-05	2.11E-05
1016	320	3.148994	3.17E-05	1.46E-04	6.14E-05	2.89E-05	1.09E-05

A stretched value of 20 means that the smallest value of ΔS_j is around 20 times smaller than the largest value. The transformation parameters have been set to provide the smallest mesh interval at the strike and the largest at $S = S_{max}$.

Figure 3-7 below shows the best least squares fit on a log-log plot. The slope of the straight line fit shows the overall rate of convergence. Since the jump terms are treated fully explicitly and hence are only first order accurate we do not expect to see second order convergence for this refinement strategy, unlike in (d' Halluin, 2005). For the coarsest stretch of $tau = 5$ we get close to second order convergence. This implies that the integral term error is subordinate in this discretization.

Figure 3-7 Strike Error convergence rate for explicit integral treatment



Note that the error for $tau = 40$ is smaller for $N_S = 127$ than the one in (d' Halluin, 2005) for $N_S = 1016$, though the rate of convergence is highly suboptimal. The rate of convergence for a fixed value of tau is determined by the order of accuracy of the finite difference approximation, since the timestep errors are only first order with $\theta_C = 0$, the estimated combined convergence rate is increasingly suboptimal as the spatial errors are reduced. The results for a stretch of $tau = 40$ have been calculated for value of M, the number of timesteps, around 50% greater than the values quoted for the other computations in order to remove oscillations in Gamma. In all cases, a sequence of 10 small timesteps (starting at $\frac{\Delta t}{10}$) has been used.

Table 3-2 presents the next set of results where the integral term is treated based on $\boldsymbol{\theta_C = 0.5}$ for the same set of meshes discussed previously.

Table 3-2 Price accuracy for implicit integral treatment

Discretization		(d' Halluin, 2005)		Thetta_C=0.5			
NS	M	Price	error	tau=5	tau=10	tau=20	tau=40
127	40	3.146666	2.36E-03	7.74E-03	1.93E-03	8.84E-06	2.37E-04
254	80	3.148498	5.28E-04	1.89E-03	4.88E-04	4.70E-06	5.73E-05
508	160	3.148897	1.28E-04	4.65E-04	1.21E-04	9.57E-07	1.42E-05
1016	320	3.148994	3.17E-05	1.16E-04	3.01E-05	7.95E-08	3.57E-06

It can be seen that the errors in this case exhibit a more uniform behavior in relation to the mesh refinement for all values of coordinate stretch. The optimal stretch is around $tau = 20$ as the errors for $tau = 40$ are larger. Although the errors are reduced locally at the strike, with excessive stretching, they eventually start to dominate in coarsening regions away from the strike and compromise overall accuracy. This becomes apparent in the case of $tau = 40$. This type of behavior is typical with coordinate stretching - the error initially decreases with increasing tau and at some point begins to increase, i.e. there is an optimal value.

Figure 3-8 Strike Error convergence rate for implicit integral treatment

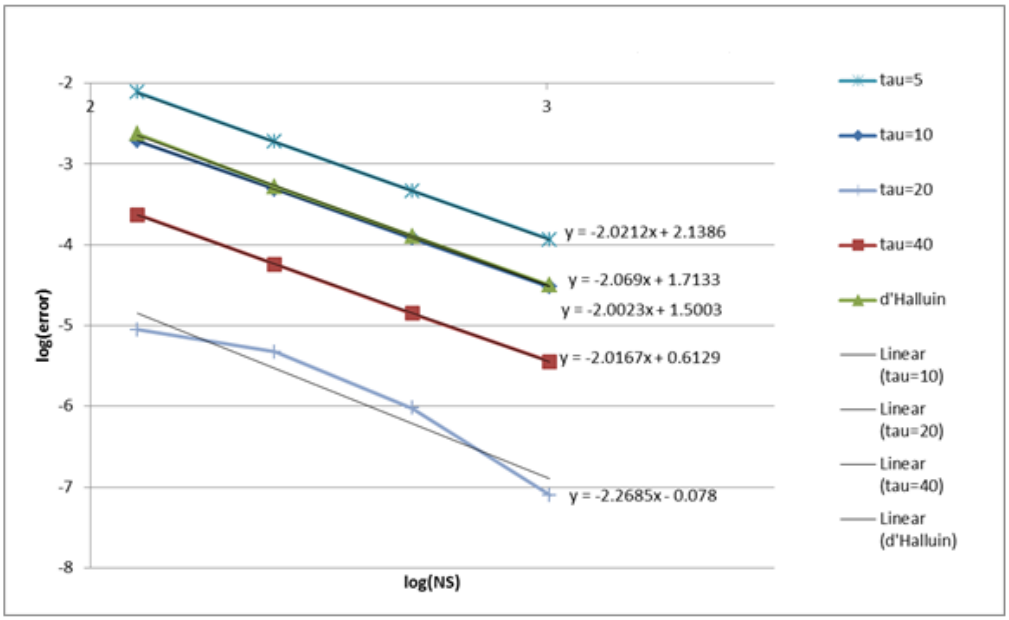
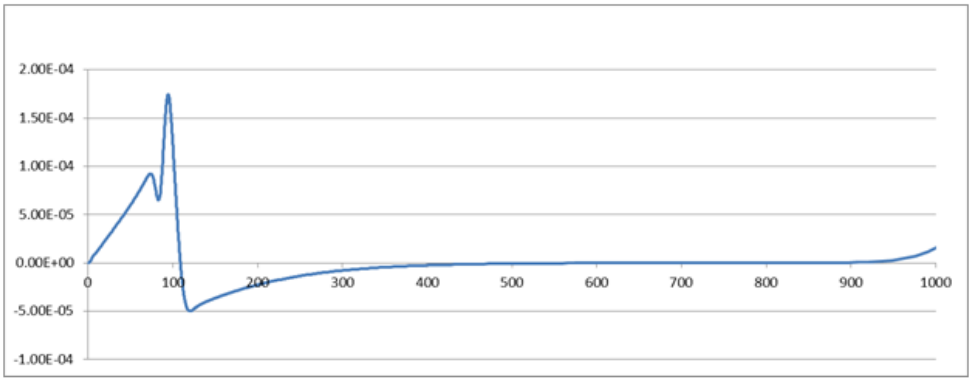


Figure 3-9 Price error ($N_s = 508, tau = 10$)



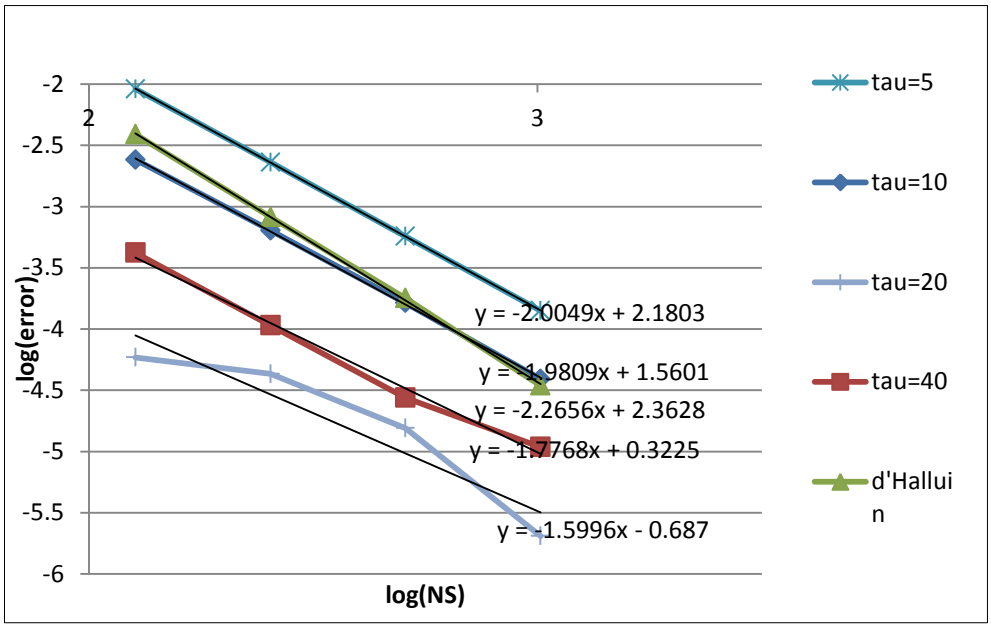
American option case results

Table 3-3 contains results for the American put option in the case of $\theta_c = 0.5$. The overall behavior is similar to the European case, with second order convergence for $tau = 5$ and $tau = 10$. Lower convergence occurs for $tau = 20$ and $tau = 40$ but with significantly smaller strike errors.

Table 3-3 American option price accuracy

Discretization		(d' Halluin, 2004)		$\theta_C = 0.5$			
NS	M	Price	error	tau=5	tau=10	tau=20	tau=40
127	40	3.23735	3.89E-03	9.13E-03	2.41E-03	5.86E-05	4.20E-04
254	80	3.240423	8.21E-04	2.30E-03	6.38E-04	4.30E-05	1.07E-04
508	160	3.241065	1.79E-04	5.71E-04	1.63E-04	1.55E-05	2.75E-05
1016	320	3.241209	3.45E-05	1.41E-04	3.91E-05	2.04E-06	1.09E-05

Figure 3-10 Strike Error convergence rate for the American case



Impact of Merton jump-diffusion parameters on the FD accuracy

The Merton jump distribution is lognormal in the jump ratio $Y = \frac{S_t^+}{S_t}$ with parameters a, b and p.d.f:

$$\tilde{v}(y) = \frac{1}{y\sqrt{2\pi}b} e^{-\frac{(\ln(y)-a)^2}{2b^2}} \quad (3.88)$$

where a is the mean jump size and b is the jump variance. The initial asset price is assumed to be $S_t = 100$ and the p.d.f is graphed against the asset price after a jump from $S_t = 100$ to $S_t^+ = S_t y = 100y$.

The implication for the approximation and localization of the convolution integral:

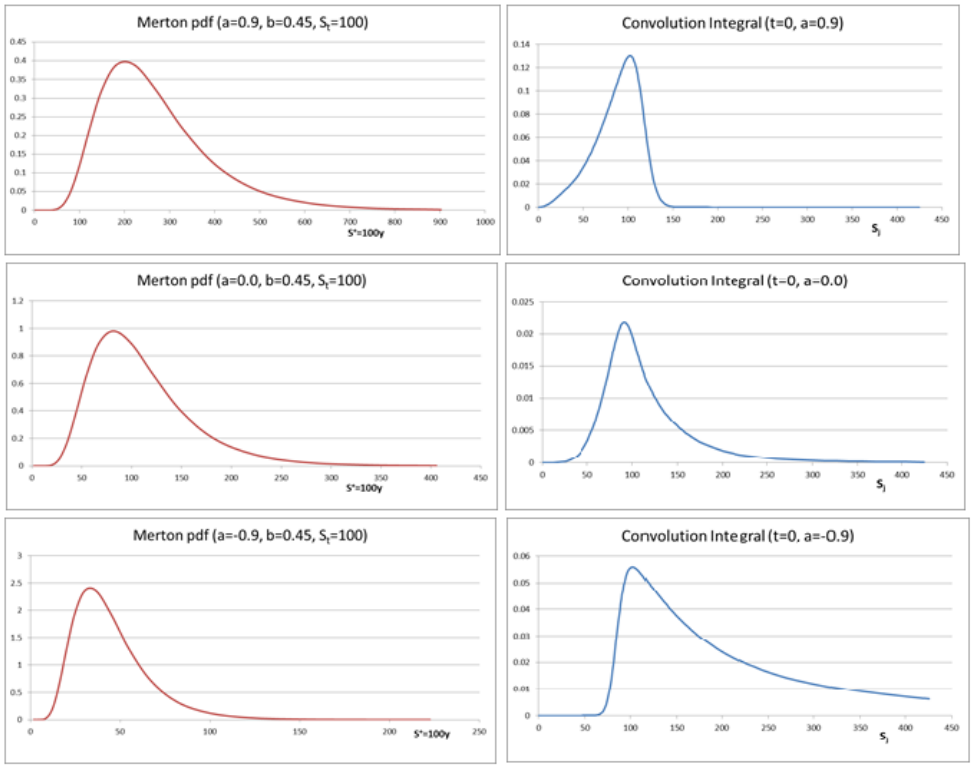
$$I_C(S_j, t) = \int_0^\infty \left(V(S_j y, t^m) - V(S_j, t) - S_j(y-1)\frac{\partial V}{\partial S} \right) \tilde{v}(y)\, dy \quad (3.89)$$

$$= F_j^m(y), \text{ (as defined in (3.18))}$$

depends on the product of the term in the brackets and the p.d.f part of the PIDE.

Figure 3-11 shows the plots of the Merton distribution function for three different mean jump ratio values ($a = -0.9, 0, 0.9$) showing probabilities of jumps from the strike price $S = 100$, alongside the graphs of the corresponding convolution integrals for $t = 0$ and for the case of a European put option.

Figure 3-11 Impact of jump ratio on Merton's distribution function and the convolution integral



The pricing calculations were performed with parameters $T = 1, K = 100, r = 0.05, \sigma = 0.15$ for a range of asset price (S_j) values. Since the payoff function in this case is $g(S,t) = (100 - S)^+$ all the convolutions have a maximum at $S = 100$. This is due to the sharp twist in the payoff function, that is, the term $f(S_j, t)$ is zero when $V(S,t)$ is a continuous linear function. Therefore its distribution in this case will have a maximum near $S = 100$.

***Case:* $\boldsymbol{a = 0.9}$** The p.d.f is biased towards the increased asset values, so for $S > 100$ we have that $f(S_j, t) \to 0$ and the convolution integral will cut off shortly after $S = 100$, while it is non-zero close to $S = 0$ since small asset values may exhibit jumps up close the strike as there is a long tail to the right.

Case: $\boldsymbol{a = 0.0}$ The p.d.f is unbiased so only the convolution integral for $S > 300$ still draws values of $f(S_j, t)$ close to the strike region where it is non-zero. The convolution integral is zero for $S < 25$ since jumps from these asset values are unlikely to reach the strike region.

Case: $\boldsymbol{a = -0.9}$ The p.d.f is biased towards decreased asset values for $S > 400$. The convolution still samples $f(S_j, t)$ values close to the strike region where is it non-zero. The increased negative bias of the convolution integral is zero for $S < 60$ since jumps from these asset values are unlikely to reach the strike region.

It becomes apparent from the graphs that the negative bias occurring for $a = -0.9$ poses the greatest challenge in the localization.

Table 3-4 to Table 3-6 show via three accuracy measures the effect of varying truncation values on accuracy. Those measures are the error at the strike, the discrete L_2 error and the L_∞ error. All computations used a stretched mesh with $tau = 12$.

Table 3-4 Effect on accuracy for varying truncation values ($\boldsymbol{a = 0.9}$)

a=0.9			
Smax	Strike error	L2 Error	Lmax Error
400	4.25E+00	3.57E+01	6.13E+00
600	9.25E-01	1.02E+01	2.08E+00
800	1.73E-01	2.83E+00	6.84E-01
1000	1.77E-02	6.77E-01	2.12E-01
1200	3.76E-02	4.84E-01	1.49E-01
1500	4.32E-02	4.38E-01	1.28E-01

Table 3-5 Effect on accuracy for varying truncation values ($\boldsymbol{a = 0}$)

Table 2: a=0.0			
Smax	Strike error	L2 Error	Lmax Error
400	1.27E-02	9.47E-02	1.52E-02
600	3.51E-04	1.35E-02	2.64E-03
800	1.18E-03	1.21E-02	2.73E-03
1000	1.15E-03	1.25E-02	2.93E-03

Table 3-6 Effect on accuracy for varying truncation values ($\boldsymbol{a = -0.9}$)

Table 3: a=-0.9			
Smax	Strike error	L2 Error	Lmax Error
400	6.41E-03	1.24E-01	5.53E-02
600	6.25E-03	1.06E-01	2.62E-02
800	6.12E-03	1.04E-01	2.70E-02

1000	5.98E-03	1.04E-01	2.79E-02

It can be seen that the truncation when using a truncation value of $S_{max} = 400$ for both the PDE and the convolution term, is only effective when the distribution preferentially jumps to the left ($a = -0.9$) and that much larger S_{max} values are needed for improved strike errors for neutral or positive weightings. This is the case, despite of the convolution integral being zero for $S > 400$ in this case.

