PDE Option Pricing: Analysis and Application to Stochastic Correlation

by

Nat Chun-Ho Leung

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Computer Science
University of Toronto

Abstract

PDE Option Pricing: Analysis and Application to Stochastic Correlation

Nat Chun-Ho Leung
Doctor of Philosophy
Graduate Department of Computer Science
University of Toronto
2017

This thesis is a study of numerical Partial Differential Equation (PDE) methods in financial derivatives pricing. The first part of the thesis is concerned with the behaviour of a numerical PDE solution when the initial condition is not smooth. The second part of the thesis develops computational PDE methods for option pricing problems with stochastic correlation.

In the first part of this thesis, we provide an analysis of the error arising from a non-smooth initial condition when solving a pricing problem modelled by a parabolic PDE with a finite difference method. We build our framework on the sharp error estimate in [25], and study three types of non-smoothness that are of financial interest. Wheareas the framework in [25] focuses on the effect of Rannacher timestepping, we utilize their techniques to study the numerical error with focus on the error due to spatial non-smoothness of the initial condition. We show that the error of the numerical solution under Crank-Nicolson-Rannacher timestepping with central spatial differences can be decomposed into two components. The first component is a second order discretization error primarily resulting from the approximation to the heat kernel by a discrete operator. The second component is a *quantization error* that depends on the relative position of the point of non-smoothness on the grid. We obtain explicit expressions of the two kinds of errors. From this viewpoint, we discuss how mesh positioning relative to the point of non-smoothness of the initial condition affects the quality of the numerical solution, and the possibility of an optimal positioning of the point of non-smoothness. We also study explicitly the effect of smoothing on the error of the numerical solution.

The second part of the thesis focuses on the pricing of European options using a stochastic correlation model. We derive a time-dependent PDE for the pricing problem under stochastic

correlation, with the correlation variable giving rise to an extra dimension, and develop computational approaches for its solution. The first approach we develop is a finite difference scheme. We study the effect of localization of the domain, formulate appropriate boundary conditions, and study discretization schemes and stability in the l^∞-norm. Of particular interest is the formulation of boundary conditions in the correlation dimension. The second approach we develop is an asymptotic solution of the PDE, appropriate for cases when the correlation process exhibits fast mean reversion and when a numerical PDE solution is considered costly. Numerical experiments demonstrate the effectiveness of our methods, and the agreement among the two solutions and Monte Carlo simulations results. We also experimentally study the effect of smoothing on the quality of the numerical solution, as an application of the work presented in the first part of the thesis. We verify second order convergence for the price and derivatives of various types of options. Furthermore, we present numerical results that illustrate the effect of certain problem parameters on the approximate solution. For these results, we also give a mathematical explanation based on the asymptotic solution.

This thesis is dedicated to my family, and Ms. Poon Suet Fan.

Acknowledgements

First and foremost, I would like to express my most sincere gratitude to my PhD supervisor Professor Christina Christara, whose passion and professionalism have had a huge impact on my work. I have benefited tremendously from her insights and immense knowledge. Outside of our academic discussion, Christina has also provided very useful advice and support during my difficult times. I am fortunate to have her as my supervisor. Without her, this thesis would not have been possible. More importantly, she has provided both the necessary training and freedom for me to become an independent researcher. This is a skill that, I believe, will prove useful beyond this thesis.

I would like to thank my committee members, Professor Kenneth Jackson and Professor Wayne Enright, whose insightful and invaluable inputs have polished and strengthened this thesis. I would also like to thank University of Waterloo faculty Professor Peter Forsyth for agreeing to be my external examiner, and University of Toronto faculty Professor Thomas Fairgrieve for joining my final oral examination. This thesis has benefited enormously from their helpful comments and suggestions.

My gratitude also goes to my MMath supervisors Professor Justin Wan and Professor Spiro Karigiannis in University of Waterloo. My training with Spiro has been invaluable and integral to my later academic pursuits. I am grateful to Justin for his inspirational guidance, and for introducing me to the wonderful field of Scientific Computing, which has since changed the course of my career.

My gratitude extends to my collaborators, Professor Serge D'Alessio and Professor Duy-Minh Dang. The intellectual conversations we had were both enjoyable and inspiring. I am grateful to Dr. Denglin Zhou, Dr. Anthony Vaz and Dr. Henry Li for giving me the chance to work with them on interesting projects at Manulife Financial.

Throughout my PhD studies, I have enjoyed support from many friends and colleagues. I would like to thank Chris Wong, Vincent Yip, Eric Kwan, Lorraine Chan, Ivy Yip, Kalen Chan, Roki Lam, Dylan Kopechanski, Adam Ng, Xuezhi Liu, Nathan Zhou, Shantell Lo, Clement Kwok, Stella Lin, Vida Heidarpour, Michael Chiu, Ken Chan, among others, for their enduring friendship and the fun we have had. Special thanks to my girlfriend, Chui-Man Yip, for her love and encouragement.

Finally, I would like to thank my parents Chiu-Wah Leung and Kam-Tim Leung, my sister Wing-Suet Leung and my brother Chun-Fung Leung for their sacrifice, unconditional love, and faith in me. I am also indebted to my high school Mathematics teacher, Ms. Poon Suet Fan, whose work ethics and uncompromising integrity have had a strong and long lasting influence on me.

Contents

List of Tables

List of Figures

List of Symbols

General Notations

$\mathbb{Q}$ A particular choice of a risk-neutral probability measure

μ_S Drift of asset price

σ_{S_j} Volatility of the j-th asset

σ_S Volatility of asset price

τ Backward time $\tau = T - t$

$B(\cdot)$ Brownian motion

$B_j(\cdot)$ The j-th component of a multidimensional, possibly correlated Brownian motion

$g(\cdot)$ Payoff function of a contingent claim

K Strike of a call/put/forward

q Dividend yield of an asset

r Risk-free rate in a constant interest rate world

S Asset price as a state variable

$S(\cdot)$ Asset price process

S_j Price of the j-th asset as a state variable

$S_j(\cdot)$ Process of the j-th asset

T Maturity time of a claim

t Forward time

$V(\cdot)$ Value of a contingent claim as a function of time and state variables

$W_j(\cdot)$ The j-th component of a multidimensional independent Brownian motion adapted to the filtration generated by $B_j(\cdot)$

Chapter 1

$p(T, X(T) \mid t, X(t))$ Transition density from $(t, X(t))$ to $(T, X(T))$

η_r Mean reversion level of the short rate process

η_v Mean reversion level of variance in Heston model

$\hat{\Omega}$ Range space of risk factors

$\Lambda_v(\cdot)$ Market price of variance risk

λ_r Mean reversion rate of short rate

λ_v Mean reversion rate of variance in Heston model

ρ Instantaneous correlation between two correlated Brownian motions

σ_r Volatility of the short-rate process

σ_v Volatility of variance in Heston model

$F(t)$ Time T Forward price of an asset as determined at time t

$P(t,T)$ Time t price of a risk-free zero-coupon bond maturing at time T

$r(\cdot)$ Short rate process in a stochastic interest rate model

$X(\cdot)$ Process of risk factors including asset prices

Chapter 2

α Relative position of the point of non-smoothness in the grid

δ As subscript, Dirac-delta function

$\delta(\cdot)$ Dirac-delta function

$\hat{f}$ Discrete-time Fourier transform of a function f

κ Generic variable in frequency space with scaling $\kappa = \frac{\theta}{h}$

λ The ratio $\frac{k}{h}$

$\mathcal{H}(\cdot)$ Heaviside function

Φ_μ Smoothing operator of order μ

σ Volatility of asset

θ Generic variable in frequency space

$\tilde{f}$ Continuous Fourier transform of a function f

Υ Rate of convergence estimated based on successive errors

ζ_1, ζ_2 Constant purely imaginary complex numbers

a Speed of convection

C (As subscript, non-exponential) call

d The ratio $\frac{k}{h^2}$

$E^{(D)}_{(\cdot)}$ Numerical error from the discretization operator

$E^{(Q)}_{(\cdot)}$ Quantization error

EC As subscript, exponential call

EP As subscript, exponential put

$G(\cdot,\cdot)$ Green's function for the model problem

H As subscript, Heaviside function

h Spatial stepsize

i Canonical choice of complex number such that $i^2 = -1$

k Time stepsize

m The ratio $\frac{1}{k}$

P (As subscript, non-exponential) put

p_μ Polynomial corresponding to smoothing operator of order μ

R Number of Rannacher timesteps, assumed to be 2

x State variable in log price space

x^* Point of evaluation of both numerical and exact solutions

z Generic variable on the complex plane

Chapters 3-5

$\alpha(\cdot)$ Drift function of correlation process

$\bar{\rho}$ Mean of ρ with respect to the invariant distribution

$\beta(\cdot)$ Volatility function of correlation process

$\eta, \hat{\eta}$ Mean reversion level of correlation

$\Lambda(\cdot)$ Market price of correlation risk

$\lambda, \hat{\lambda}$ Mean reversion rate of correlation

$\mathcal{L}$ Spatial part of the pricing PDE in original price variables

$\mathcal{L}_{\log}$ Spatial part of the pricing PDE in log price variables

$\mu_R(\cdot)$ Drift function of the exchange rate

Φ Invariant distribution of the correlation process

$\pm R_{\log}$ Boundary of the log price variables in the localized problem

$\pm S_j^{\max}$ Boundary of the j-th price variable in the localized problem

ρ Correlation as a state variable

$\rho(\cdot)$ Correlation process

$\sigma_\rho, \hat{\sigma}_\rho$ A scalar that determines volatility of correlation

σ_R Volatility of the exchange rate

θ Parameter that controls the degree of implicitness in the θ-timestepping

Ξ Rate of convergence estimated based on successive changes in value

$a_j(\cdot)$ Number of shares in the j-th asset to hedge the claim

$b(\cdot)$ Number of shares to short in the money-market account to hedge the claim

h_1 Spatial step-size of $x = \log(S_1)$

h_2 Spatial step-size of $y = \log(S_2)$

h_3 Spatial step-size of ρ

n_1 Number of subintervals in the x direction

n_2 Number of subintervals in the y direction

n_3 Number of subintervals in the ρ direction

$p(\cdot), p_m^{(0)}(\cdot)$ Joint transition density of (S_1, S_2) with constant correlation η

$p_m(\cdot)$ Marginal transition density of (S_1, S_2)

$p_m^{\epsilon,1}(\cdot)$ Perturbed version of $p_m(\cdot)$

R Exchange rate as a state variable

$R(\cdot)$ Exchange rate process

r_d Domestic risk-free rate

r_f Foreign risk-free rate

$u_{\log}^{(l)}$	Vectorized numerical solution at the l-th timestep	x	State variable in $\log(S_1)$-space
		y	State variable in $\log(S_2)$-space

Chapter 1

Introduction

1.1 The pricing problem

This thesis focuses on the computational aspects of pricing problems. Classically, pricing seeks to determine the present value of a contingent claim in such a way that it is impossible to create arbitrage by trading the claim at this price and other available instruments in the market. In other words, it is impossible to start with zero capital and end up with almost surely non-negative investment gain and a profit with positive probability.

For some financial contracts, there is a simple way to determine their arbitrage-free values. One example is the agreement to purchase a stock at a predetermined price K at a maturity time T. It is not hard to show that (under constant interest and dividend rates) this value at time t is equal to

$$S(t)e^{-q(T-t)} - Ke^{-r(T-t)},$$

where $S(t)$ is the current price of the stock, q is the dividend rate over the time period and r is the risk-free rate. Similarly, under simplifying assumptions, there are model-free prices for interest rate swaps, government bonds, or currency forwards among others.

For some other derivatives, modelling is necessary to determine its "arbitrage-free" price. These include European options with a more complex payoff function or financial arrangements that are path-dependent. Some of these problems do not have a known closed-form expression of solution using elementary functions, and numerical methods become necessary. The approximate values as determined by this approach are model-dependent, but the models are calibrated in a way that best match the prices of available instruments publicly traded in the market.

Financial modelling is a rich and complex subject that involves advanced mathematics and probability theory. A branch of pricing theory uses machinery from the theory of partial dif-

ferential equations (PDEs) with great success. This thesis is an investigation into the computational and numerical aspects of this PDE approach.

1.2 Price as an expectation

In the discrete time case, it is the content of the fundamental theorem of asset pricing that arbitrage does not exist if and only if there is a "risk-neutral" measure, equivalent to the real-world (or "physical") measure, such that every tradeable asset has the risk-free rate r as the expected return. This risk-neutral measure is not necessarily unique. The generalization of this statement to continuous time processes is non-trivial and requires more advanced definition of the concept of "no-arbitrage" (see for example [16]).
In this framework, once a risk-neutral measure $\mathbb{Q}$ is fixed, and assuming prices are Markovian in the risk factors, the time-t price of a claim V, payable at time T in the amount of $g(X(T))$ is given by

$$V(t, X(t)) = \mathbf{E}_t^{\mathbb{Q}}\left[e^{-r(T-t)}g(X(T))\right],$$

where:

- $\mathbf{E}_t^{\mathbb{Q}}$ denotes the conditional expectation with respect to the filtration at time t, and
- $X(s)$ (where $0 \leq s \leq T$) is the relevant risk factors, possibly multi-dimensional.

A corollary is that if $g(X(T))$ is non-negative and is positive on a set with non-zero probability, then $V(t, X(t))$ has to be positive by properties of the expectation operator. This holds true for any constant linear combination of tradeable instruments as well. This can be seen to prevent arbitrage by "static" trades.
If a density function $p(T, X(T) \mid t, X(t))$ exists, we can write the price as follows:

$$V(t, X(t)) = e^{-r(T-t)} \int_{\hat{\Omega}} g(X(T))p(T, X(T) \mid t, X(t))dX(T),$$

where $\hat{\Omega}$ is the range space of $X(T)$.

1.3 Popular models

Brownian motions are popular building blocks for risk processes in financial models. In this section we outline a few such basic models that are built from them. These models are extensively studied and will be useful in the later parts of this thesis.

A first model for the stock price process is *Bachelier*'s model, which assumes that the T-forward price $F(t)$ at time t of an asset is governed by the stochastic differential equation (SDE) in the T-forward measure (risk-neutral measure with the zero coupon bond maturing at time T as the numéraire, see [51])

$$dF(t) = \sigma_F dB(t), \tag{1.1}$$

where $\big(B(t)\big)_{\{0\leq t\leq T\}}$ denotes the standard Brownian motion. In this model, assuming temporarily that the interest rate is constant, and dividend rate is zero, we have that $S(t) = e^{-r(T-t)}F(t)$ is the price of the underlying asset, and therefore this SDE is same as modelling the asset price at any future time by a normal distribution.

Consequently, in this model, asset prices have positive probability of being negative, which is an undesirable feature for some applications. On the other hand, the SDE (1.1) from Bachelier's model is extensively used as a quotation model in the swaption market, due to the (increasingly common) possibility that interest rates, or in this case the forward swap rate, can in fact be negative.
European option prices in this model are given by

$$V^{\text{Bachelier}}(t, F(t)) = \frac{P(t,T)}{\sqrt{2\pi\sigma_F^2(T-t)}} \int_{-\infty}^{\infty} g(F(T)) e^{-\frac{(F(T)-F(t))^2}{2\sigma_F^2(T-t)}} dF(T), \tag{1.2}$$

where $P(t,T)$ is the time-t price of a risk-free zero-coupon bond maturing at time T, and is equal to $e^{-r(T-t)}$ for constant interest rate r.

The celebrated Black-Scholes-Merton (BSM) model is given by the following SDE for the asset price, in the physical measure:

$$\frac{dS(t)}{S(t)} = \mu_S dt + \sigma_S dB(t). \tag{1.3}$$

This is also called the *geometric Brownian motion* (GBM). In this model, asset prices remain positive and their logarithms follow a normal distribution (at a fixed future time). From Itô's lemma, we have

$$d\log(S(t)) = (\mu_S - \frac{\sigma_S^2}{2})dt + \sigma_S dB(t). \tag{1.4}$$

Similarly in the PDE framework, this logarithmic transformation converts the Black-Scholes PDE (1.13) into a constant coefficient convection-diffusion equation.

The Black-Scholes-Merton model assumes a dynamics of asset prices in the physical measure as opposed to a risk-neutral measure. The expected return in $[0,T]$ under this measure is equal to $e^{\mu_S T}$ instead of e^{rT}. Therefore, one has to find an equivalent measure such that

the expected return is e^{rT}. In the BSM model one can always accomplish this by a "hedging" argument.

European option prices in this model for a non-dividend-paying asset can be written as

$$V^{\text{BS}}(t, S(t)) = \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_S^2(T-t)}} \int_{-\infty}^{\infty} g(e^x) e^{-\frac{(x-\log(S(0))-(r-\frac{\sigma_S^2}{2})(T-t))^2}{2\sigma_S^2(T-t)}} dx. \tag{1.5}$$

When there are two or more assets relevant to a particular problem, it is straightforward to use the GBM to model each asset price process separately, then combine with a correlation/covariance matrix. Mathematically, the price process of the j-th asset can be written as

$$\frac{dS_j(t)}{S_j(t)} = \mu_{S_j} dt + \sigma_{S_j} dB_j(t), \tag{1.6}$$

where $\{B_j(t)\}$ are correlated Brownian motions. One of the problems we study in this thesis is the computational challenges when correlation is also assumed to be stochastic.

The limitations of using (1.3) to describe asset price processes are well-documented in the literature. In particular, it does not account for the fact that European options traded in the market exhibit a non-flat implied volatility surface. Various models are designed to address this limitation.

Stochastic volatility models make an additional assumption that the instantaneous volatility is itself a stochastic process. The Heston model ([29]) is one of the most popular models favoured by researchers and practitioners. In this model, the asset price and its instantaneous variance are jointly specified:

$$\begin{aligned} \frac{dS(t)}{S(t)} &= \mu_S dt + \sqrt{v(t)} dB_1(t) \\ dv(t) &= \lambda_v(\eta_v - v(t)) + \sigma_v \sqrt{v(t)} dB_2(t) \\ dB_1(t) dB_2(t) &= \rho dt. \end{aligned} \tag{1.7}$$

Without parameter restrictions, there is no guarantee that $v(t)$ will stay positive. The Feller condition requires

$$2\lambda_v \eta_v > \sigma_v^2 \tag{1.8}$$

in order that the variance is strictly positive.

Stochastic volatility models alone are known to insufficiently describe the steepness of the volatility surface at short expiries. Jump diffusion models (see e.g. [43] or [35]), on the other hand, are able to capture this short-term skew given that the mean jump size is large enough. In this thesis, we will focus on diffusion-based models.

In the interest rate world, short rate models find their use in many pricing problems. In these models, the interest rate $r = r(t)$ is no longer constant and deterministic. Below we introduce two such models:

$$dr(t) = \lambda_r(\eta_r - r(t)) + \sigma_r dB(t) \quad \text{(Vasicek model)} \tag{1.9}$$

$$dr(t) = \lambda_r(\eta_r - r(t)) + \sigma_r \sqrt{r(t)} dB(t) \quad \text{(Cox-Ingersoll-Ross model)} \tag{1.10}$$

In these models, the short rate is directly modelled by SDEs. The discount process is no longer deterministic, and is instead given by $e^{-\int_0^t r(s)ds}$. In the Vasicek model, the short rate has a positive probability of going negative. In the CIR model, the short rate can stay positive under a parameter restriction similar to (1.8), and in this case the short rate process has a natural lower boundary at 0.

1.4 The PDE approach

Let $X(t)$ satisfy

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dB(t),$$

where $B(t)$ is a Brownian motion, and the functions $\mu(\cdot)$, $\sigma(\cdot)$ are globally Lipschitz in x and are of linear growth in x. Let $f(t,x) = \mathbf{E}_{t,x}\left[e^{\int_t^T c(s,X_s)ds} g(X(T))\right]$. Then the Feynman-Kac formula states that, if $f \in C^{1,2}([0,T) \times \mathbb{R})$ and g is of polynomial growth, we have

$$\frac{\partial f}{\partial t} + \frac{\sigma(t,x)^2}{2}\frac{\partial^2 f}{\partial x^2} + \mu(t,x)\frac{\partial f}{\partial x} + c(t,x)f = 0, \tag{1.11}$$

satisfying the terminal condition $f(T,x) = g(x)$. Moreover, the conditional expectation $f(t,x)$ is the only solution to the Cauchy problem with polynomial growth. The proof can be found in [34].

The significance of this formula is that, "prices" can also be interpreted as solutions to partial differential equations. This connection between a conditional expectation and a PDE opens the door to PDE approaches in pricing. Each of the stochastic models in Section 1.3 has a corresponding PDE interpretation. The integral representation of prices (1.2) or (1.5) can be regarded as the convolution of the terminal condition with the fundamental solution to the corresponding partial differential equation. Both the PDE and the probabilistic interpretation are popular and classical approaches to pricing financial instruments.

In our presentation of the Bachelier model (1.1), the state variable is the T-forward price $F(t)$, and the corresponding pricing PDE for the value function $V(t,F)$ (with $c(t,x) = -r$ in

(1.11)) is

$$\frac{\partial V}{\partial t} + \frac{\sigma_F^2}{2}\frac{\partial^2 V}{\partial F^2} - rV = 0. \tag{1.12}$$

In the Black-Scholes-Merton model (1.3), the asset price process is specified in the *physical* measure, in which its drift is μ_S. In a risk-neutral measure with constant risk-free rate r, its drift should be r. As prices are given by risk-neutral expectations, the corresponding PDE according to (1.11) is the familiar Black-Scholes PDE

$$\frac{\partial V}{\partial t} + \frac{\sigma_S^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{1.13}$$

As another example, the Heston model (1.7) has the following pricing PDE, according to a multi-dimensional version of (1.11):

$$\begin{aligned} &\frac{\partial V}{\partial t} + \frac{vS^2}{2}\frac{\partial^2 V}{\partial S^2} + \rho\sigma_v vS\frac{\partial^2 V}{\partial S \partial v} + \frac{\sigma_v^2 v}{2}\frac{\partial^2 V}{\partial v^2} \\ &+(\lambda_v(\eta_v - v(t)) - \Lambda_v(S,v,t))\frac{\partial V}{\partial v} + rS\frac{\partial V}{\partial S} - rV \;=\; 0. \end{aligned} \tag{1.14}$$

The variable $\Lambda_v(S,v,t)$ is the so-called *market price of variance risk*. It arises from the fact that there is more than one risk-neutral measure, intuitively because all we require is that $S(t)$ grows at the risk-free rate r, and no such requirement is imposed on $v(t)$ as it is not a tradeable asset. This gives us freedom in choosing its drift, and as a result each choice of $\Lambda(S,v,t)$ corresponds to one risk-neutral measure.

While the PDEs (1.12), (1.13) and (1.14) can be interpreted as an instance of (1.11), it is also possible to obtain them through a hedging argument.

In practice, the PDE approach is effective in handling options in low dimensions, even for options with path-dependency features. The pricing of American options can be posed as a linear complementarity problem, while that of Asian options can be formulated as a PDE with an auxiliary dimension. In addition, the calculation of sensitivities to risk factors is straightforward in the PDE method as solutions are obtained for the entire grid as opposed to a single point.

While for simple models such as the multi-dimensional geometric Brownian motion (1.6) the fundamental solution to the corresponding PDE has a known explicit expression using elementary functions, this is not necessarily the case for models with more complex structures. In addition, the approach of fundamental solution is not generally applicable to path-dependent options. In those cases, numerical methods become necessary. The investigation of numerical methods to such PDEs is the primary subject of this thesis.

1.5 Solution space, incompleteness and viscosity solutions

To develop a convergent numerical scheme, it is necessary to first identify the target *solution space*. Classical option pricing as described in Section 1.4 identifies the price of a European option as a conditional expectation, and if the process is generated by a diffusion, then the Feynman-Kac formula implies that it is also the solution to a linear PDE similar to (1.11). By the parabolic smoothing property, as long as the diffusion coefficient is positive, one can expect that the solution is C^2 in space and C^1 in time for positive time.

This has been summarized in [21], which proves that, under certain growth assumptions, as long as this "stochastic solution" is continuous, it is also C^2 in space and C^1 in time. We will revisit this theorem when we work with a model with stochastic correlation. For such solutions, a large class of classical numerical schemes are available.

It is not always the case that the solution to a pricing problem enjoys such smoothness. While in this thesis we are primarily concerned with solutions that are C^2 in space and C^1 in time, we briefly mention another space, naturally arising from other pricing considerations, that is also numerically interesting.

As mentioned in Section 1.2, the risk-neutral measure is a mathematical construction that seeks to price all contingent claims by an expectation operator, and is not necessarily unique. When this probability measure is not unique, the market is said to be *incomplete*. Each of these possible measures corresponds to a PDE, and there can be more than one possible price for the particular instrument that prevents arbitrage.

There are various approaches to pricing in an incomplete market. One possibility is to select certain special risk-neutral probability measures for pricing - for example, the *minimum entropy martingale measure*, and consider the conditional expectation of the payoff under this measure as the price. This will lead to a unique pricing PDE, and the solution can be expected to be smooth when the underlying process is a diffusion.

It is also possible to consider the set of all equivalent martingale measures and define the price as the supremum (or infimum) of these conditional expectations, depending on the position taken. One example is the uncertain volatility model, where the asset price is specified like (1.3), but the volatility is uncertain and can be anywhere in an interval $[\underline{\sigma}, \overline{\sigma}]$. The worst case path at each point in time depends on the *Gamma* (second derivative with respect to the spot price) of the option. The resulting price from considering the worst case path of volatility is also called the *superreplication cost*.

Another possibility is to take into account the investor's utility function. Roughly speaking, the utility indifference (buy) price identifies the amount of cash price such that the investor is indifferent between paying the price and obtaining a certain amount of the claim, and not

paying and not receiving the claim at maturity, in the sense that his or her expected utility under optimal trading is unchanged. This is also a popular approach to pricing in an incomplete market.

Both the worst-case price approach and the utility indifference pricing can be cast in control form, and the associated PDE is of *Hamilton-Jacobi-Bellman* type. They are also nonlinear pricing methods, meaning that the price of the sum of two claims does not necessarily equal the sum of the individual prices. The suitable solution space to such problems is that of the *viscosity solutions*. Numerical schemes, aside from consistency and stability, have to be monotone in order that convergence to such solutions is guaranteed [2].

In this thesis, we are primarily concerned with linear PDEs and solutions that are C^2 in space, and C^1in time. The Lax equivalence theorem applies, that is, a consistent finite difference method is convergent if and only if it is stable.

1.6 Thesis contributions

In this thesis, we will explore the option pricing problem from a PDE standpoint. The first part of the thesis is an analysis of a model problem, and our focus is on understanding the effect of non-smoothness on the numerical solution. The second part of the thesis focuses on developing computational methods for option pricing problems with stochastic correlation.

As mentioned, a risk-neutral expectation has a PDE representation (1.11). The initial condition for the PDE is simply the payoff function, which is often non-smooth. For example, the call option payoff $g(S(T)) = \max(S(T) - K, 0)$ is not C^1 at the strike K. The first part of the thesis is an analysis of the numerical solution under Crank-Nicolson-Rannacher timestepping with central spatial differences when the initial condition is not smooth. Our contributions are summarized below.

- We develop a general framework to analyze the error arising from non-smoothness of the initial condition for a finite difference scheme, in relation to the relative position of the non-smoothness in the grid.

- We show that for the types of non-smoothness we consider, the error of the numerical solution can be decomposed into a timestepping error typical of Crank-Nicolson timestepping, and another component called *quantization error* which depends (linearly or quadratically) on the relative position of the point of non-smoothness on the grid, which allows us to study possible minimization of error by varying the position of placement.

- We demonstrate explicitly that for the unsmoothed Heaviside initial condition, the leading error is first order, except when the point of discontinuity is placed at a midpoint. While placing the discontinuity at a mid-point is a known technique, we will review this from a different viewpoint. It is an immediate corollary of our analysis that, for our choice of finite difference with an unsmoothed Heaviside initial condition, a first order *quantization* error proportional to $(\alpha - \frac{1}{2})$ arises, where α is the relative position of the non-smoothness in the grid, explaining the inverse relationship between the error and the distance of the discontinuity from a mid-point in the grid.

- Our analysis shows that an unstable convergence estimate can result when the relative position of the non-smoothness, α, is not maintained during grid refinement. We also studied the possibility of choosing an optimal α.

- We demonstrate explicitly the effect of smoothing operators on numerical solutions and how the optimal order is restored. In addition, we show how the dependence of the leading error on α can be removed by smoothing.

The next part of the thesis studies option pricing in the presence of stochastic correlation, from a computational viewpoint. We derive the pricing PDE, where the correlation variable gives rise to an extra dimension. We then develop two approaches to computing option values in this setting. Our contributions are listed below.

- We develop a numerical PDE solution to the problem. We discuss such issues as localization, boundary conditions, discretization and stability of the numerical solution in the l^∞-norm. Of particular interest is the formulation of boundary conditions in the correlation dimension.

- When the correlation process exhibits fast mean-reversion, we propose a second approach, based on singular perturbation ([46]). The asymptotic solution involves a correction to the (multi-asset) Black-Scholes price under a constant correlation.

- For options where the values or derivatives for a constant correlation under the Black-Scholes multi-dimensional framework do not have known closed-form expressions using elementary functions, we have studied a quadrature method based on the asymptotic density to approximate the price. Explicit expressions of the required density corrections are provided.

- We demonstrate the effectiveness and agreement of our proposed solutions to the problem. We also experimentally study the effect of smoothing on the quality of the numerical

solution, as an application of the work presented in the first part of the thesis. Moreover, we numerically study the effects of certain problem parameters on the solution.

1.7 Thesis outline

The thesis is organized as follows.

Chapter 2 is devoted to a detailed analysis of the Crank-Nicolson-Rannacher method, in the presence of a point of non-smoothness in the initial conditions. This study is motivated by the fact that, when applying the PDE approach to pricing, initial conditions are often not smooth. We derive the leading error term of such problems, and from this point of view, study some useful strategies in mitigating the non-smoothness.

Chapter 3 presents the stochastic correlation problem studied in this thesis. The correlation parameter used for pricing and valuation is often treated as a constant. We study a model that introduces stochasticity in correlation. We also prove that the option price defined as a risk-neutral expectation is C^2 in space and C^1 in time, and satisfies a Black-Scholes type PDE.

Chapter 4 is devoted to developing a finite difference scheme for the problem. By Lax equivalence theorem, a consistent finite difference scheme is convergent if and only if it is stable. Our numerical scheme is developed in such a way that we are able to prove that, under some conditions on the mesh, fully implicit timestepping together with a certain spatial discretization scheme that we develop is unconditionally stable in the l^∞-norm.

Chapter 5 develops an alternative approximation to the solution. The motivation is that an implementation of the scheme developed in Chapter 4 could be costly in some computational scenarios, given today's computing power. For computational scenarios where efficiency is preferred to a high degree of accuracy, we develop an asymptotic solution under large mean reversion speed of the correlation process. We also provide justification for payoffs that satisfy certain smoothness and growth assumptions. Based on this asymptotic solution, we also propose a heuristic computational approach, which is demonstrably effective for options that do not have a known closed-form expression of values or derivatives using elementary functions in the Black-Scholes world.

Chapter 6 presents numerical experiments that demonstrate effectiveness of our solutions developed in both Chapters 4 and 5. We also demonstrate the effect of smoothing the initial condition, which is an effective method for non-smooth payoffs used by many researchers and which was studied from a different angle in Chapter 2.

Chapter 2

Mesh Error Analysis of the Crank-Nicolson-Rannacher Method

For many financial pricing problems, exact solutions based on elementary functions are often unknown, and numerical solutions to the Black-Scholes equation and its variants is required. As mentioned in Section 1.5, for diffusion-based linear problems one can expect the solution to be at least C^2 in the interior of the spatial domain and at least C^1 in time. In fact, for the problems we consider in this chapter, the solutions are C^∞ in both space and time away from the initial time. Local analysis of leading error terms, common in numerical analysis textbooks, shows that, under certain smoothness assumptions that include the initial time, the Crank-Nicolson timestepping method combined with central differencing in space should yield second order convergence.

However, special difficulties arise in applying classical PDE timestepping methods to pricing European contracts whose payoffs are not smooth in space. The European call option with payoff given by $\max(S(T) - K, 0)$, considered as a function of the terminal asset price $S(T)$, does not have a continuous first derivative at the strike K. The non-smoothness is known to cause high frequency errors under a classical Crank-Nicolson time discretization ([25]).

The Rannacher timestepping method has been proposed ([48]) to address the difficulty with non-smooth initial data. In this method, the first few timesteps of the Crank-Nicolson timestepping are replaced by fully implicit timesteppings to restore optimal convergence order. It has been shown for various non-smooth initial conditions that the Rannacher start-up is able to suppress the high frequency error associated with the non-smoothness.

While this simple modification in time discretization effectively restores theoretical second order convergence, without special tricks in the spatial direction such as smoothing the initial conditions, "erratic" and suboptimal convergence estimates are still observed in numerical experiments ([47]).

