UNIVERSITY OF SUSSEX

Nonlinear Parabolic Stochastic Partial Differential Equation with Application to Finance

by

Rabab Alharbi

Supervised by Omar Lakkis

Submitted for the degree of Doctor of Philosophy

University of Sussex

December 2020

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Rabab Alharbi

UNIVERSITY OF SUSSEX

Rabab Alharbi, Doctor of Philosophy

Volatility Study

Summary

We study a nonlinear parabolic stochastic partial differential equation (SPDE) with multiplicative space–time white noise. The noise coefficient is the square root of the unknown, with respect to which such nonlinearity is Hölder but not Lipschitz. The SPDE's deterministic part includes also a first order quadratic nonlinearity. Mathematically speaking this space–time SPDE reduces to one of two stochastic differential equations (SDEs) appearing in the celebrated Heston and Cox–Ingersoll–Ross models describing the stock price volatility evolution; we therefore propose and test its use as a possible tool in understanding the so-called "volatility smile" observed in the implied volatility upon inverting the Black–Scholes–Merton model against option market data [Rouah, 2013] [Gatheral and Taleb, 2011]. Similar equations arise in other fields as well, for example, in surface growth modelling with or without a random forcing term. A typical example is the nonlinear stochastic Kardar–Parisi–Zhang equation (KPZ) whose unique solvability required the establishment of regularity structure methods [Hairer, 2013]. In principle, the model we herein propose lends itself to analysis via the regularity structures approach, but it exhibits better stability properties than KPZ thanks to a favourable sign in the first order quadratic nonlinearity and a resulting crucial energy identity; owing to this we are thus able to take a more straight-forward approach using energy methods and Galerkin approximations and show the SPDE is well posed in one spatial dimension, which is the relevant case in financial modelling. In line with the Galerkin approximation idea, we introduce an Euler–Maruyama numerical scheme to approximate the solution [Lord et al., 2014] which we use to close our work by looking at possible applications of the extended Heston model we propose. This extended Heston model includes a new independent variable (which acts as "space" in "space–time") which signifies the option's strike price on which the implied volatility depends.

Acknowledgements

Writing a thesis during the rapidly developing COVID-19 pandemic seems a challenge. Life tends to throw us surprises at the best of times that derail monumental plans. However COVID-19 is the biggest curveball of them all, on the top of everything else going on. Indeed, schools and universities are locked-down and life is not the same as we know it for now. However, during this crisis and even before, I got tremendous moral support and encouragement to keep working on this project. Therefore, it is my pleasure to acknowledge the roles of the people whose lack of them contribution would not make this work possible. I owe sincere and earnest gratitude to Dr Omar Lakkis for his kindly help, support and supervision during my research study and assistance in keeping my progress on schedule. Dr Lakkis was always willing to advise, providing me with the right guidance that I needed to keep working productively and reminding me to relax when the stress has become high. Special thanks go to Dr Qi Tang for his constructive suggestions and his help in collecting and analysing the financial data. My special thanks are extended to my colleague James Van Yperen for the numerous discussions we had on numerical analysis and coding. I wish to thank all the staff members of the School of Mathematical and Physical Sciences at Sussex University for a positive and productive work environment. I am grateful to Qassim University in Saudi Arabia for funding my PhD studies, and I also would like to show my gratitude to my family and friends for their encouragement and support. I also want to say thank you to all of the frontline workers for them efforts during these critical times to keeping the society safe, healthy and providing access to all of the resources that we ever needed.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this thesis, we propose an extension of the stochastic option price Heston's model that describes the evolution of the volatility u of an underlying asset S which is given by the following equations

$$\mathrm{d}S(t) = \mu(t)S(t)\mathrm{d}t + \sqrt{u(t)}S(t)\mathrm{d}W_1(t) \quad t \geq 0, \tag{1.1}$$

and

$$\mathrm{d}u(t) = -\kappa(u(t) - \gamma)\mathrm{d}t + \eta\sqrt{u(t)}\mathrm{d}W_2(t) \quad t \geq 0, \tag{1.2}$$

with

$$\langle \mathrm{d}W_1 \mathrm{d}W_2 \rangle = \rho \mathrm{d}t,$$

where κ, γ, $\eta > 0$ are the model parameters, representing speed, long-term mean and the volatility of u respectively. The parameter ρ is the correlation coefficient between two Wiener processes W_1 and W_2 or equivalently, the covariance $\rho \mathrm{d}t$. Equation (1.1) is known as the *Black–Scholes model*, not to be confused with the related *Black–Scholes (partial differential) equations*; in fact, the Black–Scholes model underpins the Black–Scholes equation (which we briefly discuss in Chapter 4). The process u that satisfies (1.2) follows a Cox–Ingersoll–Ross (CIR) diffusion (also known as the square root process), and the stochastic differential equation for the CIR diffusion satisfies the Yamada-Watanabe condition, so it admits a unique strong solution [Gatheral and Taleb, 2011], [Rouah, 2013]. Our new extension aims to reveal why options with the same asset price and time to maturity but different strike price (exercise price) have different implied volatilities. Implied volatility is the volatility input in an option pricing model (such as Black–Scholes formula and Heston's model) that generates the actual market price. It can be recovered from the real-life financial data via an inverse procedure

involving the Black–Scholes equation. This is explained in detail in Chapter 4, where we use an algorithm to produce the so-called "volatility smile", a known feature in financial modelling, which means that the implied volatility of a stock price depends on the strike price that an agent is prepared to pay for a given option on that stock. Based on this data analysis, we propose a novel mathematical model which describes the volatility as a strike-price and time dependent quantity. This phenomenon is known as the volatility smile, which means that the smile shape is a result of the fact that implied volatility can be written as a function of two variables, time and strike [Gatheral and Taleb, 2011]. Heston's model is a stochastic volatility model which is not modelling the volatility $\sqrt{u(t)}$ directly but rather through its square, that is the variance $u(t)$. The Heston variance process arises from the nonlinearly multiplicative Ornstein–Uhlenbeck process which has Markovian dynamics of the following form:

$$\mathrm{d}u(t) = -\underbrace{\kappa}_{\text{speed}}\underbrace{(u(t)-\gamma)}_{\text{mean-reverting}}\,\mathrm{d}t + \underbrace{\eta}_{\text{volatility}}\sqrt{u(t)}\mathrm{d}W(t). \tag{1.3}$$

Like all models, Black–Scholes model has its shortcomings, some of which are obviated by the many alternatives and extensions to it. In all named stochastic volatility models in general, including Heston's model, the framework is based on the fact that the volatility of the asset is randomly fluctuating and driven by its own stochastic process. In some sense, we can consider Heston's model as an extension of the Black–Scholes model by taking into account a square-root diffusion for the stochastic variance. However, Heston's model is not always able to fit the implied volatility smile very well, especially at short maturities [Rouah, 2013]. Therefore, researchers have contributed to extending Heston's model in many ways and different approaches. For example, one approach is known as the double Heston's Model proposed by Christoffersen et al. 2009, where the authors introduced a second factor of the variance independent of the first one. The idea behind this extension is that the additional volatility factor provides a more flexible approach to model the volatility surface. Another approach is to allow the parameters to be time-dependent. This latter approach is adopted by Mikhailov and Nogel (2003), Elices (2009), Benhamou, Gobet, and Miri (2010) and others [Rouah, 2013]. All of these extensions and others are aiming to provide a better fit to the volatility surface. Yet, none of them has shown the relationship between option implied volatilities and their strike prices, where the volatility surface suggests that the smile phenomenon resulting from the fact that implied volatility is a function of the strike prices [Austing, 2014].

In this thesis, we depart from previous Heston-based models by proposing a mathematical framework that includes the option's strike price, denoted by x, as the variable upon which the volatility depends. As a result, we transform the system of two SDEs forming the Heston model into a system of SPDE and SDE. We focus especially on the SPDE as it is a nonlinear

parabolic equation depending on the strike price x and time to maturity t, and a nonlinearity depending on the square of gradient of the variance with respect to x. The idea is by assuming that Heston's variance can be written in the context of the heat equation. Hence, we consider the transformation of the variance $u(t)$ to be in the following form:

$$u = \log \frac{1}{v}, \tag{1.4}$$

where v solves the homogeneous heat equation which is described by the following linear second order parabolic partial differential equation: $\partial_t v - \partial_{xx} v = 0$, some basic calculations lead to

$$\frac{1}{v} = e^u \tag{1.5}$$

$$v = e^{-u}. \tag{1.6}$$

Then we have

$$\partial_t v = e^{-u}(-\partial_t u) \tag{1.7}$$

$$\partial_{xx} v = e^{-u}(-\partial_x u)^2 - e^{-u}\partial_{xx} u. \tag{1.8}$$

plugging into the heat equation, yields

$$\partial_t u - \partial_{xx} u + |\partial_x u|^2 = 0 \tag{1.9}$$

Now, as assumed by Heston's model, we take into account a mean reversion term and a multiplicative noise. The obtained model "strike price dependent" variance process obeys the following stochastic IVP:

$$\begin{aligned} \partial_t u(x,t) - \partial_{xx} u(x,t) + \kappa(u(x,t) - \gamma) + |\partial_x u(x,t)|^2 \\ = \eta \sqrt{u(x,t)} \partial_{xt} W \quad \text{for } t \geq 0 \text{ and } x \in (0,1), \end{aligned} \tag{1.10a}$$

$$u(0,t) = -\alpha_0 \partial_x u(0,t) \quad \text{for} \quad \alpha_0 \geq 0, \tag{1.10b}$$

$$u(1,t) = \alpha_1 \partial_x u(1,t) \quad \text{for} \quad \alpha_1 \geq 0, \tag{1.10c}$$

$$u(x,0) = u_0. \tag{1.10d}$$

The spatial domain is the interval $(0,1)$, with the Robin boundary conditions (1.10b),(1.10c) and given initial condition u_0. The random term $\partial_{xt} W$ is a space–time white noise (cylindrical Wiener process) defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$, the parameters κ, γ and η are the same as Heston's model speed, long-term mean and the volatility of u (volatility of the volatility) respectively. Throughout this work we suppress the dependence on space and write $u(t)$ for $u(\cdot, t)$ and $\partial_{xt} W$ for $\partial_{xt} W(x,t)$ unless otherwise stated.

We study the stochastic IVP (1.10a)–(1.10d) analytically and numerically, which we apply to financial models of stochastic volatility. Equation (1.10a) presents specific mathematical challenges that have notf been addressed in the literature to the best of our knowledge. For instance, the nontrivial feature of equation (1.10a) appears from the nonlinear first order term $|\partial_x u|^2$ and a Hölder, yet not Lipschitz, continuous noise coefficient. Due to the nonlinear first order term in the equation (1.10a) with space–time white noise, the solution to the equation is expected to be too rough to be differentiable in space and time. Indeed, even with dropping the quadratic nonlinearity, the equation reduces to the a stochastic heat equation with the mean-reverting term, whose solution is not differentiable. In this sense equation (1.10a) is reminiscent of the Kardar–Parisi–Zhang (KPZ) equation which is a nonlinear SPDE, introduced by Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang in 1986 in [Kardar et al., 1986]. It describes the temporal change of a height field $h(x,t)$. In [Hairer, 2013], Hairer made a breakthrough in solving the KPZ equation by constructing approximations using Feynman diagrams. In 2014 he was awarded the Fields Medal for this work, along with rough paths theory and regularity structures. Equation (1.10a) fits the regularity structures framework, however in comparison to the KPZ equation (1.10a) has exhibits waning where KPZ exhibits nonlinearity, which allows us to show existence of the solution by a more elementary, straightforward approach. The solution of the KPZ equation is obtained by making use of the theory of regularity structure introduced in [Hairer, 2013]. The theory of regularity structures includes a renormalisation procedure, which consists in a way to appropriately modify the enhancement of the noise, in order for the corresponding modified sequence u^ε converges to a limit u. However, due to the quadratic nonlinearity's sign, the regularity structures is not required here. Nevertheless, proving the existence by the usual argument, i.e. contraction mapping and fixed point theorem is not possible, as the quadratic nonlinearity can not be bounded in an appropriate manner.