Focusing now on the function:

$$f(S,t) = V(S_j y, t^m) - V(S_j, t) - S_j(y-1)\frac{\partial V}{\partial S} \quad (3.90)$$

we will see how it and the convolution integral behaves (for $S_J = 395$) for the negative ($a = -0.9$) and positive ($a = 0.9$) jump ratios.

Figure 3-12 Behavior of $F(S_j, t)$, $S_j = 395$, $t = 0$, $a = -0.9$

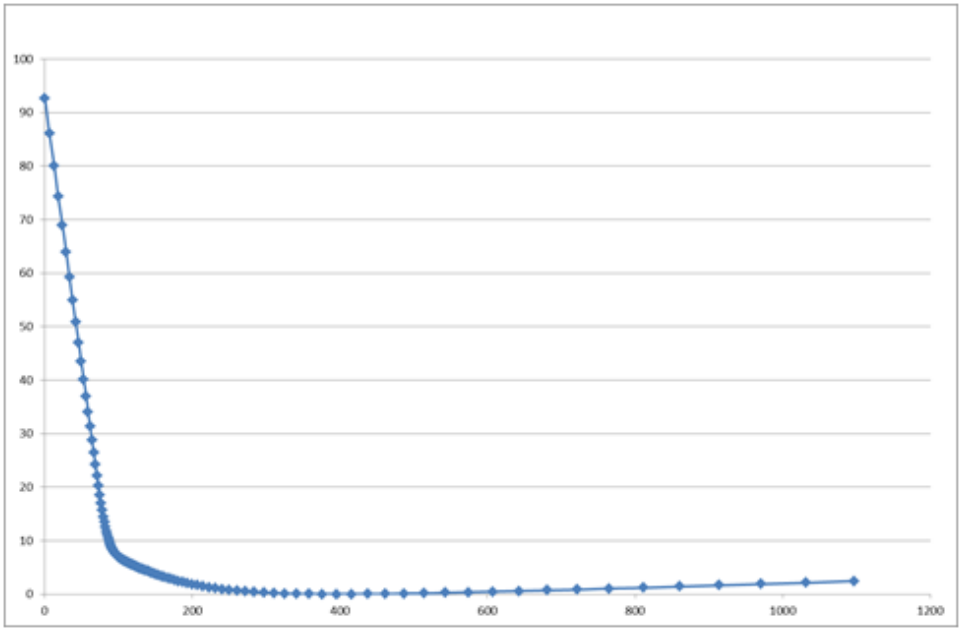


Figure 3-13 Convolution integral for $t = 0, a = -0.9$

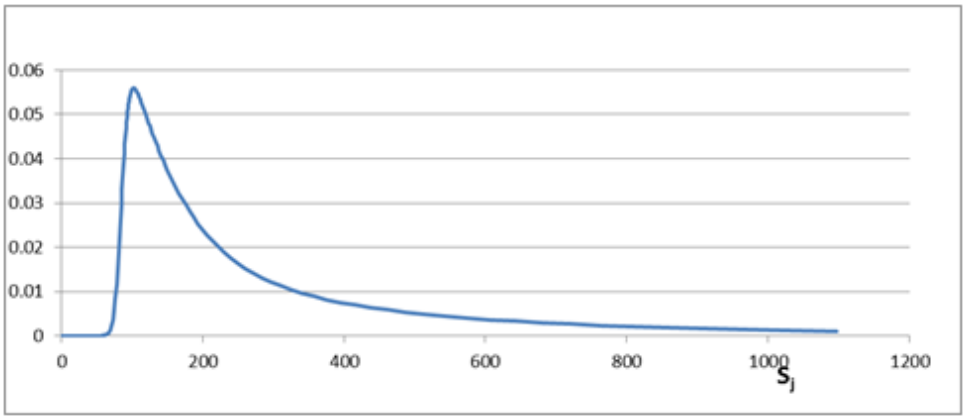
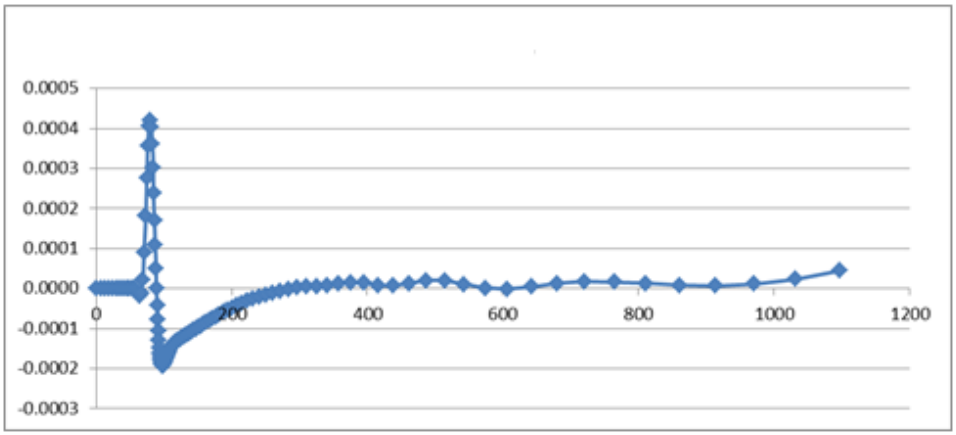


Figure 3-14 Convolution integral (I_c) error for $= 0, a = -0.9$



It can be seen from Figure 3-12 that the function closely resembles the solution since $V(395,0) \approx 0$ and $V_S(395,0) \approx 0$. It is interesting to observe that the function is non zero for $S = 800$.

Below are the equivalent graphs for $a = 0.9$.

Figure 3-15 Behavior of $F(S_j, t)$, $S_j = 395$, $t = 0$, $a = 0.9$

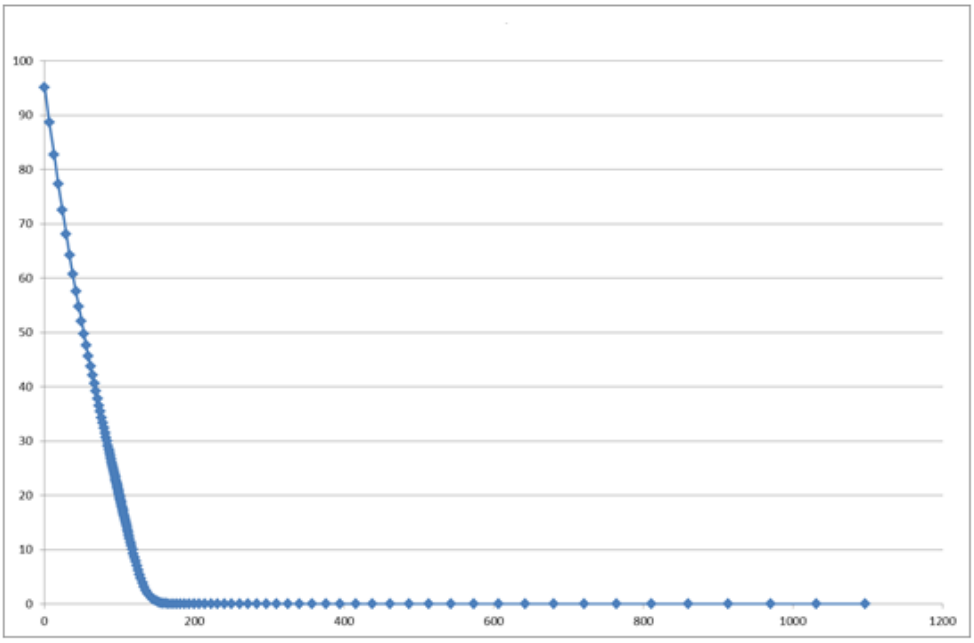
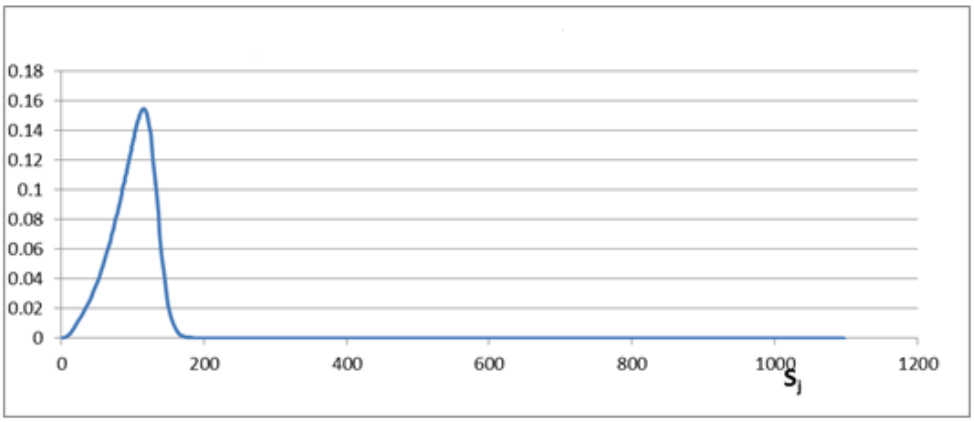
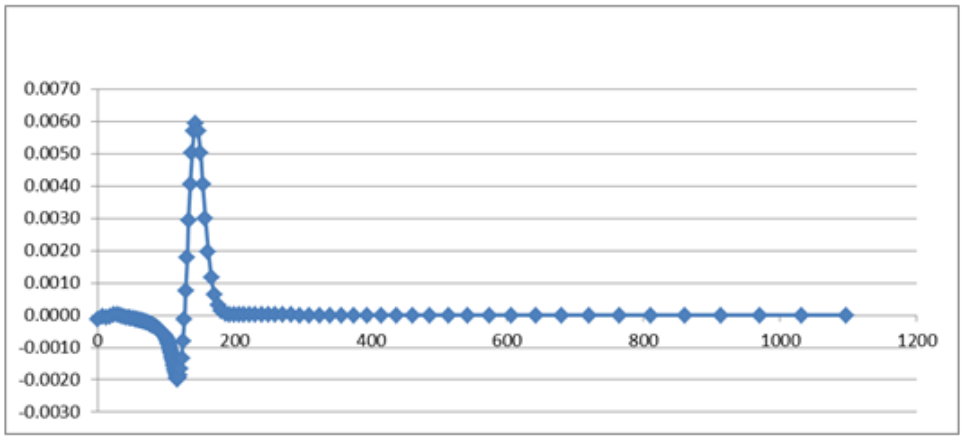


Figure 3-16 Convolution integral for $t = 0$, $a = 0.9$



We see that for $a = 0.9$ the *strange* behavior of F_j^m for large values of S disappears.

Figure 3-17 Convolution integral (I_c) error for $= 0, a = 0.9$



3.4 Jump-diffusion model calibration

We discussed in (Section 1.1) that the jump-diffusion models provide a useful tool in explaining the non-flat volatility shape observed in the market, when one is trying to reconstruct this parameter from option prices, since those types of models allow for different volatility levels for different strikes and maturities.

In the results computed and presented in this work, the volatility level was set to be the same as the one used in published work from research peers, to allow for a meaningful comparison of the numerical approximation performance.

However, one could select to use volatility levels, observed in the market, via means of calibration. That would be performed by choosing the optimal volatility of volatility levels that would minimize the error between model produced option prices (given an initial value of volatility) and option prices observed in the market, for different option types, strikes and maturities. That would enable for the efficient *backtesting* of the model when the input parameters are derived from historical observed prices. Doing so, one can assess the ability of models to "predict" the realized/ (historical) option price levels by comparing the calibrated model produced ones with the observed ones for the same time period. Interesting findings regarding the Merton's jump-diffusion models backtesting performance and other jump-diffusion models can be found in (Voltchkova, 2006).

In this chapter we presented the FD approximation for the solution of the PIDE (3.1). Results were presented and compared with published work from research peers.

Concluding remarks

1. The mesh based quadrature approach to the jump integral has been shown to be accurate when applied to the Merton jump-diffusion model.
2. The use of coordinate stretching improves both the accuracy of this quadrature as well as the accuracy of the p.d.e. approximation.
3. The combination of coordinate stretching and mesh based quadrature gives comparable or better accuracy (for the same number of degrees of freedom) than the standard approach based on a separate approximation of the p.d.e. and the jump integral.
4. Localization of the jump integral is achieved for mesh based quadrature with the standard p.d.e. mesh truncation.
5. This localization is improved by applying the quadrature to the non-convolution terms, hence retaining helpful cancellations in the integrand, and this effect can be explained by defining the linear extrapolation error.
6. The mesh refinement process for the p.d.e. approximation automatically reduces both the p.d.e. errors and the quadrature error.
7. Simple explicit weighting of the jump term has useful accuracy with this approach and is stable under the usual mesh refinement but leads to sub-optimal rates of convergence as has been noted by other authors.
8. Crank-Nicolson timestepping with the approach leads to optimal second order accuracy for both European and American options.
9. The use of a sequence of refined timesteps of the initial evolution is important for accuracy, as is placing a mesh point at the strike for vanilla payoffs examined here.

4. Singularities

In this section we present the implementation of the previously discussed numerical approximation, on the Variance Gamma and CGMY test cases (Section 4.1). Results and the effectiveness of the approach will also be discussed (Section 4.2).

4.1 Treating singularities

We saw in (Section 2.2.4) the mathematical formulation for the pricing of European options where the underlying S is following a Lévy jump-diffusion process whose Lévy measure satisfies:

$$\int_{-1}^{+1} u^2 v(u)du < \infty, \quad \int_{|u|>1} v(u)du < \infty \quad (4.1)$$

The approach needs to be modified for Variance Gamma and CGMY since now the finite difference scheme (3.44) will contain singular terms due to the singularity of $\tilde{v}(y_{jp})$ for $p = j$ i.e. $y = 1$. Hence the quadrature terms $\tilde{v}(y_j)$ are singular, namely:

$$\omega_{jj} = \frac{1}{2}(y_{jj+1} - y_{jj-1})\tilde{v}(y_{jj}) \quad (4.2)$$

where $\tilde{v}(y_{jj})$ the transformed Lévy density[13].

The approach described in (Wang, 2007) is adapted to our mesh quadrature approach; this introduces an integrable replacement of the singular integral term by a second order Taylor series.