Another novel timestepping technique has also been proposed recently in [49], where it was shown that for Dirac-delta initial condition, a square root change of variable of the time dimension restores the optimal second order convergence (for small enough time-space step-size ratio) without the need of Rannacher timestepping. Numerical experiments there also suggest that the technique is also useful for more complicated problems including the pricing of an American option.

This chapter is dedicated to a detailed study of the error due to grid resolution of the point of non-smoothness. We will focus on non-smoothness that is of most financial interest. In the course of the analysis, we will additionally develop and justify a few numerical schemes that could help achieve a stable convergence order.

Convergence of difference schemes for non-smoothness in initial data has been studied theoretically in [54]. Smoothing schemes for such initial data, as a remedy to restore optimal convergence of difference schemes, is suggested in [36]. Application of these approaches in the financial context can be found in [47], [28], [52] or [19]. In the course of our analysis, these approaches will also be discussed. Other regularization and smoothing techniques for the Dirac-delta and Heaviside functions can be found in [55], [56] or [57], among others.

2.1 Non-smooth initial data and convergence

The Black-Scholes equation is one of the most important equations in financial pricing. In its basic form, the Black-Scholes equation is

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV = 0, \tag{2.1}$$

where $V(t,S)$ is the value of the option at time t and asset price S, which is assumed to have dividend rate q. The risk-free rate is assumed to be a constant r. The volatility σ is unobservable, and in the original formulation of the Black-Scholes model, this quantity is assumed to be constant. When this quantity is deterministically dependent on time and space, the resulting model is the local volatility model due to Dupire ([20]).

Upon substitution $x = \log(S)$ and $\tau = T - t$, equation (2.1) is transformed to a convection-diffusion equation with constant coefficients

$$\frac{\partial v}{\partial \tau} = \frac{\sigma^2}{2}\frac{\partial^2 v}{\partial x^2} + (r - q - \frac{\sigma^2}{2})\frac{\partial v}{\partial x} - rv, \tag{2.2}$$

where $v(\tau, x) = V(T - \tau, e^x)$.
The payoff of the option $g(S_T)$ dependent on the terminal asset price at maturity T translates

into a terminal condition for (2.1) or an initial condition for (2.2). Numerical solutions to (2.1), (2.2) and their generalizations are important in many occasions. When more complex structures are specified, for example a parametric form of the local volatility or higher dimensional volatility models, exact solution based on elementary functions is often unknown even for basic payoff functions $g(\cdot)$. In addition, the numerical PDE approach is effective in handling options with path-dependency features, for example an Asian option or a cliquet option. Numerical solutions to these equations therefore remain important for many applications.

Many financial derivatives, however, have non-smooth payoff functions. The most representative of all are the calls and puts, which respectively have the form $\max(S-K,0)$ and $\max(K-S,0)$, where K is known as the *strike*. The first derivative with respect to S is not continuous precisely at the strike. Another common payoff that has similar difficulties is the digital option, which has payoff $\mathcal{H}(S-K)$, where $\mathcal{H}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else} \end{cases}$ is the Heaviside function. This option pays off a fixed amount if and only if the asset price is above (or alternatively, below) a certain strike. The payoff itself is not continuous.

It has been widely reported and known that applying a finite difference method with Crank-Nicolson directly to (2.1) or (2.2) with non-smooth initial data will result in erratic convergence rates and in some cases large errors in derivative approximations. The Rannacher timestepping successfully eliminates higher frequency errors and restores second order leading errors for calls and puts ([25]). However, suboptimal convergence is still observed experimentally for digital options ([47]).

As an example, we consider solving (2.2) with an initial condition equal to $\mathcal{H}(x)$ so that discontinuity occurs at $x=0$. Equivalently, this is the price of a digital option with strike 1 under the assumption of geometric Brownian motion. We use a finite difference method with central differences and Rannacher timestepping so that the two Crank-Nicolson timesteps are replaced by four fully implicit timesteps of half the step-size. We begin with a uniform grid on $[-8,8]$ with step-size $h=\frac{1}{12}$. For each successive run, we insert mid-points into the grid so that the grid remains uniform, and the step-size is halved. This is a common method of refining grids (but by no means the only one). We shall revisit this point later in the chapter. Finally, Dirichlet conditions with the exact solution are imposed on the two far end-points.

From the $(l-1)$-th mesh (coarser) to the l-th mesh (finer), we also define the quantity for $l>1$:

$$\Upsilon_l \equiv \log\left(\left|\frac{\text{error}_{l-1}}{\text{error}_l}\right|\right) / \log(2).$$

The error is defined to be the numerical approximation minus the exact value of the solution to the PDE. If the numerical scheme has first order convergence, then error is approximately

Spatial step-size h	Time step-size k	Error	Convergence rate estimate Υ
1/12	1/6	7.9320×10^{-2}	–
1/24	1/12	3.9038×10^{-2}	1.0228
1/48	1/24	1.9495×10^{-2}	1.0018
1/96	1/48	9.7551×10^{-3}	0.9989

Table 2.1: Results of solving equation (2.2) with initial condition the Heaviside function $\mathcal{H}(x)$, evaluated at 0. Volatility σ is 20%, risk-free rate r is 5%, and dividend q is 0%. Numerical method is Rannacher timestepping with central spatial difference. Each grid is refined by inserting mid-points. Strike aligned with a grid-point.

halved as the grid is refined by one level. In this case, the Υ_l's would be close to 1. On the other hand, one can expect the Υ_l's to be close to 2 for a quadratically convergent scheme.

The results from solving (2.2) with an initial condition equal to $\mathcal{H}(x)$ *using central difference with Rannacher timestepping* are shown in Table 2.1. It is evident that, even though theoretically second order convergence is to be expected, in this setting, one only observes a first order convergence experimentally. An existing technique in mitigating this sub-optimal convergence is by placing the discontinuity at a mid-point (e.g. [47]). We will revisit this technique from a different viewpoint as we develop the analysis later in the chapter.

If the strike is not a grid-point, which is a common scenario, and no additional effort is taken to align the strike to a grid-point in the numerical software, then an erratic experimental convergence using the aforementioned way of refining grids can result. This can be observed in Table 2.2. In this experiment, the first grid has grid-points $(\frac{1}{30} + \frac{j}{12})$, where $-100 \leq j \leq 92$, so that the endpoints are $(-8.3, 7.7)$, on which we impose Dirichlet boundary conditions based on the known exact solution. We refine the grid by inserting mid-points. As evident in Table 2.2, the error does not necessarily improve even as the step-sizes are halved. The experimental convergence is far from stable.

An erratic convergence could be problematic. Extrapolation, for example, is a common technique to eliminate the leading error term in order to obtain a more accurate solution using numerical solutions from a coarse grid and a finer grid. This is a useful technique when computational costs are high, for instance in a higher dimension PDE solver. However, extrapolation is only possible when the convergence is stable. With a convergence table like the one in Table 2.2, it is difficult to obtain a reliable extrapolated value.

Finally, the errors in Table 2.2 in fact are smaller than those in Table 2.1. This is an expected phenomenon and we will explain why placing the strike on a grid-point will lead to larger errors later in the chapter.

Spatial step-size h	Time step-size k	Error	Convergence rate estimate Υ
1/12	1/6	1.6067×10^{-2}	–
1/24	1/12	2.3803×10^{-2}	-0.5670
1/48	1/24	3.9294×10^{-3}	2.5988
1/96	1/48	5.8572×10^{-3}	-0.5759

Table 2.2: Results of solving equation (2.2) with initial condition the Heaviside function $\mathcal{H}(x)$, evaluated at 0 with cubic spline interpolation. Volatility σ is 20%, risk-free rate r is 5%, and dividend q is 0%. Numerical method is Rannacher timestepping with central spatial difference. Each grid is refined by inserting mid-points. Strike not aligned with a grid-point.

The error resulting from the alignment of the non-smoothness is known as the *quantization error* in [52]. In other words, this is an error that arises from the resolution of the discontinuity (or point of non-smoothness) on the grid, on top of the classical finite difference discretization errors. In this chapter, we will analyze in detail how this quantization error affects the quality of a numerical solution.

2.2 Model problem

As the logarithmic transformation converts the Black-Scholes equation to a convection-diffusion equation with constant coefficients, we work with the following model problem as in [25]:

$$\frac{\partial v}{\partial t} + a\frac{\partial v}{\partial x} = \frac{\partial^2 v}{\partial x^2}, \quad (x,t) \in (-\infty, \infty) \times [0,1]. \tag{2.3}$$

This equation has an explicit fundamental solution given by

$$G(t,x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{(x-at)^2}{4t}}.$$

Given any initial condition $g(x)$, the solution to (2.3) for $t > 0$ is

$$v(t,x) = \int_{-\infty}^{\infty} g(y) G(t, x-y) dy.$$

In particular, for Dirac-delta initial condition at 0, the solution is

$$v_\delta(t,x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{(x-at)^2}{4t}}. \tag{2.4}$$

We consider a finite difference method using second order central difference with Rannacher

timestepping. Let h be the stepsize of a spatial discretization, and k be the time stepsize. Denote $t_l = lk$ (with $l = 1, 2, \ldots, m$ and $t_m = 1$) and $x_j = (j + (1-\alpha))h$, where $j \in \{\ldots, -1, 0, 1, \ldots\} = \mathbf{Z}$, and $\alpha \in (0, 1]$. Let $v^{(l)}$ be a discretized version of v, i.e. $v_j^{(l)} = v(t_l, x_j)$. The fully implicit discretization of (2.3) with a time step-size of $\frac{k}{2}$ is

$$\frac{v_j^{(l+1)} - v_j^{(l)}}{\frac{k}{2}} = \frac{v_{j+1}^{(l+1)} - 2v_j^{(l+1)} + v_{j-1}^{(l+1)}}{h^2} - a\frac{v_{j+1}^{(l+1)} - v_{j-1}^{(l+1)}}{2h}, \tag{2.5}$$

whereas the Crank-Nicolson discretization of (2.3) with a time step-size k is as follows:

$$\begin{aligned}\frac{v_j^{(l+1)} - v_j^{(l)}}{k} &= \frac{1}{2}\Bigg(\frac{v_{j+1}^{(l)} - 2v_j^{(l)} + v_{j-1}^{(l)}}{h^2} - a\frac{v_{j+1}^{(l)} - v_{j-1}^{(l)}}{2h} \\ &\quad + \frac{v_{j+1}^{(l+1)} - 2v_j^{(l+1)} + v_{j-1}^{(l+1)}}{h^2} - a\frac{v_{j+1}^{(l+1)} - v_{j-1}^{(l+1)}}{2h}\Bigg).\end{aligned} \tag{2.6}$$

Our goal is to compare $v^{(m)}$ and $v(1, \cdot)$ and investigate the effect of non-smoothness on their discrepancy. We will also investigate how the error changes as we refine the grid by inserting mid-points into the previous mesh. As in Section 2.1, the quantity $\lambda = \frac{k}{h}$ is held constant as the grid is refined.

For Rannacher timestepping, the first R timesteps (of step-size k) are each replaced by 2 timesteps of step-size $\frac{k}{2}$. A detailed investigation in [25] concluded that $R = 2$ is the optimal choice to reduce high-frequency errors associated with non-smoothness of the initial condition while not increasing the more prominent low-frequency errors. This is known as the *Crank-Nicolson-Rannacher* (CN-Rannacher) method.

Other implementations of the Rannacher timestepping, including replacing two initial Crank-Nicolson timesteps by two fully implicit timesteps, have been studied in [25]. We refer the reader to their work for these other possible choices.

2.3 Difference equation and the discrete-continuous Fourier transform

For the rest of this chapter, the variable i denotes the canonical choice of the complex number such that $i^2 = -1$. Following [25], for a function U defined on the discretized grid such that its value at x_j is given by U_j, we define the transforms

$$\hat{U}(\theta) = h \sum_{j=-\infty}^{\infty} U_j e^{-\frac{ix_j\theta}{h}}. \tag{2.7}$$

The inverse transform is given by

$$U_j = \frac{1}{2\pi h}\int_{-\pi}^{\pi} \hat{U}(\theta)e^{\frac{ix_j\theta}{h}}d\theta = \frac{1}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \hat{U}(h\kappa)e^{ix_j\kappa}d\kappa \quad (\theta = h\kappa). \tag{2.8}$$

These are also known as *discrete-time Fourier Transform pair*. Discrete samples of a function are transformed into a continuous spectrum of frequencies, and vice versa. It is not hard to verify that the inverse transform recovers the sampled values of the function.
Multiplying both sides of (2.5) by $he^{-\frac{ix_j\theta}{h}}$ and summing across j, we obtain

$$\begin{aligned}\sum_j v_j^{(l)}e^{-\frac{ix_j\theta}{h}} &= \sum_j v_j^{(l+1)}e^{-\frac{ix_j\theta}{h}} - \frac{k}{2}\left(\frac{\sum_j v_{j+1}^{(l+1)}e^{-\frac{ix_j\theta}{h}} - \sum_j 2v_j^{(l+1)}e^{-\frac{ix_j\theta}{h}} + \sum_j v_{j-1}^{(l+1)}e^{-\frac{ix_j\theta}{h}}}{h^2}\right)\\ &\quad + \frac{k}{2}\left(a\frac{\sum_j v_{j+1}^{(l+1)}e^{-\frac{ix_j\theta}{h}} - \sum_j v_{j-1}^{(l+1)}e^{-\frac{ix_j\theta}{h}}}{2h}\right).\end{aligned}$$

Then using the transform definition in (2.7), we have

$$\begin{aligned}\hat{v}^{(l)}(\theta) &= \left(1 + \frac{k}{h^2} - \frac{k}{2h^2}(e^{i\theta} + e^{-i\theta}) + \frac{ka}{4h}(e^{i\theta} - e^{-i\theta})\right)\hat{v}^{(l+1)}(\theta)\\ &= \left(1 + \frac{k}{h^2}(1 - \cos\theta) + i\frac{ka}{2h}\sin\theta\right)\hat{v}^{(l+1)}(\theta)\\ &= \left(1 + \frac{2k}{h^2}\sin^2\frac{\theta}{2} + i\frac{ka}{2h}\sin\theta\right)\hat{v}^{(l+1)}(\theta),\end{aligned}$$

from which we get

$$\hat{v}^{(l+1)}(\theta) = \frac{1}{1 + i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2}}\hat{v}^{(l)}(\theta),$$

where $\lambda = \frac{k}{h}$ and $d = \frac{k}{h^2}$.
Working similarly with (2.6), we get

$$\hat{v}^{(l+1)}(\theta) = \frac{1 - i\frac{a\lambda}{2}\sin\theta - 2d\sin^2\theta}{1 + i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2}}\hat{v}^{(l)}(\theta).$$

After $2R$ applications of (2.5) with time step-size $\frac{k}{2}$ followed by $m-R$ applications of (2.6) with time step-size k, we have at terminal time $l=m$ (where $t_m=1$),

$$\begin{aligned}\hat{v}^{(m)}(\theta) &= \left(\frac{1}{1+i\frac{a\lambda}{2}\sin\theta+2d\sin^2\frac{\theta}{2}}\right)^{2R}\left(\frac{1-i\frac{a\lambda}{2}\sin\theta-2d\sin^2\theta}{1+i\frac{a\lambda}{2}\sin\theta+2d\sin^2\frac{\theta}{2}}\right)^{m-R}\hat{v}^{(0)}(\theta)\\ &= \left(\frac{1}{1+i\frac{a\lambda}{2}\sin\theta+2d\sin^2\frac{\theta}{2}}\right)^{2R}\left(\frac{1+i\frac{a\lambda}{2}\sin\theta+2d\sin^2\frac{\theta}{2}}{1-i\frac{a\lambda}{2}\sin\theta-2d\sin^2\theta}\right)^{R}\\ &\quad\times\left(\frac{1-i\frac{a\lambda}{2}\sin\theta-2d\sin^2\theta}{1+i\frac{a\lambda}{2}\sin\theta+2d\sin^2\frac{\theta}{2}}\right)^{m}\hat{v}^{(0)}(\theta)\end{aligned}$$

Therefore, with

$$z_1(\theta) = (1-i\frac{a\lambda}{2}\sin\theta-2d\sin^2\frac{\theta}{2})(1+i\frac{a\lambda}{2}\sin\theta+2d\sin^2\frac{\theta}{2})^{-1} \tag{2.9}$$

$$z_2(\theta) = (1-i\frac{a\lambda}{2}\sin\theta-2d\sin^2\frac{\theta}{2})^{-1}(1+i\frac{a\lambda}{2}\sin\theta+2d\sin^2\frac{\theta}{2})^{-1}, \tag{2.10}$$

we have

$$\hat{v}^{(m)}(\theta)=z_1^m(\theta)z_2^R(\theta)\hat{v}^{(0)}(\theta) \tag{2.11}$$

for $m>R$.

2.4 Error Analysis of CN-Rannacher method

2.4.1 Review of Giles-Carter analysis [25]

Our analysis relies heavily on utilizing the sharp error estimates developed in [25] for linear PDEs with Dirac-delta initial data. In this section, we summarize the relevant results in [25]. For completeness we shall also include the key parts of the proofs. We denote

$$\hat{U}^{(m)}(\theta)=z_1^m(\theta)z_2^R(\theta).$$

The variable $\hat{U}^{(m)}$ is easily seen as the numerical timestepping operator up to time $t=1$ in Fourier space given any initial $\hat{v}^{(0)}$. Algebraically, we write (here and in the rest of the chapter) $\theta=h\kappa$.

The domain of κ is $[-\frac{\pi}{h},\frac{\pi}{h}]$. Choose b such that $0<b<\frac{1}{3}$ and c such that $\frac{1}{2}<c<1$. For each h, we divide this domain of κ into three parts:

- Low frequency domain: $|\kappa|<h^{-b}$
- High frequency domain: $|\kappa|>h^{-c}$, and

- Mid frequency domain: $h^{-b} \leq |\kappa| \leq h^{-c}$.

The following propositions essentially show that the value of $\hat{U}^{(m)}(\theta)$ is more prominent in the low frequency domain than in the other two. In the low frequency domain, the value of $\hat{U}^{(m)}$ is of order $O(1)$. In the high frequency domain, the value of $\hat{U}^{(m)}$ is of order $O(h^{2R})$ where R is the number of fully implicit timesteps initially applied. Finally, the value in the mid frequency domain goes to zero faster than any polynomial in h, as $h \to 0$.

PROPOSITION 1. *For $|\kappa| < h^{-b}$ (low frequency domain), we have as $h \to 0$,*

$$\hat{U}^{(m)}(\theta) = \hat{U}^{(m)}(h\kappa) = e^{-ia\kappa-\kappa^2}\left(1 + h^2 p(\kappa, a, \lambda, R)\right) + O(h^3(\kappa^3 + \kappa^9)),$$

where

$$p(\kappa, a, \lambda, R) = \frac{1}{6}ia\kappa^3 + \frac{1}{12}\kappa^4 - \frac{1}{12}\lambda^2\kappa^3(ia+\kappa)^3 + \frac{1}{4}R\lambda^2\kappa^2(ia+\kappa)^2.$$

Proof. From the definition of $\hat{U}^{(m)}$, and that $m = \frac{1}{\lambda h}$,

$$\log(\hat{U}^{(m)}(\theta)) = m\log z_1 + R\log z_2 = \frac{1}{\lambda h}\log z_1 + R\log z_2.$$

From the definition of z_1, and the expansion $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots$ for small x,

$$\begin{aligned}\log z_1 &= \log(1 - i\frac{a\lambda}{2}\sin\theta - 2d\sin^2\frac{\theta}{2}) - \log(1 + i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2}) \\ &= 2\times\left(-(i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2}) - \frac{1}{3}(i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2})^3\right) + \text{higher order terms},\end{aligned}$$

while

$$\begin{aligned}\log z_2 &= -\log(1 - i\frac{a\lambda}{2}\sin\theta - 2d\sin^2\frac{\theta}{2}) - \log(1 + i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2}) \\ &= (i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2})^2 + \text{higher order terms}.\end{aligned}$$

Recall $d = \frac{\lambda}{h}$. We use the expansion $\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$ for small x (and similarly for $\cos(x)$)

to obtain

$$\begin{aligned} i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2} &= \sin(\frac{\kappa h}{2})(ia\lambda\cos(\frac{\kappa h}{2}) + 2\frac{\lambda}{h}\sin(\frac{\kappa h}{2})) \\ &= \lambda(\frac{\kappa h}{2} - \frac{\kappa^3 h^3}{48})(ia - \frac{iah^2\kappa^2}{8} + \kappa - \frac{\kappa^3 h^2}{24}) + \text{higher order terms} \\ &= \frac{1}{2}\lambda h(\kappa - \frac{\kappa^3 h^2}{24})(ia + \kappa - \frac{iah^2\kappa^2}{8} - \frac{\kappa^3 h^2}{24}) + \text{higher order terms} \\ &= \frac{1}{2}\lambda h(ia\kappa + \kappa^2 - \frac{ia\kappa^3 h^2}{6} - \frac{\kappa^4 h^2}{12}) + \text{higher order terms} \end{aligned}$$

It follows that, by regrouping terms up to h^2,

$$\log\hat{U}^{(m)}(h\kappa) = ia\kappa - \kappa^2 + h^2 p(\kappa, a, \lambda, R) + \text{higher order terms},$$

and so

$$\hat{U}^{(m)}(h\kappa) = e^{-ia\kappa-\kappa^2}\Big(1 + h^2 p(\kappa, a, \lambda, R)\Big) + \text{higher order terms},$$

as desired. The relation $0 < b < \frac{1}{3}$ is obtained by requiring that $h^2\kappa^6$, the highest order term in κ of p, tends to zero as $h \to 0$. □

REMARK 1. *In the original paper [25], it was stated that the higher order terms are dominated by $O(h^3(\kappa^3+\kappa^9))$. However, the Taylor series expansion seems to suggest that the higher order terms are of the form $O(h^4(R\kappa^4 + \kappa^5 + \kappa^6 + R\kappa^8 + \kappa^{10}))$. This discrepancy, however, does not affect our subsequent analysis.*

PROPOSITION 2. *For $|\kappa| > h^{-c}$ (high frequency domain), we have*

$$\hat{U}^{(m)}(\theta) = \hat{U}^{(m)}(h\kappa) = \frac{(-1)^{m-R}h^{2R}}{(2\lambda\sin^2\frac{\theta}{2})^{2R}}e^{-\frac{1}{\lambda^2\sin^2(\frac{\theta}{2})}}(1 + O(h\theta^{-2}))$$

as $h \to 0$.

Proof. We rewrite

$$\begin{aligned} z_1 &= (1 - i\frac{a\lambda}{2}\sin\theta - 2d\sin^2\frac{\theta}{2})(1 + i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2})^{-1} \\ &= (1 - i\frac{a\lambda}{2}\sin\theta - 2\frac{\lambda}{h}\sin^2\frac{\theta}{2})(1 + i\frac{a\lambda}{2}\sin\theta + 2\frac{\lambda}{h}\sin^2\frac{\theta}{2})^{-1} \\ &= -(1 + \frac{iah}{2}\cot\frac{\theta}{2} - \frac{h}{2\lambda\sin^2\frac{\theta}{2}})(1 + \frac{iah}{2}\cot\frac{\theta}{2} + \frac{h}{2\lambda\sin^2\frac{\theta}{2}})^{-1}. \end{aligned}$$

Therefore, as $h \to 0$,

$$\begin{aligned}(-z_1)^m &= (-z_1)^{\frac{1}{\lambda h}} \\ &= \frac{(1+\frac{iah}{2}\cot\frac{\theta}{2}-\frac{h}{2\lambda\sin^2\frac{\theta}{2}})^{\frac{1}{\lambda h}}}{(1+\frac{iah}{2}\cot\frac{\theta}{2}+\frac{h}{2\lambda\sin^2\frac{\theta}{2}})^{\frac{1}{\lambda h}}} \\ &\to \exp(\frac{1}{\lambda}(\frac{ia}{2}\cot(\frac{\theta}{2})-\frac{1}{2\lambda\sin^2\frac{\theta}{2}})-\frac{1}{\lambda}(\frac{ia}{2}\cot(\frac{\theta}{2})+\frac{1}{2\lambda\sin^2\frac{\theta}{2}})) \\ &= \exp(-\frac{1}{\lambda^2\sin^2\frac{\theta}{2}}).\end{aligned}$$

On the other hand, similar rearrangement of z_2 reveals that

$$(-z_2)^R = \frac{(\frac{h}{2\lambda\sin^2\frac{\theta}{2}})^{2R}}{(1+\frac{iah}{2}\cot\frac{\theta}{2}+\frac{h}{2\lambda\sin^2\frac{\theta}{2}})^R(1+\frac{iah}{2}\cot\frac{\theta}{2}-\frac{h}{2\lambda\sin^2\frac{\theta}{2}})^R}.$$

As a result, the leading error term in $(-1)^{m-R}\hat{U}^{(m)}$ is given by

$$(-1)^{m-R}\hat{U}^{(m)} = (-z_1)^m(-z_2)^R \to \frac{(-1)^{m-R}h^{2R}}{(2\lambda\sin^2\frac{\theta}{2})^{2R}}e^{-\frac{1}{\lambda^2\sin^2(\frac{\theta}{2})}}$$

as $h \to 0$. For the term $\frac{h}{2\lambda\sin^2\frac{\theta}{2}}$ to vanish as $h \to 0$, it is necessary that $|\kappa| > h^{-c}$, where $\frac{1}{2} < c < 1$ (recall $\theta = h\kappa$). □

PROPOSITION 3. *For $h^{-b} < |\kappa| < h^{-c}$ (mid frequency domain) and for any positive q, we have*

$$\hat{U}^{(m)}(\theta) = \hat{U}^{(m)}(h\kappa) = o(h^q)$$

as $h \to 0$.

Proof. Let $s = \sin^2\frac{\theta}{2}$. From the definition of z_1 (2.9), we have

$$|z_1|^2 = \frac{(1-ds)^2+a^2\lambda^2 s(1-s)}{(1+ds)^2+a^2\lambda^2 s(1-s)}.$$

By differentiation, one deduces that $|z_1|$ attains its minimum at $s = \sqrt{\frac{1}{d^2-r^2}}$. Rewriting this in terms of h, for h small, the minimum point is attained at

$$s = \pm\sqrt{\frac{h^2}{\lambda^2(1-h^2a^2)}} \sim O(h).$$

For h small, again, $s = |\sin\frac{\theta}{2}|^2 \approx |\frac{\theta^2}{4}| \sim O(h^2\kappa^2)$. So, in terms of κ, the minimum point is at $\kappa = O(h^{-\frac{1}{2}})$. This lies in the mid frequency region. It is easy to see that $|z_1|^m = o(h^q)$ for any $q > 0$ at h^{-b} and h^{-c}. As a result, we have $\hat{U}^{(m)} = o(h^q)$ for any $q > 0$ in this region, as $|z_1^m z_2^R| < |z_1|^{m-R}$. □

From this analysis, one sees that our finite difference solution for (2.3) with Dirac-delta initial data has three components. The low-frequency component is of order $O(1)$ and differs from the true frequency representation by h^2. To see this, consider the continuous Fourier transform (in x) of an L^1 function $f(t,x)$ [1]:

$$\tilde{f}(t,\Psi) = \int_{-\infty}^{\infty} f(t,x)e^{-i\Psi x}dx. \tag{2.12}$$

Its inverse transform is given by

$$f(t,x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{f}(t,\Psi)e^{i\Psi x}d\Psi. \tag{2.13}$$

Taking the continuous Fourier transform on both sides of (2.3), we have

$$\begin{aligned} \frac{\partial\tilde{v}}{\partial t} + ia\Psi\tilde{v} &= -\Psi^2\tilde{v} \\ \tilde{v}(1,\Psi) &= e^{-ia\Psi-\Psi^2}\tilde{v}(0,\Psi). \end{aligned}$$

For Dirac-delta initial data (at 0), we have $\tilde{v}(0,\Psi) = 1$ and so $\tilde{v}(1,\Psi) = e^{-ia\Psi-\Psi^2}$. This is to be compared with the low frequency region approximate in Proposition 1

$$\hat{U}^{(m)}(\theta) = \hat{U}^{(m)}(h\kappa) = e^{-ia\kappa-\kappa^2}\Big(1 + h^2 p(\kappa,a,\lambda,R)\Big) + \text{higher order terms}.$$

Substituting formally Ψ with κ, we see that the true frequency representation of the Dirac-delta function $\tilde{v}$ therefore is of $O(h^2)$ difference with the representation $\hat{U}^{(m)}$ in Proposition 1.

Finally, when $R = 2$, the high frequency component in Proposition 2 is of order h^4, which can be shown to contribute to an $O(h^3)$ value in the spatial domain after performing an inverse transform.

[1] Notationally, we denote $\tilde{f}$ to be the continuous Fourier transform of f, and $\hat{f}$ to be the discrete Fourier transform from samples of f.

2.4.2 Kernel estimates

The analysis in Section 2.4.1 provides the following insights for (2.3) in the case of Dirac-delta initial data:

- The low frequency component approximates the timestepping "kernel" of the exact solution with an error of order $O(h^2)$
- The mid frequency component decays rapidly - that is, of order $O(h^q)$ for any $q > 0$
- The high frequency component is of order h^{2R}, where R is the number of Rannacher timesteps. When $R = 2$, i.e. two initial Crank-Nicolson timesteps of size k are replaced by four fully implicit timesteps of size $\frac{k}{2}$, the high-frequency error will be of order $O(h^4)$.

Intuitively, therefore, when CN-Rannacher timestepping (with $R = 2$) is used, one could expect that the discretization error is dominated by the low frequency error. Below we present analysis based on this conclusion of [25], which leads to interesting insights about numerical solutions to functions with non-smoothness commonly found in financial applications. In the following section, we will always assume $R = 2$.

The framework in [25] is concerned primarily with the effect of CN-Rannacher timestepping on the quality of the numerical solution. In other words, the main focus of [25] is on the *time* discretization. We utilize their techniques in the *spatial* dimension, and seek to understand the explicit effect of the non-smoothness on the quality of the numerical solution.

In the following sections, we will study three types of non-smoothness of financial interest. We first illustrate our analysis for the solution of (2.3) with Dirac-delta initial condition, which is the continuous analogue of the price of an *Arrow-Debreu* security, also known as the state-price security, in finance. Next, we will consider the case when the initial condition is the Heaviside function, which is discontinuous at zero. We will demonstrate how the discontinuity gives rise to a first order error that will dominate the second order error expected of a Crank-Nicolson-Rannacher central difference method. Finally, we demonstrate the effect of the relative position of the point of non-smoothness on the leading error when the ramp function is the initial condition, even though it is continuous. In option pricing terminology, this initial condition is the payoff of a call option.

2.4.3 Dirac-delta function

We start with the analysis of the numerical solution of (2.3) with Dirac-delta initial condition. The Dirac-delta function $\delta(x)$ is a generalized function, defined formally as:

- $\delta(x) = 0$ for $x \neq 0$

- $\int_{\mathbb{R}} \delta(x)dx = 1$.

Despite the singularity, the solution to (2.3) is smooth and is given by the Gaussian (2.4). Numerically, such an initial condition requires an approximation. Recall that our discretized grid is $x_j = (j + (1-\alpha))h$, where $j \in \{\ldots, -1, 0, 1, \ldots\} = \mathbf{Z}$. We shall use the following grid-dependent approximation of the Dirac-delta function:

$$v^{(0)}_{\delta,\alpha,h}(x_j) = \begin{cases} \frac{(1-\alpha)}{h} & \text{for } j = -1 \\ \frac{\alpha}{h} & \text{for } j = 0 \\ 0 & \text{else.} \end{cases} \tag{2.14}$$

The subscript δ in $v^{(0)}_{\delta,\alpha,h}$ indicates that it is an adaptation of the Dirac-delta function, while α and h indicate dependence on the discretized grid. The point of non-smoothness is at $x = 0$. To see how this formula is obtained, recall that the Heaviside function is the weak derivative of the ramp function:

$$g_c(x) = \begin{cases} x & \text{for } x \geq 0 \\ 0 & \text{else,} \end{cases}$$

and that the Dirac-delta function is the distributional derivative of the Heaviside function. Formally applying the second order central difference operator on g_c yields (2.14).

Equation (2.14) is by no means the only way to approximate the Dirac-delta function. A more detailed study on this point can be found in [56].