One way to establish the existence and regularity of weak solutions of the SPDE is by the use of the Fourier–Galerkin method. The idea of the existence proof is to approximate $u : [0,T] \to H^1(0,1)$ by functions $u_n : [0,T] \to E_n$ that take values in a finite-dimensional subspace $E_n \subset H^1(0,1)$ of dimension n. To obtain the sequence u_n, we project the SPDE onto E_n, meaning that we require that u_n satisfies the SPDE up to an orthogonal residual. This gives a system of SODEs for u_n, which has a solution by standard SODE theory. Each u_n satisfies an energy estimate of the same form as a priori estimate for solutions of the SPDE which are uniform in n and thus we obtain a solution of the SPDE. Of course, what allows us to deal with the nonlinear term $|\partial_x u|$ and the Hölder, but not Lipschitz noise coefficient is the sign of the quadratic nonlinearity.

Studying the positivity of solutions was not palpable as much as the original Heston's model solution. However, using what is known as Feller's square root condition introduced with some weak minimum principle theorem inspiration, we managed to obtain the positivity. The idea to show the positivity is that the noise coefficient, $\eta\sqrt{v(t)}$, prevents negative solutions for positive values of κ and γ. In other words, when the solution is close to zero, the noise coefficient also becomes very small, which turns the noise off. Consequently, when the solution gets close to zero, its evolution becomes dominated by the drift factor κ and γ.

We closed our work by numerical experiments to study the convergence rate of the approximated solutions with the ability of our model to generate a "smile". The numerical scheme of the SPDE is based on the Euler–Maruyama method. A Monte–Carlo simulation is utilised to understand the mechanism of the volatility as a function of two variables (strike, time) over time.

Before developing a theory for (1.10a) we briefly compare the KPZ equation as it introduced in [Hairer, 2013], [Bruned, 2016] with [Funaki and Quastel, 2015], and (1.10a). The KPZ equation is given by

$$\partial_t h - \partial_{xx} h - |\partial_x h|^2 = \partial_{xt} W, \tag{1.11}$$

the nontrivial features of both equations appearing from the nonlinear term $|\partial_x u|^2$. The classical KPZ is solution obtained through the Cole–Hopf transformation which reduces the KPZ equation to the one dimensional linear stochastic heat equation with multiplicative noise.

$$\partial_t h = \partial_{xx} h + h\partial_{xt} W, \tag{1.12}$$

having an initial value $h(0, x) \geq 0$. The stochastic heat equation is a well-posed SPDE. In fact, its solution can be defined in a mild sense as follows:

$$h(t, x) = \int_0^1 \mathcal{K}(t, x, y) h(0, y) \mathrm{d}y + \int_0^t \int_0^1 \mathcal{K}(t - s, x, y) h(s, y) \partial_{xt} W, \tag{1.13}$$

where $\mathcal{K}(t, x, y)$ is the heat kernel on $(0, 1)$ and it admits a unique positive solution in a suitable space of adapted processes. The Cole–Hopf solution of the KPZ equation is defined from the solution of (1.12) as

$$u(t) := \log h(t), \tag{1.14}$$

which is well-defined since $h(t) > 0$. In order to link the Cole–Hopf solution to the KPZ equation, we need to deal with an infinite Itô correction term. In other words, a certain renormalisation factor which balances with this diverging term (i.e. term that diverges as the mollification vanishes) should be introduced in the KPZ equation. This is can be seen from the heuristic

derivation of the KPZ equation from the stochastic heat equation under the Cole-Hopf transformation which goes as follows, recall Itô's formula for $u = f(h)$:

$$\mathrm{d}u = f'(h)\mathrm{d}h + \frac{1}{2}f''(h)(\mathrm{d}h)^2, \tag{1.15}$$

and, from (1.12), since $\partial_{xt}W\partial_{yt}W = \delta(x-y)\mathrm{d}t$, we can compute as

$$\begin{aligned}(\mathrm{d}h(t,x))^2 &= (h\partial_{xt}W)^2 \\ &= h^2\delta_x(x)\mathrm{d}t.\end{aligned} \tag{1.16}$$

Under the Cole–Hopf transformation, noting that $(\log h)' = h^{-1}$ and $(\log h)'' = -h^{-2}$, Itô's formula proves that

$$\begin{aligned}\partial_t u &= h^{-1}\partial_t h - \frac{1}{2}h^{-2}(\partial_t h)^2 \\ &= h^{-1}(\frac{1}{2}\partial_{xx}h + h\partial_{xt}W) - \frac{1}{2}\delta_x(x) \\ &= \frac{1}{2}h^{-1}\partial_{xx}h + \partial_{xt}W - \frac{1}{2}\delta_x(x).\end{aligned} \tag{1.17}$$

However, since $u = \log h$, some computation shows that

$$h^{-1}\partial_{xx}h = \partial_{xx}u + (\partial_x u)^2. \tag{1.18}$$

This leads to the KPZ equation with renormalisation factor:

$$\partial_t u = \frac{1}{2}\partial_{xx}u + \frac{1}{2}\{(\partial_x u)^2 - \delta_x(x)\} + \partial_{xt}W. \tag{1.19}$$

The simplest approximation scheme is to introduce a cutoff, so that the equation makes usual sense. Smooth out the noise by a smooth version $\partial_{xt}W^\varepsilon$, for instance obtained by a convolution with a family of mollifiers:

$$\partial_t u = \frac{1}{2}\partial_{xx}u + \frac{1}{2}\{(\partial_x u)^2 - \delta_x^\varepsilon(x)\} + \partial_{xt}W^\varepsilon. \tag{1.20}$$

If $\partial_{xt}W^\varepsilon$ were a smooth function of space and time, we could simply solve the equation for each realisation of the noise separately then take a limit as $\varepsilon \to 0$. However, the enhancement $\partial_{xt}W^\varepsilon$ is not expected to converge in the same space as $\varepsilon \to 0$ [Bruned, 2016]. Therefore in general, the solution u_ε of the KPZ is not guaranteed to converge either, or may convergence to a different answer.

The problem with the Cole–Hopf solution as stated by [Hairer, 2013] is that it does not provide a satisfactory theory of approximations to the KPZ equation. Indeed, all approximations to the KPZ equation must first be reinterpreted as approximations to the stochastic heat equation, which is not always convenient. Therefore, [Hairer, 2013] has extended the Cole–Hopf

solution to introduce the so-called regularity structure approach, to solve the classical KPZ equation. Moreover, using Cole–Hopf transform fails in the more general situation (the generalised KPZ which is similar to equation (1.10a) in some sense) or in the case of a system of coupled KPZ equations. Hence, [Bruned, 2016] used the theory of regularity structures inspired from the rough path and introduced by Martin Hairer to solve the generalised KPZ equation in his joint work with Harier, M. The generalised KPZ equation given by

$$\partial_t u(x,t) = \partial_{xx} u + g(u)(\partial_x u)^2 + k(u)\partial_x u + h(u) + f(u)\xi, \tag{1.21}$$

where ξ denotes space-time white noise. This equation is equivalent to equation (1.10a) with $g = -1$, $h = -\kappa(u(x,t) - \gamma)$, $k(u) = 0$ and $f(u) = \eta\sqrt{u(x,t)}$. The challenge in (1.21) is not only the ill-defined nonlinearity but also the term $f(u)\xi$ is the product of the distribution ξ and the noise coefficient $f(u)$, which is however not sufficiently regular to give a classical meaning to this product; therefore this term is as tricky as $|\partial_x u|^2$ to treat. The theory of regularity structures includes a renormalisation procedure, which consists in a way to appropriately modify the enhancement of ξ_ε, in order for the corresponding modified u_ε to converge. The renormalisation procedure has an algebraic step, by which the appropriate modification of the enhancement of ξ_ε is constructed, and an analytic step, where the actual convergence of the modified enhancement is proved.

1.1 Preliminaries and notations

In this section we introduce the definitions, theorems and notations that are consistently using throughout the thesis.

1.1.1 Functional analysis

The contraction mapping theorem is used to prove the existence and uniqueness of solutions to initial-value problems.

1.1.2 Theorem (Contraction mapping [Lord et al., 2014]). *Let Y be a non-empty closed subset of the Banach space $(X, \|\cdot\|)$. Consider a mapping $J : Y \to Y$ such that, for some $\mu \in (0,1)$,*

$$\|Ju - Jv\| \leq \mu\|u - v\|, \quad \textit{for all} \quad u, v \in Y. \tag{1.22}$$

There exists a unique fixed point of J in Y; that is, there is a unique $u \in Y$ such that $Ju = u$.

1.1.3 Theorem (Arzelà–Ascoli theorem [Kelley, 1975]). *Let C be the family of all continuous functions on a regular locally compact topological space to Hausdorff uniform space X, and let*

C have the topology of uniform convergence on compact. Then a subfamily F of C is compact if and only if

1. *F is closed in C.*

2. *F(x) has a compact closure for each member x of X, and the family F is equicontinuous.*

1.1.4 Definition (Orthonormal basis [Lord et al., 2014])**.** Let H be a Hilbert space. A set $\{\chi_j : j \in \mathbb{N}\} \subset H$ is orthonormal if

$$\langle \chi_j, \chi_k \rangle = \begin{cases} 1 & \text{for } j = k, \\ 0 & \text{for } j \neq k. \end{cases} \tag{1.23}$$

This is also written as $\langle \chi_j, \chi_k \rangle = \delta_{jk}$, where δ_{jk} is known as the Kronecker delta function. If the span of $\{\chi_j\}$ is additionally dense in H then we have a complete orthonormal set and $\{\chi_j\}$ is said to be an orthonormal basis for H. In that case, $u = \sum_{j=1}^{\infty} \langle u, \chi_j \rangle \chi_j$ for any $u \in H$ and

$$\|u\|^2 = \sum_{j=1}^{\infty} |\langle u, \chi_j \rangle|^2 = \sum_{j=1}^{\infty} |\langle \chi_j, u \rangle|^2. \tag{1.24}$$

A Hilbert space is separable if it contains a countable dense subset and every separable Hilbert space has an orthonormal basis (as can be proved by applying the Gram–Schmidt orthogonalisation procedure to the countable dense subset).

1.1.5 Definition (Sobolev spaces [Lord et al., 2014])**.** Let D be a domain and Y be a Banach space. For $p \geq 1$, the Sobolev space $W_p^r(D, Y)$ is the set of functions whose weak derivatives up to order $r \in \mathbb{N}$ are in $L^p(D, Y)$. That is,

$$W_p^r(D, Y) := \{u : \mathcal{D}^\alpha u \in L^p(D, Y) \text{ if } |\alpha| \leq r\}. \tag{1.25}$$

equipped with norm

$$\|u\|_{W_p^r(D,Y)} := \Big(\sum_{0 \leq |\alpha| \leq r} \|\mathcal{D}^\alpha u\|_{L^p(D,Y)}^p \Big)^{1/p}. \tag{1.26}$$

When $p = 2$ and $r > 0$ we write

$$H^r(D, Y) := W_2^r(D, Y). \tag{1.27}$$

1.1.6 Example (Weak derivative in one dimension)*.* Suppose $D \subseteq \mathbb{R}$, We can define $H^1(D)$ in the following way: it is a subspace of $L^2(D)$ of functions u for which there exists $g \in L^2(D)$ such that

$$\int_D u\varphi' = -\int_D g\varphi \quad \text{for each } \varphi \in C^\infty(D). \tag{1.28}$$

We will denote g by u'. This definition is equivalent to the definition with distributional derivatives.

1.1.7 Theorem (Sobolev embedding theorem [Lord et al., 2014]). *Let $W_p^k(\mathbb{R}^n)$ denote the Sobolev space consisting of all real-valued functions on $\mathbb{R}^n$ whose first k weak derivatives are functions in L^p. Here k is a non-negative integer and $1 \leq p < \infty$. The first part of the Sobolev embedding theorem states that if $k > i$ and $1 \leq p < q < \infty$ are two real numbers such that*

$$\frac{1}{p} - \frac{k}{n} = \frac{1}{q} - \frac{i}{n}, \tag{1.29}$$

then

$$W_p^k(\mathbb{R}^n) \subseteq W_q^i(\mathbb{R}^n), \tag{1.30}$$

and the embedding is continuous.