The discrete integral terms:

$$I_j^m = \sum_{p=0}^{p=N_S} \omega_{jp} V_p^m - \Omega_j V_j^m - \kappa_j S_j \frac{(V_{j+1}^m - V_{j-1}^m)}{2\Delta S} \quad (4.3)$$

resulted from applying composite quadrature to the jump integral, i.e:

[13] $y\tilde{v}(y) = v(lny) = v(u)$

$$I_j\big(V(S,t^m)\big) = \int_0^{y^*} F_j^m(y)\tilde{v}(y)dy = \sum_{p=0}^{p=N_Q-1} \int_{y_{jp}}^{y_{jp+1}} F_j^m(y)\tilde{v}(y)dy \quad (4.4)$$

$$\approx \sum_{p=0}^{p=N_Q-1} \frac{1}{2}(y_{p+1}-y_p)[F_j^m(y_{p+1})\tilde{v}(y_{p+1}) + F_j^m(y_p)\tilde{v}(y_p)] \quad (4.5)$$

We remove and evaluate separately the two integrals ($\tilde{v}(y_{jj})$ is singular):

$$\int_{y_{jj-1}}^{y_{jj}} F_j^m(y)\tilde{v}(y)dy \,,\, \int_{y_{jj}}^{y_{jj+1}} F_j^m(y)\tilde{v}(y)dy \quad (4.6)$$

This is easily accounted for in the quadrature summation by setting:

$$\omega_{jj} = 0, \omega_{jj-1} = \cdots, \omega_{jj+1} = \cdots,$$

Combining the singular integrals to:

$$\begin{aligned} \int_{y_{jj-1}}^{y_{jj}} F_j^m(y)\tilde{v}(y)dy &= I_j^s \\ &= \int_{y_{jj-1}}^{y_{jj+1}} \left(V\big(S_j y, t^m\big) - V\big(S_j, t^m\big) \right. \\ &\quad \left. - S_j(y-1)\left(\frac{\partial V}{\partial S}\right)_{S=S_j}^{t=t^m} \right) \tilde{v}\,(y)dy \end{aligned} \quad (4.7)$$

The use of mesh derived quadrature points y_{jp} means that the interval removed is different for each set of quadrature intervals and is smallest where the mesh is most refined. Expanding $V\big(S_j y, t^m\big)$ in a Taylor series around $y = 1$ up to second order gives:

$$\begin{aligned} V\big(S_j y, t^m\big) \approx V\big(S_j, t^m\big) &+ S_j(y-1)\left(\frac{\partial V}{\partial S}\right)_{S=S_j}^{t=t^m} \\ &+ \frac{1}{2}\big[S_j(y-1)\big]^2 \left(\frac{\partial^2 V}{\partial S^2}\right)_{S=S_j}^{t=t^m} \end{aligned} \quad (4.8)$$

So the singular integral terms can be approximated as:

$$I_j^s \approx \int_{y_{jj-1}}^{y_{jj+1}} \left(\frac{1}{2}\big[S_j(y-1)\big]^2 \left(\frac{\partial^2 V}{\partial S^2}\right)_{S=S_j}^{t=t^m} \right) \tilde{v}\,(y)dy \quad (4.9)$$

$$= S_j^2 \left(\frac{\partial^2 V}{\partial S^2}\right)_{S=S_j}^{t=t^m} \int_{y_{jj-1}}^{y_{jj+1}} \left(\frac{1}{2}[(y-1)]^2\right) \tilde{\nu}\,(y)dy = \frac{1}{2}\sigma_{sj}^2 S_j^2 \left(\frac{\partial^2 V}{\partial S^2}\right)_{S=S_j}^{t=t^m}$$

where the mesh defined singular jump volatility σ_{sj} is given by:

$$\sigma_{sj}^2 = \int_{y_{jj-1}}^{y_{jj+1}} [(y-1)]^2 \tilde{\nu}\,(y)dy \quad (4.10)$$

Evaluating the singular volatility

Applying the change of variable $u = \ln y$ to this integral we get:

$$\sigma_{sj}^2 = \int_{u_{jj-1}}^{u_{jj+1}} [(e^u - 1)]^2\, \nu(u)du \quad (4.11)$$

In the case of CGMY Lévy density, this corresponds to:

$$\sigma_{sj}^2 = \int_{u_{jj-1}}^{u_{jj+1}} [(e^u - 1)]^2\, C \left(\frac{e^{Mu}}{u^{1+Y}} 1_{>0} + \frac{e^{M|u|}}{|u|^{1+Y}} 1_{<0}\right) du \quad (4.12)$$

with:

$$u_{jj-1} = \ln\left(\frac{S_{j-1}}{S_j}\right) = \ln\left(1 - \frac{(S_j - S_{j-1})}{S_j}\right) = \ln\left(1 - \frac{\Delta S_j}{S_j}\right) = -\frac{\Delta S_j}{S_j} + O\left(\frac{\Delta S_j}{S_j}\right)^3 \quad (4.13)$$

and similarly for:

$$u_{jj+1} = \ln\left(\frac{S_{j+1}}{S_j}\right) = \ln\left(1 + \frac{(S_{j+1} - S_j)}{S_j}\right) = \ln\left(1 + \frac{\Delta S_{j+1}}{S_j}\right) = \frac{\Delta S_{j+1}}{S_j} + O\left(\frac{\Delta S_{j+1}}{S_j}\right)^3 \quad (4.14)$$

With the absolute differences expressed as:

$$\Delta u_j^D = \frac{\Delta S_j}{S_j}, \qquad \Delta u_j^U = \frac{\Delta S_{j+1}}{S_j} \quad (4.15)$$

the integral (4.9) can be approximated as:

$$\sigma_{sj}^2 = \int_{u_{jj-1}}^{u_{jj+1}} [(e^u - 1)]^2\, C \left(\frac{e^{-Mu}}{u^{1+Y}} 1_{>0} + \frac{e^{-G|u|}}{|u|^{1+Y}} 1_{<0}\right) du \quad (4.16)$$

$$\approx \int_{-\Delta u_j^D}^{\Delta u_j^U} [(e^u - 1)]^2\, C \left(\frac{e^{-Mu}}{u^{1+Y}} 1_{>0} + \frac{e^{-G|u|}}{|u|^{1+Y}} 1_{<0}\right) du$$

$$= \int_{-\Delta u_j^D}^{0} [(e^u - 1)]^2 \frac{Ce^{-G|u|}}{|u|^{1+Y}} du + \int_0^{\Delta u_j^U} [(e^u - 1)]^2 \frac{Ce^{-Mu}}{u^{1+Y}} du$$

$$= \int_0^{\Delta u_j^D} [(e^u - 1)]^2 \frac{Ce^{-Gu}}{u^{1+Y}} du + \int_0^{\Delta u_j^U} [(e^u - 1)]^2 \frac{Ce^{-Mu}}{u^{1+Y}} du$$

Since the size of these fractional jump returns is small, expanding the exponentials gives a rapidly converging sequence of easily computable integrals:

$$= \int_0^{\Delta u_j^D} Ce^{-Gu} u^{1-Y} \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \ldots\right]^2 du + \int_0^{\Delta u_j^U} Ce^{-Mu} u^{1-Y} \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \ldots\right]^2 du \quad (4.17)$$

Variance Gamma

For Variance Gamma $Y = 0$ and $G = \lambda_n$. Then the singular volatility can be written:

$$\sigma_{sj}^2 = \int_0^{\Delta u_j^D} Ce^{-Gu} u \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \ldots\right]^2 du + \int_0^{\Delta u_j^U} Ce^{-Mu} u \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \ldots\right]^2 du \quad (4.18)$$

which for $Y < 2$ can be expressed in terms of incomplete Gamma functions.

We can then define (Section 7.3) the successive approximations:

$$\left[\sigma_{sj}^2\right]^{(1)} = \frac{C}{G^2}(1 - e^{-a} - ae^{-a}) + \frac{C}{M^2}\left(1 - e^{-b} - be^{-b}\right) \quad (4.19)$$

$$\left[\sigma_{sj}^2\right]^{(2)} = \frac{C}{G^2}(1 - e^{-a} - ae^{-a}) + \frac{2C}{G^3}\left(1 - e^{-a} - ae^{-a} - \frac{a^2}{2}e^{-a}\right) \quad (4.20)$$

$$+ \frac{C}{M^2}\left(1 - e^{-b} - be^{-b}\right) + \frac{2C}{M^3}\left(1 - e^{-b} - be^{-b} - \frac{b^2}{2}e^{-b}\right)$$

$$a = G\Delta u_j^D, b = M\Delta u_j^U$$

Adaptive upwinding criterion

The size of the singular viscosity is directly relevant to the adaptive upwinding criterion. If we assume for simplicity that $\theta = \theta_C$ then upwinding is needed if:

$$\left|\theta(\tilde{r} - D) - \theta_C \kappa_j\right| > \theta \left[\frac{\sigma^2 T(x_j)}{\Delta x}\right] \quad (4.21)$$

Since for the transformation $S(x)$ we have that:

$$S\frac{dx}{dS} = T(x), \text{ then } \frac{T(x_J)}{\Delta x} \approx \frac{S_J}{\Delta S_j} \quad (4.22)$$

and (for zero Black-Scholes volatility):

$$\frac{\sigma^2 T(x_j)}{\Delta x} \approx \frac{\sigma_{sj}^2 S_J}{\Delta S_j} \approx \frac{C}{2}\left(\frac{a^2}{G^2} + \frac{b^2}{M^2}\right)\frac{S_J}{\Delta S_j} = \frac{C}{2}\left(\left(\Delta u_j^D\right)^2 + \left(\Delta u_j^U\right)^2\right)\frac{S_J}{\Delta S_j}$$

$$= \frac{C}{2}\left(\left(\frac{\Delta S_j}{S_j}\right)^2 + \left(\frac{\Delta S_{j-1}}{S_j}\right)^2\right)\frac{S_J}{\Delta S_j} \approx C\frac{\Delta S_j}{S_j} \quad (4.23)$$

Therefore and as expected upwinding will be needed to satisfy the positivity requirement at some point while the mesh is refined, since the singular volatility term will tend to zero. The practical requirement depends on the size of C and the local refinement of the mesh.

The behavior of this singular volatility σ_{sj} is shown below for two different meshes for the VG case.

Figure 4-1 Singular volatility for $NS = 80, tau = 10$

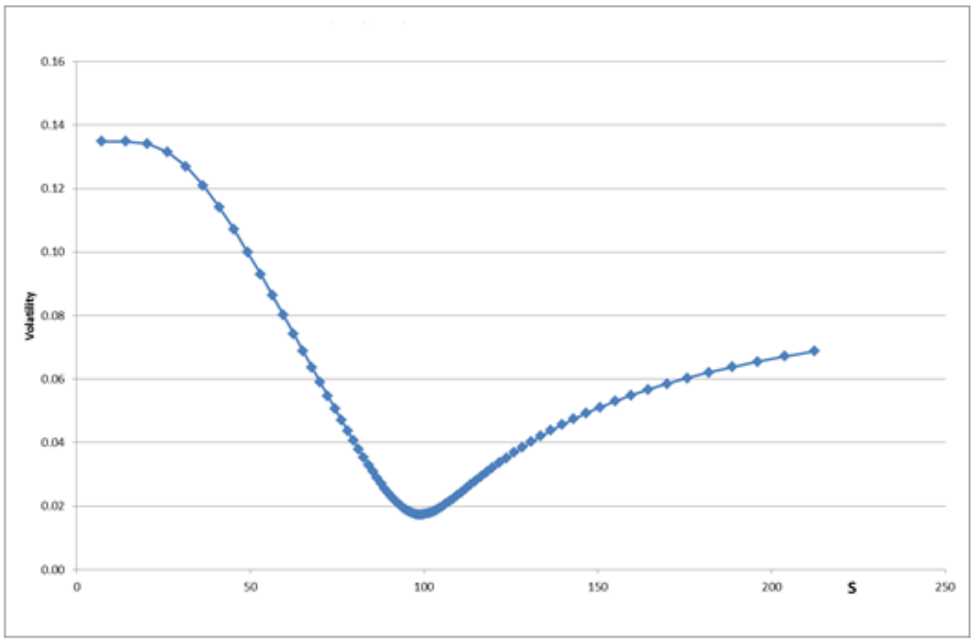
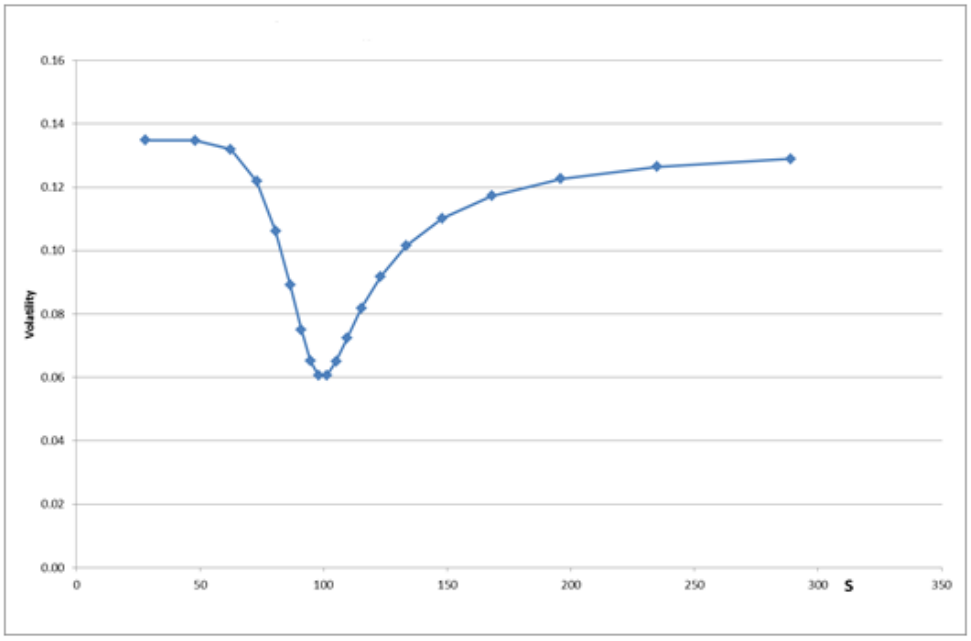


Figure 4-2 Singular volatility for $NS = 20, tau = 10$



The behavior of the singular volatility σ_{sj}^2 as shown above is consistent with what was previously discussed.

4.2 Results

In what follows we will present the results computed following the numerical approach discussed in the previous sections. The first set of results will be compared against the ones in (Wang, 2007). The VG option parameter set used for the case of a European call option is:

$$K = 98, T = 0.5, r = 0, \sigma = 0, \nu = 0.1686, \lambda_n = 20.264, \lambda_p = 39.784$$

In (Wang, 2007) the prices are quoted for a sequence of meshes for $S = 90$, $t = 0$ with the exact price being $V(90,0) = 0.6133591$.

The results from the coordinate transformed meshes require interpolating to $S = 90$ and for simplicity this was done using linear interpolation of neighbouring mesh points. Implicit timestepping is used both for the PDE ($\theta = \frac{1}{2}$) and the integral term ($\theta_C = \frac{1}{2}$). Upwinding was not used for our results in the case of the European options.

Table 4-1 Price accuracy for the VG European call option

Discretization		(Wang, 2007)		$\theta, \theta_C = 0.5$			
NS	M	Price	error	tau=5	tau=10	tau=20	tau=40
129	50	0.603428	9.93E-03	2.06E-03	2.14E-05	8.93E-05	1.49E-03
257	100	0.610918	2.44E-03	1.09E-04	5.86E-05	7.73E-05	5.52E-05
513	200	0.612863	4.96E-04	1.51E-04	2.73E-05	3.09E-05	4.28E-05
1025	400	0.613263	9.61E-05	2.51E-05	2.03E-05	1.90E-06	1.61E-05

Results in Table 4-1 show that our approach is consistently more accurate for the same number of mesh points and timesteps for all the stretch factors tested, but the convergence is non-uniform, although it is typically second order. The non-uniformity observed is most likely due to the use of interpolation (to the same order of accuracy of the finite difference method) as it disappears in the second set of results which show strike prices and errors. Please note that results from other authors on this test case are not available.

Figure 4-3 Error Convergence for VG European Call ($S = 90, \theta_C = 0.5$)

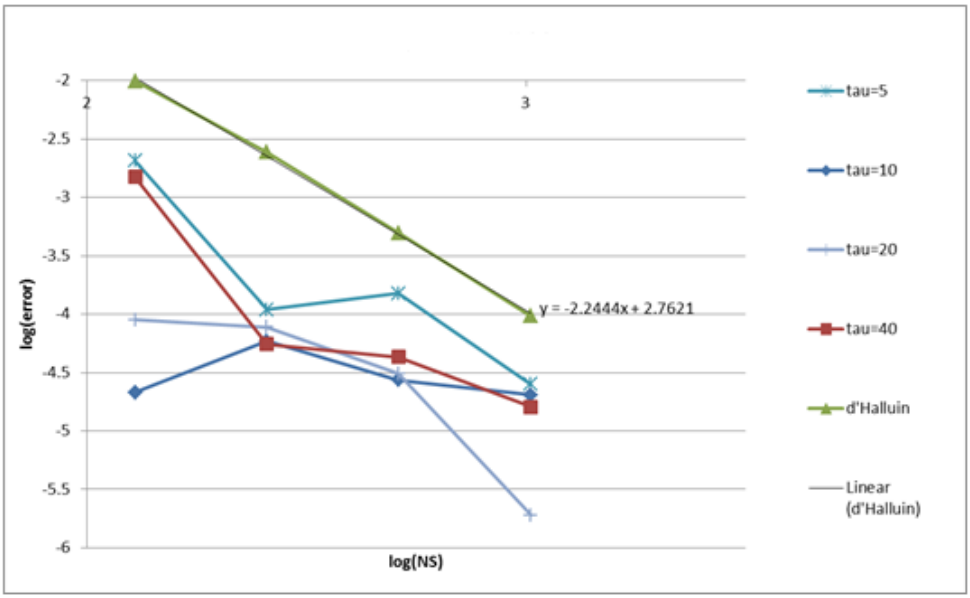


Table 4-2 Price accuracy for the VG European call option at strike price

Discretization		Theta_C=0.5			
NS	M	tau=5	tau=10	tau=20	tau=40
129	50	7.88E-03	1.52E-03	2.11E-03	4.46E-03
257	100	2.25E-03	4.83E-04	4.61E-04	1.05E-03
513	200	6.21E-04	1.36E-04	1.11E-04	2.60E-04
1025	400	1.61E-04	3.53E-05	2.77E-05	6.48E-05

The results in Table 4-2 are for the strike price $V(98,0)$. In this case interpolation is not required since the mesh is always aligned with the strike.