Applying the discrete Fourier transform (2.7) to (2.14), we obtain

$$\hat{v}^{(0)}_{\delta,\alpha,h}(\theta) = (1-\alpha)e^{i\alpha\theta} + \alpha e^{-i(1-\alpha)\theta}. \tag{2.15}$$

From (2.11) and Proposition 2, the value of $\hat{v}^{(m)}_{\delta,\alpha,h}(\theta) = \hat{v}^{(m)}_{\delta,\alpha,h}(h\kappa)$ in the high frequency component remains fourth order in h as $h \to 0$. This portion of the frequency domain then translates into an $O(h^3)$ value at any test point x^* in the spatial domain, since this high frequency domain

contributes to the inverse Fourier transform by

$$\begin{aligned}
& \frac{1}{2\pi h}\Big|\int_{|\kappa|>h^{-c}} \hat{U}^{(m)}(\theta)\hat{v}_{\delta,\alpha,h}(\theta)e^{i\kappa x^*}d\kappa\Big| \\
\leq\ & \frac{1}{2\pi h}\Big|\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \frac{(-1)^{m-2}h^4}{(2\lambda\sin^2\frac{\theta}{2})^4}e^{-\frac{1}{\lambda^2\sin^2(\frac{\theta}{2})}}(1+O(h\theta^{-2}))d\kappa\Big| \qquad (2.16)\\
\leq\ & \frac{h^3}{(2\lambda)^4\pi}\int_0^{\pi}\frac{1}{\sin^8\frac{\theta}{2}}e^{-\frac{1}{\lambda^2\sin^2(\frac{\theta}{2})}}d\theta + \text{higher order terms} \quad (\theta=\kappa h)\\
=\ & O(h^3)
\end{aligned}$$

where the second last integral is finite by Appendix A in [25]. As a result, the dominating error term is $O(h^2)$ and is given by the low-frequency component. We rewrite (2.15) as

$$\begin{aligned}
\hat{v}^{(0)}_{\delta,\alpha,h}(\theta) &= (1-\alpha)e^{i\alpha\theta}+\alpha e^{-i(1-\alpha)\theta}\\
&= 1-\frac{\alpha(1-\alpha)}{2}\kappa^2h^2+O(h^3)\\
&= \tilde{v}^{(0)}_{\delta}(\kappa)-\frac{\alpha(1-\alpha)}{2}\kappa^2h^2+O(h^3), \qquad (2.17)
\end{aligned}$$

where $\tilde{v}^{(0)}_{\delta}(\kappa)\equiv 1$ is the continuous Fourier transform of the Dirac-delta function. As discussed, up to $O(h^2)$, we are only concerned with the low frequency component of $\hat{U}^{(m)}$, for $R=2$. Therefore, using (2.8), an approximation of our finite difference solution $v^{(m)}_{\delta,\alpha,h}(x^*)$ at x^* is given by (modulo $O(h^3)$) [2]

$$\begin{aligned}
v^{(m)}_{\delta,\alpha,h}(x^*) &\approx \frac{1}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-ia\kappa-\kappa^2}\Big(1+h^2p(\kappa,a,\lambda,R)\Big)(\tilde{v}^{(0)}_{\delta}(\kappa)-\frac{\alpha(1-\alpha)}{2}\kappa^2h^2)e^{i\kappa x^*}d\kappa\\
&\approx \frac{1}{2\pi}\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\left(\tilde{v}^{(0)}_{\delta}(\kappa)+h^2p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_{\delta}(\kappa)-\frac{\alpha(1-\alpha)}{2}\kappa^2h^2\right)d\kappa\\
&\approx \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\left(\tilde{v}^{(0)}_{\delta}(\kappa)+h^2p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_{\delta}(\kappa)-\frac{\alpha(1-\alpha)}{2}\kappa^2h^2\right)d\kappa\\
&\qquad\qquad \text{for } h \text{ small} \qquad (2.18)\\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\tilde{v}^{(0)}_{\delta}(\kappa)d\kappa + E^{(D)}_{\delta}+E^{(Q)}_{\delta}\\
&= v_{\delta}(1,x^*)+E^{(D)}_{\delta}+E^{(Q)}_{\delta}, \qquad (2.19)
\end{aligned}$$

where v_δ is the exact solution to (2.3) with Dirac-delta initial data, and is given by (2.4). There-

[2] As $h\to 0$, the integral outside $[-\frac{\pi}{h},\frac{\pi}{h}]$ is arbitrarily small and can be controlled by considering an asymptotic expansion of the error function $\mathrm{erfc}(x)$. Intuitively speaking, this approximation from a finite integral to infinite integral holds as the Gaussian in the integrand $e^{-ia\kappa-\kappa^2}$ goes to zero faster than any polynomial as $h\to 0$.

Spatial step-size h	Time step-size k	FD Error	Error from (2.18)	Convergence rate estimate Υ (FD)
1/12	1/36	1.8962×10^{-4}	1.8932×10^{-4}	–
1/24	1/72	4.7349×10^{-5}	4.7329×10^{-5}	2.0017
1/48	1/144	1.1833×10^{-5}	1.1832×10^{-5}	2.0005
1/96	1/288	2.9581×10^{-6}	2.9581×10^{-6}	2.0001
1/192	1/576	7.3952×10^{-7}	7.3951×10^{-7}	2.0000

Table 2.3: Results of solving equation (2.3) with initial condition the Dirac-delta function $v^{(0)}_{\delta,\alpha,h}(x_j)$ (2.14), evaluated at $x^* = 0.3$ with cubic spline interpolation. The speed of convection a is 0.5. Numerical method is CN-Rannacher timestepping with central spatial difference. Each grid is refined by inserting mid-points. Initially, the singularity is at a grid-point ($\alpha = 1$).

fore, the error of our finite difference solution at x^* is given by $E^{(D)}_\delta + E^{(Q)}_\delta$, where

$$\begin{aligned} E^{(D)}_\delta &\approx \frac{h^2}{2\pi}\int_{-\infty}^{\infty} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_\delta d\kappa \\ E^{(Q)}_\delta &\approx -\frac{h^2}{2\pi}\frac{\alpha(1-\alpha)}{2}\int_{-\infty}^{\infty} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\kappa^2 d\kappa, \end{aligned}$$

and the subscript δ indicates that this error is pertinent to Dirac-delta initial condition, approximated as in (2.14). It is helpful to think of $E^{(D)}_\delta$ as the inherent error from a Crank-Nicolson-Rannacher discretization of the continuous problem. This error is present in the low frequency component and is invariant with respect to the positioning of the point of singularity.

The error $E^{(Q)}_\delta$ is in a similar spirit of the "quantization error" loosely defined in [52] as the error resulting from the resolution of the point of non-smoothness. This error, considered as a function of α, is a quadratic function that varies as the positioning of the singularity changes. For Dirac-delta initial condition, both of these errors can be explicitly solved by elementary integration.

To illustrate this result, we take $\alpha = 1$ and compare our finite difference (FD) results with (2.18). Results are shown in Table 2.3. Here and in subsequent tables, "FD Error" will mean the error of our finite difference approximation compared to the known exact solution of the PDE.

As α is always 1 in Table 2.3, it turns out that $E^{(Q)}$ is always zero in all runs. What remains is the error term $E^{(D)}$, which is of second order. This is the optimal convergence order of CN-Rannacher with central differencing, and is experimentally observed in Table 2.3.

More interestingly, we start with $\alpha > 0$, and refine the grid by inserting mid-points so that

Spatial step-size h	Time step-size k	FD Error	Error from (2.18)	Convergence rate estimate Υ (FD)
1/12	1/36	8.9209×10^{-5}	8.9528×10^{-5}	–
1/24	1/72	1.8841×10^{-5}	1.8818×10^{-5}	2.2433
1/48	1/144	7.0749×10^{-6}	7.0804×10^{-6}	1.4131
1/96	1/288	1.1758×10^{-6}	1.1761×10^{-6}	2.5891
1/192	1/576	4.4262×10^{-7}	4.4253×10^{-7}	1.4094

Table 2.4: Results of solving equation (2.3) with initial condition the Dirac-delta function $v^{(0)}_{\delta,\alpha,h}(x_j)$ (2.14), evaluated at $x^* = 0.3$ with cubic spline interpolation. The speed of convection a is 0.5. Numerical method is CN-Rannacher timestepping with central spatial difference. Each grid is refined by inserting mid-points. Initially, the singularity is placed at a non grid-point ($\alpha = 0.7$).

the step-sizes are halved. Results in Table 2.4 show an unstable experimental convergence. Clearly, the error does not depend only on the spatial step-size, but also on the relative position of the singularity in the grid. While the error itself is *always* $O(h^2)$, the *coefficient* of the leading error term changes from one run to the next. With this particular way of refining the grid, the second order error is not experimentally observed.

This oscillatory behavior of convergence can be understood by looking at $E^{(Q)}$, which depends quadratically on α. The usual scheme of refining the grid by inserting mid-point will result in a different α from one run to the next. More precisely, from the $(l-1)$-th run to the (l)-th, we have

$$\alpha_l = \begin{cases} 2\alpha_{l-1} - 1 & \text{if } \alpha_{l-1} > 0.5 \\ 2\alpha_{l-1} & \text{if } \alpha_{l-1} \leq 0.5. \end{cases}$$

Thus, for Dirac-delta initial condition, the discretization error depends not only on the step-sizes but also on the relative position of the singularity in the grid. We shall see that this dependence occurs for other examples we shall consider in this chapter.

2.4.4 Heaviside function

The Heaviside function[3] is defined as

$$v^{(0)}_H(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{else.} \end{cases} \tag{2.20}$$

[3]In Section 2.1, the Heaviside function is denoted by $\mathcal{H}(\cdot)$, but, in this subsection and in what follows, we use $v^{(0)}_H(\cdot)$ for consistency with other subsections.

One would run into trouble when applying (2.7) directly to (2.20). This is because the series

$$\hat{v}_{H,\alpha,h}(\theta) = h\sum_{j=0}^{\infty} e^{-i(j+(1-\alpha))\theta} \tag{2.21}$$

does not converge for any $\theta \in \mathbb{R}$. Therefore, without a Fourier transform as in (2.7), it would be difficult to apply the theory in Section 2.4.1.

Fortunately, the fix is easy. Consider instead a *complex* θ. If the imaginary part of θ, is negative (i.e. $\mathrm{Im}(\theta) < 0$), then the geometric series (2.21) will converge as $|e^{-i\theta}| < 1$.

The transforms in (2.7), (2.8), (2.12) and (2.13) extend to complex-valued θ and correspondingly $\kappa = \frac{\theta}{h}$ by considering contour integrals on horizontal lines in the complex plane. For real numbers ζ_1, ζ_2 define

$$C_{\zeta_1} = \{x + i\zeta_1, x \in [-\pi, \pi]\},$$

$$D_{\zeta_2} = \{x + i\zeta_2, x \in \mathbb{R}\}.$$

The only difference between C_ζ and D_ζ is that the former is a finite domain while the latter is infinite. Explicitly, for $\Psi \in C_\zeta$, the discrete-time Fourier transform that takes a discrete sample of a function into a continuous spectrum of frequencies is

$$\hat{U}(\Psi) = h\sum_{j=-\infty}^{\infty} U_j e^{-\frac{ix_j\Psi}{h}}. \tag{2.22}$$

Its inverse transform is given by

$$U_j = \frac{1}{2\pi h}\int_{C_\zeta} \hat{U}(\Psi) e^{\frac{ix_j\Psi}{h}} d\Psi. \tag{2.23}$$

Similarly, the continuous Fourier transform for $\Psi \in D_{\zeta_2}$ is

$$\tilde{f}(t,\Psi) = \int_{-\infty}^{\infty} f(t,x) e^{-i\Psi x} dx. \tag{2.24}$$

The inverse transform is given by

$$f(t,x) = \frac{1}{2\pi}\int_{D_\zeta} \tilde{f}(t,\Psi) e^{i\Psi x} d\Psi. \tag{2.25}$$

While the algebraic operations in Section 2.3 and Section 2.4.1 mostly apply to the case of complex θ and κ, there are a few key differences.

Firstly, we know that for $\theta \in \mathbb{R}$, the Crank-Nicolson timestepper z_1 satisfies

$$|z_1(\theta)| = |(1 - i\frac{a\lambda}{2}\sin\theta - 2d\sin^2\frac{\theta}{2})(1 + i\frac{a\lambda}{2}\sin\theta + 2d\sin^2\frac{\theta}{2})^{-1}| \leq 1.$$

This is no longer true for complex θ. We have, however, the following bound. Recall $\kappa = \frac{\theta}{h}$.

PROPOSITION 4. *Let $\theta \in C_{h\zeta}$ (in other words, $\zeta = \mathrm{Im}(\kappa)$ is fixed and independent of h). If the scaling $\frac{k}{h} = \lambda$ is maintained, then $|z_1(\theta)|^n$ is bounded independently of n and h.*

Proof. Write

$$\theta = \mathrm{Re}(\theta) + i\,\mathrm{Im}(\theta)$$

$$\kappa = \mathrm{Re}(\kappa) + i\,\mathrm{Im}(\kappa),$$

where the two variables are again related by $\theta = h\kappa$.
From (tedious) differentiation, the function $z_1(\theta)$, considered as a function of $\mathrm{Re}(\theta)$, attains its maximum at θ^* characterized by $\sin(\mathrm{Re}(\theta^*)) = 0$. As a result, the complex number $\sin(\theta^*)$ is purely imaginary. As $\zeta = \mathrm{Im}(\kappa)$ is assumed to be fixed, we have that $\sin(\theta^*) = \pm\frac{e^{-h\zeta} - e^{h\zeta}}{2i}$. For simplicity, take $\sin(\theta^*) = \frac{e^{-h\zeta} - e^{h\zeta}}{2i}$. Therefore,

$$\begin{aligned}
|z_1(\theta)|^n &\leq |z_1(\theta^*)|^n \\
&= |(1 - \frac{1}{2}ia\lambda\sin\theta^* - 2d\sin^2\frac{1}{2}\theta^*)|^n|(1 + \frac{1}{2}ia\lambda\sin\theta^* + 2d\sin^2\frac{1}{2}\theta^*)|^{-n} \\
&= |1 - \frac{1}{2}a\lambda\frac{e^{-h\zeta} - e^{h\zeta}}{2} + 2d\frac{e^{-h\zeta} + e^{h\zeta} - 2}{4}|^n \\
&\quad |1 + \frac{1}{2}a\lambda\frac{e^{-h\zeta} - e^{h\zeta}}{2} - 2d\frac{e^{-h\zeta} + e^{h\zeta} - 2}{4}|^{-n} \\
&= |1 + \frac{1}{2}a\lambda h\zeta + \frac{\lambda}{h}\frac{h^2\zeta^2 + O(h^4)}{2}|^{\frac{1}{\lambda h}}|1 - \frac{1}{2}a\lambda h\zeta - \frac{\lambda}{h}\frac{h^2\zeta^2 + O(h^4)}{2}|^{\frac{1}{\lambda h}} \\
&\to \exp(a\zeta + \zeta^2),
\end{aligned}$$

as $h \to 0$. □

The analysis in Section 2.4.1 goes through for complex θ and correspondingly $\kappa = \frac{\theta}{h}$, with the following modifications:

- The Taylor series for the logarithm could have an additional term which would be an integral multiple of $2\pi i$, due to the complex logarithm being a multi-valued function.

This does not affect the argument as the subsequent exponentiation will yield the same result regardless ($e^{2\pi i} = 1$).

- Following the proof of Proposition 4, the maximum and the minimum points of $z_1(\theta)$ as a function of $\mathrm{Re}(\theta)$ can be similarly identified. The rest of the argument goes through.

We fix $\zeta = \mathrm{Im}(\kappa) < 0$ and consider $\theta = h\kappa$. As $\mathrm{Im}(\theta) < 0$,

$$\hat{v}^{(0)}_{H,\alpha,h}(\theta) = h\sum_{j=0}^{\infty} e^{-i(j+(1-\alpha))\theta}(\theta) = \frac{he^{-i(1-\alpha)\theta}}{1-e^{-i\theta}} \tag{2.26}$$

The continuous Fourier transform (2.12) of the Heaviside function is given by

$$\tilde{v}^{(0)}_H(\kappa) = \int_0^{\infty} e^{-i\kappa x}dx = \frac{1}{ik}. \tag{2.27}$$

Substituting $\theta = h\kappa$ in (2.26), Taylor series expansion yields

$$\hat{v}^{(0)}_{H,\alpha,h}(h\kappa) = \tilde{v}^{(0)}_H(\kappa) + (\alpha - \frac{1}{2})h + \frac{i\kappa h^2}{2}(\alpha^2 - \alpha + \frac{1}{6}) + O(h^3).$$

It is not hard to prove that the high-frequency error is again $O(h^3)$ when two Rannacher timesteps are used ($R = 2$). As a result, up to $O(h^2)$, for h small, our finite difference solution is

$$\begin{aligned}
v^{(m)}_{H,\alpha,h}(x^*) &\approx \frac{1}{2\pi}\int_{D_\zeta} e^{-ia\kappa-\kappa^2}\Big(1 + h^2 p(\kappa,a,\lambda,R)\Big) \\
&\quad \times\Big(\tilde{v}^{(0)}_H(\kappa) + (\alpha-\frac{1}{2})h + \frac{i\kappa h^2}{2}(\alpha^2-\alpha+\frac{1}{6})\Big)e^{i\kappa x^*}d\kappa \\
&\approx \frac{1}{2\pi}\int_{D_\zeta} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\Big(\tilde{v}^{(0)}_H(\kappa) + h^2 p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_H(\kappa) \\
&\quad +(\alpha-\frac{1}{2})h + \frac{i\kappa h^2}{2}(\alpha^2-\alpha+\frac{1}{6})\Big)d\kappa \\
&= \frac{1}{2\pi}\int_{D_\zeta} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\tilde{v}^{(0)}_H(\kappa)d\kappa + E^{(D)}_H + E^{(Q)}_H \\
&= v_H(1,x^*) + E^{(D)}_H + E^{(Q)}_H,
\end{aligned} \tag{2.28}$$

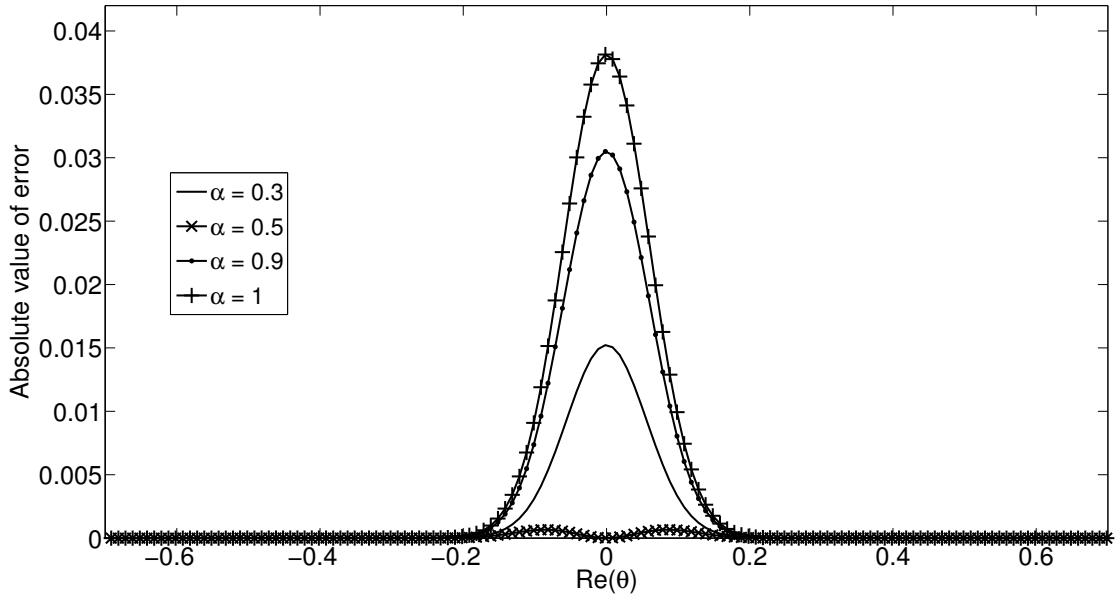

Figure 2.1: The error of our finite difference approximation in frequency space, at $t = 1$. Parameters: $a = 1, \lambda = \frac{1}{3}, h = \frac{1}{12}$. The imaginary part of κ is fixed to -0.1.

where $E_H^{(D)}$ and $E_H^{(Q)}$ are analogously given by

$$\begin{aligned} E_H^{(D)} &= \frac{h^2}{2\pi}\int_{D_\zeta} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}p(\kappa,a,\lambda,R)\tilde{v}_H^{(0)}d\kappa \\ E_H^{(Q)} &= \frac{h}{2\pi}(\alpha-\frac{1}{2})\int_{D_\zeta} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}d\kappa \\ &\quad +\frac{ih^2}{4\pi}(\alpha^2-\alpha+\frac{1}{6})\int_{D_\zeta} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\kappa d\kappa. \end{aligned}$$

In other words, the quantization error[4] is first order in h. The relative position of the discontinuity on the grid has a more prominent effect than the "usual" timestepping error from Crank-Nicolson-CN-Rannacher timestepping, and cannot be damped by the initial backward Euler integrations. In the lower end of the frequency space, it corresponds to a shift by a Gaussian. Figure 2.1 shows this phenomenon.

Again, it is straightforward to obtain the integrals exactly or numerically in $E_H^{(Q)}$ or $E_H^{(D)}$. In Table 2.5, we show the agreement between the numerical solution error and the error as

[4]To be precise, $E^{(Q)}$ also contains the inherent difference between a discrete Fourier transform and a continuous one.

Spatial step-size h	Time step-size k	FD Error	Error from (2.18)	Convergence rate estimate Υ (FD)
1/12	1/24	1.0504×10^{-2}	1.0492×10^{-2}	–
1/24	1/48	5.2241×10^{-3}	5.2227×10^{-3}	1.0076
1/48	1/96	2.6057×10^{-3}	2.6055×10^{-3}	1.0035
1/96	1/192	1.3013×10^{-3}	1.3013×10^{-3}	1.0017
1/192	1/384	6.5029×10^{-4}	6.5029×10^{-4}	1.0008

Table 2.5: Results of solving equation (2.3) with initial condition the Heaviside function $v_H^{(0)}(x)$ (2.20), evaluated at $x^* = 0$. The speed of convection a is 0.7. Numerical method is CN-Rannacher timestepping with central spatial difference. Each grid is refined by inserting mid-points. Initially, the discontinuity is at a grid-point ($\alpha = 1$).

approximated in (2.28). As expected, the convergence is only linear when the point of discontinuity is placed at a grid-point.

Considered as a function in α, the $O(h)$-term in the quantization error $E_H^{(Q)}$ is directly proportional to $(\alpha - \frac{1}{2})$, and vanishes when $\alpha = \frac{1}{2}$. A corollary is that, placing the discontinuity at grid-point is the worst possible choice in terms of minimizing error. The farther the discontinuity is away from the mid-point, the larger the first order error will be. This is illustrated in Table 2.6. In each refinement, we use a mesh that has the required α and spatial step-size h, and compute our finite difference solution based on such a grid. Table 2.6 shows that, with essentially the same computational effort, the grid placement has a direct and prominent effect on the efficiency of the numerical method.

This particular form of $E_H^{(Q)}$ also explains why the errors in Table 2.1 are larger than the errors in Table 2.2, despite the more stable convergence of the former. As $|\alpha - \frac{1}{2}|$ is maximized when $\alpha = 0$ or $\alpha = 1$, the error of our finite difference approximation is also maximized when the discontinuity is placed at a grid-point, other things equal.

2.4.5 Call and put type initial conditions

We consider the following functions:

$$v_C^{(0)}(x) = \max(x, 0) \quad \text{(Call)} \tag{2.29}$$

$$v_P^{(0)}(x) = \max(-x, 0) \quad \text{(Put)} \tag{2.30}$$

$$v_{EC}^{(0)}(x) = \max(e^x - 1, 0) \quad \text{(Exponential Call)} \tag{2.31}$$

$$v_{EP}^{(0)}(x) = \max(1 - e^x, 0) \quad \text{(Exponential Put)} \tag{2.32}$$

These functions are continuous but not continuously differentiable. The exponential call

Spatial step-size h	Time step-size k	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.9$	$\alpha = 1$
1/12	1/24	-4.1349×10^{-3}	1.7457×10^{-5}	8.3946×10^{-3}	1.0504×10^{-2}
1/24	1/48	-2.0730×10^{-3}	4.3549×10^{-6}	4.1772×10^{-3}	5.2241×10^{-3}
1/48	1/96	-1.0381×10^{-3}	1.0882×10^{-6}	2.0840×10^{-3}	2.6057×10^{-3}
1/96	1/192	-5.1949×10^{-4}	2.7201×10^{-7}	1.0409×10^{-3}	1.3013×10^{-3}
1/192	1/384	-2.5986×10^{-4}	6.7999×10^{-8}	5.2020×10^{-4}	6.5029×10^{-4}
Approximated Convergence		Linear	Quadratic	Linear	Linear

Table 2.6: Results of solving equation (2.3) with initial condition the Heaviside function $v_H^{(0)}(x)$ (2.20), evaluated at $x^* = 0$. The speed of convection a is 0.7. Numerical method is CN-Rannacher timestepping with central spatial difference. The relative position α is maintained at each run.

and put functions are related to solving for the value of a call/put option under geometric Brownian motion, after a log transform.

Similarly to Section 2.4.4, we can consider the complex extension of the Fourier transform, i.e. (2.22) to (2.25). In order for the series to converge, we require that

$$\mathrm{Im}(\theta) < 0 \quad \Leftrightarrow \quad \mathrm{Im}(\kappa) < 0 \quad \text{for call} \tag{2.33}$$

$$\mathrm{Im}(\theta) < -h \quad \Leftrightarrow \quad \mathrm{Im}(\kappa) < -1 \quad \text{for exponential call} \tag{2.34}$$

$$\mathrm{Im}(\theta) > 0 \quad \Leftrightarrow \quad \mathrm{Im}(\kappa) > 0 \quad \text{for put/exponential put.} \tag{2.35}$$

2.4.5.1 Call and put

For θ such that $\mathrm{Im}(\theta) < 0$, the discrete Fourier transform of the ramp function (2.29) is

$$\hat{v}^{(0)}_{C,\alpha,h}(\theta) = h^2 \sum_{j=0}^{\infty} (j+(1-\alpha)) e^{-i(j+(1-\alpha))\theta} = h^2 e^{-i(1-\alpha)\theta} \left(\frac{1-\alpha}{1-e^{-i\theta}} + \frac{e^{-i\theta}}{(1-e^{-i\theta})^2} \right). \tag{2.36}$$

This is to be compared with the continuous Fourier transform of (2.29), which for $\mathrm{Im}\,\kappa < 0$ is given by

$$\tilde{v}^{(0)}_C(\kappa) = \int_0^{\infty} x e^{-i\kappa x} dx = -\frac{1}{\kappa^2}. \tag{2.37}$$

Substituting $\theta = h\kappa$ in (2.36), Taylor series expansion yields

$$\hat{v}^{(0)}_{C,\alpha,h}(h\kappa) = \tilde{v}^{(0)}_C(\kappa) + h^2 (-\frac{\alpha^2}{2} + \frac{\alpha}{2} - \frac{1}{12}) + O(h^3). \tag{2.38}$$

Let $\zeta_1 < 0$. By repeating the argument in Section 2.4.4, we have the expression of our finite difference solution

$$\begin{aligned} v^{(m)}_{C,\alpha,h}(x^*) &\approx \frac{1}{2\pi}\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}\Big(1+h^2p(\kappa,a,\lambda,R)\Big) \\ &\quad \times\Big(\tilde{v}^{(0)}_C(\kappa)+h^2(-\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12})\Big)e^{i\kappa x^*}d\kappa \\ &= \frac{1}{2\pi}\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\tilde{v}^{(0)}_C(\kappa)d\kappa + E^{(D)}_C + E^{(Q)}_C \\ &= v_C(1,x^*)+E^{(D)}_C+E^{(Q)}_C, \end{aligned} \tag{2.39}$$

where $E^{(D)}_C$ and $E^{(Q)}_C$ are given by

$$\begin{aligned} E^{(D)}_C &= \frac{h^2}{2\pi}\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_C d\kappa \\ E^{(Q)}_C &= \frac{h^2}{2\pi}(-\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12})\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}d\kappa. \end{aligned}$$

As a result, even though a second order error is to be expected from a Crank-Nicolson-Rannacher discretization, the coefficient of the error depends (quadratically) on the placement of the point of non-smoothness in the grid. In both the frequency space and the original mesh, this error corresponds to a shift by a Gaussian.

Incidentally, for $R = 2$, the spatial error due to high frequency component for the call is not $O(h^3)$, but in fact $O(h^5)$. This is because

$$\tilde{v}^{(0)}_C(\kappa) = -\frac{1}{\kappa^2} = -\frac{h^2}{\theta^2},$$

which adds two orders in h to the high frequency component, in a calculation similar to (2.16):

$$\begin{aligned} & \frac{1}{2\pi h}\Big|\int_{|\kappa|>h^{-c}} \hat{U}^{(m)}(\theta)\hat{v}_{\delta,\alpha,h}(\theta)e^{i\kappa x^*}d\kappa\Big| \\ \le\ & \frac{1}{2\pi h}\Big|\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} \frac{(-1)^{m-2}h^6}{(2\lambda\sin^2\frac{\theta}{2})^4\theta^2}e^{-\frac{1}{\lambda^2\sin^2(\frac{\theta}{2})}}(1+O(h\theta^{-2}))d\kappa\Big| \\ \le\ & \frac{1}{(2\lambda)^4\pi}\int_0^{\pi}\frac{h^5}{\theta^2\sin^8\frac{\theta}{2}}e^{-\frac{1}{\lambda^2\sin^2(\frac{\theta}{2})}}d\theta + \text{higher order terms} \quad (\theta=\kappa h) \\ =\ & O(h^5). \end{aligned} \tag{2.40}$$

We compute the discrete and continuous Fourier transforms of (2.30) for $\mathrm{Im}(\theta) > 0$ and

$\mathrm{Im}(\kappa) > 0$. It turns out that

$$\hat{v}^{(0)}_{P,\alpha,h}(\theta) = h^2 \sum_{j=-\infty}^{-1} (j + (1-\alpha))e^{-i(j+(1-\alpha))\theta} = h^2 e^{-i(1-\alpha)\theta}\left(\frac{-(1-\alpha)e^{i\theta}}{1-e^{i\theta}} + \frac{e^{i\theta}}{(1-e^{i\theta})^2}\right), \tag{2.41}$$

and that for $\mathrm{Im}\,\kappa > 0$,

$$\tilde{v}^{(0)}_P(\kappa) = -\int_{-\infty}^{0} xe^{-i\kappa x}dx = -\frac{1}{\kappa^2}. \tag{2.42}$$

Substituting $\theta = h\kappa$ in (2.41), Taylor series expansion yields

$$\hat{v}^{(0)}_{P,\alpha,h}(h\kappa) = \tilde{v}^{(0)}_P(\kappa) + h^2(-\frac{\alpha^2}{2} + \frac{\alpha}{2} - \frac{1}{12}) + O(h^3). \tag{2.43}$$

Interestingly, the initial conditions (2.38) and (2.43) have the same transform, even though they are *defined* on different regions on the complex plane.
Let $\zeta_2 > 0$.Our finite difference solution under Crank-Nicolson-Rannacher timestepping is

$$\begin{aligned} v^{(m)}_{P,\alpha,h}(x^*) &\approx \frac{1}{2\pi}\int_{D_{\zeta_2}} e^{-ia\kappa-\kappa^2}\left(1 + h^2 p(\kappa,a,\lambda,R)\right) \\ &\quad \times\left(\tilde{v}^{(0)}_P(\kappa) + h^2(-\frac{\alpha^2}{2} + \frac{\alpha}{2} - \frac{1}{12})\right)e^{i\kappa x^*}d\kappa \\ &= \frac{1}{2\pi}\int_{D_{\zeta_2}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\tilde{v}^{(0)}_P(\kappa)d\kappa + E^{(D)}_P + E^{(Q)}_P \\ &= v_P(1,x^*) + E^{(D)}_P + E^{(Q)}_P, \end{aligned} \tag{2.44}$$

where $E^{(D)}_P$ and $E^{(Q)}_P$ are given by

$$\begin{aligned} E^{(D)}_P &= \frac{h^2}{2\pi}\int_{D_{\zeta_2}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_P d\kappa \\ E^{(Q)}_P &= \frac{h^2}{2\pi}(-\frac{\alpha^2}{2} + \frac{\alpha}{2} - \frac{1}{12})\int_{D_{\zeta_2}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}d\kappa. \end{aligned}$$

We also have that

$$p(\kappa,a,\lambda,R)\times(-\frac{1}{\kappa^2}) = -\frac{1}{6}ia\kappa - \frac{1}{12}\kappa^2 + \frac{1}{12}\lambda^2\kappa(ia+\kappa)^3 - \frac{1}{4}R\lambda^2(ia+\kappa)^2$$

is analytic as a function of κ. As a result,

$$\begin{aligned} E^{(D)}_C &= E^{(D)}_P, \quad \text{and} \\ E^{(Q)}_C &= E^{(Q)}_P. \end{aligned}$$

In other words, at least up to second order, the error of Crank-Nicolson-Rannacher is the same for the call and the put. This is to be expected, as it is easy to prove that

$$v_C(t,x) - v_P(t,x) = x - at,$$

and that our numerical scheme is exact on linear functions. This numerical phenomenon does not occur for the exponential call and put, as we shall see in the next section.