1.1.8 Lemma (Lax–Milgram [Evans, 2013]). *Let H be a real Hilbert space with norm $\|\cdot\|$ and let l be a bounded linear functional on H. Let $a : H \times H \to R$ be a bilinear form that is bounded and coercive. There exists a unique $u_l \in H$ such that $a(u_l, x) = l(x)$ for all $x \in H$.*

1.1.9 Definition. [Lord et al., 2014] Let H, U be separable Hilbert spaces with norms $\|\cdot\|, \|\cdot\|_U$ respectively. For an orthonormal basis $\{\chi_j : j \in \mathbb{N}\}$ of U, define the Hilbert–Schmidt norm

$$\|L\|_{\mathrm{HS}(U,H)} := \left(\sum_{j=1}^{\infty} \|L\chi_j\|_U^2 \right)^{1/2}. \tag{1.31}$$

The set $\mathrm{HS}(U, H) := \{L \in \mathcal{L}(U, H) : \|L\|_{\mathrm{HS}(U,H)} < \infty\}$ is a Banach space with the Hilbert–Schmidt norm and $L \in \mathrm{HS}(U, H)$ is known as a Hilbert–Schmidt operator.

1.1.10 Definition (Positive definite [Lord et al., 2014]). A function $G : D \times D \to R$ is positive definite if, for any $x_j \in D$ and $a_j \in \mathbb{R}$ for $j = 1, \cdots, N$, we have

$$\sum_{j,k=1}^{N} a_j a_k G(x_j, x_k) \geq 0. \tag{1.32}$$

A linear operator $L \in \mathcal{L}(H)$ on a Hilbert space H is non-negative definite (sometimes called positive semi-definite) if $\langle u, Lu \rangle \geq 0$ for any $u \in H$ and positive definite if $\langle u, Lu \rangle > 0$ for any $u \in H \backslash \{0\}$.

1.1.11 Definition (Trace class [Lord et al., 2014]). For a separable Hilbert space H, a non-negative definite operator $L \in \mathcal{L}(H)$ is of trace class if $\mathrm{Tr}(L) < \infty$, where the trace is defined by $\mathrm{Tr}(L) := \sum_{j=1}^{\infty} \langle L\chi_j, \chi_j \rangle$, for an orthonormal basis $\{\chi_j : j \in \mathbb{N}\}$.

1.1.12 Probability Theory

1.1.13 Definition (Stochastic process [Lord et al., 2014])**.** Given a set $\mathcal{T} \subset \mathbb{R}$, a measurable space $(H, \mathcal{H})$, and a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, a H-valued stochastic process is a set of H-valued random variables $\{X(t) : t \in \mathcal{T}\}$. We write $X(t)$ to denote the process.

1.1.14 Definition (Filtration [Lord et al., 2014])**.** Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space.

1. A filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is a family of sub σ-algebras of $\mathcal{F}$ that are increasing; that is, $\mathcal{F}_s$ is a sub σ-algebra of $\mathcal{F}_t$ for $s \leq t$. Each $(\Omega, \mathcal{F}, \mathbf{P})$ is a measure space and we assume it is complete.

2. A filtered probability space is a quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$, where $(\Omega, \mathcal{F}, \mathbf{P})$ is a probability space and $\{\mathcal{F}_t\}_{t\geq 0}$ is a filtration of $\mathcal{F}$.

Stochastic processes that conform to the notion of time described by the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ are known as adapted processes.

1.1.15 Definition (Adapted [Lord et al., 2014])**.** Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ be a filtered probability space. A stochastic process $\{X(t) : t \in [0, T]\}$ is a $\mathcal{F}_t$-adapted if the random variable $X(t)$ is a $\mathcal{F}_t$-measurable for all $t \in [0, T]$.

To stress the relationship between Brownian motion and a filtration, we give the following definition.

1.1.16 Definition ($\mathcal{F}_t$-Brownian motion)**.** A real-valued process $\{W(t) : t \geq 0\}$ is a $\mathcal{F}_t$-Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ if

1. $W(0) = 0$ a.s.,

2. $W(t)$ is continuous as a function of t,

3. $W(t)$ is $\mathcal{F}_t$-adapted and $W(t) - W(s)$ is independent of $\mathcal{F}_s$, $s < t$,

4. $W(t) - W(s) \sim N(0, t - s)$ for $0 \leq s \leq t$, where $N(0, t - s)$ denotes the normal distribution with expected value 0 and variance $t - s$.

1.1.17 Definition (Predictable [Lord et al., 2014])**.** A stochastic process $\{X(t) : t \in [0, T]\}$ is predictable if there exists $\mathcal{F}_t$-adapted and left-continuous processes $\{X_n(t) : t \in [0, T]\}$ such that $X_n(t) \to X(t)$ as $n \to \infty$ for $t \in [0, T]$.

1.1.18 Lemma (Banach space $\mathcal{L}_T^2$ [Lord et al., 2014]). *Let $\mathcal{L}_T^2$ denote the space of predictable real-valued processes $\{X(t) : t \in [0,T]\}$ with*

$$\|X\|_{\mathcal{L}_T^2} := \mathbb{E}\Big[\int_0^T |X(s)|^2 \mathrm{d}s\Big]^{1/2} < \infty. \tag{1.33}$$

Then, $\mathcal{L}_T^2$ is a Banach space with norm $\|\cdot\|_{\mathcal{L}_T^2}$.

1.1.19 Noise process

Let $(\Omega, F, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ be a filtered complete probability space and $(D, \mathcal{B}(D), \mu)$ a measure space for some domain $D = [0,1] \times \mathbb{R}^+$ and a Lebesgue measure μ. In [Walsh, 1986], Walsh defined a space–time white noise $\partial_{xt}W$ on D, based on μ, to be a measurable mapping such that for $A \in \mathcal{B}(D)$,

$$W(A) = \int_A \partial_{xt}W, \tag{1.34}$$

where $W(A)$ is a Gaussian random variable with mean zero and variance $\mu(A)$, provided $\mu(A) < \infty$. If A and $B \in \mathcal{B}(D)$ are disjoint, then the random variables $W(A)$ and $W(B)$ are independent and $W(A \cup B) = W(A) + W(B)$. The covariance structure of W is given by

$$\mathrm{Cov}(W(A), W(B)) = E(W(A)W(B)) = \mu(A \cap B), \tag{1.35}$$

as a consequence of L^2–isometry, which is a fundamental property of the stochastic integral and holds for the Itô integral, that is

$$\mathbb{E}\left[\Big(\int_{\mathbb{R}^+}\int_0^1 f(x,t)\partial_{xt}W\Big)^2\right] = \mathbb{E}\left[\int_{\mathbb{R}^+}\int_0^1 f(x,t)^2 \mathrm{d}x\mathrm{d}t\right], \tag{1.36}$$

for any $\mathcal{F}_t$–measurable $f \in L^2([0,1] \times \mathbb{R}^+)$, where $\{\mathcal{F}_t\}_{t\geq 0} = \sigma(W(x,s) : x \in [0,1], 0 \leq s \leq t)$ is sigma–algebra generated by W up to time t. Consider a right continuous filtration $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s, \forall t$, for which each $\mathcal{F}_t$ contains all the nullsets of F and satisfies:

1. $W(A)$ is $\mathcal{F}_t$–measurable whenever $A \in \mathcal{B}_s$, where $\mathcal{B}_s$ is the set of Borel–subsets of the domain D.

2. For $t \geq 0, \{W(A); A \in \mathcal{B}_s\}$ is independent of the $\mathcal{F}_t$.

According to [Walsh, 1986] the process $\{W(A,t) := W(A \times [0,t]), \mathcal{F}_t;\ t \geq 0,\ A \in \mathcal{B}_s\}$ is a martingale measure, i.e.

1. $W(A, 0) = 0$;

2. if $t \geq 0, W(\cdot, t)$ is σ–finite $L^2(\Omega)$–valued measure;

3. $\{W(A,t),\mathcal{F}_t;\ t\geq 0\}$ is a martingale.

Furthermore, the white noise process $W(A\times[0,t])$ is orthogonal martingale measure, where for any two disjoint sets A and $B\in\mathcal{B}_s$ the martingales $W(A\times[0,t])$ and $W(B\times[0,t])$ are orthogonal (i.e. the product of $W(A)$ and $W(B)$ is a martingale). In addition, we can say about the white noise process *worthy* with μ on D as dominating measure. Moreover, if f is a real valued predictable function with a norm defined by

$$||f||_W := \left(\mathbb{E}\left\{\int_0^t\int_0^1 f^2(x,t)\mathrm{d}x\mathrm{d}t\right\}\right)^{\frac{1}{2}} < +\infty, \tag{1.37}$$

we can define a space–time integral of f with respect to W as a worthy martingale measure

$$\left\langle \int\int f(x,t)\partial_{xt}W, \int\int f(x,t)\partial_{xt}W \right\rangle_t = \int\int f(x,t)\mathrm{d}x\mathrm{d}t, \tag{1.38}$$

where $\langle,\rangle_t$ stands for the least-squares product over the probability space over $\mathcal{F}_t$, the filtration's algebra at time t. Walsh theory, emphasises integration with respect to worthy martingale measures where the solutions are random fields. However, in this project we address an analogous approach which the Da Prato and Zabczyk theory of stochastic integrals with respect to Hilbert-space-valued Wiener processes to study the well-posedness of (1.10a)-(1.10d). According to [Dalang and Lluis, 2011], it is well-known that in certain cases, the Hilbert-space-valued integral is equivalent to a martingale-measure stochastic integral. For instance, it is pointed out in ([Da Prato and Zabczyk, 2014], Section 4.3) that when the random perturbation is space-time white noise, then Walsh's stochastic integral in [Walsh, 1986] is equivalent to an infinite-dimensional stochastic integral as in [Da Prato and Zabczyk, 2014]. Next we introduce the Q-Wiener process. Assume that W is a H-valued Wiener process with covariance operator Q. This process may be defined in terms of its Fourier series provided we build the right functional set-up to make such a series actually converge. Suppose that Q has eigenvalues $q_i>0$ and corresponding eigenfunctions χ_i. Then the *Q-Wiener process* is given by

$$W(x,t)=\sum_{i=0}^{\infty}\sqrt{q}_i\chi_i(x)\beta_i(t), \tag{1.39}$$

where $\beta_i, i=0,1,\ldots$ are i.i.d. $\mathcal{F}_t$-Brownian motions. If Q is trace class, then W is an L^2-valued process. However, if Q is not a trace class, for example $Q=I$ then the series which defines $W(t)$ does not converge in $L^2(\Omega,H)$, for a fixed $t>0$, since the strong law of large numbers ensures that when $N\to\infty$

$$\sum_{i=1}^{N}|\beta_i(t)|^2\sim\chi^2_{Nt}, \tag{1.40}$$

where χ_k^2 stands for the the chi-square distribution ([Ross, 2009], [Bréhier, 2014]). To overcome this issue we introduce a larger Hilbert space U such that $L^2 \subset U$ and the embedding is Hilbert-Schmidt.