Figure 4-4 shows the strike error convergence for the VG European call option. Is it important to note, that although the finite difference method is clearly not a positive difference scheme as the meshes refine, there is not oscillation observed in the price, the delta or Gamma price sensitivities. A typical Gamma plot is shown in Figure 4-5 for $N_s = 512$.

Figure 4-4 Error Convergence for VG European Call ($S = S_K, \theta_C = 0.5$)

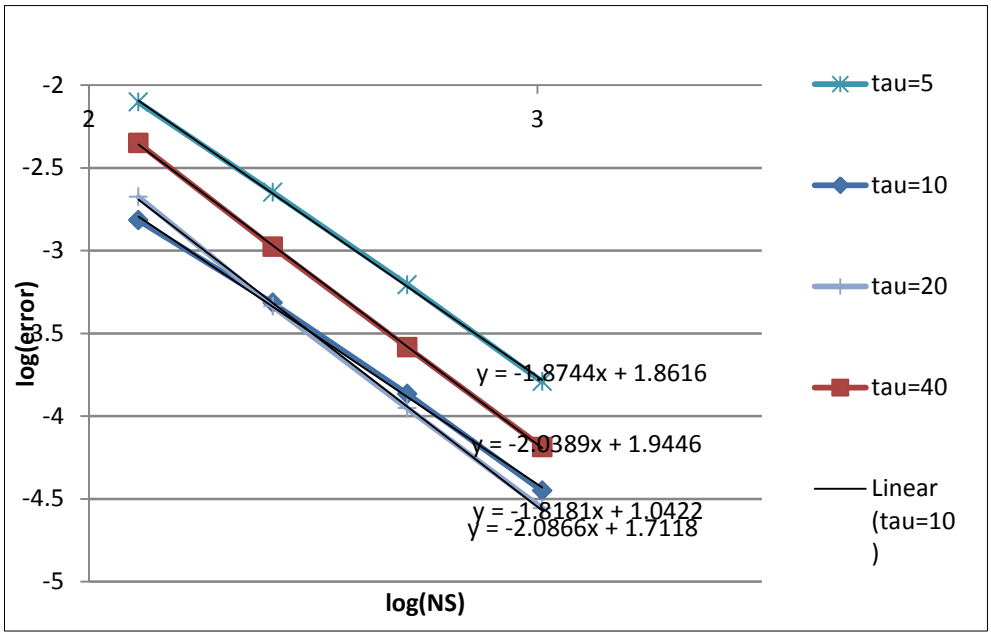
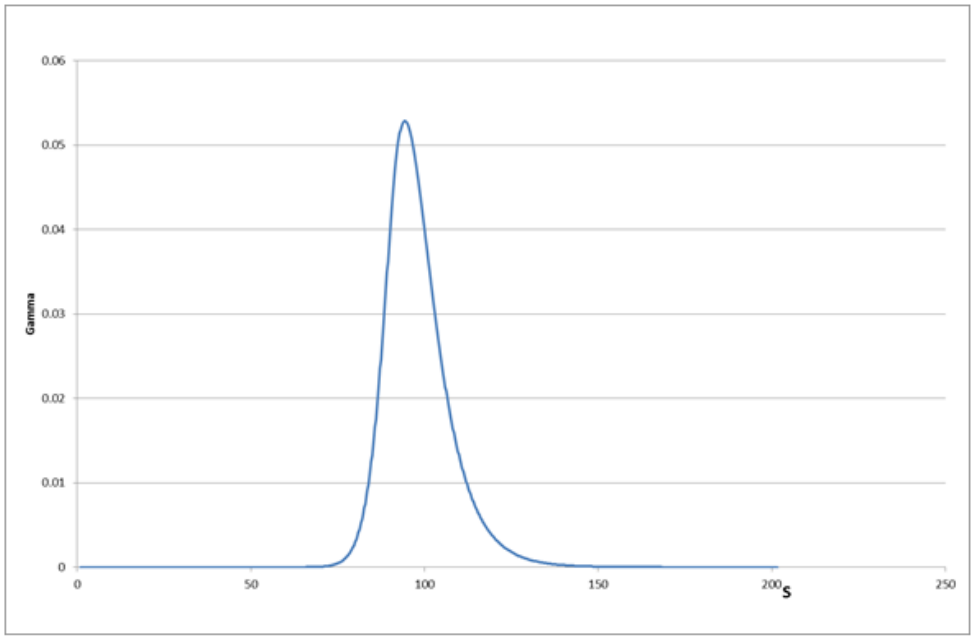
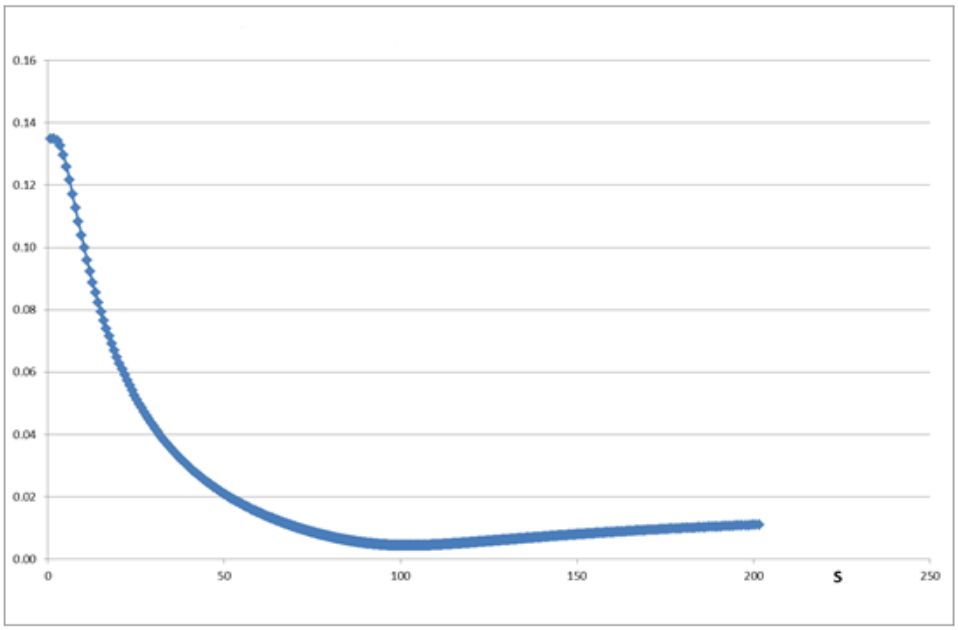


Figure 4-5 Finite differences Gamma for VG European Call option



The corresponding singular volatility function is shown in Figure 4-6 below.

Figure 4-6 Singular volatility function for European Call option ($N_S = 512, tau = 5$)



American options with Variance Gamma

When the underlying is following a VG distribution, solving for option prices with the early exercise feature can be problematic, since smooth pasting need not apply. Table 4-3 shows our results. Second order accuracy is shown for the strike price with $\theta = \theta_C = 0.5$. However we observe severe oscillations in the Gamma sensitivity local to the exercise boundary.

The nature of the solution here appears to involve a delta function like behaviour in Gamma that proves to be challenging to compute. Incorporating adaptive upwinding leads to a very similar but smoother functional behaviour for Gamma. In this case though, the numerical approximation is first order accurate as shown by the results in Table 4-3 (for exact price 2.90347) and the convergence graph shown in Figure 4-7. No sufficiently detailed comparison results are easily available for American options.

Table 4-3 Price Accuracy for the VG early exercise

Discretization		$\theta_C = 0.5$			
NS	M	tau=5	tau=10	tau=20	Tau=20(upwind)
64	20	3.54E-02	1.70E-02	7.91E-03	9.94E-02
127	40	1.06E-02	4.96E-03	2.30E-03	4.94E-02
254	80	3.01E-03	1.33E-03	5.79E-04	2.48E-02
508	160	7.43E-04	2.73E-04	6.63E-05	1.24E-02
1016	320				6.28E-03

Figure 4-7 Error Convergence for VG American Put ($S = S_K, \theta_C = 0.5$)

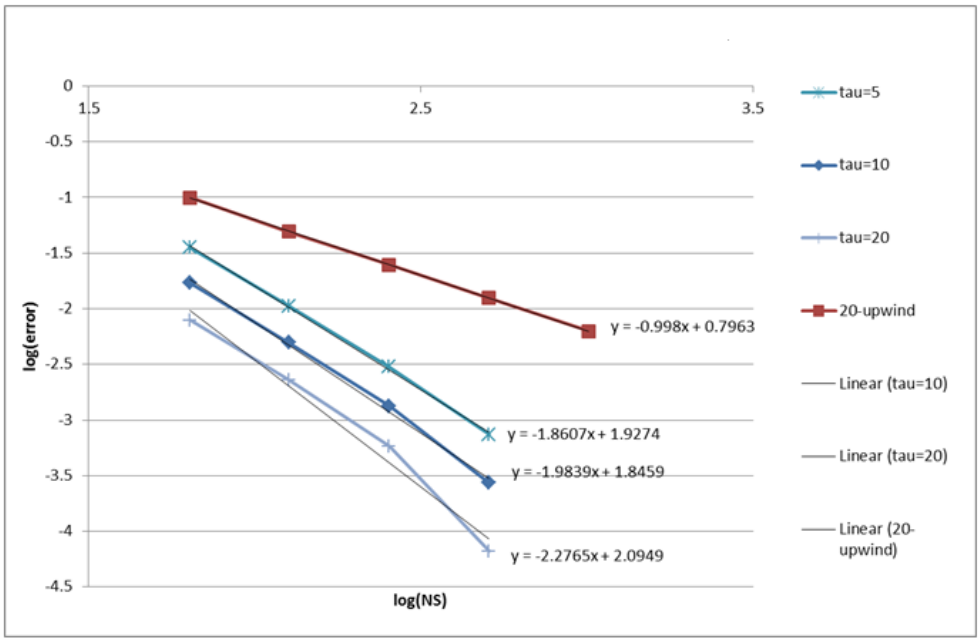


Figure 4-8 Finite differences Gamma for VG Early exercise (upwinding)

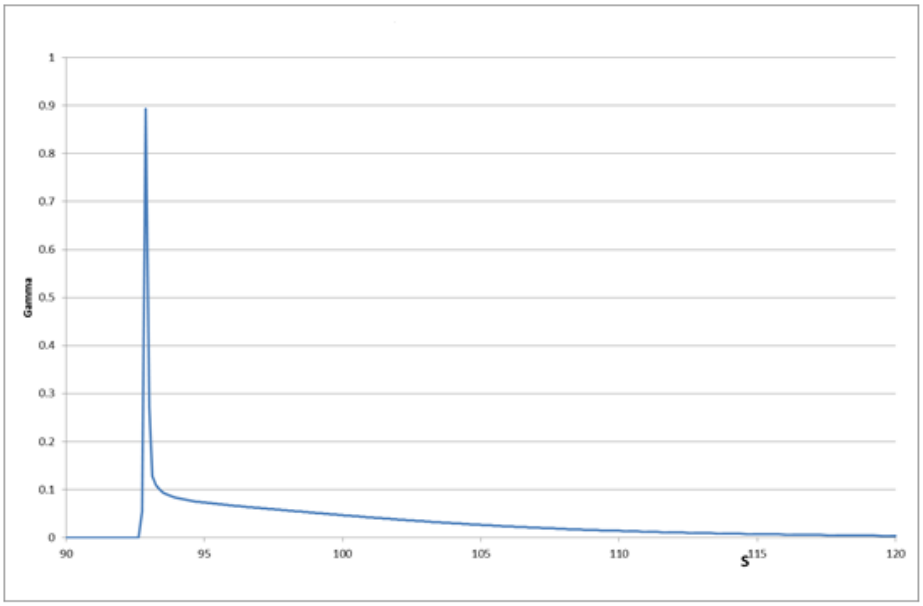
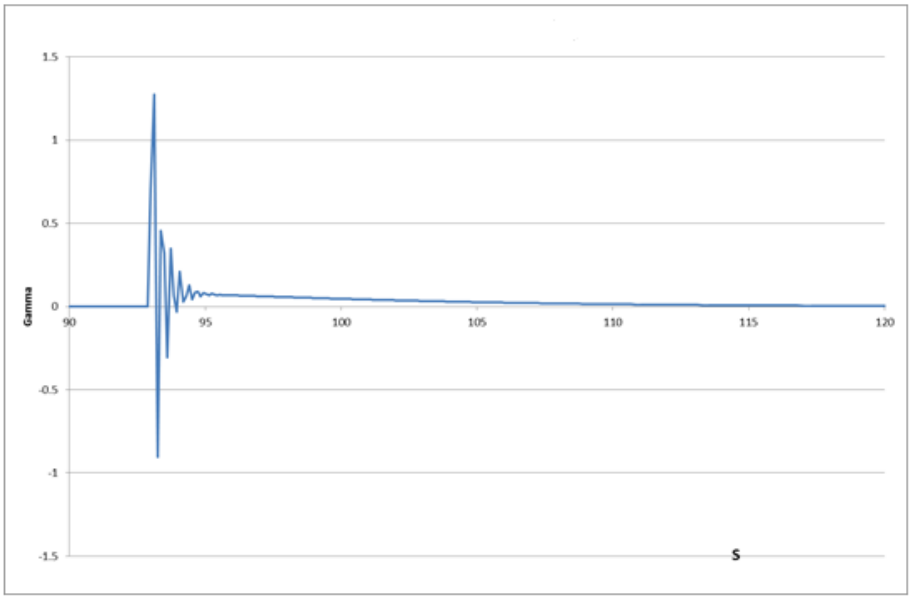


Figure 4-9 Finite differences Gamma for VG Early exercise (no upwinding)



CGMY

In what follows we present the results for a European call option where the underlying is following a CGMY process. The parameters used are:

$$\theta = \theta_C = 0.5, \quad S = 90, \quad K = 98,, \quad \sigma = 0, \quad T = 0.25, \quad r = 0.06,$$

$$C = 0.42, \quad G = 4.37, \quad M = 191.2, \quad Y = 1.0102$$

The need for interpolation to $S = 90$ again leads to some noise in the rate of convergence but it can roughly be seen to achieve second order accuracy and the errors are approximately 5 times smaller than the ones reported in (Wang, 2007) for the same number of mesh points. In the case of a European option, Gamma is always smoother but the SOR iteration did not always converge when high stretching was used and without increasing the number of timesteps (those were excluded from the Table 4-4 below).