REMARK 2. *In [48], it is established that for a finite element discretization with linear basis functions, second order L_2 convergence is obtained if Rannacher timestepping is used (in the presence of discontinuities in the initial condition). In that setting, the payoff function is projected on the space of basis functions, which is a type of smoothing. As the basis functions are piecewise linear, projection does not alter the call/put payoff if there is a node at the strike. A corollary is that to obtain second order convergence with this finite element method, no smoothing is necessary for call and put payoffs if Rannacher timestepping is used, and if there is a node at the strike.*
Similarly, for our finite difference method, smoothing is not necessary for call/put type initial conditions to obtain $O(h^2)$ errors. However, for Heaviside initial conditions, if the strike is not aligned at mid-point, an $O(h)$ error will result even if Rannacher timestepping is used.
It would be interesting to obtain similar analysis for finite element methods with the results we obtain in this chapter as a starting point.

2.4.5.2 Exponential call and put

Consider now the exponential call as the initial condition to (2.3), given by (2.31). Its discrete Fourier transform for $\mathrm{Im}(\theta) < -h$ is

$$\hat{v}^{(0)}_{EC,\alpha,h}(\theta) = h\sum_{j=0}^{\infty}(e^{(j+(1-\alpha))h} - 1)e^{-i(j+(1-\alpha))\theta)} = he^{-i(1-\alpha)\theta}\left(\frac{e^{(1-\alpha)h}}{1-e^{-i\theta+h}} - \frac{1}{1-e^{-i\theta}}\right). \tag{2.45}$$

Its continuous Fourier transform is, for $\mathrm{Im}\,\kappa < -1$,

$$\tilde{v}^{(0)}_{EC}(\kappa) = \int_0^{\infty}(e^x - 1)e^{-i\kappa x}dx = \frac{1}{i\kappa(i\kappa - 1)}. \tag{2.46}$$

Substituting $\theta = h\kappa$ in (2.45), Taylor series expansion yields

$$\hat{v}^{(0)}_{EC,\alpha,h}(h\kappa) = \tilde{v}^{(0)}_{EC}(\kappa) + h^2(-\frac{\alpha^2}{2} + \frac{\alpha}{2} - \frac{1}{12}) + O(h^3). \tag{2.47}$$

This is the same relation as (2.38). As a result, the E_Q-component will be the same as a (non-exponential) call. Let $\zeta_1 < -1$. By repeating the argument in Section 2.4.5.1, we have the following expression of our finite difference solution

$$\begin{aligned} v^{(m)}_{EC,\alpha,h}(x^*) &\approx \frac{1}{2\pi}\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}\Big(1+h^2p(\kappa,a,\lambda,R)\Big) \\ &\quad \times\Big(\tilde{v}^{(0)}_{EC}(\kappa)+h^2(-\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12})\Big)e^{i\kappa x^*}d\kappa \\ &= \frac{1}{2\pi}\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\tilde{v}^{(0)}_{EC}(\kappa)d\kappa + E^{(D)}_{EC}+E^{(Q)}_{EC} \\ &= v_{EC}(1,x^*)+E^{(D)}_{EC}+E^{(Q)}_{EC}, \end{aligned} \tag{2.48}$$

where $E^{(D)}_{EC}$ and $E^{(Q)}_{EC}$ are given by

$$\begin{aligned} E^{(D)}_{EC} &= \frac{h^2}{2\pi}\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_{EC}d\kappa \\ E^{(Q)}_{EC} &= \frac{h^2}{2\pi}(-\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12})\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}d\kappa. \end{aligned}$$

Similarly, for $\mathrm{Im}(\theta) > 0$ and $\mathrm{Im}(\kappa) > 0$, the discrete and continuous transforms for the exponential put are

$$\hat{v}^{(0)}_{EP,\alpha,h}(\theta) = h\sum_{j=-\infty}^{0}(1-e^{(j+(1-\alpha))h})e^{-i(j+(1-\alpha))\theta} = he^{-i(1-\alpha)\theta}\Big(\frac{e^{i\theta}}{1-e^{-i\theta}}-\frac{e^{(1-\alpha)h+i\theta-h}}{1-e^{i\theta-h}}\Big), \tag{2.49}$$

and

$$\tilde{v}^{(0)}_{EP}(\kappa) = \int_{-\infty}^{0}(1-e^{x})e^{-i\kappa x}dx = \frac{1}{i\kappa(i\kappa-1)}. \tag{2.50}$$

Substituting $\theta = h\kappa$ into (2.49), once again Taylor series expansion yields

$$\hat{v}^{(0)}_{EP,\alpha,h}(h\kappa) = \tilde{v}^{(0)}_{EP}(\kappa)+h^2(-\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12})+O(h^3). \tag{2.51}$$

For $\zeta_2 > 0$, we have the following expression of our finite difference solution for the exponen-

tial put

$$\begin{aligned} v^{(m)}_{EP,\alpha,h}(x^*) &\approx \frac{1}{2\pi}\int_{D_{\zeta_2}} e^{-ia\kappa-\kappa^2}\Big(1+h^2p(\kappa,a,\lambda,R)\Big) \\ &\quad \times\Big(\tilde{v}^{(0)}_{EP}(\kappa)+h^2(-\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12})\Big)e^{i\kappa x^*}d\kappa \\ &= \frac{1}{2\pi}\int_{D_{\zeta_2}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\tilde{v}^{(0)}_{EP}(\kappa)d\kappa + E^{(D)}_{EP}+E^{(Q)}_{EP} \\ &= v_{EP}(1,x^*)+E^{(D)}_{EP}+E^{(Q)}_{EP}, \end{aligned} \tag{2.52}$$

where $E^{(D)}_{EP}$ and $E^{(Q)}_{EP}$ are given by

$$\begin{aligned} E^{(D)}_{EP} &= \frac{h^2}{2\pi}\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}p(\kappa,a,\lambda,R)\tilde{v}^{(0)}_{EP}d\kappa \\ E^{(Q)}_{EP} &= \frac{h^2}{2\pi}(-\frac{\alpha^2}{2}+\frac{\alpha}{2}-\frac{1}{12})\int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2}e^{i\kappa x^*}\kappa. \end{aligned}$$

Obviously, as their corresponding integrands are analytic, we have

$$E^{(Q)}_{EC} = E^{(Q)}_{EP}.$$

However, because of a pole at $\kappa = -i$, it holds that $E^{(D)}_{EC} \neq E^{(D)}_{EP}$. To see this, consider a positively oriented contour Γ consisting of the following segments:

$$\begin{aligned} \Gamma_1 &= \{x+i\zeta_1 | -M \leq x \leq M\} \\ \Gamma_2 &= \{M+iy | \zeta_1 \leq y \leq \zeta_2\} \\ \Gamma_3 &= \{x+i\zeta_2 | -M \leq x \leq M\} \\ \Gamma_4 &= \{-M+iy | \zeta_1 \leq y \leq \zeta_2\}. \end{aligned}$$

By Cauchy's residue theorem, we have

$$\int_\Gamma e^{-iaz-z^2}e^{izx^*}\frac{p(z,a,\lambda,R)}{iz(iz-1)}dz = -2\pi i\Big[e^{-iaz-z^2}e^{izx^*}\frac{p(z,a,\lambda,R)}{z}\Big]_{z=-i}.$$

The last quantity is readily computable as $\frac{p(z,a,\lambda,R)}{z}$ itself is a polynomial in z. Finally, as

Spatial step-size h	Time step-size k	FD Error	Error from (2.53)	Convergence rate estimate Υ (FD)
1/12	1/24	-2.0221×10^{-4}	-2.0174×10^{-4}	–
1/24	1/48	-5.0466×10^{-5}	-5.0434×10^{-5}	2.0025
1/48	1/96	-1.2610×10^{-5}	-1.2609×10^{-5}	2.0007
1/96	1/192	-3.1523×10^{-6}	-3.1521×10^{-6}	2.0001
1/192	1/384	-7.8804×10^{-7}	-7.8803×10^{-7}	2.0001

Table 2.7: Results of solving equation (2.3) with initial condition the exponential forward $v_F^{(0)}(x)$ (2.54), evaluated at $x^* = 0$. The speed of convection a is 0.7. Numerical method is CN-Rannacher timestepping with central spatial difference. Each grid is refined by inserting mid-points. Initially, we set $\alpha = 0.7$.

$M \to \infty$, we note that the contribution from Γ_2 and Γ_4 vanish and

$$\int_{\Gamma_1} e^{-ia\kappa-\kappa^2} e^{i\kappa x^*} p(\kappa, a, \lambda, R) \tilde{v}_{EC}^{(0)} d\kappa \to \int_{D_{\zeta_1}} e^{-ia\kappa-\kappa^2} e^{i\kappa x^*} p(\kappa, a, \lambda, R) \tilde{v}_{EC}^{(0)} d\kappa$$

and similarly

$$-\int_{\Gamma_3} e^{-ia\kappa-\kappa^2} e^{i\kappa x^*} p(\kappa, a, \lambda, R) \tilde{v}_{EP}^{(0)} d\kappa \to -\int_{D_{\zeta_2}} e^{-ia\kappa-\kappa^2} e^{i\kappa x^*} p(\kappa, a, \lambda, R) \tilde{v}_{EP}^{(0)} d\kappa$$

Therefore, we have

$$\begin{aligned} E_{EC}^{(D)} - E_{EP}^{(D)} &= -h^2 i \left[e^{-iaz-z^2} e^{izx^*} \frac{p(z, a, \lambda, R)}{z} \right]_{z=-i} \\ &= -h^2 e^{x-a+1} \left(\frac{a}{6} - \frac{1}{12} + \frac{1}{12}\lambda^2 (a-1)^3 - \frac{1}{4} R\lambda^2 (a-1)^2 \right). \end{aligned} \quad (2.53)$$

As $E_{EC}^{(Q)} = E_{EP}^{(Q)}$, the quantity $E_{EC}^{(D)} - E_{EP}^{(D)}$ is in fact the second order error of solving (2.3) with the initial condition

$$v_F^{(0)}(x) = e^x - 1 \quad (2.54)$$

under Crank-Nicolson-Rannacher timestepping[5] . In financial context, this initial condition is the payoff of a forward contract under the geometric Brownian motion model. As the quantization error is cancelled out, the relative position of the strike on the grid is no longer relevant in the second order error, and the leading error depends (computationally) only on the time and spatial step size. This is illustrated in Table 2.7.

[5]This connection between the values of put, call and forward via integration across complex poles is a form of put-call parity [37].

Spatial step-size h	Time step-size k	FD Error	Convergence rate estimate Υ (FD)
1/12	1/24	9.3332×10^{-7}	–
1/24	1/48	1.1988×10^{-7}	2.9608
1/48	1/96	1.5140×10^{-8}	2.9851
1/96	1/192	1.8924×10^{-9}	3.0001
1/192	1/384	2.3323×10^{-10}	3.0205

Table 2.8: Results of solving equation (2.3) with initial condition the exponential put $v_{EP}^{(0)}(x)$ (2.32), evaluated at $x^* = 0$ with cubic spline interpolation. The speed of convection a is -0.3. Numerical method is CN-Rannacher timestepping with central spatial difference. The relative position of the strike is maintained at $\alpha = 0.37853$.

2.5 Maintaining the relative position of non-smoothness on the grid

It is evident from the discussion in Section 2.4.3 through 2.4.5 that, when solving (2.3) with those initial conditions, the parameter α is required to be maintained in order to obtain a stable convergence. In [52] and [47], mesh shifting techniques, mostly aligning the strike on a mid-point, are suggested to restore convergence order.

The analysis from Sections 2.4.3 to 2.4.5 suggests that as long as the relative position of the point of non-smoothness on the grid is maintained, the convergence order is stable. The next question is to determine an optimal α such that the error is minimized.

This is complicated by the fact that, while α directly influences $E^{(Q)}$, the other term in the error $E^{(D)}$ is independent of α. It is possible to use the quantization error $E^{(Q)}$ to our advantage. For the initial conditions considered in this chapter, one could consider the error $E^{(D)} + E^{(Q)}$ as a quadratic function in α. In some cases, the leading error term could be completely eliminated by a good choice of α, leading to super-convergence by a second order finite difference scheme (see Table 2.8).

This technique of choosing α to obtain a superconvergence does not seem to be possible in practical situations, as a detailed study of E_D and E_Q seems necessary to determine the α for which superconvergence occurs. In addition, such an α that cancels the leading second order term may not exist. Instead, we proceed to minimize merely $E^{(Q)}$. Consider $E^{(Q)}$ as a function in α in itself, one can minimize its absolute value and obtain the estimates as listed in Table 2.9. For the case of call and put, often the combined error $E^{(Q)} + E^{(D)}$ has no root, considered as a function of α. In those cases, the mid-point minimizes the overall error. We remark that these numbers seem to confirm the empirical findings of [42], in which the authors found

Initial condition	Optimal α to eliminate the leading term of $E^{(Q)}$	Point of the extremum of the quadratic $E^{(D)} + E^{(Q)}$
Dirac-delta	0 or 1	0.5
Heaviside	0.5	not applicable (linear)
Call/Put/Exponential Call/ Exponential Put	0.2113 or 0.7887	0.5

Table 2.9: Special choices of α.

experimentally that the optimal value of α lies in $(0.2, 0.3)$ or $(0.7, 0.8)$ for the call option, and 0.5 for the bet option (Heaviside initial condition).

2.6 Smoothing

Smoothing has long been a popular approach to obtain stable convergence and in some cases restore optimal order of convergence in the presence of non-smoothness in the initial data. In the financial context, a very popular approach is averaging ([47], [52], [28], [19]). This technique has been used successfully in the case of digital options (the initial condition being the Heaviside function).
In this section, we will take a closer look at the smoothing technique in the context we developed in the earlier parts of the chapter.
We start with the family of smoothing operators suggested in [36]. Their idea is to consider operators of the convolution type, which in frequency space corresponds to pointwise multiplication. In frequency space, define

$$\hat{\Phi}_\mu(h\kappa) = \frac{p_\mu(\sin\frac{h}{2}\kappa)}{(\frac{h}{2}\kappa)^\mu}, \tag{2.55}$$

where $p_\mu(\sin\omega)$ is a polynomial in $\sin\omega$ that satisfies

$$p_\mu(\sin\omega) = \omega^\mu + O(\omega^{2\mu}), \quad \text{as } \omega \to 0.$$

The idea is that high frequency (large κ) components, which are often the cause for non-smoothness, in the initial condition can be damped simply by multiplication with $\hat{\Phi}_\mu$. The integer μ is considered the order of the smoothing operator, as from the definition of p_μ we have

$$\hat{\Phi}_\mu(\omega) = 1 + O(\omega^\mu), \quad \text{as } \omega \to 0, \text{ and}$$

$$\hat{\Phi}_\mu(\omega) = O(|\omega - 2l\pi|^\mu), \text{as } \omega \to 2l\pi, l \in \mathbb{Z}.$$

The first two polynomials are particularly simple:

$$\begin{aligned} p_1(\sin\omega) &= \sin\omega \\ p_2(\sin\omega) &= \sin^2\omega. \end{aligned}$$

The first smoothing operator $\hat{\Phi}_1$ is the familiar averaging technique. To see this, it suffices to compute its inverse Fourier transform at a spatial point x:

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\sin\frac{h}{2}\kappa}{\frac{h}{2}\kappa} e^{i\kappa x} d\kappa &= \int_{-\infty}^{\infty} \frac{e^{i\frac{h}{2}\kappa} - e^{-i\frac{h}{2}\kappa}}{ih\kappa} e^{i\kappa x} d\kappa \\ &= \int_{-\infty}^{\infty} \frac{e^{i\kappa(\frac{h}{2}+x)} - e^{i\omega(-\frac{h}{2}+x)}}{i\kappa} d\kappa \\ &= \begin{cases} 0 & \text{if } |x| > \frac{h}{2} \\ \frac{1}{h} & \text{else} \end{cases} \end{aligned}$$

As a result, the convolution operator that $\hat{\Phi}_1$ induces in the spatial domain is of the form

$$(\Phi_1 * v)(x) = \frac{1}{h} \int_{-\frac{h}{2}}^{\frac{h}{2}} v(x-y) dy, \tag{2.56}$$

Similarly, the inverse transform of $\hat{\Phi}_2$ is

$$\Phi_2(x) = \begin{cases} 0 & \text{if } |x| > h \\ \frac{1}{h}\left(1 - \frac{|x|}{h}\right) & \text{else.} \end{cases}$$

In convolution form, the second order smoothing takes the form

$$(\Phi_2 * v)(x) = \frac{1}{h} \int_{-h}^{h} (1 - \frac{|y|}{h}) v(x-y) dy, \tag{2.57}$$

We shall apply these operators to the initial conditions we have studied in Section 2.4.3 to Section 2.4.5 and analyze how errors are affected by these techniques.

2.6.1 Dirac-delta function

As the Dirac-delta function is a generalized function, it can only be approximated on our numerical grid $x_j = (j + (1 - \alpha))h$. If we replace formally the Dirac-delta function by the first order smoothed version of it (2.56), then we obtain the following approximation of the Dirac-delta initial condition (we leave out the case $\alpha = 0.5$ to avoid ambiguity):

$$v^{(0)}_{\Phi_1,\delta}(x_j) = \begin{cases} \frac{1}{h} & \text{if } (\alpha < 0.5 \text{ or } \alpha = 1) \text{ and } j = 0 \\ \frac{1}{h} & \text{if } \alpha > 0.5 \text{ and } j = 1 \end{cases}$$

Its discrete Fourier transform is

$$\hat{v}^{(0)}_{\Phi_1,\delta,\alpha,h}(h\kappa) = \begin{cases} e^{i\alpha h\kappa} & \text{if } \alpha < 0.5 \text{ or } \alpha = 1 \\ e^{-i(1-\alpha)h\kappa} & \text{if } \alpha > 0.5 \end{cases}.$$

Clearly then

$$\hat{v}^{(0)}_{\Phi_1,\delta}(h\kappa) = \begin{cases} 1 + ih\alpha\kappa + O(h^2) & \text{if } \alpha < 0.5 \text{ or } \alpha = 1 \\ 1 - i(1-\alpha)h\kappa + O(h^2) & \text{if } \alpha > 0.5. \end{cases}$$

In other words, had we started our analysis with this approximation of the Dirac-delta function, then we will end up with a first order error of our finite difference solution.

In fact, one can show that (2.14) is in fact the second order smoothing operator (2.57) applied formally to the Dirac-delta function. The results in Section 2.4.3 show that only the second order error term will remain, although the second order error depends quadratically on the relative position of the singularity on the grid.

2.6.2 Heaviside function

Applying (2.56) to the Heaviside function, we obtain the following modified initial condition:

$$v^{(0)}_{\Phi_1,H}(x_j) = \begin{cases} \frac{1-2\alpha}{2} & \text{if } \alpha < 0.5 \text{ and } j = 0 \\ \frac{3-2\alpha}{2} & \text{if } \alpha \geq 0.5 \text{ and } j = 1 \\ v^{(0)}_H(x_j) & \text{else.} \end{cases}$$

In other words, first order smoothing involves modifying only one point of the sampled function given any α. When $\alpha = 0.5$, the function is identical to the original sample of the

unsmoothed Heaviside function $v_H^{(0)}(x_j)$. It is not surprising that the smoothing technique restores an error of second order in h. In fact, its discrete Fourier transform (for κ suitably defined on the complex plane) is

$$\hat{v}^{(0)}_{\Phi_1,H}(h\kappa) = \begin{cases} \tilde{v}_H^{(0)}(\kappa) + ih^2\kappa(-\frac{\alpha^2}{2} + \frac{1}{12}) + O(h^3) & \text{if } \alpha < 0.5 \\ \tilde{v}_H^{(0)}(\kappa) + ih^2\kappa(-\frac{\alpha^2}{2} + \alpha - \frac{5}{12}) + O(h^3) & \text{if } \alpha \geq 0.5. \end{cases}$$

The first order term, proportional to $(\alpha - \frac{1}{2})$ in (2.28) is removed by the first order smoothing technique. This observation has been noted in [47] and [52].
Although the first order error is successfully removed by smoothing, it is interesting to see what effect the second order smoothing operator Φ_2 would have on the Heaviside function. After applying (2.57) to the Heaviside function $v_H^{(0)}(x)$, one obtains

$$v^{(0)}_{\Phi_2,H}(x_j) = \begin{cases} \frac{(1-\alpha)^2}{2} & \text{if } j = 0 \\ \frac{2-\alpha^2}{2} & \text{if } j = 1 \\ v_H^{(0)}(x_j) & \text{else.} \end{cases}$$

Namely, the second order smoothing modifies two points on the sampled function. Its discrete Fourier transform is given by

$$\hat{v}^{(0)}_{\Phi_1,H}(h\kappa) = \tilde{v}_H^{(0)}(\kappa) + \frac{i\kappa h^2}{12}.$$

Therefore, the second order smoothing not only removes the first order error that would be present with a non-smooth Heaviside initial condition, it also removes the dependence of the second order error on α. The relative position of the grid no longer affects the dominant error term.

2.6.3 Call and put

The first order smoothing of the call and put gives the following modifications:

$$v^{(0)}_{\Phi_1,C}(x_j) = \begin{cases} \frac{(1-2\alpha)^2 h}{8} & \text{if } \alpha < 0.5 \text{ and } j = 0 \\ \frac{(3-2\alpha)^2 h}{8} & \text{if } \alpha \geq 0.5 \text{ and } j = 1 \\ v_C^{(0)}(x_j) & \text{else.} \end{cases}$$

Type of initial condition	Unsmoothed	Φ_1 smoothing	Φ_2 smoothing
Dirac-delta	Not applicable	$O(h)$ error	$O(h^2)$ error, dependent on α
Heaviside	$O(h)$ error	$O(h^2)$ error, dependent on α	$O(h^2)$ error, independent of α
Usual Call, Put and Exponential Call, Put	$O(h^2)$ error, dependent on α	$O(h^2)$ error, independent of α	–

Table 2.10: Summary of the effect of smoothing techniques on CN-Rannacher error under different types of non-smooth initial conditions.

$$v^{(0)}_{\Phi_1,P}(x_j) = \begin{cases} \frac{(1+2\alpha)^2 h}{8} & \text{if } \alpha < 0.5 \text{ and } j = 0 \\ \frac{(1-2\alpha)^2 h}{8} & \text{if } \alpha \geq 0.5 \text{ and } j = 1 \\ v^{(0)}_P(x_j) & \text{else.} \end{cases}$$

For κ suitably defined, the discrete Fourier transforms give

$$\hat{v}^{(0)}_{\Phi_1,C}(h\kappa) = \tilde{v}^{(0)}_C(\kappa) + \frac{h^2}{24}, \quad \text{and}$$

$$\hat{v}^{(0)}_{\Phi_1,P}(h\kappa) = \tilde{v}^{(0)}_P(\kappa) + \frac{h^2}{24}.$$

As a result, the first order smoothing successfully removes the dependence on α in the second order error. Removing the dependence on α is favorable, as the only computational parameters that affect the error will be step-sizes. This can be found convenient in some implementations. We summarize these discussions in Table 2.10.

2.7 General recommendation

Before we conclude the chapter, it would be appropriate to include a summary of analysis that gives advice for the general user who solves a similar problem. For the sake of analysis, the problems we consider in this chapter have known closed-form solutions based on elementary functions, and therefore do not generally require numerical solutions. However, certain results from our analysis are expected to hold for a similarly defined problem.

The analysis in [48], [25] and this chapter suggest that Rannacher timestepping, which replaces the first two Crank-Nicolson timesteps with four fully implicit timesteps with half the time step-size, is important in preventing spurious oscillations in the solution and its deriva-

Initial condition	Second order error	"Stable" convergence
Dirac-delta type	Second order smoothing	Second order smoothing, and maintain α
Heaviside type	Placement at midpoint, or first order smoothing	Second order smoothing, or maintain α
Usual Call, Put and Exponential Call, Put type	–	First order smoothing, or maintain α

Table 2.11: Summary of recommendations on how to obtain second order error and stable convergence with non-smooth initial conditions.

tives, and in some cases restoring the optimal order of convergence. Alternatively, variable change in the time dimension may also be useful [49].

In addition to Rannacher timestepping, some simple strategies are useful in restoring second order convergence or maintaining a stable order of convergence for extrapolation purposes. Based on the analysis in this chapter, our recommendations are summarized in Table 2.11.

2.8 Summary

In this chapter, we have studied extensively the effect of mesh positioning on the error of the Crank-Nicolson-Rannacher method in the presence of non-smoothness. We have specifically studied three types of non-smoothness of financial interest. Our contributions are listed below.

- We develop a general framework to analyze the quantization error for a finite difference scheme in relation to the relative position of the non-smoothness in the grid.

- We demonstrate that in the presence of discontinuity/non-smoothness, the leading error of our numerical solution depends not only on the spatial and time stepsizes, but also on the relative positioning of the point of non-smoothness in the grid. For CN-Rannacher method with central differencing, the error can be decomposed into a "normal" timestepping error component and a quantization error component, and it is the latter that is relevant to the positioning of the non-smoothness on the grid.

- We demonstrate that while the Crank-Nicolson-Rannacher method is formally second order, suboptimal convergence can result from the placement of a discontinuity. While this is a known result, we reviewed this from a different viewpoint. It is an immediate corollary of our analysis that, for our choice of finite difference with an unsmoothed Heaviside initial condition, a first order *quantization* error proportional to $(\alpha - \frac{1}{2})$ arises,

explaining the inverse relationship between the error and the distance of the discontinuity from a mid-point in the grid.

- Our analysis shows that an unstable convergence estimate can result when the relative position of the non-smoothness, α, is not maintained during grid refinement. We also studied the possibility of choosing an optimal α. For our choice of finite difference with an unsmoothed ramp (call or put) initial condition, the quantization error is second order with a $(\alpha^2 + \alpha - \frac{1}{6})$ coefficient, which gives two α values that result in minimum quantization error.

- We demonstrate explicitly that smoothing operators can recover optimal convergence, which was proved in [36]. In addition, we show how the dependence of the leading error on α can be removed by smoothing.

Chapter 3

Stochastic Correlation Model

3.1 Background

In many areas of financial modeling such as pricing and risk reporting, correlation between random variables is a critical input. A sound modeling of correlation is therefore necessary for capturing the relationship between asset returns, particularly when the quantity of concern is sensitive to correlations. It has been well-documented in the literature that correlation is not a constant variable, but a time-varying one (see, for example, [3, 45, 18], among many others). In particular, during periods of financial crisis, it is observed that correlations between asset returns increase (e.g. [38, 6, 44])

In the options market, there is considerable empirical evidence of a large correlation risk premium (e.g. [18, 17, 5]). In particular, in [18], it is established that correlation risk constitutes the missing link between (empirically) un-priced individual variance premium risk and priced market variance risk. Therefore, proper modelling of correlation is important in estimating correlation risk exposure.

There are several methods in the literature that seek to model the stochasticity of correlations. The dynamic conditional correlation approach (see [23, 58]) proposes a class of multivariate GARCH models that have time varying correlations. This class of discrete-time model enjoys popularity, especially in econometric analysis. In the continuous-time literature, the Wishart process [4], sometimes considered a generalization of the Heston model [29], is often used to capture stochastic variances-covariances. This is studied in the context of derivative pricing in such works as [12, 13, 26]. Being an affine specification, the Wishart model has the advantage of a known expression of call and put prices in terms of elementary functions, a feature desired in many situations. However, in this model, variances and covariances must be specified, calibrated and evolved jointly, unlike the dynamic conditional correlation approach. This could be an inconvenient feature when a specific model of the volatility is required.

A third approach to modeling stochastic correlation is to directly specify the dynamics of the correlation variable by a stochastic process (see, for example, [53, 17, 59]). This is also the modeling approach of stochastic correlation that we adopt in this thesis. Compared to Wishart processes, this approach has an advantage of possible separate calibration of parameters. However, under this class of models, expressions of option prices in terms of elementary functions are often unknown, and numerical approximations become necessary.

In the domain of numerical methods, the Monte Carlo (MC) simulation is a popular choice. However, depending on the use case, this approach can have disadvantages, such as slow convergence for problems in low-dimensions, and the limitation that the price is obtained only at a single point in the domain, as opposed to the global character of the PDE approach. In addition, unlike PDE methods, MC simulations usually suffer from difficulty in computing accurate hedging parameters. To our best knowledge, a PDE approach for pricing contingent claims under stochastic correlation has not been investigated in the literature. This forms the motivation for this part of the thesis.

In the remaining parts of the thesis, we will explore the computational aspects of pricing contingent claims when the correlation variable is directly modeled by a mean-reverting stochastic process, with focus on the case of two correlated underlying risk factors. The contributions of this part of the thesis are:

- We derive a time-dependent PDE in three space dimensions of the pricing problem, with the correlation variable being the third dimension. We show that the solution is C^1 in time and C^2 in space.

- We propose a numerical solution to the PDE. We analyze the truncation of the unbounded domain, prove a stability result, and study the boundary conditions and their associated numerical issues, especially those arising from the correlation variable.

- Using singular perturbation theory, we develop and analyze an asymptotic solution of the PDE as the mean reversion rate of the correlation process becomes large. In cases where the price of the contingent claim and its derivatives do not have a known closed-form solution under a constant correlation model, we propose a novel asymptotic solution based on the transition density.

- Through numerical results, we illustrate the accuracy of the numerical PDE and asymptotic solution. We study the effect of certain problem parameters on the solution.

The outline of the remainder of the thesis is as follows. In Section 3.2, we present a pricing model with stochastic correlation and its corresponding PDE formulation. The numerical solution to the PDE is discussed in detail in Chapter 4. In Chapter 5, we discuss an asymptotic

solution using singular perturbation theory. Numerical experiments and results are discussed in Chapter 6.

3.2 Formulation

3.2.1 Model problem: contingent claims on two assets

As a model problem, we consider the pricing of a contingent claim on two (non-dividend-paying) risky assets, whose price processes, denoted by $S_1(t)$ and $S_2(t)$, under the physical measure evolve as follows:

$$\begin{aligned} dS_1(t)/S_1(t) &= \mu_{S_1} dt + \sigma_{S_1} dB_1(t), \\ dS_2(t)/S_2(t) &= \mu_{S_2} dt + \sigma_{S_2} dB_2(t), \\ dB_1(t) dB_2(t) &= \rho(t) dt. \end{aligned} \tag{3.1}$$

Here, $B_1(t)$ and $B_1(t)$ are two correlated Brownian motions, and μ_{S_1}, μ_{S_2}, σ_{S_1}, σ_{S_2} are positive constants. The correlation variable $\rho(t)$ is unobservable, and is assumed to evolve stochastically as

$$d\rho(t) = \alpha(t, \rho(t))dt + \beta(t, \rho(t))dB_3(t), \tag{3.2}$$

where $\alpha(t, \rho(t))$ and $\beta(t, \rho(t))$ are functions that ensure a strong solution to the stochastic differential equation (SDE), and are such that $\rho(t)$ is bounded in $[-1, 1]$ with probability 1. In this specification, the Brownian motion $B_3(t)$ driving the correlation process is assumed to be independent of $B_1(t)$ and $B_2(t)$. While it is possible to include a second layer of correlation structure between the correlation process and the random shocks in asset prices, we shall restrict ourselves to the independence assumption [1].

We are interested in pricing a contingent claim with terminal payoff $g(S_1(T), S_2(T))$, where T is the maturity of the contract. We denote by $V = V(t, S_1(t), S_2(t), \rho(t))$ the time-t value of the contingent claim, $0 \leq t \leq T$. We assume that the value of the contingent claim is Markovian in $(S_1(t), S_2(t), \rho(t))$. We now derive the PDE that governs the price V.

Following a usual "no-arbitrage" argument, we consider a self-financing portfolio consisting of one long unit position in V, (algebraically) short $a_1(t)$ shares of S_1, $a_2(t)$ shares of S_2, and $\Delta(t)$ units of another derivative W on S_1, S_2. We assume also the existence of a money-

[1] A second layer of correlation is considered in [17]. From a computational viewpoint, the effect of this will be extra cross terms in the pricing PDE. While this is seldom a difficulty in practice with a numerical PDE solver, it is not straightforward to interpret and specify second layer correlations. If one would like to specify an instantaneous correlation between $dB_3(t)$ and $dB_i(t)$ (say $(\psi_i, i = 1, 2)$), then clearly $\rho(t)$ has to be bounded within an interval depending on ψ_i in order that the instantaneous correlation matrix is positive semi-definite.

market account, whose value at time t is denoted by $M(t)$, that pays instantaneous interest with rate r. For convenience, we assume $r > 0$, however, our arguments are valid with minor modifications even if r is negative. Denote by $b(t)$ the (algebraically short) position in money-market account.