1.1.20 Lemma (White noise convergence)**.** *Let χ_j be an orthonormal basis of $L^2(0,1)$ where $\chi_j(x) = \{c_j\, \mathrm{e}^{i2\pi jx}, j = 1, \cdots, c_j : \int_0^1 |\chi_j|^2 = 1\}$ then the cylindrical Wiener process given by (1.39) with $Q = I$ converges in $L^2((0,1), H^1(0,1)')$ if the inclusion $\iota : L^2 \to H^1(0,1)'$ is Hilbert–Schmidt.*

Proof. It is enough to show that the inclusion map is a Hilbert–Schmidt on $H^1(0,1)'$ where according to [Lord et al., 2014] the series (1.39) with $Q = I$ converges in $L^2(\Omega, U)$ if the inclusion $\iota : L^2 \to U$ is Hilbert–Schmidt.

$$\|\chi_j\|_{H^1(0,1)'} = \sup_{\varphi \neq 0} \frac{\langle \chi_j | \varphi \rangle}{\|\varphi\|_{H^1(0,1)}}, \tag{1.41}$$

or equivalently,

$$= \sup_{\|\varphi\|_{H^1(0,1)}=1} \langle \chi_j | \varphi \rangle. \tag{1.42}$$

where $\langle\, |\, \rangle$ denote the duality paring between $H^1(0,1)$ and $H^1(0,1)'$. By the Riesz representation theorem,

$$\langle \chi_j | \varphi \rangle = \int_0^1 \chi_j(x)\varphi(x)\mathrm{d}x, \quad \forall \varphi \in H^1(0,1), \tag{1.43}$$

applying integration by parts which yields

$$\Big[\int_0^x \chi_j(\xi)\mathrm{d}\xi\varphi(x)\Big]_0^1 - \int_0^1 \varphi'(x)\Big(\int_0^x \chi_j(\xi)\mathrm{d}\xi\Big)\mathrm{d}x. \tag{1.44}$$

Estimating $\int_0^x \chi_j(\xi)\mathrm{d}\xi$ for $\chi_j(x) = c_j\, \mathrm{e}^{i2\pi jx}$ gives

$$\begin{aligned} \int_0^x \chi_j(\xi)\mathrm{d}\xi &= \int_0^x c_j\, \mathrm{e}^{i2\pi j\xi}\, \mathrm{d}\xi \\ &= \frac{c_j}{2\pi ij}(\mathrm{e}^{2\pi ijx} - 1). \end{aligned} \tag{1.45}$$

Then,

$$\langle \chi_j | \varphi \rangle = \Big| \frac{c_j}{2\pi ij}(\mathrm{e}^{2\pi ij} - 1)\varphi(1) - \int_0^1 \varphi'(x)(\frac{c_j}{2\pi ij}(\mathrm{e}^{2\pi ijx} - 1))\mathrm{d}x \Big| \tag{1.46}$$

$$\leq c_{(1.47)}. \tag{1.47}$$

Hence,

$$\sum_{j=1}^{\infty} \|\chi_j\|^2_{H^1(0,1)'} < \infty. \tag{1.48}$$

□

1.1.21 Stochastic integral

We want to define the stochastic integral where the noise $W(t)$ is given by a L^2-cylindrical Wiener process, and taking values in U space. Let $B : t \in [0,T] \to \mathcal{L}(L^2, U)$, and W be a cylindrical Wiener process in L^2.

1.1.22 Lemma (Stochastic integral [Bréhier, 2014]). *If $\int_0^T \|B(t)\|^2_{\mathcal{L}_2(L^2,U)} \mathrm{d}t < \infty$, then we can define the stochastic integral as follows*

$$\int_0^T B(t)\partial_{xt}W = \sum_{i,j\in I,J} \int_0^T \langle B(t)e_j, e_i^*\rangle_U \mathrm{d}\beta_j(t) e_i^*, \tag{1.49}$$

where $(e_j)_{j\in J}$ and $(e_i^*)_{i\in I}$ are complete orthonormal systems of L^2 and U respectively. The result does not depend on the choice of these systems. Moreover we have the Itô isometry property

$$\mathbb{E}\Big| \int_0^T B(t)\partial_{xt}W \Big|^2_U = \int_0^T \|B(t)\|^2_{\mathcal{L}_2(L^2,U)} \mathrm{d}t. \tag{1.50}$$

Proof. Check the Itô isometry property:

$$\begin{aligned}
\mathbb{E}\Big| \int_0^T B(t)\partial_{xt}W \Big|^2_U &= \sum_{i\in I} \mathbb{E}\Big| \sum_{j\in J} \int_0^T \langle B(t)e_j, e_i^*\rangle \mathrm{d}\beta_j(t) \Big|^2 \\
&= \sum_{i\in I}\sum_{j\in J} \mathbb{E}\Big| \int_0^T \langle B(t)e_j, e_i^*\rangle \mathrm{d}\beta_j(t) \Big|^2 \\
&= \sum_{i\in I}\sum_{j\in J} \Big| \int_0^T \langle B(t)e_j, e_i^*\rangle \mathrm{d}t \Big|^2 \\
&= \sum_{j\in J} \int_0^T \Big| B(t)e_j \Big|^2 \mathrm{d}t \\
&= \int_0^T \big\| B(t) \big\|^2_{\mathcal{L}_2(L^2,U)} \mathrm{d}t.
\end{aligned} \tag{1.51}$$

□

Chapter 2

Existence and positivity of solutions

We study existence of mild solutions to (1.10a) with cylindrical Wiener process, where the noise coefficient is Hölder but not Lipschitz and the nonlinearity is genuinely not linear. In this chapter, we start by recalling the existence of a positive solution $u(t)$ of the SODE (1.2) which we will use crucially in the prove of the existence of solution to the SPDE (1.10a) and positivity of the approximate weak solution. In section 2.2 we show the existence for (1.10a) where the quadratic nonlinearity term is dropped. We follow [Lord et al., 2014] and [Da Prato and Zabczyk, 2014] to consider the SPDE as a SODE on a Hilbert space. A fixed point argument is used here to prove the existence and uniqueness of mild solutions for the stochastic equation via a Banach contraction argument. In section 2.3, we establish the existence and the regularity of the weak solution of equations (1.10a)-(1.10d) by using energy-based Galerkin approximation methods. A fixed point argument akin to the one used in section 2.2 is harder to apply here and our approach is based on defining the solution u as the limit of its Fourier-Galerkin approximation u_n. Finally the positivity of the solution is considered in section 2.4.

2.1 CIR process

We review next the properties of the CIR process which is the departure point for our extension. We start with an short–time existence result in Lemma 2.1.1 and then turn our attention to the role of the Feller condition and show that it guarantees long time positive solution in Theorem 2.1.3.

2.1.1 Lemma (Existence and uniqueness)**.** *For each $\varepsilon > 0$ and $T > 0$ there exists a $v_\varepsilon(t)$ which satisfies the following stochastic differential equation*

$$\mathrm{d}v_\varepsilon(t) = -\kappa(v_\varepsilon(t) - \gamma)\mathrm{d}t + G_\varepsilon(v_\varepsilon)\mathrm{d}W(t),\ v_\varepsilon(0) = v_0 > 0,\quad 0 \le t \le T. \tag{2.1}$$

where, $G_\varepsilon(r)$ is the square root's cut off function given by

$$G_\varepsilon(r) := \begin{cases} \eta r/\sqrt{r} & \text{if } r > \varepsilon, \\ \eta r/\sqrt{\varepsilon} & \text{if } r \le \varepsilon. \end{cases} \tag{2.2}$$

Proof. We follow an argument found in [Øksendal, 1985] Theorem 5.2.1 to show that there exists a unique solution of the equation (2.1). Define $v_\varepsilon^0(t) = v_0$ and $v_\varepsilon^n(t)$ inductively as follows

$$v_\varepsilon^{n+1}(t) = v_0 - \int_0^t \kappa(v_\varepsilon^n(s) - \gamma)\mathrm{d}s + \int_0^t G_\varepsilon(v_\varepsilon^n)\mathrm{d}W(s), \quad 0 \le t \le T,\ n \ge 0. \tag{2.3}$$

Put $a(t) := \kappa(v_\varepsilon^n(t) - \gamma) - \kappa(v_\varepsilon^{n-1}(t) - \gamma)$, $b(t) := G_\varepsilon(v_\varepsilon^n) - G_\varepsilon(v_\varepsilon^{n-1})$ and denote $\hat{v}_0$ to be an initial value of $v_\varepsilon^n(t)$. By using the initial condition, Jensen's inequality and Lipschitz continuity we have

$$\begin{aligned}
\mathbb{E}\Big[\big|v_\varepsilon^{n+1}(t) - v_\varepsilon^n(t)\big|^2\Big] &\le \mathbb{E}\Big[\big(v_0 - \hat{v}_0 + \int_0^t a(s)\mathrm{d}s + \int_0^t b(s)\mathrm{d}W(s)\big)^2\Big] \\
&\le 3\mathbb{E}[|v_0 - \hat{v}_0|^2] + 3\mathbb{E}\Big[\big(\int_0^t a(s)\mathrm{d}s\big)^2\Big] + 3\mathbb{E}\Big[\big(\int_0^t b(s)\mathrm{d}W(s)\big)^2\Big] \\
&\le 3\mathbb{E}[|v_0 - \hat{v}_0|^2] + 3t\mathbb{E}\Big[\int_0^t a(s)^2\mathrm{d}s\Big] + 3\mathbb{E}\Big[\int_0^t b(s)^2\mathrm{d}s\Big] \\
&\le 3\mathbb{E}[|v_0 - \hat{v}_0|^2] + (1+T)3c_{(A.1)}^2 \int_0^t \mathbb{E}[|v_\varepsilon^n(s) - v_\varepsilon^{n-1}(s)|^2]\mathrm{d}s.
\end{aligned}$$

note that $v_0 = \hat{v}_0$. Hence,

$$\mathbb{E}\Big[\big|v_\varepsilon^{n+1}(t) - v_\varepsilon^n(t)\big|^2\Big] \le (1+T)3c_{(A.1)}^2 \int_0^t \mathbb{E}[|v_\varepsilon^n(s) - v_\varepsilon^{n-1}(s)|^2]\mathrm{d}s, \tag{2.4}$$

for $n \ge 1, t \le T$ and by using linear growth condition we obtain

$$\begin{aligned}
\mathbb{E}[|v_\varepsilon^{(1)}(t) - v_\varepsilon^{(0)}(t)|^2] &\le 2c_{(A.2)}^2 t^2(1 + \mathbb{E}[|v_0|^2]) + 2c_{(A.2)}^2 t(1 + \mathbb{E}[|v_0|^2]) &(2.5)\\
&\le c_{(2.6)} t, &(2.6)
\end{aligned}$$

where $c_{(2.6)}$, is constant, depends on $c_{(A.2)}, T$ and $\mathbb{E}[|v_0|^2]$. So by induction on n we obtain

$$\mathbb{E}[|v_\varepsilon^{n+1}(t) - v_\varepsilon^n(t)|^2] \le \frac{c_{(2.7)}^{n+1} t^{n+1}}{(n+1)!}; \quad n \ge 0, t \in [0,T], \tag{2.7}$$

for some suitable constant $c_{(2.7)}$ depending only on $c_{(A.2)}, c_{(A.1)}, T$ and $\mathbb{E}[|v_0|^2]$. Now

$$\begin{aligned}
\sup_{0\le t\le T} |v_\varepsilon^{n+1}(t) - v_\varepsilon^n(t)| \le \int_0^T \big|\kappa v_\varepsilon^n(s) - \kappa v_\varepsilon^{n-1}(s)\big|\mathrm{d}s \\
+ \sup_{0\le t\le T} \Big|\int_0^t G_\varepsilon(v_\varepsilon^n) - G_\varepsilon(v_\varepsilon^{n-1})\mathrm{d}W(s)\Big|.
\end{aligned} \tag{2.8}$$

By Doob's martingale inequality we obtain

$$\begin{aligned}
\mathbf{P}\Big[\sup_{0\le t\le T}|v_\varepsilon^{n+1}(t)-v_\varepsilon^n(t)|>2^{-n}\Big] &\le \mathbf{P}\Big[(\int_0^T\big|\kappa v_\varepsilon^n(s)-\kappa v_\varepsilon^{n-1}(s)\big|\mathrm{d}s)^2>2^{-2n-2}\Big]\\
&+\Big[\sup_{0\le t\le T}\big|\int_0^t G_\varepsilon(v_\varepsilon^n)-G_\varepsilon(v_\varepsilon^{n-1})\mathrm{d}W(s)\big|>2^{-n-1}\Big]\\
&\le 2^{2n+2}T\int_0^T\mathbb{E}(\big|\kappa v_\varepsilon^n(s)-\kappa v_\varepsilon^{n-1}(s)\big|^2)\mathrm{d}s+2^{2n+2}\int_0^T\mathbb{E}\big[|G_\varepsilon(v_\varepsilon^n)-G_\varepsilon(v_\varepsilon^{n-1})|^2\big]\mathrm{d}s\\
&\le 2^{2n+2}c_{(A.3)}^2(T+1)\int_0^T\frac{c_{(2.7)}^n t^n}{n!}dt\\
&\le \frac{(4c_{(2.7)}T)^{n+1}}{(n+1)!},\quad \text{if } c_{(2.7)}\ge c_{(A.1)}(T+1).
\end{aligned}$$