Table 4-4 Price accuracy for the CGMY European call option

Discretization		(Wang, 2007)		$\theta_C = 0.5$			
NS	M	Price	error	tau=5	tau=10	tau=20	tau=40
64	-			1.65E-02	1.87E-02	2.40E-02	3.23E-02
127	-	2.26919	3.85E-02	6.93E-03	6.65E-03	7.14E-03	8.61E-03
254	-	2.24117	1.05E-02	2.98E-03	2.23E-03	2.11E-03	2.14E-03
508	-	2.23341	2.71E-03	8.33E-04	5.56E-04	5.32E-04	5.97E-04
1016	-	2.23135	6.50E-04	2.57E-04	1.36E-04	1.30E-04	1.52E-04

Figure 4-10 Error Convergence for CGMY European Call ($S = 90$, $\theta_C = 0.5$)

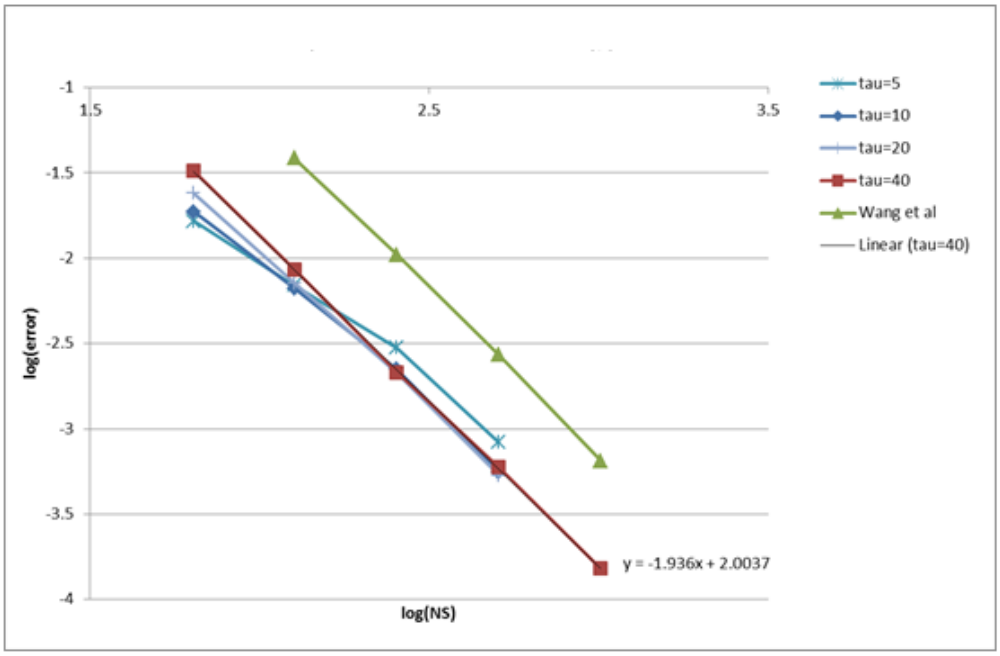
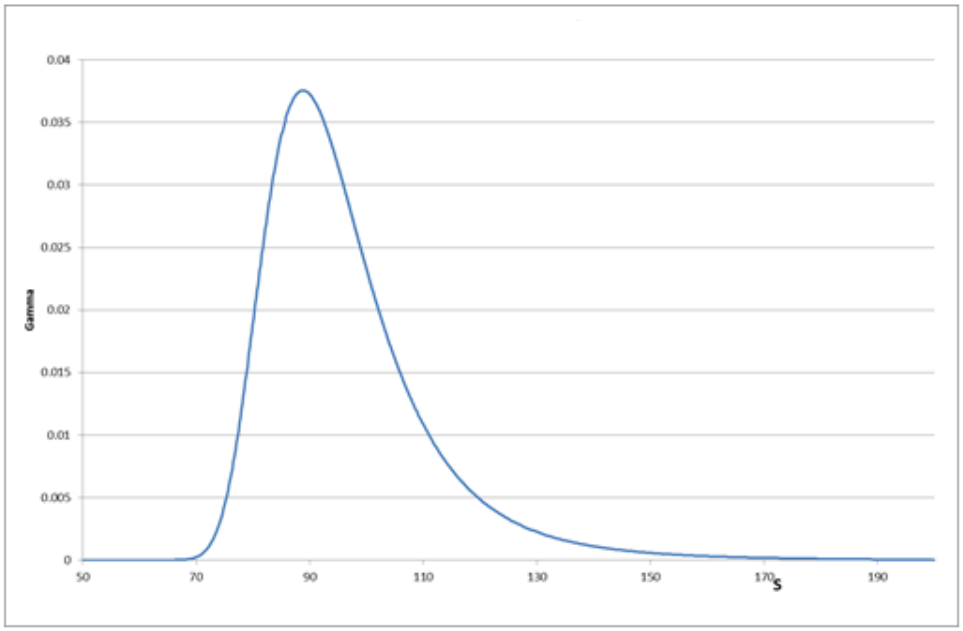


Figure 4-11 Finite differences Gamma for CGMY European (no-upwinding)



Results for an American put under CGMY process are computed and compared again with (Wang, 2007) Table 4-5 below.

The overall rate of convergence is shown in Figure 4-12 and is similar to the convergence achieved by (Wang, 2007). However, the interpolation required has introduced again noise.

Table 4-5 Price accuracy for the CGMY American put option

Discretization		(Wang, 2007)		$\theta_C = 0.5$			
NS	M	Price	error	tau=5	tau=10	tau=20	tau=40
64	50			2.26E-02	2.39E-02	2.39E-02	3.14E-02
128	100	9.2639	3.84E-02	5.97E-03	6.28E-03	8.94E-03	1.27E-02
256	200	9.23635	1.09E-02	9.88E-04	1.48E-03	2.35E-03	3.74E-03
512	400	9.22836	2.88E-03	3.17E-04	4.78E-04	6.76E-04	9.79E-04
1024	800	9.22619	7.10E-04	4.74E-05	1.35E-04	1.81E-04	1.81E-04

The localised oscillation in Gamma (Figure 4-13) can be removed by incorporating adaptive upwinding, which comes with a great cost on the accuracy. However, that does not appear to affect the accuracy at $S = 90$ which is superior to the one reported in (Wang, 2007) for the same number of mesh points for all the different extends of stretch.

Figure 4-12 Error Convergence for CGMY American put ($S = 90$, $\theta_C = 0.5$)

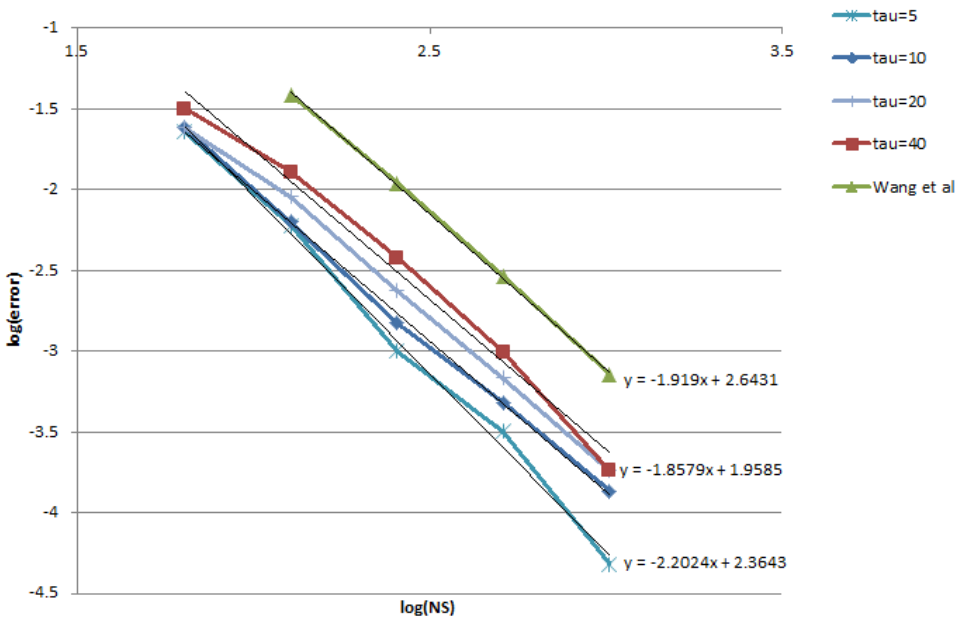
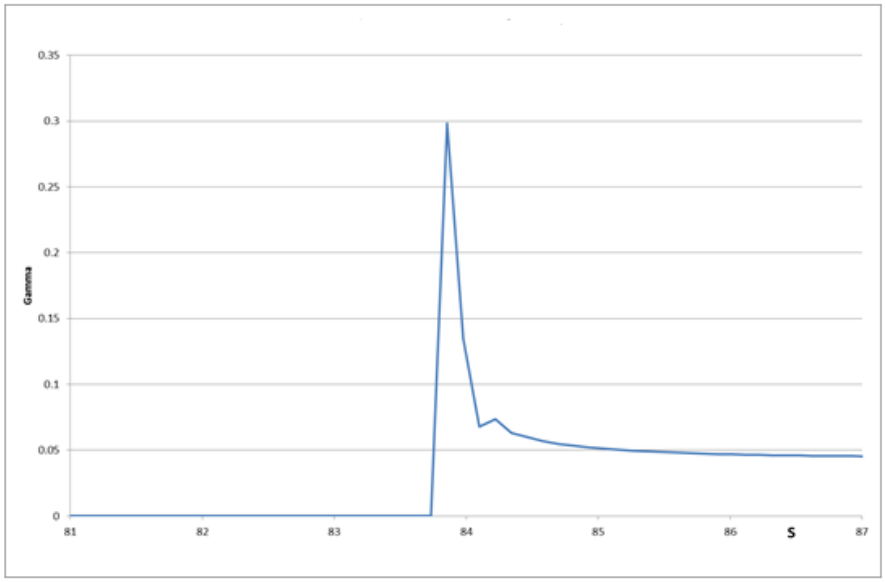


Figure 4-13 Finite differences Gamma for CGMY early exercise (no-upwinding)



In this chapter we presented the implementation of the FD approximation to test cases with singularities (VG, CGMY). Results were presented and compared with published work from research peers.

Concluding remarks

1. The use of mesh-based quadrature is easily extended to singular Lévy densities.
2. The singular volatility introduced leads to accurate prices for pure jumps despite the lack of positivity in the difference scheme.
3. The singular volatility σ_{sj}^2 varies across the mesh with smallest values where the mesh is refined most by the coordinate transformation. Since it is based on a second order approximate Taylor series this should support the accuracy of the method. The term $\sigma_{sj}^2 S_j^2$ goes to zero as the mesh is refined.
4. The use of adaptive upwinding introduces a numerical dissipation term at the strike and reduces the accuracy of the method to first order.
5. Without upwinding the method gives second order accuracy for European options for both VG and CGMY and with oscillation free Gamma, however the lack of positivity leads to problems using SOR as a solver. LU factorization would be straightforward.
6. The first VG European test case used a coordinate transformation focused at the strike ($S = 98$) and evaluated the accuracy at a point away from the strike ($S = 90$). The accuracy was still superior to the comparison results suggesting that the early evolution at the strike is crucial for these problems.

7. Without upwinding the method gives second order accuracy with PSOR for American options for both VG and CGMY. However there is unacceptable oscillation in Gamma for VG and a small localised oscillation for CGMY. The improvement with CGMY is consistent with its greater intensities for small diffusive-like jumps with CGMY for $Y > 1$.
8. There is a need for a higher order numerical dissipation term to use with this method to deal with the numerical oscillations in Gamma for the early exercise cases for pure jumps.

5. IT splitting for American options with jumps

In this chapter we incorporate the IT splitting iterative method for the approximation of the integral term and investigate its effectiveness in combination with coordinate stretching for the solution of an American option for the classical Merton model, VG and CGMY.

We saw in Section 3.2 that under certain assumptions, the Option price V, where the underlying follows an exponential Lévy process, is shown to follow the PIDE (3.1).

Introducing the auxiliary variable $\lambda(S,t)$ permits the following reformulation:

$$-\frac{\partial V}{\partial t} = \lambda + LV + I(V) \qquad (5.1)$$

$$V \geq g, \qquad \lambda \geq 0, \qquad (V-g)\lambda = 0$$

It is easy to check by elimination of λ that this formulation is identical to the original LCP but with the advantage that the PIDI is now a PIDE. Employing the usual finite difference discretization leads to:

$$V_j^{m-1} = \Delta t \lambda_j^{m-1} + V_j^m + \theta \Delta t L_{\Delta S} V_j^{m-1} + (1-\theta)\Delta t L_{\Delta S} V_j^m + \theta_C \Delta t I_j^{m-1} + (1-\theta_C)\Delta t I_j^m \qquad (5.2)$$

$$V_j^{m-1} \geq g\left(S_j, t^{m-1}\right), \qquad \lambda_j^{m-1} \geq 0, \qquad \left(V_j^{m-1} - g\left(S_j, t^{m-1}\right)\right)\lambda_j^{m-1} \geq 0$$

following the approach in Toivanen (Toivanen, 2004) where the term in λ has been made fully implicit. Toivanen describes a fractional step approximation to the discretization above, as follows:

Define the solution to the first step as $\left\{\tilde{V}_j^{m-1}, \tilde{\lambda}_j^{m-1}\right\}$, and then the first fractional step is given by:

$$\tilde{V}_j^{m-1} = \Delta t\tilde{\lambda}_j^{m-1} + V_j^m + \theta\Delta t L_{\Delta S}\,\tilde{V}_j^{m-1} + (1-\theta)\Delta t L_{\Delta S}\,V_j^m + \theta_C\Delta t\tilde{I}_j^{m-1} + (1-\theta_C)\Delta t I_j^m \quad (5.3)$$

where $\tilde{I}_j^{m-1} = I_j^{m-1}\left(\tilde{V}_p^{m-1}\right)$, and we choose $\tilde{\lambda}_j^{m-1} = \lambda_j^m$- which implies that the early exercise boundary has not moved during the first fractional step.

In the second fractional step, for $\{V_j^{m-1},\ \lambda_j^{m-1}\}$, we have:

$$\left(V_j^{m-1} - \tilde{V}_j^{m-1}\right) - \Delta t\left(\lambda_j^{m-1} - \tilde{\lambda}_j^{m-1}\right) = 0 \quad (5.4)$$

$$\left(V_j^{m-1} - g\left(S_j, t^{m-1}\right)\right)\lambda_j^{m-1} = 0 \quad (5.5)$$

$$V_j^{m-1} \geq g\left(S_j, t^{m-1}\right), \qquad \lambda_j^{m-1} \geq 0 \quad (5.6)$$

It is easily verified that the solution to step 2 can be stated explicitly as:

$$V_j^{m-1} = \max\left[\tilde{V}_j^{m-1} - \Delta t\tilde{\lambda}_j^{m-1}, g\left(S_j, t^{m-1}\right)\right] \quad (5.7)$$

$$\lambda_j^{m-1} = \max\left[0, \qquad \tilde{\lambda}_j^{m-1} + \frac{g\left(S_j, t^{m-1}\right) - \tilde{V}_j^{m-1}}{\Delta t}\right] \quad (5.8)$$

Note that adding the first equation of step 2 to the equation in step 1 gives:

$$V_j^{m-1} = \Delta t\,\tilde{\lambda}_j^{m-1} + V_j^m + \theta\Delta t L_{\Delta S}\,\tilde{V}_j^{m-1} + (1-\theta)\Delta t L_{\Delta S}\,V_j^m + \theta_C\Delta t\tilde{I}_j^{m-1} + (1-\theta_C)\Delta t I_j^m \quad (5.9)$$

so that the implicit differential and integral terms approximate the exercise boundary as fixed during the timestep and are not updated till the following timestep.

Solution procedure for Step 1

This step is solved by applying LU factorization to the tridiagonal matrix resulting from the difference approximation to the PDE terms (including the singular volatility terms for VG and CGMY) and applying a simple fixed point iteration to update the implicit jump term. As Wang et al (Wang, 2007) have shown this converges extremely rapidly (typically in 3 to 4 iterations).

It is feasible to include a Step 2 solution within the fixed point iteration and this has been done for the calculations presented later in this chapter (this appears to improve the accuracy on very refined meshes but a systematic study has not been done). The use of LU factorization in Step 1 is advantageous as it has an optimal operation count in contrast to PSOR where the convergence deteriorates on fine meshes.

Variance Gamma IT Splitting/ PSOR comparison

The first set of results described are for Variance Gamma and are a comparison with those of (Wang, 2007) and (Oosterlee, 2007). The variance Gamma option dataset is for an American call option:

$$K = 100, T = 9.0, r = 0.1, q = 0.1, \sigma = 0, \nu = 1.0, \lambda_n = 1/9.5085, \lambda_p = 1/5.2585$$

The parameters were chose to show the lack of smooth pasting for the early exercise boundary and this is visible from the plot of the option price Delta. The plot from (Wang, 2007) is reproduced below together with a plot obtained using our method combined with IT splitting, with $N_s = 240, \tau = 20, S^* = 140$.

Figure 5-1 shows the plot of Delta against asset price for $\sigma_{BS} = 0, 0.1$ showing discontinuity in Delta, where the number of mesh points is $N = 1025$ and timesteps $M = 7500$ (from (Wang, 2007)).

Figure 5-1 Discontinuity in Delta at $S = 150$.