We denote by $\Pi(t)$ the value of the portfolio. The portfolio value process can be written as

$$\begin{aligned}\Pi(t) &= V(t, S_1(t), S_2(t), \rho(t)) - \Delta(t)W(t, S_1(t), S_2(t), \rho(t)) \\ &\quad -a_1(t)S_1(t) - a_2(t)S_2(t) - b(t)M(t).\end{aligned}$$

As the portfolio is self-financing, its infinitesimal change is

$$\begin{aligned}d\Pi(t) &= dV(t, S_1(t), S_2(t), \rho(t)) - \Delta(t)dW(t, S_1(t), S_2(t), \rho(t)) \\ &\quad -a_1(t)dS_1(t) - a_2(t)dS_2(t) - b(t)dM(t).\end{aligned}$$

By Itô's lemma, $d\Pi(t)$ can be expanded as

$$\begin{aligned}d\Pi(t) &= \left(\left(\frac{\partial}{\partial t} + \tilde{\mathcal{L}}\right)V - \Delta(t)\left(\frac{\partial}{\partial t} + \tilde{\mathcal{L}}\right)W - a_1\mu_{S_1}S_1 - a_2\mu_{S_2}S_2 - b(t)rM(t)\right)dt \\ &+ \sigma_{S_1}S_1\left(\frac{\partial V}{\partial S_1} - \Delta(t)\frac{\partial W}{\partial S_1} - a_1(t)\right)dB_1(t) \\ &+ \sigma_{S_2}S_2\left(\frac{\partial V}{\partial S_2} - \Delta(t)\frac{\partial W}{\partial S_2} - a_2(t)\right)dB_2(t) \\ &+ \beta(t,\rho(t))\left(\frac{\partial V}{\partial \rho} - \Delta(t)\frac{\partial W}{\partial \rho}\right)dB_3(t), \qquad (3.3)\end{aligned}$$

where

$$\begin{aligned}\tilde{\mathcal{L}}V &\equiv \frac{\sigma_{S_1}^2 {S_1}^2}{2}\frac{\partial^2 V}{\partial S_1^2} + \frac{\sigma_{S_2}^2 {S_2}^2}{2}\frac{\partial^2 V}{\partial S_2^2} + \rho\sigma_{S_1}\sigma_{S_2}S_1S_2\frac{\partial^2 V}{\partial S_1 \partial S_2} + \frac{\beta(t,\rho)^2}{2}\frac{\partial^2 V}{\partial \rho^2} \\ &+ \mu_{S_1}S_1\frac{\partial V}{\partial S_1} + \mu_{S_2}S_2\frac{\partial V}{\partial S_2} + \alpha(t,\rho)\frac{\partial V}{\partial \rho},\end{aligned}$$

and similarly for $\tilde{\mathcal{L}}W$. In the above and for the rest of the thesis, where applicable, the dependence on t of $S_1(t)$, $S_2(t)$ and $\rho(t)$ are suppressed for notational convenience.

We choose $a_1(t), a_2(t)$ and $\Delta(t)$ such that the following holds:

$$\begin{aligned} \frac{\partial V}{\partial S_1} - \Delta(t)\frac{\partial W}{\partial S_1} - a_1(t) &= 0, \\ \frac{\partial V}{\partial S_2} - \Delta(t)\frac{\partial W}{\partial S_2} - a_2(t) &= 0, \\ \frac{\partial V}{\partial \rho} - \Delta(t)\frac{\partial W}{\partial \rho} &= 0. \end{aligned} \tag{3.4}$$

These three equations can be solved for $a_1(t), a_2(t), \Delta(t)$. Using these choices, the terms involving real-world drifts μ_{S_1} and μ_{S_2} are cancelled. Also, from the construction of $a_1(t), a_2(t)$ and $\Delta(t)$, the terms involving dB_i, $i = 1, 2, 3$, also disappear from $d\Pi$ in (3.3). As a result, the portfolio is instantaneously riskless. In the presence of the risk-free money-market account paying instantaneous interest r, the following equation has to hold:

$$d\Pi(t) = r\Pi(t)dt = r(V - \Delta(t)W - a_1(t)S_1(t) - a_2(t)S_2(t) - b(t)M(t))dt. \tag{3.5}$$

Therefore, from (3.3)-(3.5), we have

$$\frac{\partial V}{\partial t} + \mathcal{L}V = \Delta(t)\left(\frac{\partial W}{\partial t} + \mathcal{L}W\right), \tag{3.6}$$

where

$$\begin{aligned} \mathcal{L}V \equiv\ & \frac{\sigma_{S_1}^2 {S_1}^2}{2}\frac{\partial^2 V}{\partial S_1^2} + \frac{\sigma_{S_2}^2 {S_2}^2}{2}\frac{\partial^2 V}{\partial S_2^2} + \rho\sigma_{S_1}\sigma_{S_2}S_1S_2\frac{\partial^2 V}{\partial S_1 \partial S_2} + \frac{\beta(t,\rho)^2}{2}\frac{\partial^2 V}{\partial \rho^2} \\ & + rS_1\frac{\partial V}{\partial S_1} + rS_2\frac{\partial V}{\partial S_2} + \alpha(t,\rho)\frac{\partial V}{\partial \rho} - rV, \end{aligned}$$

and similarly for $\mathcal{L}W$. Assuming $\frac{\partial W}{\partial \rho} \neq 0$, we have $\Delta(t) = \frac{\frac{\partial V}{\partial \rho}}{\frac{\partial W}{\partial \rho}}$. As a result, the quantity

$$\frac{\frac{\partial V}{\partial t} + \mathcal{L}V}{\frac{\partial V}{\partial \rho}}$$

is invariant for every derivative of V that has non-zero sensitivity to ρ. Therefore, there exists $\Lambda(t, S_1, S_2, \rho)$ such that

$$\frac{\frac{\partial V}{\partial t} + \mathcal{L}V}{\frac{\partial V}{\partial \rho}} = \Lambda(t, S_1, S_2, \rho).$$

Consequently, we obtain the pricing PDE

$$\begin{aligned}\frac{\partial V}{\partial t}+\frac{\sigma_{S_1}^2 {S_1}^2}{2}\frac{\partial^2 V}{\partial S_1^2}+\frac{\sigma_{S_2}^2 {S_2}^2}{2}\frac{\partial^2 V}{\partial S_2^2}+\rho\sigma_{S_1}\sigma_{S_2}S_1S_2\frac{\partial^2 V}{\partial S_1\partial S_2}+\frac{\beta(t,\rho)^2}{2}\frac{\partial^2 V}{\partial \rho^2}\\ +rS_1\frac{\partial V}{\partial S_1}+rS_2\frac{\partial V}{\partial S_2}+(\alpha(t,\rho)-\Lambda(t,S_1,S_2,\rho))\frac{\partial V}{\partial \rho}=rV.\end{aligned} \tag{3.7}$$

The quantity $\Lambda(t,S_1,S_2,\rho)$ is related to a drift adjustment from the physical measure to a risk-neutral measure. This is seen heuristically as follows. For illustration purposes only, assume for now that $\Lambda(t,S_1,S_2,\rho)$ is of the special form $\Lambda(t,S_1,S_2,\rho)=\phi\beta(t,\rho)$. We have from Itô's lemma that

$$\begin{aligned}&d\left(e^{-rt}V(t,S_1(t),S_2(t),\rho(t))\right)\\ &=e^{-rt}\left(\left(\left(\frac{\partial}{\partial t}+\tilde{\mathcal{L}}\right)V-rV\right)dt+\sigma_{S_1}S_1\frac{\partial V}{\partial S_1}dB_1(t)+\sigma_{S_2}S_2\frac{\partial V}{\partial S_2}dB_2(t)+\beta(t,\rho(t))\frac{\partial V}{\partial \rho}dB_3(t)\right)\\ &=\left(\left(\mu_{S_1}-r\right)S_1e^{-rt}\frac{\partial V}{\partial S_1}+(\mu_{S_2}-r)S_2e^{-rt}\frac{\partial V}{\partial S_2}+\phi\beta(t,\rho)e^{-rt}\frac{\partial V}{\partial \rho}\right)dt\\ &+\sigma_{S_1}S_1e^{-rt}\frac{\partial V}{\partial S_1}dB_1(t)+\sigma_{S_2}S_2e^{-rt}\frac{\partial V}{\partial S_2}dB_2(t)+\beta(t,\rho(t))e^{-rt}\frac{\partial V}{\partial \rho}dB_3(t)\\ &=\sigma_{S_1}S_1e^{-rt}\frac{\partial V}{\partial S_1}\left(dB_1(t)+\frac{\mu_{S_1}-r}{\sigma_{S_1}}dt\right)+\sigma_{S_2}S_2e^{-rt}\frac{\partial V}{\partial S_2}\left(dB_2(t)+\frac{\mu_{S_2}-r}{\sigma_{S_2}}dt\right)\\ &+\beta(t,\rho(t))e^{-rt}\frac{\partial V}{\partial \rho}(dB_3(t)+\phi dt).\end{aligned}$$

We can decompose $dB_2=\rho(t)dB_1(t)+\sqrt{1-\rho(t)^2}dB_2'(t)$ for some $dB_2'(t)$ such that $dB_1(t)$, $dB_2'(t)$ and $dB_3(t)$ are independent Brownian motions. The market price of risk process $\bar{\gamma}=[\gamma_i], i=1,2,3$, which we will explain shortly, is given component-wise by the following:

$$\gamma_1(t)=\frac{\mu_{S_1}-r}{\sigma_{S_1}},\qquad \gamma_2(t)=\frac{1}{\sqrt{1-\rho(t)^2}}\left(\frac{\mu_{S_2}-r}{\sigma_{S_2}}-\rho(t)\gamma_1\right),\qquad \gamma_3(t)=\phi.$$

In turn, the function Λ is related to an equivalent local martingale measure, where the third market price of risk process is given by ϕ (subject to technical conditions that guarantee that the stochastic exponential of the market price of risk processes is a true martingale).

This is to say that γ_1,γ_2 and γ_3 connect the physical measure to a pricing measure. As exlpained in Section 1.2, in a pricing measure we require that $e^{-rt}V$ is a martingale (driftless).

In our case, a necessary condition is that the coefficient of dt in $d(e^{-rt}V)$ is zero. Define now

$$\begin{aligned} d\tilde{B}_1(t) &= dB_1(t) + \gamma_1(t) \\ d\tilde{B}_2(t) &= dB_2'(t) + \gamma_2(t) \\ d\tilde{B}_3(t) &= dB_3(t) + \gamma_3(t). \end{aligned}$$

It is easy to check algebraically that

$$\begin{aligned} d\left(e^{-rt}V(t, S_1(t), S_2(t), \rho(t))\right) &= \sigma_{S_1} S_1 e^{-rt} \frac{\partial V}{\partial S_1} d\tilde{B}_1(t) \\ &\quad + \sigma_{S_2} S_2 e^{-rt} \frac{\partial V}{\partial S_2} \left(\rho(t) d\tilde{B}_1(t) + \sqrt{1-\rho(t)^2} d\tilde{B}_2(t)\right) \\ &\quad + \beta(t, \rho(t)) e^{-rt} \frac{\partial V}{\partial \rho} d\tilde{B}_3(t). \end{aligned}$$

While $\tilde{B}_1, \tilde{B}_2, \tilde{B}_3$ are not necessarily driftless in the original physical measure, if the new probability measure is such that $\tilde{B}_j(t), j = 1, 2, 3$ are independent Brownian motions, then in this probability measure $e^{-rt}V$ is driftless. As a result, we could price using the expectation operator in this measure (subject to other technical conditions). Thus, each triplet $\{\gamma_j(t)\}$ corresponds to a pricing measure, which is not necessarily unique (see Section 1.5). This triplet is sometimes termed the *market price of risk*.

3.2.2 Quanto options

Correlation is also important in other settings. As an illustration, we also present a popular type of contingent claims that decouples equity and FX risk in the terminal payoff, referred to as quanto options. The pricing of such instruments in the presence of stochastic correlation is also studied in [53, 39]. In this context, we denote by $S(t)$ the underlying asset priced in the foreign currency, and by $R(t)$ the spot FX rate which is defined as the number of units of domestic currency per one unit of foreign currency. In a quanto option, the payoff depends on $S(T)$ and a fixed strike, and is paid in the domestic currency. For this kind of options, it is also important to realistically capture the correlation between $S(t)$ and $R(t)$. This is because the currency mismatch in a quanto option gives rise to a "quanto adjustment" that depends heavily on the correlation parameter between $S(t)$ and $R(t)$. Let $S(t)$ and $R(t)$ evolve as

$$\begin{aligned} dS(t)/S(t) &= \mu_S dt + \sigma_S dB_1(t), \\ dR(t)/R(t) &= \mu_R(t, R(t))dt + \sigma_R dB_2(t), \\ dB_1(t)dB_2(t) &= \rho(t)dt, \end{aligned} \tag{3.8}$$

where $\rho(t)$ is as specified before, and μ_S, σ_S, μ_R, and σ_R are positive constants. Under the model (3.8), it can be shown that the price $V(t, S(t), \rho(t))$ of a quanto option satisfies the PDE

$$\frac{\partial V}{\partial t} + \frac{\sigma_S^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + \frac{\beta(t,\rho)^2}{2}\frac{\partial^2 V}{\partial \rho^2} + (r_f - \sigma_S\sigma_R\rho)S\frac{\partial V}{\partial S} + (\alpha(t,\rho) - \Lambda(t,S,R,\rho))\frac{\partial V}{\partial \rho} = r_d V, \tag{3.9}$$

where r_d and r_f are positive constant domestic and foreign risk-free interest rates, respectively.

3.2.3 Correlation process

We now discuss the choice for the correlation process $\rho(t)$. The Jacobi process is a popular choice for modeling stochastic correlation ([17], [59], [39] and [40]). If the correlation is assumed to attain values anywhere in $(-1, 1)$, then the following is a candidate choice of α and β:

$$\alpha(\rho) = \lambda(\eta - \rho), \qquad \beta(\rho) = \sigma_\rho\sqrt{1-\rho^2}. \tag{3.10}$$

Here, λ, η, and σ_ρ are positive constants.

The parameter restriction $\lambda \geq \frac{\sigma_\rho^2}{1\pm\eta}$ is needed for the process to remain in $(-1, 1)$ with probability 1 (see [59]). This parameter restriction also has a PDE interpretation. Given a PDE problem

$$\begin{aligned} \frac{\partial P}{\partial t} &= a(x)\frac{\partial^2 P}{\partial x^2} + b(x)\frac{\partial P}{\partial x} + c(x)P \\ x &\in [x_{\min}, x_{\max}] \end{aligned}$$

where $a(x) \geq 0$ and $a(x_{\min}) = a(x_{\max}) = 0$, define the Fichera function

$$\mathcal{F}(x) = b - \frac{\partial a}{\partial x}.$$

At the boundary $x = x_{\min}$, if $\mathcal{F}(x_{\min}) \geq 0$, then no boundary condition is needed (or allowed) there. Similarly, at the boundary $x = x_{\max}$, if $\mathcal{F}(x_{\max}) \leq 0$, then no boundary condition should be supplied at that point for the formulation of the problem.

Substituting $b(\rho) = \lambda(\eta - \rho)$ and $a(\rho) = \frac{\sigma_\rho^2(1-\rho^2)}{2}$, we have $\mathcal{F}(\rho) = \lambda(\eta - \rho) + \sigma_\rho^2\rho$ and the boundary points relevant to our problem are $\rho_{\max} = 1$ and $\rho_{\min} = -1$. The condition $\mathcal{F}(\rho_{\max}) = \mathcal{F}(+1) \leq 0$ can be written as

$$\lambda(\eta - 1) + \sigma_\rho^2 \leq 0,$$

and the condition $\mathcal{F}(\rho_{\min}) = \mathcal{F}(-1) \geq 0$ can be written as

$$\lambda(\eta + 1) - \sigma_\rho^2 \geq 0.$$

These are identical to the Feller condition $\lambda \geq \frac{\sigma_\rho^2}{1 \pm \eta}$.

Unlike the Wishart process (see for example [11]), this class of models is not affine. In general, a closed-form solution based on elementary functions to (3.7) or (3.9) for a general payoff function is not known, and numerical methods are needed to approximate the solution.

If the correlation process is specified to be bounded in $(\underline{\rho}, \overline{\rho}) \subset (-1, 1)$, then one can consider a process of the form

$$d\rho(t) = \hat{\lambda}(\hat{\eta} - \rho(t)) + \hat{\sigma}_\rho \sqrt{(\rho(t) - \underline{\rho})(\overline{\rho} - \rho(t))} dB_3(t). \tag{3.11}$$

With appropriate parameter restrictions, the process will remain in $(\underline{\rho}, \overline{\rho})$ with probability 1. It is easy to show that (3.11) is equivalent to (3.10) by an affine transformation, and as a result it is sufficiently general to consider (3.10) as our base model.

In the analysis, we will focus on the model PDE (3.7). For pricing purposes and simplicity, for the rest of the thesis, we assume that the specification in (3.10) is risk-neutral, i.e. $\Lambda \equiv 0$.

While this thesis is not concerned with calibration or filtering, we mention that calibration of the specification (3.10) to historical data has been studied in [59].

The following theorem shows that the concept of classical solutions suffices for our purpose.

THEOREM 1. *With $\alpha(\rho)$ and $\beta(\rho)$ chosen as in (3.10), the price of a European contingent claim (with bounded payoff $g(S_1(T), S_2(T))$), given by the discounted risk-neutral expectation*

$$e^{-r(T-t)} \mathbf{E}_{t,x,y,\rho} \left[g(S_1(T), S_2(T)) \right],$$

satisfies (3.7) and is $C^{1,2,2,2}$ on $(0, T] \times (0, \infty)^2 \times (-1, 1)$.

Proof. This theorem establishes that the solution we shall seek with our numerical method is C^2 in space and C^1 in time. Consider the European option price given by

$$v(t, x, y, \rho) = e^{-r(T-t)} \mathbf{E}_{t,x,y,\rho} \left[g(S_1(T), S_2(T)) \right],$$

where we take expectation in the risk-neutral measure induced by ((3.10)). We follow [21] and start by first proving that the option price is continuous on $(0, T) \times (0, \infty)^2 \times (-1, 1)$.

Let (t_n, x_n, y_n, ρ_n) be a sequence of points converging to (t, x, y, ρ). Let S_1^n, S_2^n, and ρ^n be solutions to their corresponding SDEs (3.1) and (3.2) (having parameter choices (3.10)) with

initial conditions $S_1^n(t_n) = x_n$, $S_2^n(t_n) = y_n$, and $\rho^n(t_n) = \rho_n$, respectively. The explicit expressions for $S_1^n(T)$, $S_2^n(T)$ are

$$S_1^n(T) = x_n \exp\left(\frac{\left(r - \sigma_{S_1}^2\right)(T-t_n)}{2} + \sigma_{S_1}\int_{t_n}^T dW_1(s)\right)$$
$$S_2^n(T) = y_n \exp\left(\frac{\left(r - \sigma_{S_2}^2\right)(T-t_n)}{2} + \sigma_{S_2}\int_{t_n}^T \rho^n(s)dW_1(s) + \sigma_{S_2}\int_{t_n}^T \sqrt{1-(\rho^n(s))^2}dW_2(s)\right), \tag{3.12}$$

where $dW_1(s)$ and $dW_2(s)$ are independent Brownian motions adapted to the same filtrations generated by $dB_1(s)$ and $dB_2(s)$. Similarly, define S_1, S_2, and ρ accordingly with the initial conditions $S_1(t) = x$, $S_2(t) = y$, and $\rho(t) = \rho$, respectively. Note that $S_1(T)$, $S_2(T)$ can also be expressed in the same form as (3.12).

By the Yamada-Watanabe theorem, pathwise uniqueness holds for our choice of the correlation model (3.10). It follows from [1] that

$$\mathbf{E}\left[\sup_{s\le T}\left|\rho^n(s) - \rho(s)\right|^2\right] \longrightarrow 0 \quad \text{as} \quad n \to \infty. \tag{3.13}$$

Therefore, using Itô isometry,

$$\begin{aligned}
&\mathbf{E}\left[\left(\int_{t_n}^T \rho^n(s)dW_1(s) - \int_t^T \rho(s)dW_1(s)\right)^2\right] \\
&\qquad \le 2\mathbf{E}\left[\left(\int_{t_n}^T (\rho^n(s) - \rho(s))dW_1(s)\right)^2\right] + 2\mathbf{E}\left[\left(\int_t^{t_n} \rho(s)dW_1(s)\right)^2\right] \\
&\qquad = 2\mathbf{E}\left[\int_{t_n}^T (\rho^n(s) - \rho(s))^2 ds\right] + 2\mathbf{E}\left[\int_t^{t_n} \rho(s)^2 ds\right] \longrightarrow 0
\end{aligned}$$

as $n \to \infty$. We also have

$$\begin{aligned}
&\mathbf{E}\left[\left(\int_{t_n}^T \sqrt{1-(\rho^n(s))^2}dW_2(s) - \int_t^T \sqrt{1-(\rho(s))^2}dW_2(s)\right)^2\right] \\
&\le \mathbf{E}\left[\left(\int_{t_n}^T \sqrt{|(\rho^n(s))^2 - (\rho(s))^2|}dW_2(s)\right)^2\right] + \mathbf{E}\left[\left(\int_t^{t_n} \sqrt{1-(\rho(s))^2}dW_2(s)\right)^2\right] \\
&\le 2\mathbf{E}\left[\left(\int_{t_n}^T |\rho^n(s) - \rho(s)|\, ds\right)\right] + \mathbf{E}\left[\left(\int_{t_n}^t \left(1-(\rho(s))^2\right) ds\right)\right] \longrightarrow 0,
\end{aligned}$$

as $n \to \infty$. This is due to Itô isometry, Lipschitz continuity of the function $f(y) = y^2$ on the domain $[-1, 1]$, and Jensen's inequality. In view of (3.12) above, it follows that $S_1^n(T)$ and $S_2^n(T)$ respectively converge to $S_1(T)$ and $S_2(T)$ in probability, as $n \to \infty$. Since $g(\cdot, \cdot)$ is

bounded, it follows that

$$\mathbf{E}_{t_n,x_n,y_n,\rho_n}\left[g(S_1^n(T),S_2^n(T))\right] \to \mathbf{E}_{t,x,y,\rho}\left[g(S_1(T),S_2(T))\right], \quad \text{as} \quad n\to\infty.$$

Hence, the option price v is continuous.

Next, following Theorem 2.7 in [32], and since v is a continuous (stochastic) solution, it can be shown that v is also a classical solution (in $C^{1,2,2,2}$) that satisfies (3.7).

□

Chapter 4

Finite Difference Approximation

In this chapter, we describe our finite difference approximation to the pricing problem arising from the stochastic correlation model. We discuss some technicalities and develop a stable numerical scheme.

Recall that we assume that the parameters are risk-neutral, i.e. $\Lambda \equiv 0$. Since we solve the PDE (3.7) backward in time, the change of variable $\tau = T - t$ is used. Under this change of variable and with our specification $\alpha(\rho) = \lambda(\eta - \rho)$ and $\beta(\rho) = \sigma_\rho\sqrt{1-\rho^2}$, the PDE (3.7) becomes

$$\frac{\partial V}{\partial \tau} = \mathcal{L}V, \tag{4.1}$$

where $\mathcal{L}$ now takes the form

$$\begin{aligned}\mathcal{L}V \;=\; & \frac{\sigma_{S_1}^2 {S_1}^2}{2}\frac{\partial^2 V}{\partial S_1^2} + \frac{\sigma_{S_2}^2 {S_2}^2}{2}\frac{\partial^2 V}{\partial S_2^2} + \rho\sigma_{S_1}\sigma_{S_2}S_1S_2\frac{\partial^2 V}{\partial S_1 \partial S_2} + \frac{\sigma_\rho^2(1-\rho^2)}{2}\frac{\partial^2 V}{\partial \rho^2} \\ & + rS_1\frac{\partial V}{\partial S_1} + rS_2\frac{\partial V}{\partial S_2} + \lambda(\eta-\rho)\frac{\partial V}{\partial \rho} - rV,\end{aligned}$$

and (4.1) is solved forward in τ. The pricing problem is defined in an unbounded domain

$$\{(\tau, S_1, S_2, \rho) \in (0, T] \times (0, \infty)^2 \times (-1, 1)\}, \tag{4.2}$$

subject to the initial condition $g(\cdot, \cdot)$.

While an implementation of our finite difference scheme may seem straightforward, caution must be taken to ensure proper discretization due to the structure of this problem. Most of the techniques of this section can be modified to apply to other choices of $\alpha(t, \rho)$ and $\beta(t, \rho)$.

4.1 Localization

Localization refers to the truncation of the unbounded domain (in our case, that of the asset prices) to a bounded one for computational purposes. If we knew the *exact* expression of the solution based on elementary functions along the boundaries of the truncated domain , then it would be reasonable to impose a Dirichlet boundary condition. However, this is rarely the case and approximations are to be made at these boundaries. Localization then involves studying additionally the error in the final solution associated with the approximations made on these boundaries.

On the other hand, for financial problems one is usually interested in the solution around a certain region, so it is pointwise error instead of the L^∞ error (in the entire domain) that is of interest. One can take advantage of the fact that our processes are generated by diffusion, and it takes a certain time for the boundary error to propagate to the region of interest[1]. This is the strategy we shall take in this section.

Localization estimates are well-studied in the literature. See, for example, [33] (multi-dimensional Black-Scholes equation), [9] (exponential Lévy models and jump diffusion processes), and [8] (two-asset jump diffusion models), among many others. In [9] and also in [31], for various models (with constant correlations) and different assumption on payoffs, it has been proved that the price of a European option is approximated exponentially well by that of a corresponding barrier option. We now extend this result to the context of stochastic correlation.

For the statement of the result, we will switch to log scaling. We denote $X_t = \log(S_1(t))$ and $Y_t = \log(S_2(t))$, and generic variables $x = \log(S_1), y = \log(S_2)$. Let also $-R^{\log}$ and $R^{\log}$, where $R^{\log} > 0$, respectively denote generic lower and upper barriers for the processes X_t and Y_t.

PROPOSITION 5. *Let $u_{\log}(\tau, x, y, \rho) = e^{-r\tau}\mathbf{E}_{x,y,\rho}\left(g\left(e^{X_\tau}, e^{Y_\tau}\right)\right)$ be the option price with the bounded payoff function g in log scaling ($||g||_\infty < \infty$). Define $M_\tau^x = \sup_{\hat{\tau}\in[0,\tau]}|X_{\hat{\tau}}|$, $M_\tau^y = \sup_{\hat{\tau}\in[0,\tau]}|Y_{\hat{\tau}}|$ and $M_\tau^{x,y} = \max(M_\tau^x, M_\tau^y)$.*

Furthermore, denote by $\hat{\theta}$ is the first exit time of $(X_{\hat{\tau}}, Y_{\hat{\tau}})$, $\hat{\tau} \in [0,\tau]$ from the region $[-R^{\log}, R^{\log}] \times [-R^{\log}, R^{\log}]$.

Let

$$u^1_{R^{\log}}(\tau, x, y, \rho) = e^{-r\tau}\mathbf{E}_{x,y,\rho}\left[g\left(e^{X_\tau}, e^{Y_\tau}\right)\mathbf{1}_{\{M_\tau^{x,y} < R^{\log}\}}\right],$$

$$u^2_{R^{\log}}(\tau, x, y, \rho) = e^{-r\tau}\mathbf{E}_{x,y,\rho}\left[g\left(e^{X_\tau}, e^{Y_\tau}\right)\mathbf{1}_{\{M_\tau^{x,y} < R^{\log}\}} + g\left(e^{X_{\hat{\theta}}}, e^{Y_{\hat{\theta}}}\right)\mathbf{1}_{\{M_\tau^{x,y} \geq R^{\log}\}}\right].$$

[1] This statement is not true for processes with jumps. However, similar techniques do in fact apply to localization estimates for models with jumps.

Then, for $\gamma > 0$, there exists constant $C(\gamma, \sigma_{S_1}, \sigma_{S_2}, r, \tau)$ independent of $R^{\log}$ such that, for $i = 1, 2$,

$$|u_{\log}(\tau, x, y, \rho) - u^i_{R^{\log}}(\tau, x, y, \rho)| \leq C(\gamma, \sigma_{S_1}, \sigma_{S_2}, r, \tau)||g||_\infty(e^{-\gamma(R^{\log}-|x|)} + e^{-\gamma(R^{\log}-|y|)}),$$

pointwise in $(0, T] \times [-R^{\log}, R^{\log}] \times [-R^{\log}, R^{\log}] \times (-1, 1)$.

Proof. By construction

$$|u_{\log}(\tau, x, y, \rho) - u^1_{R^{\log}}(\tau, x, y, \rho)| \leq ||g||_\infty \mathbf{Q}\left(\{M^{x,y}_\tau \geq R^{\log}\}\right),$$

where $\mathbf{Q}$ is the pricing measure. Similarly, we have

$$|u_{\log}(\tau, x, y, \rho) - u^2_{R^{\log}}(\tau, x, y, \rho)| \leq 2||g||_\infty \mathbf{Q}\left(\{M^{x,y}_\tau \geq R^{\log}\}\right).$$

We can write $X_{\hat{\tau}} = x + U_{\hat{\tau}}$ and $Y_{\hat{\tau}} = y + \tilde{U}_{\hat{\tau}}$, $\hat{\tau} \in [0, \tau]$, where $U_{\hat{\tau}}$, and $\tilde{U}_{\hat{\tau}}$ start from 0 and have drifts $r - \frac{\sigma^2_{S_1}}{2}$ and $r - \frac{\sigma^2_{S_2}}{2}$, respectively. We have $M^x_\tau = \sup_{\hat{\tau}\in[0,\tau]} |x + U_{\hat{\tau}}|$. Theorem 25.18 of [50] implies that for any $\gamma > 0$, $C_1(\gamma, \sigma_{S_1}, r, \tau) = \mathbf{E}\left[e^{\gamma \sup_{\hat{\tau}\in[0,\tau]} |U_{\hat{\tau}}|}\right] < \infty$.

Therefore, by the exponential Chebyshev's inequality, for every $R^{\log}_1 > 0$,

$$\mathbf{Q}\left(\left\{\sup_{\hat{\tau}\in[0,\tau]} |U_{\hat{\tau}}| \geq R^{\log}_1\right\}\right) \leq C_1 e^{-\gamma R^{\log}_1}.$$

As a result,

$$\mathbf{Q}\left(\left\{M^x_\tau \geq R^{\log}\right\}\right) \leq C_1(\gamma, \sigma_{S_1}, r, \tau) e^{-\gamma\left(R^{\log}-|x|\right)}.$$

A similar bound can be obtained for $\mathbf{Q}\left(\{M^y_\tau \geq R^{\log}\}\right)$, i.e.

$$\mathbf{Q}\left(\left\{M^y_\tau \geq R^{\log}\right\}\right) \leq C_2(\gamma, \sigma_{S_2}, r, \tau) e^{-\gamma\left(R^{\log}-|y|\right)}.$$

The result follows by noting that

$$\mathbf{Q}\left(\{M^{x,y}_\tau \geq R^{\log}\}\right) \leq \mathbf{Q}\left(\{M^x_\tau \geq R^{\log}\}\right) + \mathbf{Q}\left(\{M^y_\tau \geq R^{\log}\}\right).$$

□

Essentially, Proposition 5 shows that one can set the value at the truncated boundary to zero (which is the value of a barrier option), or, even better, to the discounted payoff function evaluated at that point. If this boundary is far enough from our region of interest, then the error is "small".

More precisely, Proposition 5 shows that the price of a continuously monitored barrier option approximates that of a European option exponentially well by extending the log barrier $R^{\log}$. As only a bounded domain is required to compute the price of a barrier option, in the present case of European options, truncation of the domain is effective when the truncation boundary is far enough from the points of interest.

REMARK 3. *Proposition 5 is formulated in log price scaling. Technically, the equation in log space corresponding to (4.1) is*

$$\frac{\partial u_{\log}}{\partial \tau} = \mathcal{L}_{\log} u_{\log}, \tag{4.3}$$

defined on $(0,T] \times \mathbf{R}^2 \times (-1,1)$ *with initial condition* $u_{\log}(\tau = 0, x, y, \rho) = g(e^x, e^y)$*, where*

$$\begin{aligned} \mathcal{L}_{\log} \equiv\ & \frac{\sigma_{S_1}^2}{2}\frac{\partial^2}{\partial x^2} + \frac{\sigma_{S_2}^2}{2}\frac{\partial^2}{\partial y^2} + \rho\sigma_{S_1}\sigma_{S_2}\frac{\partial^2}{\partial x \partial y} + \frac{\beta(\rho)^2}{2}\frac{\partial^2}{\partial \rho^2} \\ & + \left(r - \frac{\sigma_{S_1}^2}{2}\right)\frac{\partial}{\partial x} + \left(r - \frac{\sigma_{S_2}^2}{2}\right)\frac{\partial}{\partial y} + \alpha(\rho)\frac{\partial}{\partial \rho} - r. \end{aligned}$$

However, if the exact solution based on elementary functions is known explicitly along the $S_i = 0$ *boundaries,* $i = 1, 2$*, such as in the case of spread/basket/exchange options, then the price of the option can be similarly approximated by the price of the corresponding barrier option. Specifically, this can be achieved by solving the original Black-Scholes equation (4.1) in a truncated domain, with approximation error of the order* $O\left(\frac{1}{\left(R^{\log}\right)^{\gamma}}\right)$ *pointwise in the domain. In the case of non-negative risk-neutral drift of the asset prices, this can also be seen from Doob's martingale inequality.*

REMARK 4. *The boundedness condition imposed on the payoff function* g *may seem restrictive. In particular, the analysis in Proposition 5 is applicable to only put payoffs, and not to call ones. However, it is possible to extend the results of Proposition 5 to more general, unbounded payoffs, such as those given in [31]. For example, by invoking put-call parity, the analysis in Proposition 5 can be extended to certain European options with call-type payoffs. In our numerical experiments, in cases with unbounded payoffs, such as a European call option, we do not observe a difficulty, and notice good agreement of the numerical PDE price with that obtained from Monte Carlo simulations. We conjecture that Proposition 5 also holds for contracts with linear growth.*

4.2 Boundary conditions in ρ

To solve the PDE (4.1) numerically by FD methods, we need to truncate the unbounded domain (4.2) into a finite-sized computational one

$$\{(\tau, S_1, S_2, \rho) \in (0, T] \times [0, S_1^{\max}] \times [0, S_2^{\max}] \times (-1, 1)\} \equiv (0, T] \times \Omega, \tag{4.4}$$

where $S_1^{\max}$ and $S_2^{\max}$ are sufficiently large (see Proposition 5). Denote $\partial_{S_1}\Omega = \{(S_1, S_2, \rho) \in \{0, S_1^{\max}\} \times [0, S_2^{\max}] \times [-1, 1]\}$ and similarly for $\partial_{S_2}\Omega$, and $\partial_\rho\Omega$. In the same fashion, we can define a localized spatial domain in the log scaling $\Omega^{\log} = \{(x, y, \rho) \in [-R_1^{\log}, R_1^{\log}] \times [-R_2^{\log}, R_2^{\log}] \times [-1, 1]\}$ and its boundaries.