Therefore, by the Borel-Cantelli lemma,

$$\mathbf{P}\Big[\sup_{0\le t\le T}|v_\varepsilon^{n+1}(t)-v_\varepsilon^n(t)|>2^{-n}\ \text{ for infinitely many } n\ \Big]=0. \tag{2.9}$$

Thus, for almost all ω there exists n_0 such that

$$\sup_{0\le t\le T}|v_\varepsilon^{n+1}(t)-v_\varepsilon^n(t)|>2^{-n}\ \text{ for } n\ge n_0. \tag{2.10}$$

Therefore the sequence

$$v_\varepsilon^k(t)=v_\varepsilon^{(0)}(t)+\sum_{n=0}^{k-1}(v_\varepsilon^{n+1}(t)-v_\varepsilon^n(t)) \tag{2.11}$$

is uniformly convergent on $[0,T]$, for almost all ω. Denote the limit by $v_\varepsilon(t)$. Then $v_\varepsilon(t)$ is t-continuous for almost all ω since $v_\varepsilon^k(t)$ is t-continuous for all k. Moreover, $v_\varepsilon(t,\cdot)$ is $\mathcal{F}_t$-measurable for all t, since $v_\varepsilon^k(t,\cdot)$ has this property for all k.

Next, note that for $m>k\ge 0$ we have by (2.7)

$$\begin{aligned}
\mathbb{E}[|v_\varepsilon^m(t)-v_\varepsilon^k(t)|^2]^{1/2}&=\|v_\varepsilon^m(t)-v_\varepsilon^k(t)\|_{L^2(\mathbf{P})}=\big\|\sum_{n=k}^{m-1}v_\varepsilon^{n+1}(t)-v_\epsilon^n(t)\big\|_{L^2(\mathbf{P})}\\
&\le\sum_{n=k}^{m-1}\big\|v_\varepsilon^{n+1}(t)-v_\varepsilon^n(t)\big\|_{L^2(\mathbf{P})}\le\sum_{n=k}^{\infty}\Big[\frac{(c_{(2.7)}t)^{n+1}}{(n+1)!}\Big]^{1/2}\to 0\ \text{as } k\to\infty.
\end{aligned} \tag{2.12}$$

So $\{v_\varepsilon^k(t)\}$ converges in $L^2(\mathbf{P})$ to a limit $\hat{v}_\varepsilon(t)$, say. A subsequence of $v_\varepsilon^k(t,\omega)$ will then converge ω-pointwise to $\hat{v}_\varepsilon(t,\omega)$ and therefore we must have $\hat{v}_\varepsilon(t)=v_\varepsilon(t)$ almost surely. It remains to show that $v(t)$ satisfies (2.1). For all k we have

$$v_\varepsilon^{k+1}(t)=v_0-\int_0^t\kappa(v_\varepsilon^k(s)-\gamma)\mathrm{d}s+\int_0^t G_\varepsilon(v_\varepsilon^k)\mathrm{d}W(s). \tag{2.13}$$

Now $v_\varepsilon^{k+1}(t)\to v_\varepsilon(t)$ as $k\to\infty$, uniformly in $t\in[0,T]$ for almost all ω. By (2.12) and the Fatou lemma we have

$$\mathbb{E}\Big[\int_0^T|v_\varepsilon(t)-v_\varepsilon^k(t)|^2\mathrm{d}t\Big]\le\lim_{m\to\infty}\sup\mathbb{E}\Big[\int_0^T|v_\varepsilon^m(t)-v_\varepsilon^k(t)|^2\mathrm{d}t\Big]\to 0 \tag{2.14}$$

as $k \to \infty$. It follows by Itô isometry that

$$\int_0^t G_\varepsilon(v_\varepsilon^k)\mathrm{d}W(s) \to \int_0^t G_\varepsilon(v_\varepsilon)\mathrm{d}W(s), \tag{2.15}$$

and by the Hölder inequality that

$$\int_0^t \kappa(v_\varepsilon^k(t)-\gamma)\mathrm{d}s \to \int_0^t \kappa(v_\epsilon(t)-\gamma)\mathrm{d}s, \tag{2.16}$$

in $L^2(\mathbf{P})$. Therefore, taking the limit of (2.13) as $k \to \infty$ we obtain (2.1) for $v_\varepsilon(t)$ where $t \in [0,T]$. This proof works for Lipschitz noise coefficient. □

2.1.2 Lemma. *A sequence $v_\varepsilon(t)$ satisfying (2.1) has a subsequence $(v_{\varepsilon_n})_{n\in\mathbb{N}}$ converging uniformly to a continuous limit function $v(t)$ which takes values in $C_0([0,T], L^2(\Omega\times[0,1]))$. The limit v is an Itô solution of the following equation*

$$\mathrm{d}v(t) = -\kappa(v(t)-\gamma)\mathrm{d}t + \eta\sqrt{v(t)}\mathrm{d}W(t). \tag{2.17}$$

Proof. We sketch the proof here (as we are giving a similar proof in section 2.2). First we need to establish some uniform bounds and continuity properties, in order to apply a weak convergence theorem, which allows to extract a subsequence. Then we show that the subsequence converges to a limit which is a solution to the original equation. □

2.1.3 Theorem (Feller's condition [Gikhman, 2011])**.** *Let κ, γ, η be positive constants and $v_0 > 0$. If Feller's condition $2\kappa\gamma > \eta^2$ is satisfied then there exists a unique solution which is positive on each finite time interval $[0,t]$ to the following equation*

$$\mathrm{d}v(t) = -\kappa(v(t)-\gamma)\mathrm{d}t + \eta\sqrt{v(t)}\mathrm{d}W(t), \tag{2.18}$$

where $W(t)$ is a Wiener process on a complete probability space.

Proof. Take $\varepsilon > 0$ and denote $T_\varepsilon = \min\{t : v(t) \le \varepsilon\}$. Then there exists a unique solution of the equation (2.18) on the interval $[0, T_\varepsilon \wedge t]$. We want to show that $\mathbf{P}\{T_\varepsilon \wedge t < t\} \to 0$ when $\varepsilon \to 0$. Define a positive constant

$$r = \frac{2\kappa\gamma - \eta^2}{\eta^2}. \tag{2.19}$$

Applying Itô formula for $g(x) = x^{-r}$ we note that

$$\begin{aligned} v(T_\varepsilon \wedge t)^{-r} =& v(0)^{-r} + \int_0^{T_\varepsilon \wedge t} r\kappa(v(s)-\gamma)v(s)^{-(r+1)}\mathrm{d}s \qquad (2.20)\\ &- \int_0^{T_\varepsilon \wedge t} r\eta\sqrt{v(s)}v(s)^{-(r+1)}\mathrm{d}W_s \\ &+ \frac{1}{2}\int_0^{T_\varepsilon \wedge t} r(r+1)\eta^2 v(s)v(s)^{-(r+2)}\mathrm{d}s \\ &= v^{-r}(0) + r\kappa\int_0^{T_\varepsilon \wedge s} v(s)^{-r}\mathrm{d}s - \int_0^{T_\varepsilon \wedge s} r\eta v(s)^{-(r+1/2)}\mathrm{d}W_s \qquad (2.21)\\ &+ \int_0^{T_\varepsilon \wedge s} [\frac{r(r+1)}{2}\eta^2 - r\kappa\gamma]v(s)^{-(r+1)}\mathrm{d}s. \end{aligned}$$

Bearing in mind (2.19) and taking expectation of the stochastic integral, we arrive at the estimate

$$\mathbb{E}[v(T_\varepsilon \wedge t)^{-r}] \leq v(0)^{-r} + r\kappa\int_0^t \mathbb{E}[v(T_\varepsilon \wedge s)^{-r}]\mathrm{d}s. \qquad (2.22)$$

Applying Gronwall's inequality, we get estimate

$$\mathbb{E}[v(T_\varepsilon \wedge t)^{-r}] \leq v(0)^{-r}\,\mathrm{e}^{r\kappa t}\,. \qquad (2.23)$$

Next using Chebyshev's inequality we note that

$$\mathbf{P}\{T_\varepsilon \leq t\} = \mathbf{P}\{v(T_\varepsilon)^r \leq \varepsilon^r\} = \mathbf{P}\{v(T_\varepsilon)^{-r} \geq \varepsilon^{-r}\} \leq \varepsilon^r v(0)^{-r}\,\mathrm{e}^{rt}\,. \qquad (2.24)$$

The right hand side tends to zero when ε tends to zero for each $t \in [0,t]$. Hence we proved that the probability of $v(t)$ becoming equal or less than an arbitrarily small positive threshold is correspondingly small. □

2.1.4 Remark. The motivation behind Theorem 2.1.3 is that the noise coefficient, $\eta\sqrt{v(t)}$, prevents negative solutions for a positive values of κ and γ. In other words, when the solution is close to zero, the noise coefficient also becomes very small, which turns the noise off. Consequently, when the solution gets close to zero, its evolution becomes dominated by the drift factor κ and γ.

2.2 A CIR type equation with diffusion

Removing the quadratic nonlinearity term from (1.10a) leaves us with the following mean-reverting heat equation, driven by a multiplicative noise, with Robin boundary conditions

$$\partial_t u(x,t) + A(u(x,t)) = \eta\sqrt{u(x,t)}\partial_{xt}W + \kappa\gamma,\ t \geq 0\ x \in (0,1), \qquad (2.25a)$$

$$u(0,t) = -\alpha_0\partial_x u(0,t) \quad \text{for} \quad \alpha_0 \geq 0, \qquad (2.25b)$$

$$u(1,t) = \alpha_1 \partial_x u(1,t) \quad \text{for} \quad \alpha_1 \geq 0, \tag{2.25c}$$

$$u(0) = u_0, \tag{2.25d}$$

where A is linear operator given by

$$A\phi = -\partial_{xx}\phi + \kappa\phi. \tag{2.26}$$

We recall that $A : H^1(0,1) \to H^1(0,1)'$ satisfies the assumption in Lemma A.0.6 and hence $-A$ is the infinitesimal generator of the semigroup e^{-tA} [Lord et al., 2014]. We show existence of a mild solution of equations (2.25a)-(2.25d) by constructing a sequence of approximate solutions $(u_\varepsilon)_{\varepsilon>0}$ which are the mild solution by making use of the contraction mapping theory. Then we establish some uniform bounds and continuity properties, in order to apply a weak convergence theorem, which allows to extract a subsequence. Finally, we show that the subsequence converges to a limit which is a solution to the original equation. We suppress the dependence on space and write $u(t)$ for $u(x,t)$.

2.2.1 Definition (Mild solution [Lord et al., 2014])**.** A predictable, $L^2(0,1)$–valued process $\{u_\varepsilon(\cdot,t) : t \in [0,T]\}$ is called a *mild solution* of the following SPDE

$$\partial_t u_\varepsilon(t) = -A(u_\varepsilon(t)) + G_\varepsilon(u_\varepsilon(t))\partial_{xt}W + \kappa\gamma, \tag{2.27}$$

if for all $t \in [0,T]$

$$u_\varepsilon(t) = \mathrm{e}^{-tA}\, u_0 + \int_0^t \mathrm{e}^{-(t-s)A}\, G_\varepsilon(u_\varepsilon(s))\partial_{xs}W + \kappa\gamma \int_0^t \mathrm{e}^{-(t-s)A}\, \mathbb{1}_{(0,1)}(s)\mathrm{d}s, \tag{2.28}$$

where e^{-tA} denotes the semigroup generated by $-A$.