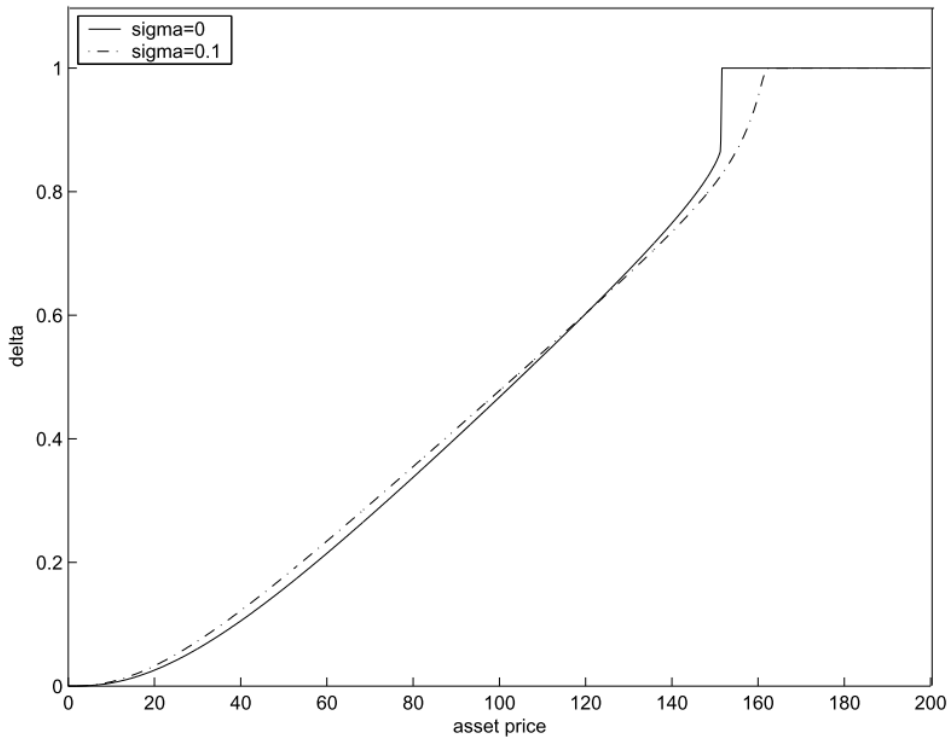
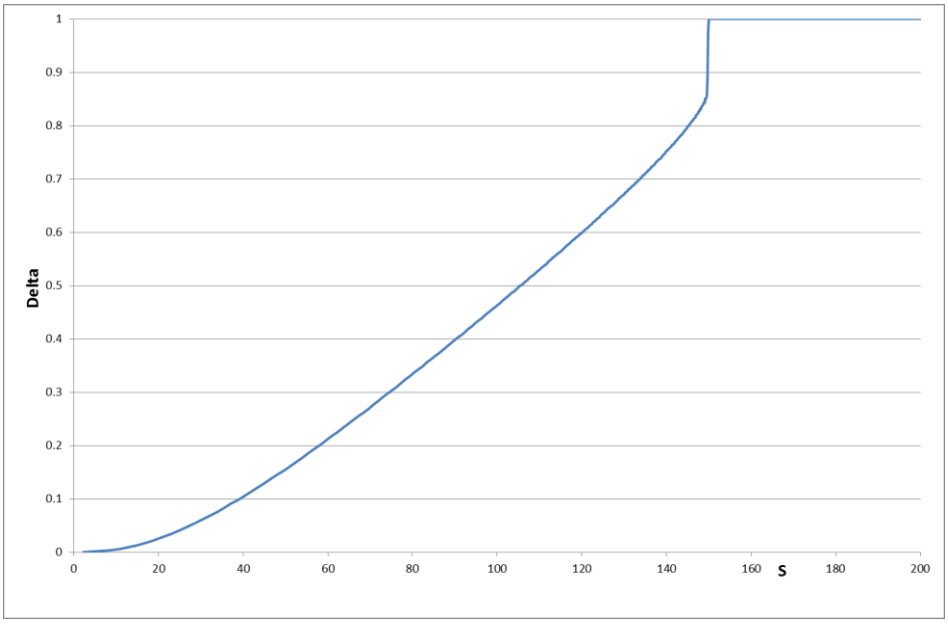


Figure 5-2 shows the comparison plot of Delta against asset price for $\sigma_{BS} = 0$, using: $\tau = 20, S^* = 145, N_s = 480, M = 240, \theta = \theta_C = 0.5, S_{max} = 300$ showing the discontinuity in Delta at $S = 150$.

Figure 5-2 Discontinuity in Delta



The second dataset is the one used with PSOR in Section 4.2 for the Variance Gamma American put option pricing problem is:

$$S = 100, \quad K = 100, \quad T = 0.5, \quad r = 0.05, \quad \sigma = 0,$$

$$\nu = 0.1686, \quad \lambda_n = 20.264, \quad \lambda_p = 39.784$$

Results with early exercise with Variance Gamma are challenging since smooth pasting need not apply and although the results below show approximately second order accuracy for the strike price with $\theta = \theta_C = 0.5$, there are severe oscillations in Gamma local to the exercise boundary. The results with IT splitting are almost identical to those obtained using PSOR; convergence to the strike price is uniform at all values of the stretch and the rate of convergence is approximately 1.8 suggesting that the oscillation in Gamma around $S \in [93,95]$ is degrading the rate of convergence at $S = 100$.

Table 5-1 shows the comparison of the results for the strike price, for the IT splitting versus PSOR, in the case of an American put option where the underlying is following a VG process, while Table 5-2 shows the strike error for different stretches using IT splitting. The exact price is approximately 2.90360 (from Richardson extrapolation).

Table 5-1 Strike prices comparison (VG American put)

Discretisation		IT Splitting			PSOR		
NS	M	tau=5	tau=10	tau=20	tau=5	tau=10	tau=20
64	20	2.867982	2.885238	2.892441	2.868059	2.886475	2.895564
127	40	2.89248	2.897572	2.899435	2.892845	2.898512	2.901173
254	80	2.900245	2.901735	2.902245	2.900463	2.902136	2.902891
508	160	2.902648	2.903069	2.903217	2.902727	2.903197	2.903404
1016	320	2.903343	2.903461	2.903504			
2035	640			2.90359			

Table 5-2 Strike error for IT splitting (VG American put)

Discretisation		$\theta_C = 0.5$			
NS	M	tau=5	tau=10	tau=20	tau=40
64	20	3.56E-02	1.84E-02	1.12E-02	9.93E-03
127	40	1.11E-02	6.03E-03	4.16E-03	4.08E-03
254	80	3.35E-03	1.87E-03	1.35E-03	1.34E-03
508	160	9.52E-04	5.31E-04	3.83E-04	3.72E-04
1016	320	2.57E-04	1.39E-04	9.59E-05	8.93E-05

Figure 5-3 American VG strike error ($\theta_C = 0.5$, IT splitting)

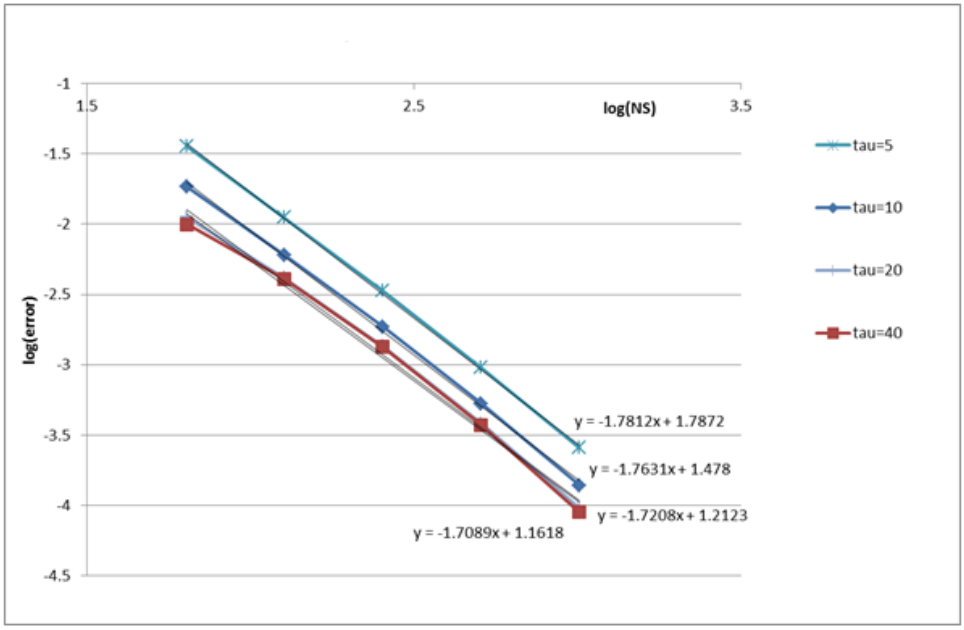


Figure 5-3 shows the best least square fit to the convergence rate for the strike errors using the IT splitting method (shown as the slope of the trend lines).

CGMY IT Splitting/ PSOR comparison

The Gamma of the option price is a sensitive feature of the solution for early exercise option and so a comparison was done for the standard PSOR implementation and the IT splitting implementation for the same mesh and number of timesteps.

The second comparison was for CGMY and comparing again with the published results in (Wang, 2007). The first set of results for CGMY are for a European call option with:

$$\theta = \theta_C = 0.5, \quad S = 90, \quad K = 98, \quad \sigma = 0, \quad T = 0.25, \quad r = 0.06,$$

$$C = 0.42, \quad G = 4.37, \quad M = 191.2, \quad Y = 1.0102$$

The need to interpolate to $S = 90$ again leads to some variability in the rate of convergence but it can roughly be seen to be second order and the accuracy is approximately 5 times smaller than (Wang, 2007) for the same number of mesh points.

Table 5-3 shows the price errors in the case of CGMY for an American option (using linear interpolation between mesh points either side of $S = 90$). The IT splitting results obtained with Crank-Nicolson, while the (Wang, 2007) results were obtained using adaptive timestepping with $M/4$ timesteps.

Table 5-3 Price accuracy results with IT splitting

Discretisation		(Wang, 2007)		$\theta_C = 0.5$, IT Splitting			
NS	M	Price	error	tau=5	tau=10	tau=20	tau=40
64	50			2.23E-02	2.36E-02	3.13E-02	4.28E-02
128	100	9.2639	3.84E-02	5.87E-03	6.28E-03	9.18E-03	1.47E-02
256	200	9.23635	1.09E-02	9.52E-04	1.53E-03	2.32E-03	3.77E-03
512	400	9.22836	2.88E-03	3.20E-04	4.63E-04	6.69E-04	1.64E-03
1024	800	9.22619	7.11E-04	8.51E-05	1.37E-04	1.84E-04	3.83E-04

Table 5-4 shows option prices in the case of CGMY for an American put option (using linear interpolation between mesh points either side of $S = 90$). Again, IT splitting results were obtained with Crank-Nicolson.

Table 5-4 Price accuracy comparison (CGMY American put)

Discretisation		$\theta_C = 0.5$, IT Splitting				$\theta_C = 0.5$, PSOR			
NS	M	tau=5	tau=10	tau=20	tau=40	tau=5	tau=10	tau=20	tau=40
64	50	9.203223	9.2019	9.194214	9.18268	9.202909	9.201556	9.201556	9.194038
128	100	9.219606	9.219199	9.216296	9.210769	9.219508	9.219199	9.216541	9.2128
256	200	9.224527	9.223952	9.223157	9.221705	9.224492	9.224003	9.223129	9.221739
512	400	9.225158	9.225016	9.22481	9.223834	9.225163	9.225002	9.224804	9.224501
1024	800	9.225394	9.225341	9.225295	9.225096	9.225433	9.225345	9.225299	9.225299

The rate of convergence for IT splitting is similar to PSOR, however this may reflect poor control of the fixed point iteration residuals as the meshes were refined and a more robust iteration control loop may improve this. The rate for $tau = 5$ is optimal and almost identical to the PSOR results, with the same notable improvement in accuracy compared to (Wang, 2007). The price Gamma has a small localized oscillation just after the peak for PSOR and this may affect the accuracy at $S = 90$. There is no observable difference in the Gamma profiles between the two methods, as it can been seen from Figure 5-6 and Figure 5-7 below.

Figure 5-4 Error Convergence IT splitting ($S = 90, \theta_C = 0.5$)

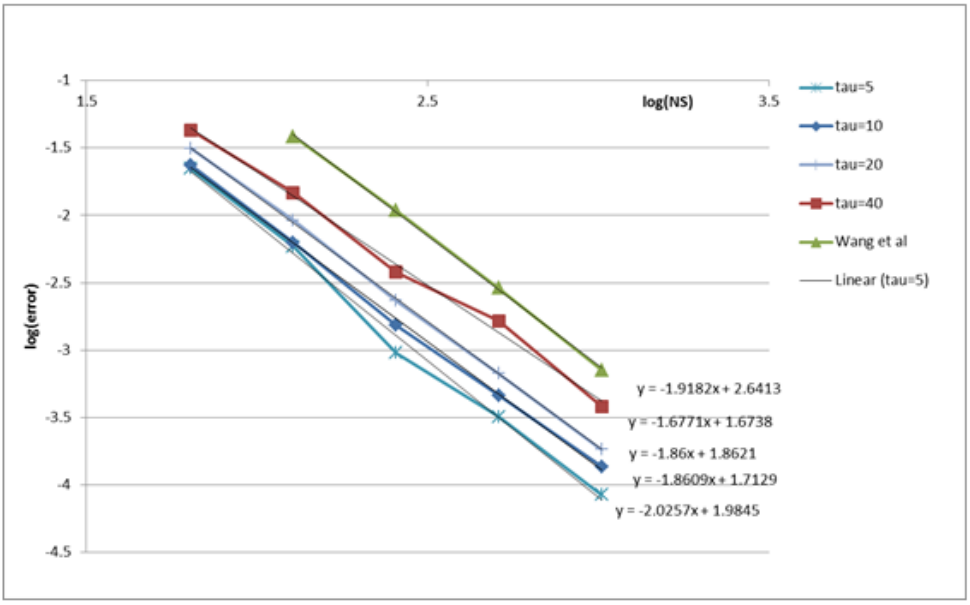


Figure 5-4 shows the strike price errors for the case of CGMY process for an American option, against the mesh size, together with the best straight line fits, showing the rate of convergence for the IT splitting algorithm and a range of coordinate transformations.

Figure 5-5 Error Convergence PSOR ($S = 90, \theta_C = 0.5$)

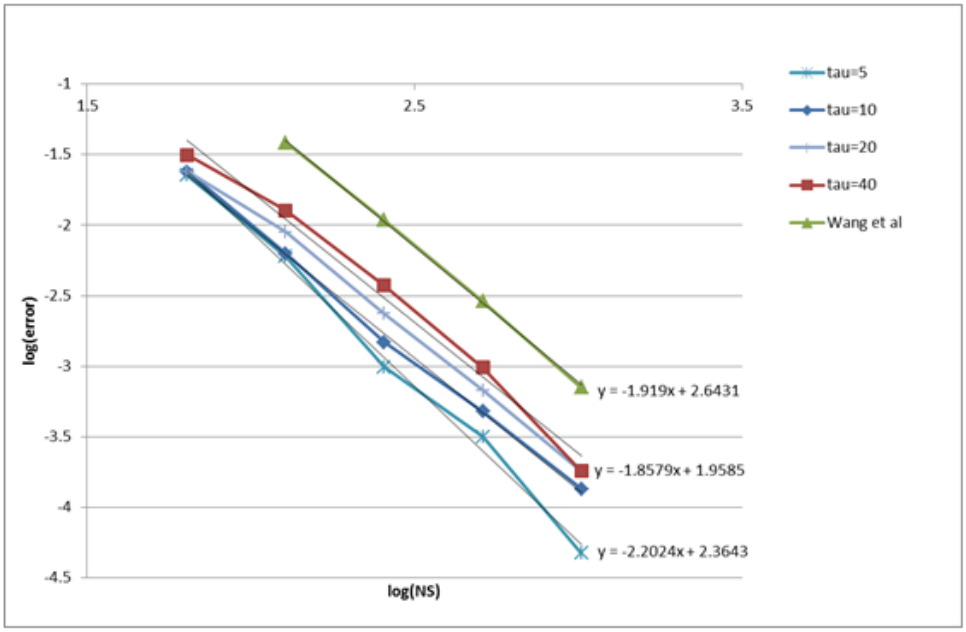


Figure 5-5 shows the strike price errors again for the CGMY American case against the mesh size, together with best straight line fits, showing the rate of convergence, for the PSOR algorithm and a range of coordinate transformations.