From Proposition 5, we know the localization errors on the boundaries $\partial_{S_1}\Omega$, and $\partial_{S_2}\Omega$ can be made negligible if $S_1^{\max}$ and $S_2^{\max}$ are chosen sufficiently large. In our experiments, we choose a Dirichlet condition for these boundaries. However, there is a difficulty with choosing the boundary conditions on $\partial_\rho\Omega$, as for an arbitrary option payoff, they are not known as $\rho \to \pm 1$, and unlike the S_1- and S_2- directions we cannot extend the truncation boundary in ρ. This is the focus of the rest of this subsection.

We note that under the choice (3.10) for $\alpha(\cdot)$ and $\beta(\cdot, \cdot)$, the correlation process ρ is of the Jacobi type, which is similar to a CIR process [10]. This type of process is commonly used in modelling interest rate and volatility. In particular, it is reported in [30] that there can be multiple solutions to the CIR bond pricing PDE under certain parameter choices that violate the Fichera condition. As mentioned in Section 3.2.3, boundary conditions should be supplied where the Fichera condition fails. In their case, the reason for having "multiple solutions" to the PDE is that the Fichera condition fails, and no boundary conditions are imposed. However, the result in [30] points out the need to correctly specify a boundary condition when the parameter restrictions are not satisfied.

In [21] (resp. [22]), the authors have studied the problems of boundary behaviors of the discounted risk neutral expectation under stochastic volatility (resp. one factor term structure) models, where at the zero boundary of the variance variable (resp. the short rate variable), a boundary condition is unclear for the specification of the PDE. In their works, they proved that under regularity assumptions such as bounded smooth payoff, linear growth of coefficients, square of volatility of the variance (resp. the short rate) being continuously differentiable with a Hölder continuous derivative, and the vanishing of the volatility of variance (resp. the short rate) as variance (resp. the short rate) goes to zero, the discounted risk neutral expectation is C^1 everywhere in the solution domain up to including the boundary of 0 in volatility (resp. the short rate). Moreover using interior Schauder estimates, it was shown that the solution satisfies a reduced PDE at the boundary, corresponding to the pricing equation with vanishing second

derivative term with respect to the variance (resp. the short rate).

From Theorem 1, the solution we seek is C^2 in the interior of space and C^1 in time. We do not intend to carry out a similar analysis to [21] or [22] here, but instead, we assume that the C^1-ness extends to the ρ-boundaries, noting that the coefficients of the correlation process in our case satisfy the assumptions in [21, 22] . Hence, a similar boundary condition holds for our European option pricing problem:

$$\lim_{\rho\to\pm1} \quad \left(-\frac{\partial V}{\partial \tau} + \frac{\sigma_{S_1}^2 {S_1}^2}{2}\frac{\partial^2 V}{\partial S_1^2} + \frac{\sigma_{S_2}^2 {S_2}^2}{2}\frac{\partial^2 V}{\partial S_2^2} + \rho\sigma_{S_1}\sigma_{S_2}S_1S_2\frac{\partial^2 V}{\partial S_1 \partial S_2} \right. \tag{4.5}$$
$$\left. + rS_1\frac{\partial V}{\partial S_1} + rS_2\frac{\partial V}{\partial S_1} + \lambda(\eta-\rho)\frac{\partial V}{\partial \rho} - rV \right) = 0.$$

Note that, in log-price variables, this becomes

$$\lim_{\rho\to\pm1} \quad \left(-\frac{\partial u_{\log}}{\partial \tau} + \frac{\sigma_{S_1}^2}{2}\frac{\partial^2 u_{\log}}{\partial x^2} + \frac{\sigma_{S_2}^2}{2}\frac{\partial^2 u_{\log}}{\partial y^2} + \rho\sigma_{S_1}\sigma_{S_2}\frac{\partial^2 u_{\log}}{\partial x \partial y} \right. \tag{4.6}$$
$$\left. + \left(r - \frac{\sigma_{S_1}^2}{2}\right)\frac{\partial u_{\log}}{\partial x} + \left(r - \frac{\sigma_{S_2}^2}{2}\right)\frac{\partial u_{\log}}{\partial y} + \lambda(\eta-\rho)\frac{\partial u_{\log}}{\partial \rho} - ru_{\log} \right) = 0.$$

The choice of finite difference discretization of (4.5) (or (4.6)) is to be further elaborated in Section 4.3, with consideration given to numerical stability. It should be noted that the authors of [27], where they consider the Heston-Hull-White three-dimensional PDE, use a similar PDE-based boundary condition for the variance variable at zero.

For the mathematical formulation of the problem, a boundary condition is not needed where the Fichera condition is satisfied (despite being necessary from a numerical perspective). When the Fichera condition is not satisfied, a boundary condition is necessary for specifying the problem. However, as noted in [21, 22], for numerical purposes, we always need to specify a condition on the boundary, even if the Fichera condition is satisfied. We hypothesize that the classical solution would satisfy the limit of the PDE from the interior. In this case, regardless of whether the Fichera condition is satisfied, this is the "correct" boundary condition to use for the numerical approximation.

In summary, the localized problem in the original price scaling is

$$\frac{\partial V}{\partial \tau} = \mathcal{L}V \tag{4.7}$$

on $(0,T] \times (0, S_1^{\max}) \times (0, S_2^{\max}) \times (-1,1)$ subject to the terminal and boundary conditions

$$\begin{aligned} V(\tau = 0, S_1, S_2, \rho) &= g(S_1, S_2) \\ V(\tau, S_1, S_2, \rho) &= V_{Dir}(\tau, S_1, S_2) \text{ on } \partial_{S_1}\Omega \cup \partial_{S_2}\Omega, \end{aligned}$$

where $V_{Dir}(\tau, S_1, S_2)$ is a Dirichlet condition of choice. And finally, V satisfies (4.5) along the spatial ρ-boundary $\partial\Omega \setminus (\partial_{S_1}\Omega \cup \partial_{S_2}\Omega)$. Similar boundary conditions hold for the formulation in log-price space (equation (4.3)).

Implementation-wise, this boundary condition corresponds, to each point along the $\rho = \pm 1$, one additional equation of a discrete version of (4.5) (resp. (4.6)), discretized using one-sided difference in ρ. This is essentially the same equation as (4.1) (resp. (4.3)) applied at $\rho = \pm 1$. Details of discretization are given in the next section.

4.3 Discretization and stability

To obtain a provably monotone discretization, we switch to log scaling (4.3). Let h_1, h_2 and h_3 be step-sizes of a uniform spatial discretization of $\Omega^{\log}$, and $i \in I_1 = \{0, 1, \ldots, n_1\}$, $j \in I_2 = \{0, 1, \ldots, n_2\}$, $k \in I_3 = \{0, 1, \ldots, n_3\}$. Let

$$\Omega^{\Delta,\log} = \{(x_i, y_j, \rho_k) = (-R_1^{\log} + ih_1, -R_2^{\log} + jh_2, -1 + kh_3), i \in I_1, j \in I_2, k \in I_3\}.$$

Recall that we solve the equation (4.3) which is in the log space. We use a finite difference method to obtain a discrete representation $\mathcal{L}^{\Delta}_{\log}$ of $\mathcal{L}_{\log}$ on the discretized grid $\Omega^{\Delta,\log}$. The time dimension is discretized using the θ-timestepping. Let $u^{(l)}_{\log}$ be the vectorized numerical solution at the l-th timestep. At the $(l+1)$-th timestep, the timestepping reads as follows:

$$\frac{u^{(l+1)}_{\log} - u^{(l)}_{\log}}{\Delta\tau} = \theta\mathcal{L}^{\Delta}_{\log}u^{(l+1)}_{\log} + (1-\theta)\mathcal{L}^{\Delta}_{\log}u^{(l)}_{\log}. \tag{4.8}$$

When $\theta = 0.5$, the scheme is known as Crank-Nicolson (CN), and the choice $\theta = 1$ is known as fully implicit timestepping.

Let A be the discretization matrix arising from $\mathcal{L}^{\Delta}_{\log}$ including the boundary conditions. For every $i' \in I_1$, every $j' \in I_2$, and every $k' \in I_3$, we denote $A^{i',j',k'}_{i,j,k}$ to be the matrix element where the index (i,j,k) corresponds to the column index, while (i',j',k') corresponds to the row index in a vectorized ordering of (i,j,k). We require that the discretization matrix satisfies the following:

$$
\begin{aligned}
A^{i,j,k}_{i,j,k} &\leq 0 \text{ for all } (i,j,k), && (4.9)\\
A^{i',j',k'}_{i,j,k} &\geq 0 \text{ for } (i',j',k') \neq (i,j,k) \text{ where equality is component-wise, and} && (4.10)\\
\sum_{i',j',k'} A^{i',j',k'}_{i,j,k} &\leq 0 \text{ for all } (i,j,k). && (4.11)
\end{aligned}
$$

To ensure these properties are satisfied, first and cross derivatives have to be carefully discretized. For first derivatives, one could choose between forward, backward and central differences in such a way that the signs of the matrices are correctly obtained. For cross derivatives, a 7-point stencil can be used according to the sign of the correlation variable, which we will describe shortly.

While in this work we are primarily concerned with classical solutions, maintaining these conditions has the advantage that the discretization matrix arising from the fully implicit timestepping is monotone, a key requirement for convergence towards viscosity solutions. This is particularly relevant for options with early exercise features. We plan to investigate this in a future work. Furthermore, for such a discretization, fully implicit timestepping methods give rise to matrices that have bounded inverse in l^∞ norm.

We define a $(n_3+1)\times(n_3+1)$ matrix ω in the ρ-direction by first defining the three $(n_3+1)\times(n_3+1)$ matrices

$$
\omega^{(c)}_{kl} = \begin{cases}
\frac{\sigma_\rho^2(1-\rho_k^2)}{2h_3^2} - \frac{\lambda(\eta-\rho_k)}{2h_3} & \text{for } l = k-1,\ 0<k<n_3\\
-\frac{\sigma_\rho^2(1-\rho_k^2)}{h_3^2} & \text{for } l = k,\ 0<k<n_3\\
\frac{\sigma_\rho^2(1-\rho_k^2)}{2h_3^2} + \frac{\lambda(\eta-\rho_k)}{2h_3} & \text{for } l = k+1,\ 0<k<n_3\\
0 & \text{else}
\end{cases} \qquad (4.12)
$$

$$
\omega^{(b)}_{kl} = \begin{cases}
\frac{\sigma_\rho^2(1-\rho_k^2)}{2h_3^2} - \frac{\lambda(\eta-\rho_k)}{h_3} & \text{for } l = k-1,\ 0<k\leq n_3\\
-\frac{\sigma_\rho^2(1-\rho_k^2)}{h_3^2} + \frac{\lambda(\eta-\rho_k)}{h_3} & \text{for } l = k,\ 0<k\leq n_3\\
\frac{\sigma_\rho^2(1-\rho_k^2)}{2h_3^2} & \text{for } l = k+1,\ 0<k<n_3\\
0 & \text{else}
\end{cases} \qquad (4.13)
$$

$$\omega_{kl}^{(f)} = \begin{cases} \frac{\sigma_\rho^2(1-\rho_k^2)}{2h_3^2} & \text{for } l = k-1,\ 0 \le k < n_3 \\ -\frac{\sigma_\rho^2(1-\rho_k^2)}{h_3^2} - \frac{\lambda(\eta-\rho_k)}{h_3} & \text{for } l = k,\ 0 \le k < n_3 \\ \frac{\sigma_\rho^2(1-\rho_k^2)}{2h_3^2} + \frac{\lambda(\eta-\rho_k)}{h_3} & \text{for } l = k+1,\ 0 \le k < n_3 \\ 0 & \text{else} \end{cases} \tag{4.14}$$

Then, starting with $\omega = \omega^{(c)}$, rows of ω are modified one-by-one (chosen from any of the three $\omega^{(\cdot)}$) so that off-diagonals are non-negative. The first row of ω (corresponding to $k = 0$) should be based on forward differencing while the last row (corresponding to $k = n_3$) should be based on backward differencing.

Denote $u_{i,j,k} = u(x_i, y_j, \rho_k)$. Now following [8], the cross derivatives are discretized as follows:

$$\begin{aligned} \frac{\partial^2 u}{\partial x \partial y}(x_i, y_j, \rho_k) \approx\ & \frac{1}{2h_1 h_2}(2u_{i,j,k} + u_{i+1,j+1,k} + u_{i-1,j-1,k} \\ & - u_{i+1,j,k} - u_{i-1,j,k} - u_{i,j+1,k} - u_{i,j-1,k}) \quad \text{for } \rho_k \ge 0, \end{aligned} \tag{4.15}$$

$$\begin{aligned} \frac{\partial^2 u}{\partial x \partial y}(x_i, y_j, \rho_k) \approx\ & \frac{1}{2h_1 h_2}(-2u_{i,j,k} - u_{i+1,j-1,k} - u_{i-1,j+1,k} \\ & + u_{i+1,j,k} + u_{i-1,j,k} + u_{i,j+1,k} + u_{i,j-1,k}) \quad \text{for } \rho_k < 0. \end{aligned} \tag{4.16}$$

Intuitively, the case for $\rho \ge 0$ corresponds to the discretization of the cross derivative operator by

$$\frac{1}{2}(\partial_{x-}\partial_{y-} + \partial_{x+}\partial_{y+}),$$

where $\pm$ indicates the direction of the one-sided difference. Similarly, the case for $\rho < 0$ corresponds to the discretization scheme

$$\frac{1}{2}(\partial_{x+}\partial_{y-} + \partial_{x-}\partial_{y+}).$$

Denote by $\chi^{(p)}$ and $\chi^{(m)}$ the induced cross derivative discretization matrices from the $\rho \ge 0$ and $\rho < 0$ cases respectively. The cross derivative discretization can be written as $\frac{\partial^2 u}{\partial x \partial y}(x_i, y_j, \rho_k) \approx \sum_{i',j'} \chi^{(*)}_{ij,i'j'} u_{i',j',k}$. We now describe the discretization in x and y directions.

Denote $\mu_i = r - \frac{\sigma_{S_i}^2}{2}$, $i = 1, 2$. With the discretized domain $\Omega^{\Delta,\log}$ define the following $(n_1+1) \times (n_1+1)$ matrices:

$$\phi_{il}^{(c)} = \begin{cases} \frac{\sigma_{S_1}^2}{2h_1^2} - \frac{\mu_1}{2h_1} & \text{for } l = i-1,\ 1 \le i < n_1 \\ -\frac{\sigma_{S_1}^2}{h_1^2} & \text{for } l = i,\ 1 \le i < n_1 \\ \frac{\sigma_{S_1}^2}{2h_1^2} + \frac{\mu_1}{2h_1} & \text{for } l = i+1,\ 1 \le i < n_1 \\ 0 & \text{else} \end{cases} \tag{4.17}$$

$$\phi_{il}^{(b)} = \begin{cases} \frac{\sigma_{S_1}^2}{2h_1^2} - \frac{\mu_1}{h_1} & \text{for } l = i-1,\ 1 \le i < n_1 \\ -\frac{\sigma_{S_1}^2}{h_1^2} + \frac{\mu_1}{h_1} & \text{for } l = i,\ 1 \le i < n_1 \\ \frac{\sigma_{S_1}^2}{2h_1^2} & \text{for } l = i+1,\ 1 \le i < n_1 \\ 0 & \text{else} \end{cases} \tag{4.18}$$

$$\phi_{il}^{(f)} = \begin{cases} \frac{\sigma_{S_1}^2}{2h_1^2} & \text{for } l = i-1,\ 1 \le i < n_1 \\ -\frac{\sigma_{S_1}^2}{h_1^2} - \frac{\mu_1}{h_1} & \text{for } l = i,\ 1 \le i < n_1 \\ \frac{\sigma_{S_1}^2}{2h_1^2} + \frac{\mu_1}{h_1} & \text{for } l = i+1,\ 1 \le i < n_1 \\ 0 & \text{else} \end{cases} \tag{4.19}$$

The matrix $\phi^{(c)}$ corresponds to the spatial discretization of terms involving $\frac{\partial^2}{\partial x^2}$ and $\frac{\partial}{\partial x}$ using central differences. The matrices $\phi^{(f)}$ and $\phi^{(b)}$ correspond to the discretization of the same terms using forward and backward differences of the first derivatives, respectively.

For each ρ, we define a $(n_1+1)\times(n_1+1)$ matrix $\phi(\rho)$ in the x-direction ($x = \log(S_1)$) which is equal to either $\phi^{(c)}$, $\phi^{(b)}$ or $\phi^{(f)}$ such that non-negativity of the off-diagonals of $\mathcal{L}^{\Delta}$ will be preserved. We can define correspondingly a matrix, say $\psi(\rho)$, for the y-direction ($y = \log(S_2)$).

The discretization matrix A is specified by the action

$$\begin{aligned} (Au)_{i,j,k} &= \sum_{i'=0}^{n_1} \phi_{ii'}(\rho_k) u_{i',j,k} + \sum_{j'=0}^{n_2} \psi_{jj'}(\rho_k) u_{i,j',k} + \sigma_{S_1}\sigma_{S_2}\rho_k (\mathbf{1}_{\rho_k \ge 0} \sum_{i',j'} \chi_{ij,i'j'}^{(p)} u_{i',j',k} \\ &+ \mathbf{1}_{\rho_k < 0} \sum_{i',j'} \chi_{i',j'}^{(m)} u_{i',j',k}) + \sum_{k'=0}^{n_3} \omega_{kk'} u_{i,j,k'} - r u_{i,j,k}, \end{aligned} \tag{4.20}$$

for $i \neq 0, n_1$ and $j \neq 0, n_2$. For $i = 0, n_1$ or $j = 0, n_2$, one can specify $(Au)_{i,j,k} = 0$ and

modify the right-side of the time-stepping equation to the imposed Dirichlet condition. From this discretization one obtains a sufficient condition for (4.10) to hold. Consider the coefficients of $u_{i+1,j,k}$ and $u_{i-1,j,k}$ in (4.20). Clearly, if

$$\frac{\sigma_{S_1}^2}{2}\frac{1}{h_1^2} - \frac{\sigma_{S_1}\sigma_{S_2}|\rho_k|}{2h_1h_2} \geq 0,$$

then one can choose between central, backward and forward differences in the x-direction such that the coefficients of $u_{i+1,j,k}$ and $u_{i-1,j,k}$ are non-negative. This is the case if we space the grid in such a way that

$$\frac{h_1}{h_2} \leq \frac{\sigma_{S_1}}{\sigma_{S_2}}.$$

By considering the non-negativity of the coefficients of $u_{i,j+1,k}$ and $u_{i,j-1,k}$, we get, as a sufficient condition,

$$\frac{\sigma_{S_2}^2}{2}\frac{1}{h_2^2} - \frac{\sigma_{S_1}\sigma_{S_2}|\rho_k|}{2h_1h_2} \geq 0,$$

which is ensured by

$$\frac{h_1}{h_2} \geq \frac{\sigma_{S_1}}{\sigma_{S_2}}.$$

In a similar way, we can study the sign of the other coefficients in (4.20). In summary, if $\frac{h_1}{h_2} = \frac{\sigma_{S_1}}{\sigma_{S_2}}$, and central, backward or forward differences are chosen appropriately in each direction, and cross-derivatives are discretized by (4.16) or (4.17) appropriately, conditions (4.9)-(4.10) are satisfied. It is also easy to see that (4.11) is trivially satisfied for $r > 0$, irrespectively of the central, backward or forward differences in each direction, and the cross-derivative discretization.

THEOREM 2. *Assume fully implicit timestepping ($\theta = 1$) is used for* (4.3)*, $\mathcal{L}^\Delta$ is discretized as above, and that the (x, y) grid satisfies $\frac{h_1}{h_2} = \frac{\sigma_{S_1}}{\sigma_{S_2}}$. Then at each timestep l, we have*

$$||u_{\log}^{(l+1)}||_\infty \leq \max(||u_{\log}^{(l)}||_\infty, ||u_{Dir}^{(l+1)}||_\infty),$$

where $u_{Dir}^{(l+1)}$ is the Dirichlet condition imposed on $\partial_x\Omega^{\log} \cup \partial_y\Omega^{\log}$ during this timestep. The same conclusion holds for Crank-Nicolson timestepping ($\theta = \frac{1}{2}$) with a timestep restriction that scales with $O(\min(h_1^2, h_2^2))$.

Proof. The result follows from the discussion above and more details can be found in [8]. □

Solving a time-dependent PDE in three space dimensions could be expensive on traditional computing architecture. The computational time can be improved by using computing techniques such as implementing an ADI scheme on graphic processing units (see e.g. [14, 15]).

While we have presented Theorem 2 in log-price scaling and for fully implicit timestepping, in practice we have not observed a numerical stability problem when a finite difference scheme is implemented with Crank-Nicolson timestepping, even on the original PDEs (3.7) or (3.9). Note that convection is strong in the ρ direction away from the mean η, and upwind differencing should be utilized where necessary.

Chapter 5

Asymptotic Solution

In practical cases, the implementation of the finite difference approximation could be computationally intensive. It is sometimes possible to develop approximation formulas that are computationally very efficient, even if they make certain simplifying assumptions to reduce the problem. Approximation formulas enjoy the advantage of rapid computation, which is suitable for calibration or in certain pricing situations.

Asymptotic solutions are to be interpreted differently than numerical approximations (such as the finite difference approximation in Chapter 4). While numerical solutions converge to the correct solution as computational efforts increase, asymptotic solutions do not enjoy such convergence. Instead, asymptotic solutions converge to the correct solution when one or more *model parameters* are sufficiently close to a limit. When these parameters are not sufficiently close, the asymptotic solution will have an error that cannot be diminished by increasing computing power.

5.1 Derivation and proof

Following [46], we assume that the mean reversion speed λ in (3.10) is fast, i.e. $\lambda = 1/\epsilon$, where $\epsilon \to 0$. We scale σ_ρ such that the variance of the correlation process' invariant distribution is finite and fixed. Therefore, the volatility of correlation has the corresponding scale $\sigma_\rho = \frac{\tilde{\sigma}_\rho}{\sqrt{\epsilon}}$. Our goal is to obtain an asymptotic solution when λ is large.

Define the differential operators

$$\begin{aligned}\mathcal{A}_0 &= (\eta-\rho)\frac{\partial}{\partial\rho}+\frac{\tilde{\sigma_\rho}^2(1-\rho^2)}{2}\frac{\partial^2}{\partial\rho^2}, \quad \text{and}\\ \mathcal{A}_1 &= \frac{\partial}{\partial t}+\frac{\sigma_{S_1}^2{S_1}^2}{2}\frac{\partial^2}{\partial S_1^2}+\frac{\sigma_{S_2}^2{S_2}^2}{2}\frac{\partial^2}{\partial S_2^2}+\rho\sigma_{S_1}\sigma_{S_2}S_1S_2\frac{\partial^2}{\partial S_1\partial S_2}\\ &\quad +rS_1\frac{\partial}{\partial S_1}+rS_2\frac{\partial}{\partial S_2}-r\mathcal{I},\end{aligned}$$

where $\mathcal{I}$ is the identity operator. Note that $\frac{\partial}{\partial t}+\mathcal{L}=\mathcal{A}_1+\frac{1}{\epsilon}\mathcal{A}_0$.

Intuitively, $\mathcal{A}_1$ is a classical Black-Scholes-type operator (although ρ should be interpreted as a state variable rather than a constant), while $\mathcal{A}_0$ is an operator that is the extra part from the stochastic correlation model.

The pricing equation (4.1) can be written as

$$\mathcal{A}_1V+\frac{1}{\epsilon}\mathcal{A}_0V=0. \tag{5.1}$$

Let V^ϵ be a power series expansion of V in ϵ

$$V^\epsilon=V^{(0)}+\epsilon V^{(1)}+\epsilon^2V^{(2)}+\ldots.$$

We will determine $V^{(0)}$ and $V^{(1)}$. We impose the terminal condition $V^{(0)}(T,S_1,S_2,\rho)=g(S_1,S_2)$. Upon substitution of V^ϵ into (5.1), the following equations are obtained from setting the lower order terms to zero:

$$O\left(\frac{1}{\epsilon}\right) : \mathcal{A}_0V^{(0)}=0 \tag{5.2}$$

$$O(1) : \mathcal{A}_1V^{(0)}+\mathcal{A}_0V^{(1)}=0 \tag{5.3}$$

$$O(\epsilon) : \mathcal{A}_1V^{(1)}+\mathcal{A}_0V^{(2)}=0 \tag{5.4}$$

Equation (5.2) implies that we could choose $V^{(0)}=V^{(0)}(t,S_1,S_2)$, i.e. independent of ρ. This is not the only solution that satisfies the condition, but we shall shortly show that this type of solutions works for our purpose. For $\lambda\geq\frac{2\sigma_\rho^2}{1\pm\eta}$ (equivalently $\tilde{\sigma}^2\leq\frac{1\pm\eta}{2}$), Equation (5.3) implies a centering condition

$$\langle\mathcal{A}_1V^{(0)}\rangle=0,$$

where $\langle\cdot\rangle$ denotes expectation with respect to the invariant distribution Φ of ρ. This is because

(5.3) leads to

$$\langle \mathcal{A}_1 V^{(0)} \rangle = -\int \mathcal{A}_0 V^{(1)} \Phi(\rho) d\rho = \int V^{(1)} \mathcal{A}_0^* \Phi(\rho) d\rho = 0,$$

where $\mathcal{A}_0^*$ is the adjoint of $\mathcal{A}_0$, taking also into account that Φ satisfies the stationary form of the Fokker-Planck equation $\mathcal{A}_0^* \Phi = 0$. The vanishing of the boundary terms requires that $\Phi(\pm 1) = \Phi'(\pm 1) = 0$. This can be satisfied by imposing $\lambda \geq \frac{2\sigma_\rho^2}{1 \pm \eta}$. The explicit form of Φ is given in Appendix A.

As $V^{(0)}$ is independent of ρ, the centering condition implies that

$$\mathcal{A}_{BS}(\bar{\rho}) V^{(0)} = 0,$$

where $\bar{\rho} = \eta$ is the mean of ρ with respect to the invariant distribution (see Appendix A), and that $\mathcal{A}_{BS}(\bar{\rho})$ is the same as $\mathcal{A}_1$, except that ρ is changed to the constant $\bar{\rho}$. Therefore, the zeroth order approximation $V^{(0)}$ is given by the solution to the two-dimensional Black-Scholes equation with constant correlation equal to $\bar{\rho} = \eta$.

We proceed to find $V^{(1)}$. From (5.3) we have

$$\mathcal{A}_0 V^{(1)} = -\mathcal{A}_1 V^{(0)} = -\sigma_{S_1} \sigma_{S_2} S_1 S_2 (\rho - \bar{\rho}) \frac{\partial^2 V^{(0)}}{\partial S_1 \partial S_2},$$

where the equality $\langle \mathcal{A}_{BS}(\bar{\rho}) V^{(0)} \rangle = 0$ is used to eliminate the non-cross derivative terms. Consequently,

$$V^{(1)} = -(\phi(\rho)) \sigma_{S_1} \sigma_{S_2} S_1 S_2 \frac{\partial^2 V^{(0)}}{\partial S_1 \partial S_2} + C(t, S_1, S_2),$$

where $\phi(y)$ is a solution to the equation $\mathcal{A}_0 \phi = \rho - \bar{\rho}$ and $C(t, S_1, S_2)$, arising from the integration of this ODE, is independent of ρ. We take, as a particular solution,

$$\phi(\rho) = \bar{\rho} - \rho = \eta - \rho.$$

Again, we shall shortly show that this particular choice works for our purposes.

Next, we determine $C(t, S_1, S_2)$. Define the differential operator $D_{1,1} \equiv S_1 S_2 \frac{\partial^2}{\partial S_1 \partial S_2}$. One sees immediately that $D_{1,1}$ commutes with $\mathcal{A}_1$. Again, when $\lambda \geq \frac{2\sigma_\rho^2}{1 \pm \eta}$, the solvability of the Poisson equation (5.4) implies

$$\langle \mathcal{A}_1 V^{(1)} \rangle = 0. \tag{5.5}$$

As $C(t, S_1, S_2)$ is independent of ρ, we have by algebraic rearrangements

$$\begin{aligned}
\langle \mathcal{A}_1 V^{(1)} \rangle &= 0 \\
\langle \mathcal{A}_1 C(t, S_1, S_2) \rangle &= \langle \mathcal{A}_1((\eta - \rho)\sigma_{S_1}\sigma_{S_2} D_{1,1} V^{(0)}) \rangle \\
\langle \mathcal{A}_1 \rangle C(t, S_1, S_2) &= \langle (\mathcal{A}_1 - \langle \mathcal{A}_1 \rangle)((\eta - \rho)\sigma_{S_1}\sigma_{S_2} D_{1,1} V^{(0)}) \rangle + \langle \langle \mathcal{A}_1 \rangle ((\eta - \rho)\sigma_{S_1}\sigma_{S_2} D_{1,1} V^{(0)}) \rangle \\
&= \langle (\rho - \eta)(\eta - \rho) \rangle \sigma_{S_1}^2 \sigma_{S_2}^2 D_{1,1}^2 V^{(0)} \\
&= -E\sigma_{S_1}^2 \sigma_{S_2}^2 D_{1,1}^2 V^{(0)},
\end{aligned}$$

where $E = \frac{(1-\eta^2)\tilde{\sigma}_\rho^2}{2+\tilde{\sigma}_\rho^2}$ (see Appendix A). The operator $D_{1,1}^2$ is $D_{1,1}$ applied twice. We specify $C(T, S_1, S_2) = 0$.

By commutativity (or by brute force computation), one can verify that

$$C(t, S_1, S_2) = (T - t) E \sigma_{S_1}^2 \sigma_{S_2}^2 D_{1,1}^2 V^{(0)}$$

is a solution, because $\langle \mathcal{A}_1 \rangle = \mathcal{A}_{BS}(\eta)$, and $\mathcal{A}_{BS}(\eta) V^{(0)} = 0$. Therefore, we obtain the approximation $V^{\epsilon,1}$ to V given by

$$V^{\epsilon,1} = V^{(0)} + \epsilon(\rho - \eta)\sigma_{S_1}\sigma_{S_2} D_{1,1} V^{(0)} + \epsilon(T - t)\frac{(1 - \eta^2)\tilde{\sigma}_\rho^2 \sigma_{S_1}^2 \sigma_{S_2}^2}{2 + \tilde{\sigma}_\rho^2} D_{1,1}^2 V^{(0)}, \tag{5.6}$$

where, as defined earlier, $V^{(0)}$ is the solution to the two-dimensional Black-Scholes equation with constant correlation η. The error of the approximation (5.6) is given in the following theorem.