To show existence of the mild solution (2.28), we use a contraction mapping theorem on the following Banach space.

2.2.2 Definition. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ be a filtered probability space. $\mathcal{H}_{2,T}$ is the set of L^2–valued predictable processes $u(t), 0 \leq t \in [0,T]$, such that

$$\|u\|_{\mathcal{H}_{2,T}} := \sup_{0\leq t\leq T} \|u(t)\|_{L^2(\Omega\times(0,1))} < \infty. \tag{2.29}$$

2.2.3 Theorem (Existence and uniqueness of the approximate mild solution [Lord et al., 2014])**.** *For each $\varepsilon > 0$ and $T > 0$, there exists a mild solution u_ε in $\mathcal{H}_{2,T}$ to (2.27). Furthermore, there exists $c_{(2.30)} > 0$ depending on T, such that*

$$\sup_{t\in[0,T]} \|u_\varepsilon(t)\|_{L^2(\Omega\times(0,1))} \leq c_{(2.30)} \Big(1 + \|u_0\|_{L^2(\Omega\times(0,1))}\Big). \tag{2.30}$$

Proof. For $u_\varepsilon \in \mathcal{H}_{2,T}$, define the following process,

$$\Gamma(u_\varepsilon)(t) := \mathrm{e}^{-tA}\, u_0 + \int_0^t \mathrm{e}^{-(t-s)A}\, G_\varepsilon(u_\varepsilon(s))\partial_{xs} W + \kappa\gamma \int_0^t \mathrm{e}^{-(t-s)A}\, \mathbb{1}_{(0,1)}(s)\mathrm{d}s. \tag{2.31}$$

A fixed point u_ε of Γ is an L^2–valued predictable process and obeys (2.28) and is hence a mild solution of (2.27). To show existence and uniqueness of the fixed point, we show Γ is a contraction by applying the contraction mapping theorem $\mathcal{H}_{2,T}$ to itself. First, we show Γ maps into $\mathcal{H}_{2,T}$. $\Gamma(u_\varepsilon)$ is a predictable process because u_0 is $\mathcal{F}_0$–measurable and the stochastic integral is a predictable process. It remains to show the $\mathcal{H}_{2,T}$ norm of $\Gamma(u_\varepsilon)$ is finite, let

$$\begin{aligned}
\mathrm{I}(t) &= \mathrm{e}^{-tA}\, u_0, \\
\mathrm{II}(t) &= \int_0^t \mathrm{e}^{-(t-s)A}\, G_\varepsilon(u_\varepsilon(s))\partial_{xs} W, \\
\mathrm{III}(t) &= \kappa\gamma \int_0^t \mathrm{e}^{-(t-s)A}\, \mathbb{1}_{(0,1)}(s)\mathrm{d}s,
\end{aligned}$$

so that $\Gamma(u_\varepsilon) = \mathrm{I}(t) + \mathrm{II}(t) + \mathrm{III}(t)$. Then,

$$\begin{aligned}
\|\Gamma(u_\varepsilon)\|_{\mathcal{H}_{2,T}} &= \sup_{0\le t\le T} \|\Gamma(u_\varepsilon)\|_{L^2(\Omega\times(0,1))} \\
&= \sup_{0\le t\le T} \|\mathrm{I}(t) + \mathrm{II}(t) + \mathrm{III}(t)\|_{L^2(\Omega\times(0,1))}.
\end{aligned} \tag{2.32}$$

We estimate each term separately,

$$\begin{aligned}
\|\mathrm{I}(t)\|_{L^2(\Omega\times(0,1))} &= \|\, \mathrm{e}^{-tA}\, u_0\|_{L^2(\Omega\times(0,1))} \\
&\le \|\, \mathrm{e}^{-tA}\, \|_{\mathcal{L}(L^2(0,1))} \|u_0\|_{L^2(\Omega\times(0,1))} \\
&\le \|u_0\|_{L^2(\Omega\times(0,1))} \\
&< \infty,
\end{aligned} \tag{2.33}$$

where $\|\, \mathrm{e}^{-tA}\, \|_{\mathcal{L}(L^2(0,1))} \le 1$. Next by using a linear growth condition on G_ε and Itô isometry for the diffusion,

$$\begin{aligned}
\big\|\mathrm{II}(t)\big\|^2_{L^2(\Omega\times(0,1))} &= \Big\| \int_0^t \mathrm{e}^{-(t-s)A}\, G_\varepsilon(u_\varepsilon(s))\partial_{xs} W \Big\|^2_{L^2(\Omega\times(0,1))} \\
&= \int_0^t \mathbb{E}\Big[\Big\|\, \mathrm{e}^{-(t-s)A}\, G_\varepsilon(u_\varepsilon)\Big\|^2_{L^2(0,1)} \Big]\mathrm{d}s && (2.34) \\
&\le \int_0^t \|\, \mathrm{e}^{-(t-s)A}\, \|^2_{\mathcal{L}(L^2(0,1))}\mathrm{d}s && (2.35) \\
&\quad \times c_{(A.2)}{}^2 \Big(1 + \sup_{0\le s\le t} \|u_\varepsilon(s)\|_{L^2(\Omega\times(0,1))}\Big)^2 \\
&\le c_{(2.37)}{}^2 t c_{(A.2)}{}^2 \Big(1 + \sup_{0\le s\le T} \|u_\varepsilon(s)\|_{L^2(\Omega\times(0,1))}\Big)^2, && (2.36)
\end{aligned}$$

where we used a semigroup estimate for some $c_{(2.37)} > 0$ as follows,

$$\int_0^T \| \mathrm{e}^{-tA} \|^2_{\mathcal{L}(L^2(0,1))} \mathrm{d}t \leq {c_{(2.37)}}^2 T, \tag{2.37}$$

where we use Exercise 10.8 in [Lord et al., 2014]. Hence,

$$\big\|\mathrm{II}(t)\big\|_{L^2(\Omega\times(0,1))} \leq c_{(2.37)} t^{1/2} c_{(A.2)} \Big(1 + \sup_{0\leq s\leq T} \|u_\varepsilon(s)\|_{L^2(\Omega\times(0,1))}\Big). \tag{2.38}$$

Finally we have,

$$\begin{aligned}
\big\|\mathrm{III}(t)\big\|_{L^2(\Omega\times(0,1))} &= \big\|\kappa\gamma \int_0^t \mathrm{e}^{-(t-s)A}\,\mathrm{d}s\big\|_{L^2(\Omega\times(0,1))} \\
&\leq \kappa\gamma \int_0^t \big\| \mathrm{e}^{-(t-s)A} \big\|_{L^2(\Omega\times(0,1))} \mathrm{d}s \\
&\leq c_{(2.37)} \kappa\gamma T^{1/2}.
\end{aligned} \tag{2.39}$$

Then, for $u_\varepsilon \in \mathcal{H}_{2,T}$, all I, II and III terms are uniformly bounded over $[0,T]$ in $L^2(\Omega\times(0,1))$ and $\|\Gamma(u_\varepsilon)\|_{\mathcal{H}_{2,T}} < \infty$. Hence Γ maps from $\mathcal{H}_{2,T}$ into itself. For the contraction property, we may show, owing to the Lipschitz properties of the mean reverting and the diffusion terms that

$$\begin{aligned}
&\|\Gamma(u_\varepsilon) - \Gamma(v_\varepsilon)\|^2_{L^2(\Omega\times(0,1))} \\
&\qquad \leq 2t \int_0^t {c_{(A.1)}}^2 \|(u_\varepsilon(s) - v_\varepsilon(s))\|^2_{L^2(\Omega\times(0,1))} \mathrm{d}s &(2.40)\\
&\qquad + 2{c_{(2.37)}}^2 t {c_{(A.1)}}^2 \|u_\varepsilon(s) - v_\varepsilon(s)\|^2_{\mathcal{H}_{2,T}}. &(2.41)
\end{aligned}$$

Then,

$$\|\Gamma(u_\varepsilon) - \Gamma(v_\varepsilon)\|^2_{\mathcal{H}_{2,T}} \leq 2(T^2 + {c_{(2.37)}}^2 T) {c_{(A.1)}}^2 \|u_\varepsilon(s) - v_\varepsilon(s)\|^2_{\mathcal{H}_{2,T}}. \tag{2.42}$$

and Γ is a contraction on $\mathcal{H}_{2,T}$ if $(T^2 + {c_{(2.37)}}^2 T) < 1/2{c_{(A.1)}}^2$, which is satisfied for T small enough. Repeating the argument on $[0,T], [T,2T], \ldots$, we find a unique mild solution for all $t > 0$. Finally, we verify the bound (2.30) by making use of the linear growth condition, the semigroup estimate (2.37) and Jensen's inequality,

$$\begin{aligned}
\|u_\varepsilon(t)\|_{L^2(\Omega\times(0,1))} &\leq 3\Big\{ \| \mathrm{e}^{-tA} u_0\|_{L^2(\Omega\times(0,1))} + \big\|\kappa\gamma \int_0^t \mathrm{e}^{-(t-s)A}\,\mathrm{d}s\big\|_{L^2(\Omega\times(0,1))} \\
&\quad + \big\| \int_0^t \mathrm{e}^{-(t-s)A} G_\varepsilon(u_\varepsilon(s)) \partial_{xs} W \big\|_{L^2(\Omega\times(0,1))} \Big\} \\
&\leq 3\Big\{ \|u_0\|_{L^2(\Omega\times(0,1))} + c_{(2.37)} \kappa\gamma T^{1/2} \\
&\quad + c_{2.37} T^{1/2} c_{(A.2)} \int_0^t (1 + \|u_\varepsilon\|_{L^2(\Omega\times(0,1))}) \mathrm{d}s \Big\} \\
&\leq 3\Big\{ c + c_{2.37} T^{1/2} c_{(A.2)} (T + \int_0^t \|u_\varepsilon\|_{L^2(\Omega\times(0,1))} \mathrm{d}s) \Big\},
\end{aligned}$$

which, by Gronwall's Lemma A.0.4, implies the bound (2.30).

Theorem 2.2.3 shows that for all $t > 0$ there exists a unique u_ε satisfying (2.27). The following theorem establishes regularity of the mild solution to the approximation solution u_ε in time.

2.2.4 Theorem (Regularity in time [Lord et al., 2014]). *Let $u_0 \in L^2(\Omega, \mathcal{F}_0, D(A))$, for $T > 0$ and $u_\varepsilon(t)$ satisfy the mild solution (2.28). Then for each $\delta \in (0, \frac{1}{2})$, there exists $c_{(2.43)} > 0$ such that*

$$\|u_\varepsilon(t_2) - u_\varepsilon(t_1)\|_{L^2(\Omega\times[0,1])} \le c_{(2.43)}(t_2 - t_1)^\delta, \quad 0 \le t_1 \le t_2 \le T. \tag{2.43}$$

Proof. Write $u(t_2) - u(t_1) = \mathrm{I} + \mathrm{II} + \mathrm{III}$, where

$$\mathrm{I} := (\mathrm{e}^{-t_2 A} - \mathrm{e}^{-t_1 A})u_0, \tag{2.44}$$

$$\mathrm{II} := \kappa\gamma\Big(\int_0^{t_2} \mathrm{e}^{-(t_2-s)A}\,\mathrm{d}s - \int_0^{t_1} \mathrm{e}^{-(t_1-s)A}\,\mathrm{d}s\Big), \tag{2.45}$$

$$\mathrm{III} := \int_0^{t_2} \mathrm{e}^{-(t_2-s)A} G_\varepsilon(u_\varepsilon(s))\partial_{xs}W - \int_0^{t_1} \mathrm{e}^{-(t_1-s)A} G_\varepsilon(u_\varepsilon(s))\partial_{xs}W. \tag{2.46}$$