Figure 5-6 American CGMY Gamma for PSOR ($NS = 1024, tau = 5$)

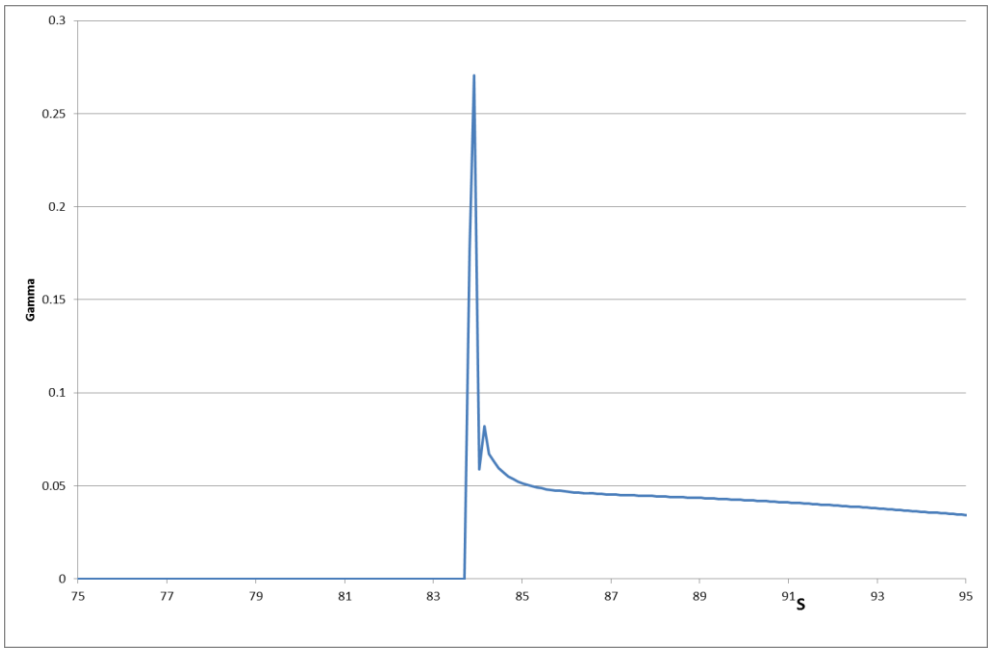
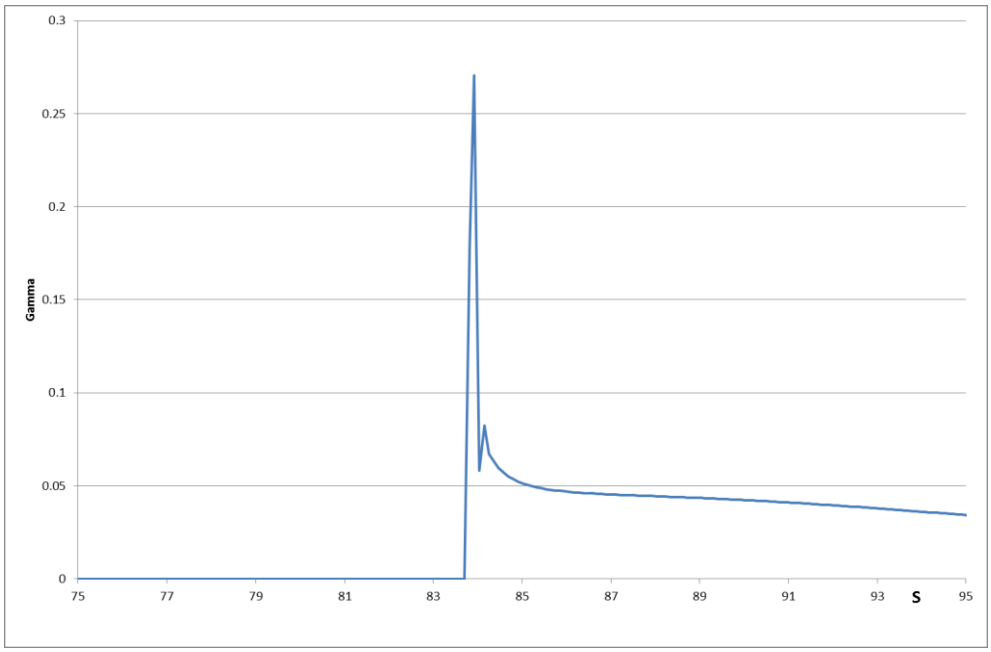


Figure 5-7 American CGMY Gamma for IT splitting ($NS = 1024, tau = 5$)



Merton's comparison

Using the Merton dataset (identical to the one used in Section 3.3):

$$S = 100, K = 100, T = 0.25, r = 0.05, \sigma = 0.15, \lambda = 0{,}1, a = -0{,}9, b = 0.45$$

with the exact value equal to 3.1490.

For this dataset (d' Halluin, 2005) use S_{max}=1000, an unspecified mesh and variable timesteps. Results for the American option comparisons are given below for $\theta_C = 0.5$. The overall picture is similar to European case with second order rates of convergence for $tau = 5{,}10$ and lower rates but for much smaller strike errors for $tau = 20{,}40$. The IT splitting method as implemented produced similar convergence rates and accuracies when compared to PSOR, though with more variation.

Figure 5-8 American Merton strike error ($\theta_C = 0.5$, IT splitting)

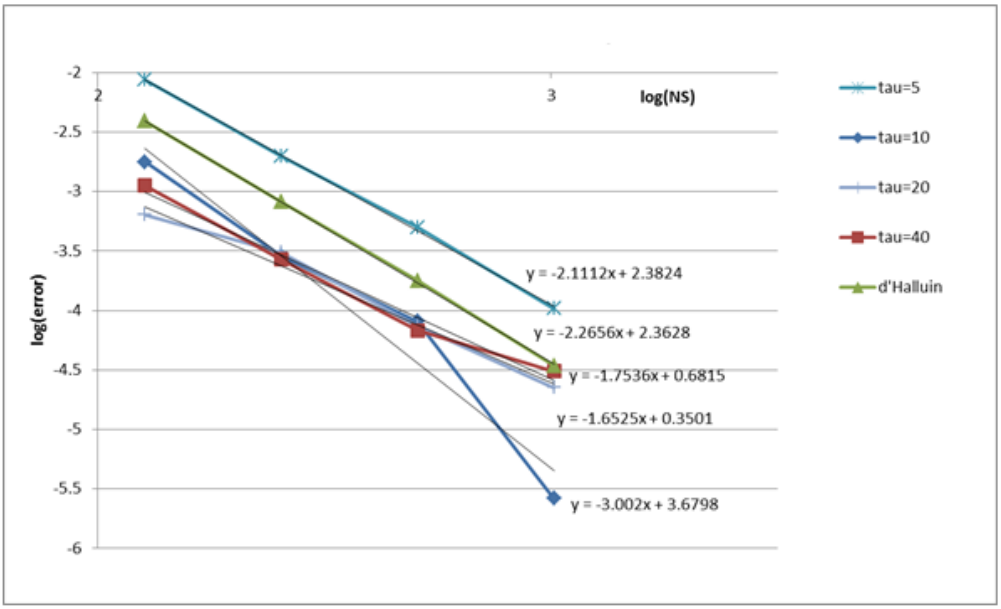


Figure 5-8 shows the strike error for an American put under Merton using IT splitting with a range of coordinate transformed mesh refinements and comparing to (d' Halluin, 2005), while Figure 5-9 shows the equivalent strike errors using PSOR.

Figure 5-9 American Merton strike error ($\theta_C = 0.5$, IT PSOR)

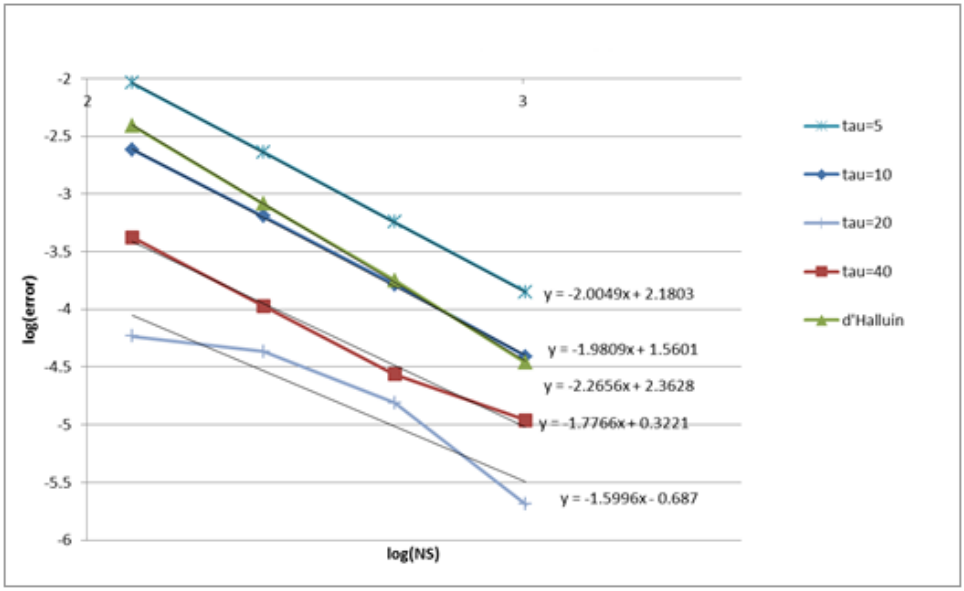


Table 5-5 Strike error comparison (Merton's American put)

Discretisation		IT Splitting				PSOR				
NS	M	tau=5	tau=10	tau=20	tau=40	M	tau=5	tau=10	tau=20	tau=40
127	160	3.23257	3.23947	3.24188	3.24237	40	3.23212	3.23883	3.24118	3.24166
254	320	3.23928	3.24096	3.24155	3.24151	80	3.23895	3.24061	3.2412	3.24135
508	640	3.24074	3.24116	3.24132	3.24135	160	3.24067	3.24108	3.24123	3.24127
1016	1280	3.24114	3.24125	3.24128	3.24129	320	3.2411	3.2412	3.24124	3.24123

Table 5-5 presents a comparison for the strike errors between IT splitting and PSOR for the case an American put option with the underlying following Merton's jump-diffusion process. Note the different number of timesteps for PSOR.

The results for IT splitting are comparable to PSOR. However, it was observed that oscillations in Gamma sometimes occurred when using the same number of timesteps as PSOR. These oscillations were removed by increasing the number of timesteps to those noted in the table above. The PSOR results for Gamma were entirely free of oscillation. The comparison suggests that the PSOR results could be more accurate (at this rate of precision) with a greater number of timesteps and the use of adaptive timestepping would be beneficial to both these issues.

Figure 5-10 Gamma function for Merton's American Put ($NS = 1016$, IT splitting)

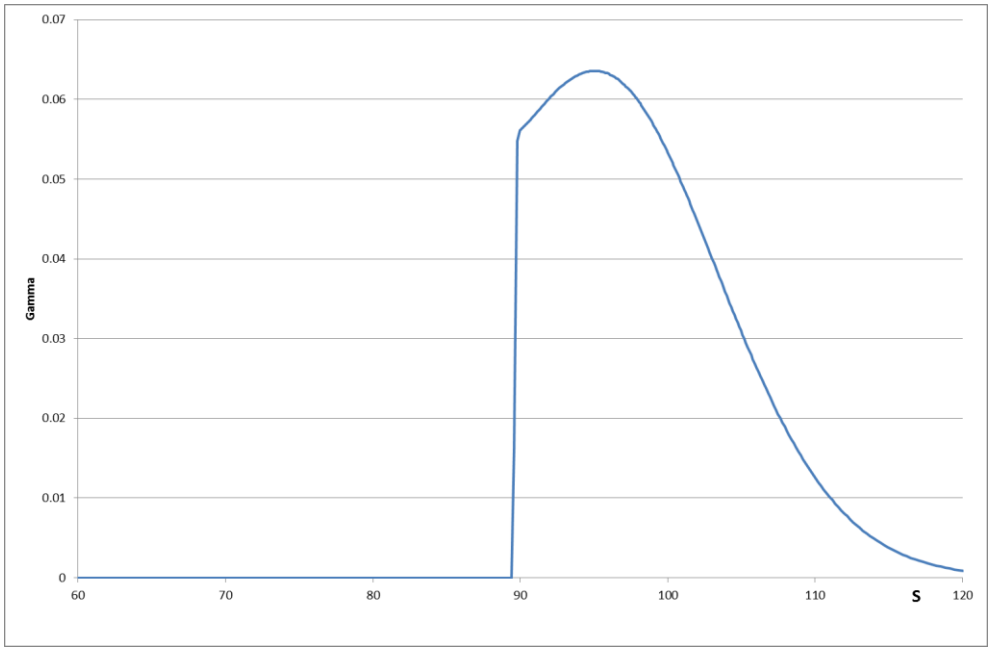
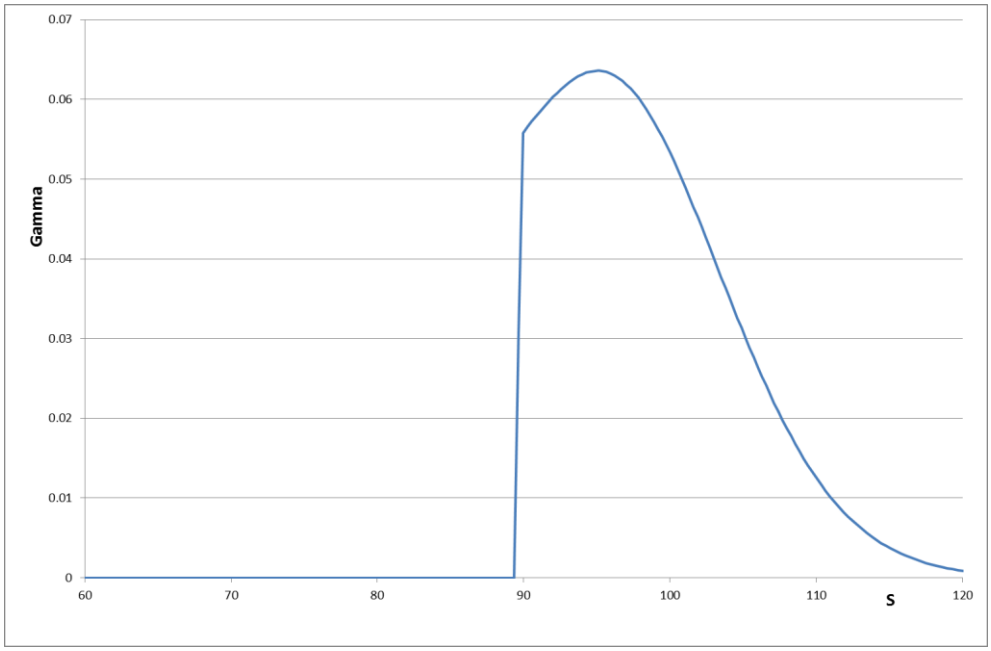


Figure 5-11 Gamma function for Merton's American Put ($NS = 1016$, PSOR)



In this chapter we investigated the performance of IT splitting, when coupled with the coordinate stretching and the compact mesh quadrature for the integral term.

Concluding remarks

1. Ikonen-Toivanen splitting adapted to the jump-diffusion linear complementarity problem works well with Merton with highly comparable results and convergence rates using mesh based quadrature and a range of coordinate transformation stretch values.
2. Ikonen-Toivanen splitting adapted to the jump-diffusion linear complementarity problem works well with Variance Gamma and CGMY, where the small jumps are replaced by a singular volatility model, are almost identical to PSOR when used with a range of coordinate stretching parameters and the mesh-based quadrature.
3. The results for Variance Gamma and CGMY would benefit from some additional second order numerical damping to remove the oscillation in Gamma near the exercise boundary. LU factorization with IT splitting is far more efficient as a solution technique compared to PSOR given the loss of the M-matrix property.
4. It would be interesting to test a version of volatility replacement for small jumps with the Merton model.

6. Conclusion

In this work we investigated the numerical solution of a European and American option pricing problem where the underlying is following a jump-diffusion process. We presented the numerical approximation focusing on the Merton's classical model, but also extended to the singular cases of VG and CGMY.

We presented the extension of coordinate stretching transformation (Parrott, 1999) to the jump-diffusion case and in combination with compact mesh based quadrature approach to the integral term we investigated the effectiveness of the approximation. An implicit time discretization was used and early exercise was treated by classical iterative projection approach. IT splitting (Toivanen, 2004) was also implemented for the American option case and its performance was compared against the PSOR approach.

The approximation was tested on the classical Merton's model as well as the singular cases of Variance Gamma and CGMY. The quadrature formulation was proved to be effective and with good localization properties and results to second-order accurate prices.

7. Appendix

7.1 Self-financing portfolio

Definition 7.15 (Self-financing portfolio) Let a portfolio π that consists of shares and bonds. $a_{n,i}$ are the units of each share n, $n = 0,1, \dots, N$ at time i contained in the portfolio. The total value of π at time i will be:

$$V_i = \sum_{n=0}^{N} a_{n,i} X_{n,i} \quad (7.1)$$

Let the portfolio be updated at time i. At that time we know only the prices of time $i-1$. This means that the random variables $a_{n,i}$ are F_{i-1} measurable. In other words, all the information that we need to estimate the values that $a_{n,i}$ take, are contained in the σ-algebra F_{i-1}. F_{i-1} contains all the information that can be derived from the evolution of the stochastic process $X_{n,i}$ up to time $t = i-1$.