THEOREM 3. *Assume the payoff function g is smooth and that g and its derivatives have at most polynomial growth as their arguments approach $\pm\infty$.*

Assume also $\lambda \geq \frac{2\sigma_\rho^2}{1 \pm \eta}$ (equivalently $\tilde{\sigma}^2 \leq \frac{1 \pm \eta}{2}$). Then for $t < T$,

$$\left| V(t, S_1, S_2, \rho) - V^{\epsilon,1}(t, S_1, S_2, \rho) \right| = O(\epsilon^2).$$

Proof. Define

$$\mathcal{A}^\epsilon V^\epsilon \equiv \mathcal{A}_1 V^\epsilon + \frac{1}{\epsilon}\mathcal{A}_0 V^\epsilon = 0,$$

and $V^\epsilon = V^{(0)} + \epsilon V^{(1)} + \epsilon^2 V^{(2)} + \ldots$. We write $Z^\epsilon = V^\epsilon - V^{(0)} - \epsilon V^{(1)} - \epsilon^2 V^{(2)}$, where $V^{\epsilon,1}$ is as defined in (5.6). At terminal time, we have

$$Z^\epsilon(T, S_1, S_2, \rho) = -\epsilon(\rho - \eta)\sigma_{S_1}\sigma_{S_2} D_{1,1} V^{(0)}(T, S_1, S_2) - \epsilon^2 V^{(2)}(T, S_1, S_2, \rho),$$

as we specified that $C(T, S_1, S_2) = 0$. In addition, we have

$$\begin{aligned}\mathcal{A}^\epsilon Z^\epsilon &= (\mathcal{A}_1 + \frac{1}{\epsilon}\mathcal{A}_0)(V^\epsilon - V^{(0)} - \epsilon V^{(1)} - \epsilon^2 V^{(2)}) \\ &= -\frac{1}{\epsilon}\mathcal{A}_0 V^{(0)} - (\mathcal{A}_1 V^{(0)} + \mathcal{A}_0 V^{(1)}) - \epsilon(\mathcal{A}_1 V^{(1)} + \mathcal{A}_0 V^{(2)}) - \epsilon^2 \mathcal{A}_1 V^{(2)} \\ &= -\epsilon^2 \mathcal{A}_1 V^{(2)},\end{aligned}$$

where we have used $\mathcal{A}^\epsilon V^\epsilon = 0$ and equations (5.2) to (5.4). Therefore, the probabilistic representation of the solution is

$$\begin{aligned}Z^\epsilon(t, S_1, S_2, \rho) = -\epsilon \mathbf{E}\Big[& e^{-r(T-t)}(\rho_T^\epsilon - \eta)\sigma_{S_1}\sigma_{S_2} D_{1,1} V^{(0)}(T, S_{1,T}^\epsilon, S_{2,T}^\epsilon) \\ & + \epsilon V^{(2)}(T, S_{1,T}^\epsilon, S_{2,T}^\epsilon, \rho_T^\epsilon) \\ & - \epsilon \int_t^T e^{-r(s-t)} \mathcal{A}_1 V^{(2)}(s, S_{1,s}^\epsilon, S_{2,s}^\epsilon, \rho_s^\epsilon) ds \mid S_{1,t}^\epsilon = S_1, S_{2,t}^\epsilon = S_2, \rho_t^\epsilon = \rho \Big],\end{aligned}$$

where subscripts t, s, T indicate time dependence. The dependence of the processes on ϵ is emphasized by superscripts. We bound each of the terms in the following lemmas. These bounds conclude the proof of Theorem 3.

□

LEMMA 1. *We have*

$$\left| \mathbf{E}\left[e^{-r(T-t)}(\rho_T^\epsilon - \eta)\sigma_{S_1}\sigma_{S_2} D_{1,1} V^{(0)}(T, S_{1,T}^\epsilon, S_{2,T}^\epsilon) \middle| S_{1,t}^\epsilon = S_1, S_{2,t}^\epsilon = S_2, \rho_t^\epsilon = \rho \right] \right| \leq C_1 e^{-C_2 \frac{1}{\epsilon}},$$

where $C_1, C_2 > 0$ are independent of ϵ.

Proof. By an argument similar to Lemmas A.1 and A.3 in [24], we can prove that

$$\sup_{\epsilon \leq 1, t \leq s \leq T} \mathbf{E}\left[|D_{1,1} V^{(0)}(s, S_{1,s}^\epsilon, S_{2,s}^\epsilon)| \middle| S_{1,t}^\epsilon = S_1, S_{2,t}^\epsilon = S_2, \rho_t^\epsilon = \rho \right] \leq C',$$

where C' is independent of ϵ. The proof consists of proving that the derivative $D_{1,1}V^{(0)}$ of the Black-Scholes price $V^{(0)}$ is of polynomial growth in S_1 and S_2, which is a consequence of the assumption on the payoff. The bound is obtained from boundedness of the marginal moments of S_1 and S_2, independent of correlation. More details can be found in [24]. The

lemma follows directly from this bound because

$$\begin{aligned}
&\left|\mathbf{E}\left[e^{-r(T-t)}(\rho_T^\epsilon-\eta)\sigma_{S_1}\sigma_{S_2}D_{1,1}V^{(0)}(T,S_{1,T}^\epsilon,S_{2,T}^\epsilon)\middle|S_{1,t}^\epsilon=S_1,S_{2,t}^\epsilon=S_2,\rho_t^\epsilon=\rho\right]\right| \\
\leq\ &\mathbf{E}\left[e^{-r(T-t)}|\rho-\eta|e^{-\frac{T-t}{\epsilon}}\sigma_{S_1}\sigma_{S_2}|D_{1,1}V^{(0)}(T,S_{1,T}^\epsilon,S_{2,T}^\epsilon)|\middle|S_{1,t}^\epsilon=S_1,S_{2,t}^\epsilon=S_2\right] \\
\leq\ &2\sigma_{S_1}\sigma_{S_2}e^{-r(T-t)}e^{-\frac{T-t}{\epsilon}}\mathbf{E}\left[|D_{1,1}V^{(0)}(T,S_{1,T}^\epsilon,S_{2,T}^\epsilon)|\middle|S_{1,t}^\epsilon=S_1,S_{2,t}^\epsilon=S_2\right].
\end{aligned}$$

Note that in the second inequality we have used (A.3). □

LEMMA 2. *We have*

$$\mathbf{E}\left[|V^{(2)}(T,S_{1,T}^\epsilon,S_{2,T}^\epsilon,\rho_T^\epsilon)| \ +|\int_t^T e^{-r(s-t)}\mathcal{A}_1V^{(2)}(s,S_{1,s}^\epsilon,S_{2,s}^\epsilon,\rho_s^\epsilon)ds| \middle| S_{1,t}^\epsilon=S_1,S_{2,t}^\epsilon=S_2,\rho_t^\epsilon=\rho\right]\leq C_3,$$

where C_3 is independent of ϵ.

Proof. By (5.4), and (5.5), we have

$$\mathcal{A}_0V^{(2)}=-\mathcal{A}_1V^{(1)}=-(\mathcal{A}_1V^{(1)}-\langle\mathcal{A}_1V^{(1)}\rangle).$$

Recall that $V^{(1)}=(\rho-\eta)\sigma_{S_1}\sigma_{S_2}D_{1,1}V^{(0)}+(T-t)\frac{(1-\eta^2)\tilde{\sigma}_\rho^2\sigma_{S_1}^2\sigma_{S_2}^2}{2+\tilde{\sigma}_\rho^2}D_{1,1}^2V^{(0)}$. As $V^{(0)}$ is independent of ρ, we have

$$\mathcal{A}_1V^{(1)}-\langle\mathcal{A}_1V^{(1)}\rangle=(\rho(\rho-\eta)-\langle\rho(\rho-\eta)\rangle)\mathcal{K}(V^{(0)}),$$

where $\mathcal{K}(V^{(0)})$ is a linear combination of derivatives of $V^{(0)}$, which similarly has polynomial growth and thereby bounded conditional expectation as in the lemma above. We now write $V^{(2)}=-\psi(\rho)\mathcal{K}(V^{(0)})$, where ψ solves the equation $\mathcal{A}_0\psi=\rho(\rho-\eta)-\langle\rho(\rho-\eta)\rangle$. Denote $g(\rho)=\rho(\rho-\eta)-\langle\rho(\rho-\eta)\rangle$. Observe that $\int_{-1}^1 g(\rho)\Phi(\rho)=\langle g\rangle=0$. Making use of (A.1), we derive

$$\psi'=\frac{2}{\tilde{\sigma}_\rho^2(1-\rho^2)\Phi}\int_{-1}^\rho g(u)\Phi(u)du.$$

We prove ψ' is bounded as $\rho\to-1$. This is seen from the L'Hôspital's rule, where we apply (A.1) once again:

$$\lim_{\rho\to-1}\frac{2}{\tilde{\sigma}_\rho^2(1-\rho^2)\Phi}\int_{-1}^\rho g(u)\Phi(u)du=\lim_{\rho\to-1}\frac{g(\rho)}{(\eta-\rho)}.$$

Similarly $\psi' = \frac{2}{\tilde{\sigma}_\rho^2(1-\rho^2)\Phi}\int_{-1}^{\rho} g(u)\Phi(u)du = \frac{2}{\tilde{\sigma}_\rho^2(1-\rho^2)\Phi}\int_{\rho}^{1} g(u)\Phi(u)du$ is bounded as $\rho \to 1$. As a result, ψ is bounded on $(-1,1)$, where the bound is independent of ϵ, i.e.

$$\psi(\rho) \leq C_4, \rho \in (-1,1). \tag{5.7}$$

Finally $\mathcal{A}_1 V^{(2)}$ and $V^{(2)}$ can be written as a polynomial combination of derivatives of $V^{(0)}$ times the bounded functions $\psi(\rho)$ and $\rho\psi(\rho)$. Hence its conditional expectation is bounded independently of ϵ as desired. □

5.2 Density approximation

The approximation (5.6) requires knowing the price of the derivative and its derivatives under a constant correlation model. This is known, however, only for a few derivatives, such as exchange options (via the Margrabe's formula [41]). For many other popular derivatives, such as spread options, a closed form solution in terms of elementary functions is currently unknown, hence the approximation (5.6) is limited in these cases. To deal with this difficulty, we propose a heuristic solution based on (5.6). The idea is simple. Since the set of European option prices fully determines the transition density, we apply heuristically (5.6) to compute the *transition density* instead, to be further explained below.

Specifically, we denote by $f(T, S_1(T), S_2(T), \rho(T)|t, S_1, S_2, \rho)$, the joint transition density function of the terminal prices $S_1(T)$ and $S_2(T)$ and the correlation value $\rho(T)$, given the asset prices S_1, S_2 and correlation value ρ at an earlier time t. The time-0 option price can be computed by

$$\begin{aligned}&\int_0^\infty \int_0^\infty \int_{-1}^{1} e^{-rT} g\left(S_1(T), S_2(T)\right) f\left(T, S_1(T), S_2(T), \rho(T)|0, S_1, S_2, \rho\right) d\rho(T) dS_1(T) dS_2(T) \\ &\quad = \int_0^\infty \int_0^\infty e^{-rT} g\left(S_1(T), S_2(T)\right) p_m\left(T, S_1(T), S_2(T)|0, S_1, S_2, \rho\right) dS_1(T) dS_2(T),\end{aligned}$$

where $p_m(\cdot|\cdot)$ is the associated marginal transition density. The idea is to approximate $p_m(\cdot|\cdot)$ by $p_m^{\epsilon,1}(\cdot|\cdot)$, a perturbed version of $p_m(\cdot|\cdot)$, obtained from formally applying (5.6) to $p_m(\cdot|\cdot)$. Explicitly, this means

$$p_m^{\epsilon,1} = p_m^{(0)} + \epsilon(\rho - \eta)\sigma_{S_1}\sigma_{S_2} D_{1,1} p_m^{(0)} + \epsilon(T-t)\frac{(1-\eta^2)\tilde{\sigma}_\rho^2 \sigma_{S_1}^2 \sigma_{S_2}^2}{2+\tilde{\sigma}_\rho^2} D_{1,1}^2 p_m^{(0)}. \tag{5.8}$$

We denote by

$$p(T, S_1(T), S_2(T)|t, S_1, S_2)$$

the joint transition density in the case of the constant correlation $\bar{\rho} = \eta$. Its explicit form is known and is given in Appendix B. It is easy to see that the zeroth order approximation $p_m^{(0)}$ of $p_m(\cdot|\cdot)$ is given by $p(\cdot|\cdot)$. The rest of the right-side on (5.8) depends on derivatives of $p(\cdot|\cdot)$. These involve some algebraic work, given also in Appendix B. Finally, we propose approximating the time-0 price of a European option by the double integral

$$\int_0^\infty \int_0^\infty e^{-rT} g\left(S_1(T), S_2(T)\right) p_m^{\epsilon,1}\left(T, S_1(T), S_2(T)|0, S_1, S_2, \rho\right) dS_1(T) dS_2(T), \tag{5.9}$$

which can be computed using quadrature methods. We present the details of this approximation in Appendix B. We demonstrate the effectiveness of this approach in Section 6.
An example algorithm that uses (5.9) for pricing is shown in Algorithm 1.

Algorithm 1 A sample pricing algorithm with the perturbed density at time $t = 0$

1: Retrieve market data: prices S_1, S_2, current correlation ρ, risk-free rate r, mean-reversion level of correlation η applicable forward-looking volatilities $\sigma_{S_1}, \sigma_{S_2}, \sigma_\rho$.
2: Retrieve European option data: payoff $g(\cdot, \cdot)$, maturity T.
3: Obtain the constant correlation (set to mean reversion level η) marginal density $p(T, S_1(T), S_2(T)|0, S_1, S_2)$, and their derivatives $D_{1,1}(p)$ and $D_{1,1}^2(p)$. These formulae are explicitly given in Appendix B.
4: Calculate $p_m^{\epsilon,1}\left(T, S_1(T), S_2(T)|0, S_1, S_2, \rho\right)$ from (5.8).
5: Choose a grid $(0, S_1^{\max}) \times (0, S_2^{\max})$ and discretize.
6: For the particular discretization chosen, calculate the price from (5.9) with truncated S_i-boundaries using quadrature.

We conclude this section by noting that, although we restrict the analysis to non-dividend-paying assets, it is relatively straightforward to generalize the asymptotic solution in the case of non-zero continuous dividend rate by adjusting the risk-neutral drifts of S_1 and S_2.

Chapter 6

Numerical Experiments

In this section, we present numerical results from the implementation of our approach on the following options:

(a) spread and basket options on two assets with stochastic correlation using the three-dimensional PDEs (4.1) and (4.3);

(b) quanto options with stochastic correlation using the two-dimensional PDE (3.9); and

(c) max options on two assets with stochastic correlation using a two-dimensional PDE obtained after a similarity reduction.

For all experiments, the S-boundary conditions are of Dirichlet type where the value on the boundary is simply the discounted payoff for the current values of the state variables. The ρ-boundary conditions are as those described in Section 4.2. See Equations (4.5) and (4.6) for the boundary conditions used in the price space and log-price space formulations, respectively. Similar conditions are used for the two-dimensional PDEs.

Unless otherwise stated, we use standard second order differences for all spatial derivatives, including the cross-derivatives. The only exception is when the convection term $\alpha(t,\rho)\frac{\partial V}{\partial \rho}$ is large, in which case we use forward or backward first order differences, depending on the sign of $\alpha(\cdot,\cdot)$. Whenever solving the log-price space formulation PDE (4.3), the discretization is carried out as described in Section 4.3, so that a monotone discretization scheme is obtained. In all cases, the timestepping is Crank-Nicolson-Rannacher, i.e. we use the fully implicit timestepping for the first few timesteps, then switch to Crank-Nicolson for the remaining timesteps.

While it is easier to carry out stability analysis in log price space, in practice there could be computational disadvantages in solving the log transformed equation. A uniform grid in log price space becomes a non-uniform grid in price space. A fine discretization around the

Volatility of first asset σ_{S_1}	30 %
Volatility of second asset σ_{S_2}	30 %
Mean reversion level of correlation η	0.0
Mean reversion speed of correlation λ	3.0
Volatility of correlation σ_ρ	50 %
Risk-free rate r	5 %

Table 6.1: Market parameters for Section 6.1

region of interest in the price space could require a much finer discretization of the log price grids, and hence higher computational cost. Moreover, for non-smooth payoffs, it could be hard to align the points of discontinuity (of derivatives) with nodes on the log-price grid. For this reason, in our numerical experiments with the log-price space PDE formulation (4.3), an averaging procedure has been applied to smooth out the payoff function (see [47]) and Chapter 2. For these log formulations, we present also results without smoothing as a comparison.

In the numerical experiments, we also report a quantity

$$\Xi_q \equiv \log\left(\left|\frac{\text{change}_{q-1}}{\text{change}_q}\right|\right) / \log(2),$$

for $q > 2$, where change_q is the absolute difference of solution value from the $(q-1)$-th grid refinement to the q-th grid refinement. For the pricing problems we shall consider, no exact solution in terms of elementary functions is known to our knowledge. Therefore, instead of estimating the order of convergence using *errors* as in Chapter 2, we estimate the convergence order using *changes* across grid refinements. From one refinement to the next, the spatial and time stepsizes are halved. For a linearly convergent numerical scheme, we can expect Ξ_q's to be close to 1. For a quadratically convergent numerical scheme, the convergence estimate Ξ_q's would be close to 2.

6.1 Options on two assets

We report numerical results of the proposed method on two types of options on two assets, namely spread and basket options. The market parameters are listed in Table 6.1.

6.1.1 Spread options

As an illustration, we will price a spread option, more specifically a call option on spread. Mathematically, the terminal payoff of such an option is given by

$$g(S_1, S_2) = \max(S_1 - S_2 - K, 0).$$

In log space, the initial condition is

$$g^{\log}(x, y, \rho) = \max(e^x - e^y - K, 0).$$

We will discuss numerical results for the pricing of a 1-year spread option with $K = 10$ for different values of spot prices and current levels of correlation $(S_1(0), S_2(0), \rho(0))$, solved in log space formulation (4.3). We will present results from two experiments, which are identical except for the smoothing of the initial condition. For the smoothing, we will use the first order smoothing operator Φ_1 (instead of the second order operator) defined in Section 2.6. Our choice is motivated by the fact that for exponential call/put type payoffs (in one dimension), first order smoothing is sufficient for removing the dependence of the leading error term on the relative position of the non-smoothness in the grid.

In the numerical experiments, instead of $[-R^{\log}, R^{\log}] \times [-R^{\log}, R^{\log}] \times [-1, 1]$ for a sufficiently large $R^{\log} > 0$ as in Proposition 5, we have localized the grid to $[1, 5] \times [1, 5] \times [-1, 1]$ in the log grid for more efficient use of computing power. As indicated, the numerical method exhibits second-order convergence.

While an analytic approach is possible by evaluating the smoothing integral similarly to the operator Φ_1 in Section 2.6, we carry out a discrete approximation to the integral by averaging the values of the payoff function at the corners of 5-point stencils, on each point of the domain of evaluation, described as follows. Given a point (x_i, y_j, ρ_k) in the grid, we divide the region $[x_i - h_1, x_i + h_1] \times [y_j - h_2, y_j + h_2] \times \{\rho_k\}$ into sub-rectangles. Let $x^*_{m_1,i} = x_i - h_1 + \frac{2m_1h_1}{J}$, and $y^*_{m_2,j} = y_j - h_2 + \frac{2m_2h_2}{J}$, where $m_1, m_2 = 0, \ldots, J$, be quadrature data points of the sub-rectangles. In our experiments, J is taken to be 30. Mathematically, the initial condition is approximated by

$$g^{\log}_{\text{smoothed}}(x_i, y_j, \rho_k) \quad = \quad \frac{1}{4h_1h_2} \int_{-h_1}^{h_1} \int_{-h_2}^{h_2} g^{\log}(x_i - u, y_j - v, \rho_k) dv du,$$

while $g^{\log}_{\text{smoothed}}$ is approximated by quadrature with

$$g^{\log}_{\text{smoothed, discrete}}(x_i, y_j, \rho_k) = \frac{1}{4J^2}\sum_{m_1=0}^{J-1}\sum_{m_2=0}^{J-1}(g^{\log}(x^*_{m_1,i}, y^*_{m_2,j}, \rho_k) + g^{\log}(x^*_{m_1+1,i}, y^*_{m_2,j}, \rho_k) + g^{\log}(x^*_{m_1,i}, y^*_{m_2+1,j}, \rho_k) + g^{\log}(x^*_{m_1+1,i}, y^*_{m_2+1,j}, \rho_k)).$$

In Chapter 2, our analysis for the one-dimensional Black-Scholes equation suggests that in the presence of such non-smoothness, smoothing by averaging helps remove the dependence of the leading error term on the relative position of the non-smoothness on the grid. This is more important in the two-dimensional problem, as there is no trivial way to maintain the line of non-smoothness on a fixed relative position on the grid, except possibly for a coordinate rotation.

Table 6.2 shows the solution values and the approximate rates of convergence, where we find again that due to the non-smoothness, it is not necessarily true that a smaller step-size leads to a smaller (approximate) error. The rate of convergence by this measure is sometimes far from the theoretical 2.0 for the unsmoothed initial condition. In Table 6.3, results are shown for the smoothed initial condition. It is observed that the error, approximated as the change from one grid resolution to the next, seems smaller for the unsmoothed run. However, the convergence rate is much closer to the theoretical 2.0 in Table 6.3. While smoothing does not necessarily lead to a better error, the resulting convergence table is more suitable for extrapolation.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
20	10	0.1	4.5747	1.6392	7.8346	4.1289	1.3861	7.3850
40	20	0.05	4.7026	1.6392	7.9954	4.2397	1.3765	7.5325
80	40	0.025	4.7321	1.6352	8.0356	4.2644	1.3679	7.5700
Ξ_3			2.12	-10.23	2.00	2.16	0.17	1.98

Table 6.2: Value of the 1-year spread option with $K = 10$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the log-price space formulation PDE (4.3). The domain is $\Omega^{\Delta,\log} = [1,5] \times [1,5] \times [-1,1]$. Nodes are placed uniformly, with n_1 (resp. n_2, n_3) being the number of subintervals in the $\log(S_1)$ (resp. $\log(S_2), \rho$) direction. Initial condition is not smoothed.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
20	10	0.1	5.2364	1.9929	8.5386	4.8107	1.7329	8.1214
40	20	0.05	4.8547	1.7250	8.1565	4.3976	1.4608	7.7022
80	40	0.025	4.7697	1.6566	8.0751	4.3036	1.3891	7.6117
Ξ_3			2.17	1.97	2.23	2.13	1.93	2.21

Table 6.3: Value of the 1-year spread option with $K = 10$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the log-price space formulation PDE (4.3). The domain is $\Omega^{\Delta,\log} = [1,5] \times [1,5] \times [-1,1]$. Nodes are placed uniformly, with n_1 (resp. n_2, n_3) being the number of subintervals in the $\log(S_1)$ (resp. $\log(S_2), \rho$) direction. Initial condition is smoothed.

In Table 6.4, we present similar results as in Table 6.3, but from solving the price space formulation PDE (4.1), with an unsmoothed initial condition. The numerical results indicate that solving directly (4.1) does not seem to pose a problem in terms of stability, although a positive discretization is no longer guaranteed.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
20	10	0.1	4.5456	1.5370	7.7828	4.0878	1.2727	7.3315
40	20	0.05	4.7035	1.6080	7.9973	4.2367	1.3385	7.5352
80	40	0.025	4.7418	1.6274	8.0478	4.2728	1.3566	7.5829
Ξ_3			2.04	1.87	2.09	2.05	1.86	2.10

Table 6.4: Value of the 1-year spread option with $K = 10$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the PDE in price space formulation (4.1). The domain is $\Omega^{\Delta} = [0,200] \times [0,200] \times [-1,1]$. Nodes are placed uniformly, with n_1 (resp. n_2, n_3) being the number of subintervals in the S_1 (resp. S_2, ρ) direction. Initial condition is not smoothed.

To validate our approaches, we compare the numerical PDE prices with those obtained by the asymptotic solution (Chapter 5) and Monte Carlo (MC) simulations. To obtain more accurate PDE prices, the PDE solutions from Tables 6.3 and 6.4 are extrapolated using Richardson extrapolation, with convergence exponent 2, as the method is supposed and has demonstrated to achieve. The numerical results are given in Table 6.5. They show good agreement among solutions under the various approaches. In particular, the PDE and asymptotic solutions all lie in the MC's 95% confidence intervals (CIs). Here, for MC simulation, 50000 scenarios and 200 timesteps are used.

	(50,50,-0.2)	(40,50,-0.2)	(50,40,-0.2)	(50,50,0.2)	(40,50,0.2)	(50,40,0.2)	Runtimes (s)
log-price space PDE (4.3)	4.7413	1.6338	8.0480	4.2722	1.3652	7.5815	4823
price space PDE (4.1)	4.7545	1.6338	8.0646	4.2848	1.3627	7.5988	3197
Asymptotic solution	4.7592	1.6412	8.0672	4.2666	1.3546	7.5797	0.40
MC 95% CI	[4.6772, 4.8543]	[1.5965, 1.6919]	[7.9666, 8.1867]	[4.2106, 4.3744]	[1.3266, 1.4111]	[7.5011, 7.7095]	0.96

Table 6.5: Value comparison and runtimes for the 1-year spread option with $K = 10$ at different values of $(S_1(0), S_2(0), \rho(0))$. Both sets of PDE prices are extrapolated from respective data in Tables 6.3-6.4, using Richardson extrapolation, assuming quadratic convergence.

In Table 6.5, we also report runtimes in seconds for a typical run on an architecture of two 12-core Intel E5-2697v2 CPUs, with a total of 128GB memory. Sparse matrices are solved with Matlab backslash. For the numerical PDE solutions, the runtimes for the finest grid in Tables 6.3 and 6.4 are reported, and it is worth noting that the values on the entire grid are computed. For Monte Carlo and asymptotic methods, the runtime is based on a single run at a given $(S_1(0), S_2(0), \rho(0))$. It is also worth nothing that solving the linear system arising from the PDE discretization with the backslash operation in Matlab is not the most optimized solution technique for this system. While this is not the topic of the thesis, in the future, we will be investigating more efficient techniques for solving the linear system, including preconditioned iterative solvers and Alternating Direction Implicit (ADI) techniques, as well as their parallel versions.

6.1.2 Effect of truncated boundary

In this section, we verify numerically, that the truncated boundary in the S_1- and S_2-directions is far enough, so that the quality of the approximation is not affected. We price again the spread option of Section 6.1.1 by solving the price space formulation PDE (4.1), this time with the truncated boundary in the S_1- and S_2-directions double as far, and with double the number of grid points in these two directions. Table 6.6 presents the results. As it can be seen, the differences between the results of Tables 6.6 and 6.4 are approximately at the level of 10^{-4}. Furthermore, the accuracy of the results of Table 6.4 seems to be at most at the level of 10^{-2}. These results indicate that the truncated boundary chosen for the experiment of Table 6.4 does not compromise the quality of the numerical PDE approximation.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
40	10	0.05	4.5504	1.5372	7.7888	4.0927	1.2723	7.3378
			4.8e-03	1.3e-04	6.0e-03	4.9e-03	-3.6e-04	6.3e-03
80	20	0.025	4.7046	1.6081	7.9987	4.2379	1.3385	7.5367
			1.2e-03	1.4e-04	1.4e-03	1.2e-03	1.5e-05	1.5e-03
160	40	0.0125	4.7420	1.6274	8.0481	4.2731	1.3566	7.5833
			2.9e-04	4.1e-05	3.4e-04	3.0e-04	8.0e-06	3.6e-04

Table 6.6: Value of the 1-year spread option with $K = 10$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the PDE in price space formulation (4.1). The domain is $\Omega^{\Delta} = [0, 400] \times [0, 400] \times [-1, 1]$. Below each line of values, the differences from the respective values of Table 6.4 are also presented. Initial condition is not smoothed.

We also note that, although the experiment corresponding to the results in Table 6.6 is costly, since the number of grid points has quadrupled compared to Table 6.4, it is a "fair" comparison as far as the effect of truncated boundary, since the grid points in Table 6.6 are a proper superset of those in Table 6.4.

6.1.3 Non-uniform mesh

In this section, we report the results for solving the price space PDE (4.1) using a non-uniform grid as opposed to a uniform grid as in the previous experiments. One of the conditions that guarantees monotonicity of the discretization of the log price PDE is a mesh ratio condition that is fairly restrictive. In this experiment, we numerically demonstrate that, while the resulting discretization may not be monotone, a non-uniform mesh in price space works well in practice.

For the experiment, we concentrate the (S_1, S_2) grid on the region around the points of evaluation ($S_1 = 50$ and $S_2 = 50$). The non-uniform grid is obtained by mapping the uniform grid with the indicated number of points into a non-uniform grid via the mapping defined in [7].

The results are reported in Table 6.7. We observe that, due to a finer grid around the region of evaluation, the numerical approximation seems to exhibit a smaller error that those in Table 6.4, measured by the change of solution value from one run to the next. This suggests that a non-uniform mesh in the (S_1, S_2)-dimensions may be more efficient in practice, despite the fact that monotonicity of discretization is not satisfied. Essentially, this experiment suggests that monotonicity is a sufficient, but not necessary condition for convergence and stability.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
20	10	0.1	4.7092	1.6169	8.0253	4.2390	1.3453	7.5611
40	20	0.05	4.7440	1.6300	8.0553	4.2743	1.3587	7.5900
80	40	0.025	4.7518	1.6329	8.0620	4.2821	1.3617	7.5963
Ξ_3			2.15	2.17	2.15	2.17	2.15	2.19

Table 6.7: Value of the 1-year spread option with $K = 10$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the PDE in price space formulation (4.1). The domain is $\Omega^{\Delta} = [0, 200] \times [0, 200] \times [-1, 1]$. The mesh is non-uniform in the (S_1, S_2)-dimensions, and concentrated around (50, 50), and uniform in the ρ-dimension, with n_1 (resp. n_2, n_3) being the number of subintervals in the S_1 (resp. S_2, ρ) direction. Initial condition is not smoothed.

6.1.4 Basket options

We also consider an equal-weighted basket call option whose payoff is

$$g(S_1, S_2) = \max(S_1 + S_2 - K, 0).$$

In Tables 6.8, 6.9 and 6.10, we present convergence results for the two versions of PDE for different values of $(S_1(0), S_2(0), \rho(0))$ with the strike $K = 100$. Again, while a monotone discretization scheme is no longer guaranteed in price space, numerically we do not observe a problem with this approach. Comparison of values using the PDE, Monte Carlo and asymptotic methods is presented in Table 6.11. Given the symmetry of the problem, the solution should be symmetric in $(S_1(0), S_2(0))$. The discrepancy in Monte Carlo solutions is due to randomness from the simulations.

Table 6.8 shows the solution values and the approximate rates of convergence for the unsmoothed initial condition. In Table 6.9, results are shown for the smoothed initial condition, with a discrete averaging procedure as in the case of spread options. The convergence rate for the smoothed initial condition run is much closer to the theoretical 2.0 than the unsmoothed. However, it is observed that the error, approximated as the change from one iteration to the next, seems smaller for the unsmoothed run.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
20	10	0.1	10.4900	5.3185	5.3185	10.9561	5.7400	5.7400
40	20	0.05	10.6550	5.3793	5.3793	11.1369	5.8185	5.8185
80	40	0.025	10.7000	5.3974	5.3974	11.1840	5.8410	5.8410
Ξ_3			1.87	1.75	1.75	1.94	1.80	1.80

Table 6.8: Value of the 1-year equal-weighted basket call option with strike $K = 100$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the log-price space formulation PDE (4.3). The domain is $\Omega^{\Delta,\log} = [1, 5] \times [1, 5] \times [-1, 1]$. Nodes are placed uniformly, with n_1 (resp. n_2, n_3) being the number of subintervals in the $x = \log(S_1)$ (resp. $y = \log(S_2), \rho$) direction. Initial condition is not smoothed.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
20	10	0.1	11.6381	6.2168	6.2168	12.0654	6.6130	6.6130
40	20	0.05	10.9276	5.5998	5.5998	11.3987	6.0315	6.0315
80	40	0.025	10.7661	5.4514	5.4514	11.2475	5.8931	5.8931
Ξ_3			2.14	2.06	2.06	2.14	2.07	2.07

Table 6.9: Value of the 1-year equal-weighted basket call option with strike $K = 100$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the log-price space formulation PDE (4.3). The domain is $\Omega^{\Delta,\log} = [1, 5] \times [1, 5] \times [-1, 1]$. Nodes are placed uniformly, with n_1 (resp. n_2, n_3) being the number of subintervals in the $x = \log(S_1)$ (resp. $y = \log(S_2), \rho$) direction. Initial condition is smoothed.

n_1, n_2	n_3	Δt	(50, 50, -0.2)	(40, 50, -0.2)	(50, 40, -0.2)	(50, 50, 0.2)	(40, 50, 0.2)	(50, 40, 0.2)
20	10	0.1	10.4788	5.2020	5.2020	10.9576	5.6262	5.6262
40	20	0.05	10.6710	5.3601	5.3601	11.1537	5.8006	5.8006
80	40	0.025	10.7165	5.3985	5.3985	11.2001	5.8428	5.8428
Ξ_3			2.08	2.04	2.04	2.08	2.05	2.05

Table 6.10: Value of the 1-year equal-weighted basket call option with strike $K = 100$ at different values of $(S_1(0), S_2(0), \rho(0))$, in three successive grid refinements, using the price space formulation PDE (4.1). The domain is $\Omega^{\Delta} = [0, 200] \times [0, 200] \times [-1, 1]$. Nodes are placed uniformly, with n_1 (resp. n_2, n_3) being the number of subintervals in the S_1 (resp. S_2, ρ) direction. Initial condition is not smoothed.