We estimate each term separately,

$$\mathrm{I} = \mathrm{e}^{-t_1 A} A^{-\epsilon}(\mathrm{e}^{-(t_2-t_1)A} - I)A^\epsilon u_0, \tag{2.47}$$

from Lemma A.0.6(i). Using properties of the operator norm in Lemma A.0.5,

$$\|\mathrm{I}\| \le \|A^{-\epsilon}(I - \mathrm{e}^{-(t_2-t_1)A})\|_{\mathcal{L}(L^2(\Omega\times(0,1)))}\|A^\epsilon u_0\|, \tag{2.48}$$

and, by Lemma A.0.6(iii), there exists $c_{(A.9)} > 0$ such that $\|\mathrm{I}\| \le c_{(A.9)}(t_2 - t_1)^\epsilon\|u_0\|$. For the second term, write $\mathrm{II} = \mathrm{II}_1 + \mathrm{II}_2$ for

$$\mathrm{II}_1 := \kappa\gamma\int_0^{t_1}(\mathrm{e}^{-(t_2-s)A} - e^{-(t_1-s)A})\mathrm{d}s, \quad \mathrm{II}_2 := \kappa\gamma\int_{t_1}^{t_2}\mathrm{e}^{-(t_2-s)A}\,\mathrm{d}s. \tag{2.49}$$

We estimate the norms of each term II_1 and II_2. For II_1,

$$\begin{aligned}\|\mathrm{II}_1\| &= \|\kappa\gamma\int_0^{t_1}(\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A})\mathrm{d}s\| \\ &\le \kappa\gamma\int_0^{t_1}\|\,\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A}\,\|_{\mathcal{L}(L^2(\Omega\times(0,1)))}\mathrm{d}s.\end{aligned} \tag{2.50}$$

Note that $\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A} = A^{-(1-\epsilon)}(\mathrm{e}^{-(t_2-t_1)A} - I)A^{1-\epsilon}\,\mathrm{e}^{-(t_1-s)A}$. Then

$$\begin{aligned}&\kappa\gamma\int_0^{t_1}\|\,\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A}\,\|_{\mathcal{L}(L^2(\Omega\times(0,1)))}\mathrm{d}s \\ &\quad\le \kappa\gamma\int_0^{t_1}\|A^{-(1-\epsilon)}(\mathrm{e}^{-(t_2-t_1)A} - I)\|_{\mathcal{L}(L^2(\Omega\times(0,1)))} \\ &\quad\quad\times\|A^{1-\epsilon}\,\mathrm{e}^{-(t_1-s)A}\,\|_{\mathcal{L}(L^2(\Omega\times(0,1)))}\mathrm{d}s \\ &\quad\le \kappa\gamma c_{(A.8)}(t_2 - t_1)^{1-\epsilon}\int_0^{t_1} c_{(A.9)}(t_1 - s)^{-(1-\epsilon)}\mathrm{d}s \\ &\quad\le \kappa\gamma c_{(A.8)}c_{(A.9)}\frac{t_1^\epsilon}{\epsilon}(t_2 - t_1)^{1-\epsilon},\end{aligned} \tag{2.51}$$

where $c_{(A.8)}, c_{(A.9)}$ are constants given by Lemma A.0.6(ii)-(iii). Thus, for $c_{(2.52)} := \kappa\gamma c_{(A.8)} c_{(A.9)} T^{\epsilon}/\epsilon$,

$$\|\mathrm{II}_1\| \leq c_{(2.52)}(t_2 - t_1)^{1-\epsilon}. \tag{2.52}$$

For II_2,

$$\begin{aligned} \|\mathrm{II}_2\| &= \|\kappa\gamma \int_{t_1}^{t_2} \mathrm{e}^{-(t_2-s)A}\,\mathrm{d}s\| \\ &\leq \kappa\gamma \int_{t_1}^{t_2} \|\,\mathrm{e}^{-(t_2-s)A}\,\|\mathrm{d}s \\ &\leq \kappa\gamma(t_2 - t_1). \end{aligned} \tag{2.53}$$

Hence,

$$\begin{aligned} \mathrm{II} &\leq \mathrm{II}_1 + \mathrm{II}_2 \\ &\leq (c_{(2.52)} + 1)(t_2 - t_1)^{1-\epsilon}. \end{aligned} \tag{2.54}$$

Finally, for the stochastic term we write $\mathrm{III} = \mathrm{III}_1 + \mathrm{III}_2$, for

$$\begin{aligned} \mathrm{III}_1 &:= \int_0^{t_1} (\mathrm{e}^{-(t_2-s)A} - e^{-(t_1-s)A}) G_\varepsilon(u_\varepsilon(s))\partial_{xs} W, \\ \mathrm{III}_2 &:= \int_{t_1}^{t_2} \mathrm{e}^{-(t_2-s)A}\, G_\varepsilon(u_\varepsilon(s))\partial_{xs} W. \end{aligned} \tag{2.55}$$

First, consider III_1. By Itô isometry

$$\mathbb{E}[\|\mathrm{III}_1\|^2] = \int_0^{t_1} \mathbb{E}\Big[\|(\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A}) G_\varepsilon(u_\varepsilon(s))\|^2_{L_0^2}\Big]\mathrm{d}s. \tag{2.56}$$

Using Lipschitz condition assumption on G_ε, we obtain

$$\begin{aligned} \mathbb{E}[\|\mathrm{III}_1\|^2] \leq \int_0^{t_1} \|(\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A})\|^2_{\mathcal{L}(L^2(0,1))}\Big]\mathrm{d}s \\ \times c^2_{(A.2)}\Big(1 + \sup_{0\leq s\leq t_1} \|u_\varepsilon(s)\|_{L^2(\Omega\times(0,1))}\Big)^2. \end{aligned} \tag{2.57}$$

Note that

$$\begin{aligned} &\int_0^{t_1} \|(\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A})\|^2_{\mathcal{L}(L^2(0,1))}\mathrm{d}s \\ &\quad = \int_0^{t_1} \|A^{1/2-\epsilon}\,\mathrm{e}^{-(t_1-s)A}\, A^{-1/2+\epsilon}(I - \mathrm{e}^{-(t_2-t_1)A})\|^2_{\mathcal{L}(L^2(0,1))}\mathrm{d}s \\ &\quad \leq \int_0^{t_1} \|A^{1/2-\epsilon}\,\mathrm{e}^{-(t_1-s)A}\,\|^2_{\mathcal{L}(L^2(0,1))}\|A^{-1/2+\epsilon}(I - \mathrm{e}^{-(t_2-t_1)A})\|^2_{\mathcal{L}(L^2(0,1))}\mathrm{d}s. \end{aligned} \tag{2.58}$$

By Lemma A.0.6(iii), there exists $c_{(A.9)}, c_{(2.37)} > 0$ such that

$$\begin{aligned} &\int_0^{t_1} \|(\mathrm{e}^{-(t_2-s)A} - \mathrm{e}^{-(t_1-s)A})\|^2_{\mathcal{L}(L^2(0,1))}\mathrm{d}s \\ &\qquad \leq \Big(c^2_{(A.9)}\frac{t_1^{2\epsilon}}{\epsilon}\Big)(c^2_{(2.37)}(t_2 - t_1)^{1-2\epsilon}). \end{aligned} \tag{2.59}$$

Then, with $c_{(2.60)} := c_{(A.9)} c_{(2.37)} c_{(A.2)} T^{\epsilon}/\sqrt{\epsilon}$,

$$\begin{aligned} \|\mathrm{III}_1\|_{L^2(\Omega\times(0,1))} &= \mathbb{E}\Big[\|\mathrm{III}_1\|^2\Big]^{1/2} \\ &\le c_{(2.60)}(t_2-t_1)^{1-2\epsilon}\Big(1+\sup_{0\le s\le T}\|u_\varepsilon(s)\|_{L^2(\Omega\times(0,1))}\Big). \end{aligned} \tag{2.60}$$

Estimating $\mathbb{E}[\|\mathrm{III}_2\|^2]$, by Itô isometry and Lipschitz condition assumption on G_ε yield

$$\begin{aligned} \mathbb{E}[\|\mathrm{III}_2\|^2] &= \int_{t_1}^{t_2} \mathbb{E}\Big[\|\,\mathrm{e}^{-(t_2-s)A}\,G_\varepsilon(u_\varepsilon(s))\|_{L_0^2}\Big]\mathrm{d}s \\ &\le \int_{t_1}^{t_2} \|\,\mathrm{e}^{-(t_2-s)A}\,\|^2_{\mathcal{L}(L^2(0,1))}\mathrm{d}s\; c^2_{(A.2)}\Big(1+\sup_{t_1\le s\le t_2}\|u_\varepsilon(s)\|_{L^2(\Omega\times(0,1))}\Big)^2. \end{aligned} \tag{2.61}$$

By applying the semigroup estimate (2.37), we find $c_{(2.37)} > 0$ such that for $c_{(2.62)} := c_{(2.37)} c_{(A.2)}$,

$$\|\mathrm{III}_2\|_{L^2(\Omega\times(0,1))} \le c_{(2.62)}(t_2-t_1)^{1/2}\Big(1+\sup_{0\le s\le T}\|u_\varepsilon(s)\|_{L^2(\Omega\times(0,1))}\Big), \tag{2.62}$$

with $c^2_{(2.63)} := (c_{(2.60)} + c_{(2.62)})(1 + c_{(2.30)})$, we find, using (2.30), that

$$\begin{aligned} \|\mathrm{III}\|_{L^2(\Omega\times(0,1))} &\le \|\mathrm{III}_1\|_{L^2(\Omega\times(0,1))} + \|\mathrm{III}_2\|_{L^2(\Omega\times(0,1))} \\ &\le c_{(2.63)}(t_2-t_1)_1^{\delta}(1+\|u_0\|_{L^2(\Omega\times(0,1))}). \end{aligned} \tag{2.63}$$

Now we have,

$$\|u(t_2)-u(t_1)\|_{L^2(\Omega\times(0,1))} \le \|\mathrm{I}\|_{L^2(\Omega\times(0,1))} + \|\mathrm{II}\|_{L^2(\Omega\times(0,1))} + \|\mathrm{III}\|_{L^2(\Omega\times(0,1))}.$$

Since $D(A)$ is continuously embedded in $H^1(0,1)$, the estimates on the norms of I, II, and III complete the time regularity of u_ϵ. □

The regularity in time condition given by (2.43) implies the equicontinuity of u_ε uniformly. Where, for every $\alpha > 0$, there exists a $\zeta > 0$ such that $\|u_\varepsilon(t_1) - u_\varepsilon(t_2)\| < \alpha$ for all $\alpha > 0$ and all $t_1, t_2 \in [0,T]$ such that $|t_1 - t_2| < \zeta$. This fact implies the relative compactness of u_ε in $\mathcal{H}_{2,T}$. Therefore, by Arzelà–Ascoli theorem, u_ε has a subsequence $(u_{\varepsilon_n})_{n\in\mathbb{N}}$ converging uniformly to a continuous limit function u which takes values in $C_0([0,T], L^2(\Omega\times[0,1]))$. We claim that u is a mild solution of the original equation (2.25a). In fact, uniform convergence implies that for each $\varrho > 0$ there is ε_0 such that for all $\varepsilon < \varepsilon_0$ we have

$$\|u_\varepsilon - u\|_{L^\infty(0,T,L^2(\Omega\times(0,1)))} < \varrho. \tag{2.64}$$

Therefore, taking the limit as $\varepsilon \to 0$ over both sides of (2.28) once has, for the stochastic term

with using Itô isometry

$$\begin{aligned}\Big\| \int_0^t \mathrm{e}^{-(t-s)A}(G_\varepsilon(u_\varepsilon(s)) - G(u(s)))\partial_{xs}W\Big\|^2_{L^\infty(0,T,L^2(\Omega\times(0,1)))} \qquad (2.65)\\ = \int_0^t \mathbb{E}\left[\| \mathrm{e}^{-(t-s)A}(G_\varepsilon(u_\varepsilon(s)) - G(u(s)))\|^2_{L^2(0,1)}\right]\mathrm{d}s\\ \leq \int_0^t \| \mathrm{e}^{-(t-s)A}\|^2_{\mathcal{L}(L^2(0,1))}\mathrm{d}s\|G_\varepsilon(u_\varepsilon(s)) - G(u(s))\|^2_{L^2(0,1)}\\ \leq c_{(2.37)}t^2 c_{(A.3)},\end{aligned}$$

which proves that $u(t)$ is indeed a solution to (2.27). □

2.3 Existence of solution to nonlinear SPDE

In this section, we establish the existence of solution to the stochastic PDE equation (1.10a)-(1.10d). Although equation (1.10a) looks similar to the KPZ equation, (1.10a) has better stability properties thanks to a favourable sign of the gradient's term (in the KPZ equation this term drives the solution away from zero, while in our case it brings the solution closer to zero). Thus we are able to follow a significantly more straightforward approach to obtain the existence. Nevertheless, proving the existence by the usual argument, i.e., fixed point theorem and square root cut off approach is not clear here as the quadratic nonlinearity $|\partial_x u|^2$ term can not be bounded in an appropriate sense. One way to establish the existence and regularity of weak solutions of the SPDEs is by the use of Fourier–Galerkin method introduced in [Evans, 2013]. The idea of Fourier–Galerkin method is to build a weak solution of the SPDE (1.10a) by constructing solutions of finite-dimensional approximations. That means that we approximate $u : [0,T] \to H^1(0,1)$ by functions $u_n : [0,T] \to E_n$ that take values in a finite subspace $E_n \subset H^1(0,1)$. To obtain u_n, we project the SPDE onto E_n, meaning that we require that u_n satisfies the SPDE up to a residual which is orthogonal to E_n. This gives a system of SODEs for u_n, which has a solution by standard SODE theory. Each u_n satisfies an energy estimate which is uniform in n, which allows us to pass the limit as $n \to \infty$ and obtain a convergence to a general solution of the SPDE (1.10a). We start this section by introducing the definition of a weak solution to SPDE's.