If the investor holds at time t, $a_{n,i}$ units of the share n, then the change in the price of that share in the time interval is $[i-1, i]$ is:

$$\Delta X_n^i = X_{n,i} - X_{n,i-1} \quad (7.2)$$

The profit that the investor makes due to that price change is:

$$G_i^n = a_{n,i} \Delta X_i^n \quad (7.3)$$

The overall profit from all the shares at time $t = i$ is:

$$G_i = \sum_{n=0}^{N} a_{n,i} \Delta X_i \quad (7.4)$$

The value of the portfolio can only change using cash gained from the changes in the value of the shares that already exist in the portfolio. Therefore there is no inflow or outflow of cash. Under this restriction the value of the portfolio at time t becomes:

$$V_t = V_0 + \sum_{i=0}^{t} G_i \quad (7.5)$$

This is called a self-financing portfolio.

From the above definition it is obvious that the value of the portfolio is a martingale transformation of the stochastic process $X_{n,i}$.

The value of the portfolio whether is self-financing or not can be derived from:

$$\sum_{n=0}^{N} \Delta a_{n,i} X_{n,i-1} = 0 \; \forall i \quad (7.6)$$

The sum of the changes that occur in the units of the shares should balance as the previous equation shows.

7.2 Integral term

$$I_j\big(V(S,t^m)\big) = \sum_{p=0}^{p=N_Q-1} \int_{y_p}^{y_{p+1}} F_j^m(\mathrm{y})\tilde{v}(\mathrm{y})\, dy \quad (7.7)$$

$$\approx I_j^m = \sum_{p=0}^{p=N_Q-1} \frac{1}{2}\big(y_{p+1}-y_p\big)\big[F_j^m\big(y_{p+1}\big)\tilde{v}\big(y_{p+1}\big) + F_j^m\big(y_p\big)\tilde{v}\big(y_p\big)\big]$$

Setting $a_p = F_j^m\big(y_p\big)\tilde{v}\big(y_p\big)$, so that:

$$\begin{aligned} 2I_j^m = {} & (y_1 - y_0)[a_1 + a_0] + (y_2 - y_1)[a_2 + a_1] \\ & + \cdots \big(y_j - y_{j-1}\big)\big[a_j + a_{j-1}\big] + \big(y_{j+1} - y_j\big)\big[a_{j+1} + a_j\big] \\ & + \cdots (y_{\mathrm{N}} - y_{N-1})[a_N + a_{N-1}] \end{aligned} \quad (7.8)$$

with:

$$2I_j^m = y_1a_1 - y_0a_1 + y_1a_0 - y_0a_0$$

$$+y_2a_2 - y_1a_2 + y_2a_1 - y_1a_1 + \cdots$$

$$+y_{\mathrm{j}-1}a_{j-1} - y_{j-2}a_{j-1} + y_{\mathrm{j}-1}a_{j-2} - y_{\mathrm{j}-2}a_{j-2}$$

$$+y_{\mathrm{j}}a_j - y_{j-1}a_j + y_{\mathrm{j}}a_{j-1} - y_{\mathrm{j}-1}a_{j-1}$$

$$+y_{\mathrm{j}+1}a_{j+1} - y_ja_{j+1} + y_{\mathrm{j}+1}a_j - y_{\mathrm{j}}a_j$$

$$+y_{j+2}a_{j+2} - y_{j+1}a_{j+2} + y_{j+2}a_{j+1} - y_{j+1}a_{j+1}$$

$$\ldots + y_{N-1}a_{N-1} - y_{N-2}a_{N-1} + y_{N-1}a_{N-2} - y_{N-2}a_{N-2}$$

$$+y_N a_N - y_{N-1}a_N + y_N a_{N-1} - y_{N-1}a_{N-1} \quad (7.9)$$

Cancelling and combining terms, we get:

$$2I_j^m = +y_1 a_0 - y_0 a_0$$

$$+y_2 a_1 - y_0 a_1 + \cdots$$

$$+y_{j-1}a_{j-1} - y_{j-2}a_{j-1} + y_{j-1}a_{j-2} - y_{j-2}a_{j-2}$$

$$+y_j a_j - y_{j-1}a_j + y_j a_{j-1} - y_{j-1}a_{j-1}$$

$$+y_{j+1}a_{j+1} - y_j a_{j+1} + y_{j+1}a_j - y_j a_j$$

$$+y_{j+2}a_{j+2} - y_{j+1}a_{j+2} + y_{j+2}a_{j+1} - y_{j+1}a_{j+1}$$

$$\ldots + y_{N-1}a_{N-1} - y_{N-2}a_{N-1} + y_{N-1}a_{N-2} - y_{N-2}a_{N-2}$$

$$+y_N a_N - y_{N-1}a_N + y_N a_{N-1} - y_{N-1}a_{N-1}$$

$$+(y_2 - y_1)[a_2 + a_1] + \cdots + (y_N - y_{N-1})[a_N + a_{N-1}] \quad (7.10)$$

$$= \sum_{p=0}^{p=N_Q} \omega_p F_j^m(y_p) \quad (7.11)$$

7.3 Details of singular integrations for Variance Gamma

Variance Gamma

$$v(u) = C\frac{e^{-M\,u}}{u^{1+Y}}1_{u>0} + C\frac{e^{-G|u|}}{|u|^{1+Y}}1_{u<0} \;\; where \;\; u = \ln y \quad (7.12)$$

With fractional jumps:

$$\Delta u_j^D = \frac{\Delta S_j}{S_j}, \qquad \Delta u_j^U = \frac{\Delta S_{j+1}}{S_j} \quad (7.13)$$

We can write the singular volatility:

$$\sigma_{sj}^2 \approx \int_0^{\Delta u_j^D} [(e^u - 1)]^2 \frac{Ce^{-Gu}}{u^{1+Y}} du + \int_0^{\Delta u_j^U} [(e^u - 1)]^2 \frac{Ce^{-Mu}}{u^{1+Y}} du \quad (7.14)$$

$$= \int_0^{\Delta u_j^D} Ce^{-Gu} u^{1-Y} \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \dots\right]^2 du + \int_0^{\Delta u_j^U} Ce^{-Mu} u^{1-Y} \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \dots\right]^2 du$$

For VG we know that $Y = 0, G = \lambda_n$ so we can write (7.14) as:

$$\sigma_{sj}^2 = \int_0^{\Delta u_j^D} Ce^{-Gu} u \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \dots\right]^2 du + \int_0^{\Delta u_j^U} Ce^{-Mu} u \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \dots\right]^2 du \quad (7.15)$$

Now,

$$\int_0^{\Delta u_j^D} Ce^{-Gu} u \left[1 + \frac{u}{2!} + \frac{u^2}{3!} \dots\right]^2 du = \int_0^{\Delta u_j^D} Ce^{-Gu} u \left[1 + u + \frac{7}{12} u^2 \dots\right] du \quad (7.16)$$

We can write the first term as:

$$\int_0^{\Delta u_j^D} Ce^{-Gu} u \, du = \frac{C}{-G} [ue^{-Gu}]_0^{\Delta u_j^D} - \frac{C}{-G} \int_0^{\Delta u_j^D} e^{-Gu} du \quad (7.17)$$

$$= \frac{C}{-G} \left(\Delta u_j^D e^{-G\Delta u_j^D}\right) - \frac{C}{G^2} \left(e^{-G\Delta u_j^D} - 1\right) = \frac{C}{G^2} \left(1 - e^{-G\Delta u_j^D} - G\Delta u_j^D e^{-G\Delta u_j^D}\right)$$

$$= \frac{C}{G^2}(1 - e^{-a} - ae^{-a}) = \frac{C}{G^2}\left(\frac{a^2}{2} - \frac{a^3}{3} + \cdots\right)$$

where $a = G\Delta u_j^D$. The second term can be written:

$$\int_0^{\Delta u_j^D} Ce^{-Gu}u^2\,du = \frac{C}{-G}[u^2e^{-Gu}]_0^{\Delta u_j^D} + \frac{2C}{G}\int_0^{\Delta u_j^D} ue^{-Gu}\,du \qquad (7.18)$$

$$= \frac{C}{-G}\left(\left(\Delta u_j^D\right)^2 e^{-G\Delta u_j^D}\right) + \frac{2C}{G^3}\left(1 - e^{-G\Delta u_j^D} - G\Delta u_j^D e^{-G\Delta u_j^D}\right)$$

$$= \frac{C}{G^3}\left(-G^2\left(\Delta u_j^D\right)^2 e^{-G\Delta u_j^D}\right) + \frac{C}{G^3}\left(2 - 2e^{-G\Delta u_j^D} - 2G\Delta u_j^D e^{-G\Delta u_j^D}\right)$$

$$= \frac{2C}{G^3}\left(1 - e^{-G\Delta u_j^D} - G\Delta u_j^D e^{-G\Delta u_j^D} - \frac{1}{2}G^2\left(\Delta u_j^D\right)^2 e^{-G\Delta u_j^D}\right)$$

$$= \frac{2C}{G^3}\left(1 - e^{-a} - ae^{-a} - \frac{a^2}{2}e^{-a}\right) = \frac{2C}{G^3}\left(1 - e^{-a} - ae^{-a} - \frac{a^2}{2}e^{-a}\right) = \frac{2C}{G^3}\left(\frac{a^3}{6} + \cdots\right)$$

Adding the two terms, we get:

$$\frac{C}{G^2}\left(\frac{a^2}{2} - \frac{a^3}{3} + \cdots\right) + \frac{2C}{G^3}\left(\frac{a^3}{6} + \cdots\right) \qquad (7.19)$$

The terms can be seen to scale in accordance to $\left(\frac{a}{G}\right)^2$ and $\left(\frac{a}{G}\right)^3$ assuming that $\left(\frac{a}{G}\right) = \Delta u_j^D \ll 1$ the terms can be usefully truncated. Hence, we can define the successive approximations:

$$\left[\sigma_{sj}^2\right]^{(1)} = \frac{C}{G^2}(1 - e^{-a} - ae^{-a}) + \frac{C}{M^2}\left(1 - e^{-b} - be^{-b}\right) \qquad (7.20)$$

$$\left[\sigma_{sj}^2\right]^{(2)} = \frac{C}{G^2}(1 - e^{-a} - ae^{-a}) + \frac{2C}{G^3}\left(1 - e^{-a} - ae^{-a} - \frac{a^2}{2}e^{-a}\right) \qquad (7.21)$$

$$+\frac{C}{M^2}\left(1-e^{-b}-be^{-b}\right)+\frac{2C}{M^3}\left(1-e^{-b}-be^{-b}-\frac{b^2}{2}e^{-b}\right)$$

$$a=G\Delta u_j^D, b=M\Delta u_j^U$$

Note that:

$$1-e^{-a}-ae^{-a}=1-\left(1-a+\frac{a^2}{2}-\frac{a^3}{3!}+\cdots\right)-a\left(1-a+\frac{a^2}{2}-\frac{a^3}{3!}+\cdots\right) \qquad (7.22)$$

$$1-e^{-a}-ae^{-a}-\frac{a^2}{2}e^{-a}=\frac{a^2}{2}-\frac{a^3}{3}+\cdots-\frac{a^2}{2}\left(1-a+\frac{a^2}{2}-\frac{a^3}{3!}+\cdots\right)=\frac{a^3}{6}+\cdots \qquad (7.23)$$

Works Cited

Andreasen, J. (2000). Andersen, L. Jump-diffusoin models: Volatility smile fitting and numerical methods for pricing. *Review of Derivative Research*, 4:231-262.

Barles, G. (1997). Convergence of numerical schemes for degenerate parabolic equations arising in finance. *Numerical methods in Finance*.

Black, M. (1973). Scholes, F. The pricing of options and corporate liabilities. *Journal of Political Economy*.

Briani, M. (2004). Chioma, C. L., Natalini, R. Convergence of numerical schemes for viscosity solutions to integro-differential degenerate parabolic problems arising in financial theory. *Numerische Mathematik,98:607-646*.

Carr, P. (2007). Mayo, A. On the numerical evaluation of option prices in jump - diffusion processes. *The European Journal of Finance,13:353-372*.

Cont, R. (2004). *Tankov, P. Financial modelling with jump processes.* Chapman & Hall.

d' Halluin, Y. (2004). Labahn, G. , Forsyth, P. A. A penalty method for american options with jump diffusion processes. *Numerische Mathematik*, 92:321-352.

d' Halluin, Y. (2005). Vetzal, K. R. , Forsyth, P. A. Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis*, 25:87-112.

Dupire, B. (1994). Pricing with a smile. *Risk Magazine,* http://www.risk.net/data/risk/pdf/technical/2007/risk20_0707_technical_volatility.pd.

Dupire, B. (2000). Singleton, K. , Pan, J. Transform analysis and option pricing for affine jump-diffusions. *Econometrica*, 68:1343-1376.

Garroni, M. (2001). Menaldi, J. Raton, B. Second order elliptic integro-differential problems. *CRC Press, 25:567-585*.

Giannakopoulos, A. N. (2003). Stochastic calculus and application in finance.

Heston, S. (1993a). A closed-form solution for options with stochastic volatility with application to bond and currency options. *The Review of Financial Studies*,6:327-343.

Hull, J. (1987). White, A. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42:281-300.

Hull, J. (1989). *Options, Futures and other derivative securities.* Prentice-Hall.

Kou, S. (2004). Wang, H. Option pricing under a double exponential jump-diffusion model. *Journal of Management science*, 50:1178-1192.

Madan, D. B. (1998). Chang, E. C., Carr, P. The variance Gamma process and option pricing. *European Finance Review*, 2:79-105.

Madan, D. B. (1999). Carr, P. Option pricing and the fast fourier transform. *Journal of Computational Finance*, 2:61-73.

Madan, D. B. (2002). Hirsa, A. Pricing american options under variance Gamma. *Journal of Computational Finance,7*.

Madan, D. B. (2002). Yor, M. , Carr, P. , Geman, H. The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75:305-332.

Marcozzi, M. (2003). On the valuation of options in jump-diffusion models by variational methods. *SIAM Journal of sientific computing*, 24:1124-1140.

Merton, R. C. (1975). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125-144.

Merton, R. C. (1975). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125-144.

Oosterlee, C. W. (2006). Almendral, A. On American options under the Variance Gamma process. *http://folk.uio.no/ariel/*.

Oosterlee, C. W. (2007). Almendral, A. Accurate evaluation of European and American options under the CGMY process. *SIAM Journal on Scientific Computing,29:93-117*.

Pananicolaou, G. (2000). Fouque, J. , Sircar, K. R. Derivatives in financial markets with stochastic volatilities. *Cambridge University Press*.

Papanicolaou, G. (2000). *Sircar, K. R., Fouque, J. P. Derivatives in financial markets with stochastic volatility.* Cambridge University Press.

Papantoleon, A. (2000). An introduction to Lévy processes with application in finance, University of Freiburg.

Parrott, K. (1999). Clark, N. Multigrid American option pricing with stochastic volatility. *Journal of Applied Mathematical Finance*, 6:177-195.

Sato, K. (1999). Lévy Processes and Infnitely Divisible Distributions. *Cambridge University Press*.

Schoutens, W. (2002). Leuven, K. U. Meixner Process: Theory and applications in finance, http://perswww.kuleuven.be/~u0009713/004wsreport.pdf.

Souganidis, P. E. (1991). Barles, G. Convergence of approximation schemes for fully non linear equations. *Asymptotic Analysis*.

Summers, L. (1989). Cutler D., Poterba J. What drives stock prices. *Journal of Portfolio Management*,15: 4-12.

Toivanen, J. (2004). Ikonen, S. Operator splitting methods for pricing american options with stochastic volatility. *Reports of the Department of Mathematical Information Technology*.

Toivanen, J. (2014). Salmi, S. IMEX schemes for pricing options under jump - diffusion. *Applied numerical mathematics,84:33-45*.

Voltchkova, E. (2005a). Cont, R. Integro-differential equations for option pricing in exponential Lévy models. *Finance Stochastic*, 9:299-325.

Voltchkova, E. (2005b). Cont, R. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM Journal of Numerical Analysis*, 43:1596-1626.

Voltchkova, E. (2006). Tankov, P. Jump-diffusion models: A practitioner's guide.

Wang, I. R. (2007). Forsyth, P. A., Wan, J. W. L. Robust numerical valuation of european and american options under CGMY process. *www.cs.uwaterloo.ca/ paforsyt/*.

Zhang, A. (2007). A secret to create a complete market from an incomplete market. *Applied Mathematics and Computation*, 191:250-262.