	(50,50,-0.2)	(40,50,-0.2)	(50,40,-0.2)	(50,50,0.2)	(40,50,0.2)	(50,40,0.2)
log-price space PDE (4.3)	10.7123	5.4019	5.4019	11.1971	5.8469	5.8469
price space PDE (4.1)	10.7317	5.4113	5.4113	11.2156	5.8569	5.8569
Asymptotic solution	10.7131	5.3945	5.3945	11.2199	5.8616	5.8616
MC 95% CI	[10.5006, 10.7668]	[5.2276, 5.4151]	[5.2429, 5.4309]	[10.9719, 11.2542]	[5.6614, 5.8629]	[5.6782, 5.8804]

Table 6.11: Value comparison for the 1-year equal-weighted basket call option with $K = 100$ at different values of $(S_1(0), S_2(0), \rho(0))$. Both sets of PDE prices are extrapolated from respective data in Tables 6.9-6.10 using Richardson extrapolation, assuming quadratic convergence.

We conclude this subsection by noting that, although call payoffs do not satisfy the boundedness required in Proposition 5, numerically, we do not observe any problem with convergence, and good agreement is achieved among different approaches. Also see Remark 4.

6.2 Quanto options

In this section, we consider the pricing of a quanto option under stochastic correlation as in model (3.8) and (3.10). Similar to previous experiments, we assume that the parameters are calibrated so that $\Lambda \equiv 0$. As an illustration, we price a 5-year quanto call option with payoff

$$g(S(T)) = \max(S(T) - K, 0),$$

where $K = 100$. The parameters to the model are given in Table 6.12.

Volatility of price σ_S	30 %
Volatility of exchange rate σ_R	10 %
Domestic risk-free rate r_d	5 %
Foreign risk-free rate r_f	3 %
Mean reversion rate of correlation λ	3.0
Mean reversion level of correlation η	-0.1
Volatility of correlation σ_ρ	30 %

Table 6.12: Market parameters for quanto option

In this case, the pricing PDE is (3.9). Because there is no cross term in the PDE, by discretizing the first derivatives carefully in the upwind direction when necessary, a positive

discretization can be obtained without grid-size restrictions. Therefore, techniques such as non-uniform spacing can be used without sacrificing a positive discretization. In our experiment, we use a non-uniform grid that concentrates around the strike value 100 (see, e.g. [7]).

Results in Table 6.13 show that the PDE method has numerically quadratic convergence. Table 6.14 shows good agreement among solutions obtained from different techniques.

n_1	n_3	Δt	$(100, -0.2)$	$(120, -0.1)$	$(90, 0.0)$
50	10	0.5	29.9609	43.9209	23.4803
100	20	0.25	29.9843	43.9513	23.5008
200	40	0.125	29.9894	43.9578	23.5052
400	800	0.0625	29.9907	43.9595	23.5064
Ξ_3			2.21	2.22	2.22
Ξ_4			1.91	1.90	1.91

Table 6.13: Value of a 5-year quanto call option with strike $K = 100$ at different values of $(S(0), \rho(0))$, in four successive grid refinements, using PDE (3.9). The domain is $\Omega^{\Delta} = [0, 500] \times [-1, 1]$. There are n_1 (resp. n_3) subintervals in the S (resp. ρ) direction.

	(100,-0.2)	(120, -0.1)	(90, 0.0)
PDE (3.9)	29.9910	43.9599	23.5067
Asymptotic solution	29.9674	43.8848	23.4952
MC 95% CI	[29.6126, 30.6397]	[43.3529, 44.6529]	[23.1074, 23.9883]

Table 6.14: Value comparison for the 5-year quanto call option with strike $K = 100$ at different values of $(S(0), \rho(0))$. The set of PDE prices is extrapolated from respective data in Table 6.13 using Richardson extrapolation, assuming quadratic convergence.

In Table 6.15, we present values of selected partial derivatives of the option price with respect to the underlying asset price S and the correlation factor ρ. Note that, although second order convergence is not guaranteed due to the choice of upwind differencing, in most cases we obtain second order convergence.

n_1	n_3	Δt	$(100, -0.2)$	$(120, -0.1)$	$(90, 0.0)$
$\partial^2 V/\partial S^2$					
50	10	0.5	4.5994×10^{-3}	3.1606×10^{-3}	5.5495×10^{-3}
100	20	0.25	4.6107×10^{-3}	3.1625×10^{-3}	5.5434×10^{-3}
200	40	0.125	4.6129×10^{-3}	3.1630×10^{-3}	5.5444×10^{-3}
400	800	0.0625	4.6134×10^{-3}	3.1631×10^{-3}	5.5445×10^{-3}
Ξ_3			2.31	2.03	2.63
Ξ_4			2.36	1.76	2.73
$\partial V/\partial \rho$					
50	10	0.5	-6.6204×10^{-1}	-8.8469×10^{-1}	-5.4737×10^{-1}
100	20	0.25	-6.6191×10^{-1}	-8.8515×10^{-1}	-5.4806×10^{-1}
200	40	0.125	-6.6194×10^{-1}	-8.8526×10^{-1}	-5.4817×10^{-1}
400	800	0.0625	-6.6196×10^{-1}	-8.8529×10^{-1}	-5.4818×10^{-1}
Ξ_3			1.80	2.05	2.70
Ξ_4			0.94	2.01	2.78

Table 6.15: Selected sensitivities of the option price with respect to S and ρ.

6.3 Effects of model parameters

In this section, the effect of correlation model parameters on option prices is studied. We will focus on the max option, which has payoff

$$V(t = T, S_1(T), S_2(T)) = \max(S_1(T), S_2(T)). \tag{6.1}$$

For options of this form, it is not necessary to solve the full three-dimensional PDE (4.1). Instead, a similarity reduction is possible because of the nature of the payoff. For $\tau = T - t > 0$, define $W(\tau, S_1, S_2)$ by

$$V(\tau, S_1, S_2) = S_1 W(\tau, S_1, S_2).$$

Introduce the similarity reduction $\xi = S_2/S_1$, corresponding to a change of numéraire. It is now straightforward to see that

$$W_\tau = \frac{(\sigma_{S_1}^2 + \sigma_{S_2}^2 - 2\rho\sigma_{S_1}\sigma_{S_2})\xi^2}{2} W_{\xi\xi} + \frac{\beta^2}{2} W_{\rho\rho} + \alpha W_\rho, \tag{6.2}$$

where α, β are chosen as in (3.10). The reduced problem has terminal condition

$$W(\tau = 0, \xi) = \max(1, \xi).$$

The diffusion coefficient in (6.2) is non-negative because $\sigma_{S_1}^2 + \sigma_{S_2}^2 - 2\rho\sigma_{S_1}\sigma_{S_2} \geq 0$. It should be noted that r, the risk-free rate, factors out of the pricing problem naturally. This similarity reduction can be easily seen to be the PDE equivalent to a measure change from the T-forward measure to the S_1-measure.

For illustration purposes we have restricted to non-dividend-paying assets. For dividend-paying assets, the same variable transformation can be carried out. In that case, a PDE two-dimensional in space similar to (6.2) will be obtained, with a convection and a discounting term.

The discretization of (6.2) is less restricted than that of (4.1). A positive discretization is ensured with usual central differences of the second derivatives and careful discretization of the first derivatives in the upwind direction where necessary. In our experiment, we solve (6.2) on a uniform grid of 801×81 nodes in $[0, 5] \times [-1, 1]$, with timestep 0.025.

In Figure 6.1, we show the effect of the long term mean reversion level η of correlation on the price of the max option. With higher η, the expected value of the total correlation experienced during the life of the option is increased. Naturally this leads to a lower value of the optionality. This effect is captured also by the asymptotic solution (5.6), which we restate here:

$$V^{\epsilon,1} = V^{(0)} + \epsilon(\rho - \eta)\sigma_{S_1}\sigma_{S_2}D_{1,1}V^{(0)} + \epsilon(T - t)\frac{(1 - \eta^2)\tilde{\sigma}_\rho^2\sigma_{S_1}^2\sigma_{S_2}^2}{2 + \tilde{\sigma}_\rho^2}D_{1,1}^2V^{(0)}. \tag{6.3}$$

Recall $V^{(0)}$ is the Black-Scholes price with constant correlation η. The Black-Scholes sensitivity $\frac{\partial V^{(0)}}{\partial \eta}$ is negative, and the dominant zeroth order term decreases in value as η increases, for $\epsilon \ll 1$. The effect of the spot-correlation is present in the first order correction (second term in the above equation), and is of order ϵ for fast mean-reversion. The sensitivity $D_{1,1}V^{(0)}$ of the Black-Scholes price is negative as it is a positive multiple of $\frac{\partial V^{(0)}}{\partial \eta}$. This explains the three decreasing set of prices for different values of $\rho(0) = \rho_0$.

In Figure 6.2, we show the effect of the mean reversion speed λ on the prices of the max option. As λ increases, it is plausible that any deviation from the mean of the correlation is more heavily punished with the stronger convection, and one should expect a price closer to the Black-Scholes price with constant correlation equal to the long-term mean. Once again this is encoded in (6.3). The terms involving ϵ decrease in absolute value and convergence towards $V^{(0)}$, the Black-Scholes price with constant correlation η, is expected.

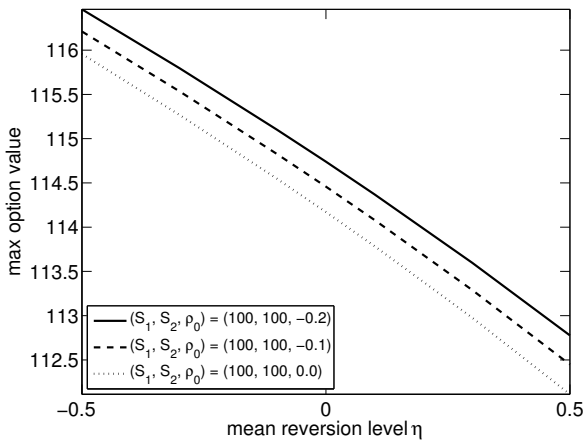

Figure 6.1: Effect of η on max option prices. Other parameters: $\sigma_{S_1} = 0.2$, $\sigma_{S_2} = 0.3$, $\lambda = 2.0$, $\sigma_\rho = 1.0$, maturity is 1 year.

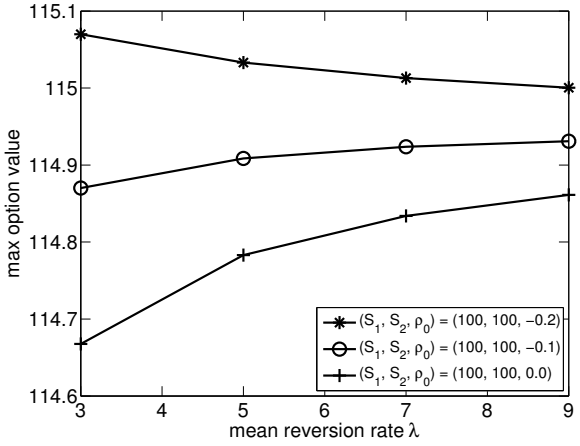

Figure 6.2: Effect of λ on max option prices. Other parameters: $\sigma_{S_1} = 0.2$, $\sigma_{S_2} = 0.3$, $\eta = -0.1$, $\sigma_\rho = 1.0$, maturity is 1 year.

When $\rho = \eta = -0.1$, the sensitivity $D^2_{1,1}V^{(0)}$ is negative. Therefore, while the second term in (6.3) vanishes, an upward trend is still predicted by the formula (as ϵ decreases), albeit of a much smaller magnitude due to the further diminishing effects of $\sigma^2_{S_1}\sigma^2_{S_2}$.

6.4 Summary

In the second part of this thesis (Chapters 3, 4, 5 and 6), we have studied the problem of option pricing in the presence of stochastic correlation from a computational viewpoint. Starting with the derivation of the pricing PDE, we have developed two approaches to computing option values in this setting. Our contributions are listed below.

- We have identified the solution space to be C^2 in space, C^1 in time (Theorem 1). The key part of the proof is to show that the discounted risk neutral expectation is a continuous function, from which the general parabolic theory applies.

- We have developed a finite difference approximation to a modified PDE corresponding to the pricing problem. We have shown that, if the assets' truncated boundary is far enough, setting the value there to zero or discounted payoff results in "small enough" error. A challenging part of the problem is the specification of the boundary *behaviour* when the correlation ρ is ± 1. The boundary condition is necessary when one uses such numerical techniques as the finite difference method. We have proposed a boundary condition defined by the PDE. Furthermore, we discuss other important numerical issues such as meshing, discretization of the cross term and numerical stability of the numerical

scheme (with this somewhat unusual boundary condition. In log price space, under fully implicit timestepping, our discretization is stable in the l^∞-norm.

- When the correlation process exhibits fast mean-reversion, a second approach, based on singular perturbation ([46]), is developed. The asymptotic solution involves a correction to the (multi-asset) Black-Scholes price under a constant correlation. This simple yet powerful solution is useful in computational scenarios where efficiency is important and where a PDE solution may be costly to implement.

- For options where the values or derivatives for the constant correlation case under the Black-Scholes multi-dimensional framework do not have known closed-form expressions using elementary functions, we have studied a quadrature method based on the asymptotic density to compute the price. Explicit expressions of the required density corrections are provided.

- Our numerical experiments demonstrate the effectiveness of our methods, and the agreement among our methods and Monte Carlo simulations. We have additionally studied the effect of smoothing on the quality of numerical solution, as an application of our analysis in Chapter 2. Finally, we have demonstrated that the asymptotic solution is able to capture effects of model parameters on prices, as shown in our numerical experiments.

Chapter 7

Conclusions and Future Work

7.1 Summary and conclusions of research

In this thesis, we have explored the option pricing problem from a PDE standpoint. The first part of the thesis is an analysis of a model problem, and our focus is on understanding the effect of non-smoothness on the numerical solution. The second part of the thesis focuses on developing computational methods for option pricing problems with stochastic correlation.

In the first part of this thesis, we have developed an analysis of the error arising from the non-smoothness in initial conditions when approximating the solution of a pricing problem with a finite difference PDE method. We have built our framework on the sharp error estimate in [25], and studied three types of non-smoothness that are of financial interest. Whereas the framework in [25] focuses on the effect of Rannacher timestepping, we utilize their techniques to study the discretization error due to spatial non-smoothness. We have shown that the error of the numerical solution under Crank-Nicolson-Rannacher timestepping with central spatial differences can be decomposed into two components. The first component is a second order discretization error primarily resulting from the approximation to the heat kernel by a discrete operator. The second component is a *quantization error*, loosely defined in [52] as the error resulting from the resolution of the non-smoothness on the grid. We have obtained explicit expressions of the two kinds of errors, and discussed how mesh positioning affects the quality of the numerical solution, as well as the possibility of an optimal positioning of the point of non-smoothness. We have also studied explicitly the effect of smoothing on the error of the numerical solution. Thus, the error analysis we developed is a powerful tool that gives the user explicit advice about how to obtain the desirable order of convergence and how to maintain a stable convergence order that can be used in extrapolation settings.

In the second part of the thesis, we considered that correlation is a stochastic variable, and derived a PDE that the price of a European contingent claim under stochastic correlation

satisfies, in which the correlation is an extra (spatial) variable. We identified that the suitable solution is C^2 in space and C^1 in time. We assumed a specific stochastic correlation process, and showed that the Feller condition for this process matches nicely the Fichera condition on the PDE correlation boundary. We studied the localization issues for the PDE domain.

Using the PDE derived, we have developed two different approaches to valuing European options under a stochastic correlation model. The first numerical method is an unconditionally stable finite difference scheme (under fully implicit timestepping). Unique to this problem is the specification of the boundary *behaviour* when the correlation ρ is ± 1. The boundary condition is necessary when one uses such numerical techniques as the finite difference method. We have proposed a boundary condition defined by the PDE. Furthermore, we discuss other important numerical issues such as meshing, discretization of the cross term and numerical stability of the numerical scheme (with this somewhat unusual boundary condition). In log price space, under fully implicit timestepping, our discretization is stable in the l^∞-norm.

A second computational method is developed based on singular perturbation theory, when the correlation process exhibits fast mean-reversion. The asymptotic solution involves a correction to the (multi-asset) Black-Scholes price under a constant correlation. This simple yet powerful solution is useful in computational scenarios where efficiency is important and where a PDE solution may be costly to implement. For options where the values or derivatives for a constant correlation under the Black-Scholes multi-dimensional framework do not have known closed-form expressions using elementary functions, we have studied a quadrature method based on the asymptotic density. Explicit expressions of the required density corrections are provided.

Through numerical experiments, we have demonstrated the effectiveness and agreement of our numerical methods. We have additionally studied the effect of smoothing on the quality of numerical solution, as an application of our analysis in Chapter 2. Finally, we have demonstrated that the asymptotic solution is able to capture effects of model parameters on prices, as shown in our numerical experiments.

The PDE with its numerical and asymptotic solution approaches allow the user to price options under a more realistic model for correlation than constant or deterministic, and under various practical settings, either targeting accuracy at the expense of computational power or targeting purely efficiency.

7.2 Future work

Some possible extensions to the work presented in this thesis are listed below.

- It would be interesting to apply the analysis in Chapter 2 to alternative differencing

schemes both in time and in space.

- A study of the effect of non-smooth initial conditions on errors for other numerical methods, such as spline collocation, is important, as mesh positioning could play a big part in the construction of the basis elements.

- Extension of the analysis in Chapter 2 to PDEs other than the convection-diffusion equation would be desirable.

- A logical next step to our research work on stochastic correlation is to develop efficient numerical solutions to multi-asset problems with stochastic correlation structure, and to include the effect of contagion in the modelling.

- It would be interesting to investigate efficient parallel solvers for the three-dimensional PDE (3.7).

- The numerical PDE approach developed for European options can be extended to American options using a non-linear penalty iteration.

Appendix A

Properties of the correlation process

The density Φ of the invariant distribution satisfies

$$\frac{\partial}{\partial\rho}(\lambda(\eta-\rho)\Phi) = \frac{1}{2}\frac{\partial^2}{\partial\rho^2}(\sigma_\rho^2(1-\rho^2)\Phi). \tag{A.1}$$

The solution that satisfies $\int \Phi = 1$ is given by

$$\Phi(x) = \frac{1}{2^{\frac{2\lambda}{\sigma_\rho^2}-1}}\frac{\Gamma(\frac{2\lambda}{\sigma_\rho^2})}{\Gamma(\frac{\lambda(1-\eta)}{\sigma_\rho^2})\Gamma(\frac{\lambda(1+\eta)}{\sigma_\rho^2})}(1-x)^{\frac{\lambda(1-\eta)}{\sigma_\rho^2}-1}(1+x)^{\frac{\lambda(1+\eta)}{\sigma_\rho^2}-1}. \tag{A.2}$$

The moments of the process (where subscript t here indicates dependency on time, $\rho_t = \rho(t)$) can be evaluated as follows:

$$\mathbf{E}(\rho_t) = \int_0^t \lambda(\eta - \mathbf{E}(\rho_u))du.$$

The solution is given by

$$\mathbf{E}(\rho_t) = \rho_0 e^{-\lambda t} + \eta(1-e^{-\lambda t}). \tag{A.3}$$

Taking limit $t \to \infty$, the mean with respect to the invariant distribution is η. This can also be verified numerically by direct integration with (A.2). Similarly, by Itô's lemma,

$$\mathbf{E}(\rho_t^2) = \int_0^t \Big(2(\lambda\eta\mathbf{E}(\rho_u)) - (2\lambda+\sigma_\rho^2)\mathbf{E}(\rho_u^2) + \sigma_\rho^2\Big)du.$$

The ODE that arises from this can be solved analytically given $\mathbf{E}(\rho_t)$ above. As $t \to \infty$, the second moment with respect to the invariant distribution is $\frac{\sigma_\rho^2+2\lambda\eta^2}{2\lambda+\sigma_\rho^2}$. Therefore, the variance is $\frac{\sigma_\rho^2(1-\eta^2)}{2\lambda+\sigma_\rho^2}$.

Appendix B

Density calculations

We denote by $p(T, S_1(T), S_2(T)|t, S_1, S_2)$, $t \leq T$, the joint transition density function in the case of constant correlation $\bar{\rho} = \eta$ of the terminal prices $S_1(T)$ and $S_2(T)$, given asset prices S_1, S_2 at an earlier time t. Note that $p(\cdot|\cdot)$ satisfies the backward Kolmogorov equation

$$\frac{\partial p}{\partial t} + \frac{\sigma_{S_1}^2 {S_1}^2}{2}\frac{\partial^2 p}{\partial S_1^2} + \frac{\sigma_{S_2}^2 {S_2}^2}{2}\frac{\partial^2 p}{\partial S_2^2} + \bar{\rho}\sigma_{S_1}\sigma_{S_2}S_1S_2\frac{\partial^2 p}{\partial S_1 \partial S_2} + rS_1\frac{\partial p}{\partial S_1} + rS_2\frac{\partial p}{\partial S_2} = 0$$

with the terminal condition $p(T, S_1(T), S_2(T)|T, S_1, S_2) = \delta(S_1(T) - S_1, S_2(T) - S_2)$, where δ denotes the Dirac delta function. By direct computation, it can be shown that

$$p(T, S_1(T), S_2(T)|t, S_1, S_2) = \frac{1}{2\pi\sqrt{\det\Sigma}S_1(T)S_2(T)}e^{-\frac{1}{2}v^T A v},$$

where

$$\Sigma = (T-t)\begin{bmatrix} \sigma_{S_1}^2 & \sigma_{S_1}\sigma_{S_2}\bar{\rho} \\ \sigma_{S_1}\sigma_{S_2}\bar{\rho} & \sigma_{S_2}^2 \end{bmatrix}, \qquad A = \Sigma^{-1},$$

$$v \equiv \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \log(\frac{S_1'}{S_1}) - (r - \frac{\sigma_{S_1}^2}{2})(T-t) \\ \log(\frac{S_2'}{S_2}) - (r - \frac{\sigma_{S_2}^2}{2})(T-t) \end{pmatrix}.$$

By straightforward differentiation,

$$\begin{aligned} D_{1,0}(p) &= S_1\frac{\partial p}{\partial S_1} = p \times (A_{1,1}v_1 + A_{1,2}v_2) \\ D_{0,1}(p) &= S_2\frac{\partial p}{\partial S_2} = p \times (A_{2,1}v_1 + A_{2,2}v_2) \\ D_{1,1}(p) &= S_1S_2\frac{\partial^2 p}{\partial S_1 S_2} = p\times\left((A_{1,1}v_1 + A_{1,2}v_2)(A_{2,1}v_1 + A_{2,2}v_2) + \frac{\bar{\rho}}{\sigma_{S_1}\sigma_{S_2}(T-t)(1-\bar{\rho}^2)}\right). \end{aligned}$$

Finally,

$$\begin{aligned}
D^2_{1,1}(p) &= D_{1,1}(D_{1,1}(p)) \\
&= D_{1,1}(p) \times \Bigg((A_{1,1}v_1 + A_{1,2}v_2)(A_{2,1}v_1 + A_{2,2}v_2) + \frac{\bar{\rho}}{\sigma_{S_1}\sigma_{S_2}(T-t)(1-\bar{\rho}^2)} \Bigg) \\
&+ D_{1,0}(p) \times \Bigg(- A_{1,2}(A_{2,1}v_1 + A_{2,2}v_2) - A_{2,2}(A_{1,1}v_1 + A_{1,2}v_2) \Bigg) \\
&+ D_{0,1}(p) \times \Bigg(- A_{1,1}(A_{2,1}v_1 + A_{2,2}v_2) - A_{2,1}(A_{1,1}v_1 + A_{1,2}v_2) \Bigg) \\
&+ p \times \Bigg(A_{1,1}A_{2,2} + A_{1,2}A_{2,1} \Bigg).
\end{aligned}$$

Incidentally, one can verify equalities such that $\frac{\partial p}{\partial \bar{\rho}} = \sigma_{S_1}\sigma_{S_2}(T-t)D_{1,1}(p)$.
It is then straightforward to apply the right-side of (5.6) to $p(\cdot|\cdot)$ to obtain $p^{\epsilon,1}_m(\cdot|\cdot)$ for use in (5.9).

Bibliography

[1] Khaled Bahlali, Brahim Mezerdi, and Youssef Ouknine. *Séminaire de Probabilités XXXII*, chapter Pathwise uniqueness and approximation of solutions of stochastic differential equations, pages 166–187. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998.

[2] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. In *29th IEEE Conference on Decision and Control*, pages 2347–2349 vol.4, Dec 1990.

[3] Tim Bollerslev, Robert F. Engle, and Jeffrey M. Wooldridge. A capital asset pricing model with time-varying covariances. *The Journal of Political Economy*, 96(1):116–131, 1988.

[4] Marie-France Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4):725–751, 1991.

[5] Adrian Buss and Grigory Vilkov. Measuring equity risk with option-implied correlations. *Review of Financial Studies*, 25(10):3113–3140, 2012.

[6] Thomas Chiang, Bang Jeon, and Huimin Li. Dynamic correlation analysis of financial contagion: Evidence from Asian markets. *Journal of International Money and Finance*, 26(7):1206–1228, 2007.

[7] C. Christara and D. M. Dang. Adaptive and high-order methods for valuing American options. *J. Comput. Finance*, 14(4):73–113, Summer 2011.

[8] Simon S. Clift and Peter A. Forsyth. Numerical solution of two asset jump diffusion models for option valuation. *Appl. Numer. Math.*, 58(6):743–782, June 2008.

[9] Rama Cont and Ekaterina Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. *SIAM J. Numer. Anal.*, 43(4):1596–1626, January 2005.

[10] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–407, 1985.

[11] José Da Fonseca, Martino Grasselli, and Claudio Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 10(2):151–180, 2007.

[12] Jos Da Fonseca, Martino Grasselli, and Claudio Tebaldi. A multifactor volatility Heston model. *Quantitative Finance*, 8(6):591–604, 2008.

[13] Jos Da Fonseca, Grasselli Martino, and Florian Ielpo. Estimating the Wishart affine stochastic correlation model using the empirical characteristic function. *Studies in Nonlinear Dynamics & Econometrics*, 18(3):253–289, 2014.

[14] Duy Minh Dang, Christina C. Christara, and Kenneth R. Jackson. An efficient graphics processing unit-based parallel algorithm for pricing multi-asset American options. *Concurrency and Computation: Practice and Experience*, 24(8):849–866, 2012.

[15] Duy Minh Dang, Christina C. Christara, and Kenneth R. Jackson. GPU pricing of exotic cross-currency interest rate derivatives with a foreign exchange volatility skew model. *Concurrency and Computation: Practice and Experience*, 16(9):1609–1625, 2014.

[16] Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(3):463–520, 1994.

[17] Joost Driessen, Pascal Maenhout, and Grigory Vilkov. Option-implied correlations and the price of correlation risk, INSEAD working paper, 2012.

[18] Joost Driessen, Pascal J. Maenhout, and Grigory Vilkov. The price of correlation risk: Evidence from equity options. *J. Finance*, 64(3):1377–1406, 2009.

[19] D.J. Duffy. *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. The Wiley Finance Series. Wiley, 2006.

[20] Bruno Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[21] Erik Ekström and Johan Tysk. The Black–Scholes equation in stochastic volatility models. *Journal of Mathematical Analysis and Applications*, 368(2):498 – 507, 2010.

[22] Erik Ekström and Johan Tysk. Boundary conditions for the single-factor term structure equation. *Ann. Appl. Probab.*, 21(1):332–350, 02 2011.

[23] Robert Engle. Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics*, 20:339–350, 2002.

[24] Jean-Pierre Fouque, Matthew Lorig, and Ronnie Sircar. Second order multiscale stochastic volatility asymptotics: stochastic terminal layer analysis and calibration. *Finance and Stochastics*, 20(3):543–588, 2016.

[25] Michael B. Giles and Rebecca Carter. Convergence analysis of Crank-Nicolson and Rannacher time-marching. *Journal of Computational Finance*, 9(4):89–112, 2006.

[26] Christian Gourieroux and Razvan Sufana. Derivative pricing with Wishart multivariate stochastic volatility. *Journal of Business & Economic Statistics*, 28(3):438–451, 2010.

[27] T. Haentjens and K.J̃. in 't Hout. Alternating direction implicit finite difference schemes for the Heston-Hull-White partial differential equation. *J. Comput. Finance*, 16(1):83–110, 2012.

[28] Steve Heston and Guofu Zhou. On the rate of convergence of discrete-time contingent claims. *Mathematical Finance*, 10(1):53–75, 2000.

[29] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Stud.*, 6:327–343, 1993.

[30] Steven L. Heston, Mark Loewenstein, and Gregory A. Willard. Options and bubbles. *Review of Financial Studies*, 20(2):359–390, 2007.

[31] N. Hilber, O. Reichmann, C. Schwab, and C. Winter. *Computational Methods for Quantitative Finance*. Springer-Verlag Berlin Heidelberg, 1 edition, 2013. An optional note.

[32] Svante Janson and Johan Tysk. Feynman-Kac formulas for Black-Scholes-type operators. *Bulletin of the London Mathematical Society*, 38:269–282, 4 2006.

[33] Raul Kangro and Roy Nicolaides. Far field boundary conditions for Black–Scholes equations. *SIAM J. Numer. Anal.*, 38(4):1357–1368, September 2000.

[34] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 1991.

[35] S. G. Kou. A jump-diffusion model for option pricing. *Management Sci.*, 48(8):1086–1101, 2002.

[36] H. O. Kreiss, V. Thome, and O. Widlund. Smoothing of initial data and rates of convergence for parabolic difference equations. *Communications on Pure and Applied Mathematics*, 23(2):241–259, 1970.

[37] Roger W. Lee. Option pricing by transform methods: extensions, unification and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.

[38] Franois Longin and Bruno Solnik. Extreme correlation of international equity markets. *J. Finance*, 56(2):649–676, 2001.

[39] Jun Ma. Pricing foreign equity options with stochastic correlation and volatility. *Annals of Economics and Finance*, 10(2):303–327, November 2009.

[40] Jun Ma. A stochastic correlation model with mean reversion for pricing multi-asset options. *Asia-Pacific Financial Markets*, 16(2):97–109, 2009.

[41] William Margrabe. The value of an option to exchange one asset for another. *J. Finance*, 33(1):177–186, 1978.

[42] Sima Mashayekhi and Jens Hugger. Kα-shifting, Rannacher Time Stepping and Mesh Grading in Crank-Nicolson FDM for Black-Scholes Option Pricing. *Communications in Mathematical Finance*, 5(1):1–31, 2016.

[43] Robert C. Merton. Option pricing when underlying stock returns are discontinuous. *J. Financial Econ.*, 3(1–2):125–144, 1976.

[44] Hong-Ghi Min and Young-Soon Hwang. Dynamic correlation analysis of US financial crisis and contagion: evidence from four OECD countries. *Applied Financial Economics*, 22(24):2063–2074, 2012.

[45] Tobias J. Moskowitz. An analysis of covariance risk and pricing anomalies. *Review of Financial Studies*, 16(2):417–457, April 2003.

[46] G. Papanicolaou, J. P. Fouque, K. Solna, and R. Sircar. Singular perturbations in option pricing. *SIAM J. Appl. Math.*, 63(5):1648–1665, 2003.

[47] D. M. Pooley, K. R. Vetzal, and P. A. Forsyth. Convergence remedies for non-smooth payoffs in option pricing. *Journal of Computational Finance*, 6(4):25–40, 2003.

[48] Rolf Rannacher. Finite element solution of diffusion problems with irregular data. *Numerische Mathematik*, 43(2):309–327, 1984.

[49] Christoph Reisinger and Alan Whitley. The impact of a natural time change on the convergence of the cranknicolson scheme. *IMA Journal of Numerical Analysis*, 34(3):1156, 2014.

[50] Ken-iti Sato. *Lévy Processes and Infinitely Divisible Distributions (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, 1st edition, November 1999.

[51] S.E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Number v. 11 in Springer Finance Textbooks. Springer, 2004.

[52] D. Tavella and C. Randall. *Pricing Financial Instruments: The Finite Difference Method*. John Wiley & Sons, Inc., 2000.

[53] L. Teng, C. van Emmerich, M. Ehrhardt, and M. Günther. A versatile approach for stochastic correlation using hyperbolic functions. *International Journal of Computer Mathematics*, 93(3):524–539, 2016.

[54] V. Thome and L. Wahlbin. Convergence rates of parabolic difference schemes for non-smooth data. *Mathematics of Computation*, 28(125):1–13, 1974.

[55] Anna-Karin Tornberg and Björn Engquist. Regularization techniques for numerical approximation of PDEs with singularities. *Journal of Scientific Computing*, 19(1):527–552, 2003.

[56] Anna-Karin Tornberg and Björn Engquist. Numerical approximations of singular source terms in differential equations. *J. Comput. Phys.*, 200(2):462–488, November 2004.

[57] John D. Towers. Finite Difference Methods for Approximating Heaviside Functions. *Journal of Computational Physics*, 228(9):3478–3489, May 2009.

[58] Y. K. Tse and Albert Tsui. A multivariate GARCH model with time-varying correlations. *Journal of Business and Economic Statistics*, pages 351–362, 2002.

[59] Cathrin van Emmerich. Modelling correlation as a stochastic process. Technical report, Bergische Universität Wuppertal, 2006. Preprint BUW-AMNA 06/03.