2.3.1 Definition. (Weak solution) A predictable process $\{u(t) : t \in [0,T]\}$ valued in a separable Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is called a *weak solution* of the SPDE (1.10a) provided

1. $u \in L^2(0,T;H)$ and $u(0) = u_0$.

2. For every $v \in H$ and almost every time $0 \leq t \leq T$,

$$\langle u, v\rangle - \langle u_0, v\rangle + \int_0^t a(u,v)\mathrm{d}s + \int_0^t \langle |\partial_x u|^2, v\rangle \mathrm{d}s = \int_0^t \langle \eta\sqrt{u}\partial_{xs}W^n, v\rangle + \int_0^t \langle \kappa\gamma, v\rangle \mathrm{d}s, \tag{2.66}$$

where $a(\cdot,\cdot)$ is the bilinear form associated with A given by

$$a(u,v) = \int_0^1 \partial_x u \partial_x v - \sum_{x=0}^{1} \frac{1}{\alpha_x} u(x)v(x) + \kappa \int_0^1 uv, \tag{2.67}$$

and the stochastic integral is given in the notation of (1.39) with $q = I$ as follows

$$\int_0^t \langle \sqrt{u}\partial_{xs}W^n, v\rangle := \int_0^s \sum_{i=1}^{n} \langle \sqrt{u}\chi_i, v\rangle \mathrm{d}\beta_i(s). \tag{2.68}$$

2.3.2 Construction of approximate solutions

We define an approximate solution in the notation of (2.66) by following [Evans, 2013]. Let E_n be the k-dimensional subspace of $H^1(0,1)$ which is spanned by the first n vectors in an orthonormal basis of $L^2(0,1)$, which we may also assume to be an orthogonal basis of $H^1(0,1)$, i.e.,

$$E_n := \langle \chi_1, \chi_2 \ldots, \chi_n\rangle. \tag{2.69}$$

We denote by $\mathcal{P}_n : L^2(0,1) \to E_n \subset L^2(0,1)$ the orthogonal projection onto E_n defined by

$$\langle \mathcal{P}_n\phi, \chi\rangle = \langle \phi, \chi\rangle \quad \text{for all} \quad \chi \in L^2(0,1), \tag{2.70}$$

where $\chi_k(x)$ are representing the eigenfunctions of the elliptic operator A subject to Robin boundary conditions on $(0,1)$. We also denote by $\mathcal{P}_n$ the orthogonal projections $\mathcal{P}_n : H^1(0,1) \to E_n \subset H^1(0,1)$ or $\mathcal{P}_n : H^1(0,1)' \to E_n \subset H^1(0,1)'$, which we obtain by restricting or extending $\mathcal{P}_n$ from $L^2(0,1)$ to $H^1(0,1)$ or $H^1(0,1)'$, respectively. Thus, $\mathcal{P}_n$ is defined on $H^1(0,1)$ by (2.70) and on $H^1(0,1)'$ by

$$\langle \mathcal{P}_n u, v\rangle = \langle u | \mathcal{P}_n v\rangle \qquad \text{for all} \quad v \in L^2(0,1), \tag{2.71}$$

where $\langle\ |\ \rangle$ denotes the duality paring between $H^1(0,1)$ and $H^1(0,1)'$.

2.3.3 Definition (approximate weak solution)**.** A predictable E_n-valued process $u_n : [0,T] \to E_n$ is an *approximate weak solution* of the SPDE (1.10a) provided

1. $u_n \in L^2(0,T;E_n)$ and $u_n(0) = \mathcal{P}_n u_0$,

2. For every $v \in E_n$ and t almost every where in $[0,T]$

$$\langle u_n(t), v\rangle - \langle u_0, v\rangle + \int_0^t a(u_n, v)\mathrm{d}s + \int_0^t \langle |\partial_x u_n|^2, v\rangle \mathrm{d}s = \int_0^t \langle \eta\sqrt{u_n}\partial_{xs} W^n, v\rangle + \int_0^t \langle \kappa\gamma, v\rangle \mathrm{d}s. \tag{2.72}$$

where the stochastic integral is given in the notation of (2.68). This is equivalent to the condition that

$$\mathrm{d}u_n = \big[- A_n u_n + \mathcal{P}_n(|\partial_x u_n|^2) + \gamma\kappa\big]\mathrm{d}t + \mathcal{P}_n(\eta\sqrt{u_n}\mathrm{d}W(t)), \quad u_n(0) = \mathcal{P}_n u_0 \tag{2.73}$$

for almost every where in $[0,T]$, meaning that u_n takes values in E_n and satisfies the projection of the PDE onto E_n. In this notation $A_n : E_n \to E_n$ defined by

$$\langle A_n w, v\rangle = a(w,v), \ \forall w, v \in E_n. \tag{2.74}$$

2.3.4 Existence of the approximate weak solution in a finite dimensional space

Since every n-dimensional vector space is isomorphic to $\mathbb{R}^n$, we can replicate the weak form introduced (2.116) in $\mathbb{R}^n$ with a slight modifications. As we are seeking to use an energy estimate in Theorem 2.3.5 by testing with u_n, we have to think very carefully in the sign of the quadratic nonlinearity. In other words, if u_n turns to be negative it will leads to some complications. Thus in order to apply an energy estimate and reserve the quadratic nonlinearity sign, we have to slightly modify our model as follows

$$\partial_t u_n = -Au_n - \mathcal{P}_n\sigma_\varepsilon(u_n)|\partial_x u_n|^2 + \eta\mathcal{P}_n\sqrt{|u_n|}\partial_{xs}W^n + \kappa\gamma, \tag{2.75}$$

where n is fixed and σ_ε is the sign function with a truncation at zero namely

$$\sigma_\varepsilon(\phi) := \begin{cases} \frac{\phi}{|\phi|} & \text{if } |\phi| > \frac{1}{\varepsilon}, \\ \frac{\phi}{\varepsilon} & \text{if } |\phi| \le \frac{1}{\varepsilon}. \end{cases} \tag{2.76}$$

The idea to prove the existence of the approximate weak solution in $\mathbb{R}^n$ is to expand ϕ as a weak solution satisfies (2.75) by its coefficient

$$\phi(x,t) := \sum_{i=1}^n \phi_i(t)\chi_i(x), \quad \text{for all } \phi \in \mathbb{R}^n \tag{2.77}$$

then we define a truncation for the drift term and the noise coefficient for each $\varepsilon \ge 1$ such that

$$F_\varepsilon(\phi,t) := \begin{cases} F(\phi,t) & \text{if } |\phi| \le \varepsilon, \\ F(\varepsilon\phi/|\varepsilon|, t) & \text{if } |\phi| > \varepsilon, \end{cases} \tag{2.78}$$

and similarly for $g_\varepsilon(\phi,t)$, where

$$F(\phi,t) := -A\phi - \mathcal{P}_n\sigma_\varepsilon(\phi)|\partial_x\phi|^2 + \kappa\gamma \tag{2.79}$$

and,

$$g(\phi,t) := \mathcal{P}_n\eta\sqrt{|\phi|} \tag{2.80}$$

in addition to the truncation (2.78) on g we have another truncation at zero defined in (2.2). Then F_ε and g_ε satisfy the local Lipschitz condition (A.10). Now we define a stopping time as follows

$$\tau_\varepsilon = T \wedge \inf\{t \in [t_0,T] : |\phi(t)| \geq \varepsilon\}. \tag{2.81}$$

Hence by Theorem A.0.7 there exists a unique solution $\phi(t)$ to equation (2.75) in $\mathcal{H}_{2,T}(\mathbb{R}^n)$, where the local Lipschitz condition guarantees that the solution exists in $[0,\tau]$, where $\tau = \lim_{\varepsilon\to\infty}\tau_\varepsilon$, but the monotone condition guarantees that $\tau = T$, i.e., the solution exists on the whole interval $[0,T]$.

2.3.5 Theorem (Energy estimates)**.** *There exists $c_{(2.82)}$, constant in n, depending on $T,\vartheta,\kappa,\gamma$ and $c_{(1.47)}$ where $\vartheta > 0$ such that*

$$\|u_n\|_{L^\infty(0,T;L^2(0,1))} + \|u_n\|_{L^2(0,T;H^1(0,1))} + \|\partial_t u_n\|_{L^2(0,T;H^1(0,1)')} \leq c_{(2.82)} + \|u_0\|_{L^2(0,1)}, \tag{2.82}$$

for $n = 1,2,\ldots$.

Proof. We multiply the equation (2.114) by a test function $v = u_n$, integrating over the domain, applying the divergence theorem, use the boundary conditions and use the expectation to obtain the following

$$\begin{aligned}&\mathbb{E}\left[\int_0^1 \partial_t u_n u_n \mathrm{d}x + \int_0^1 A u_n u_n \mathrm{d}x + \int_0^1 |\partial_x u_n|^2 u_n \mathrm{d}x\right] \\ &= \mathbb{E}\left[\int_0^1 \sqrt{u_n} u_n \partial_{xt} W^n + \int_0^1 \kappa\gamma u_n \mathrm{d}x\right].\end{aligned} \tag{2.83}$$

Integrating with respect to time and using the initial condition, we have

$$\begin{aligned}\mathbb{E}\Bigg[\frac{1}{2}\int_0^1 u_n^2(x,t)\mathrm{d}x + \int_0^t\int_0^1 |\partial_x u_n|^2 \mathrm{d}x\mathrm{d}s + \kappa\int_0^t\int_0^1 u_n^2 \mathrm{d}x\mathrm{d}t \\ + \int_0^t\int_0^1 u_n|\partial_x u_n|^2\mathrm{d}x\mathrm{d}s\Bigg] = \mathbb{E}\Bigg[\underbrace{\int_0^t\int_0^1 \kappa\gamma u_n \mathrm{d}x\mathrm{d}s}_{\mathrm{I}} \\ + \int_0^t\int_0^1 u_n^{3/2}\partial_{xt}W^n + \frac{1}{2}\int_0^1 u_0^2\mathrm{d}x\Bigg].\end{aligned} \tag{2.84}$$

For $0 \le t \le T$, we have from Cauchy inequality with ϑ for I that

$$\begin{aligned}\int_0^t \int_0^1 \kappa\gamma u_n \mathrm{d}x\mathrm{d}s &\le \Big(\kappa\gamma \int_0^t \int_0^1 u_n^2 \mathrm{d}x\mathrm{d}s\Big)^{1/2} \\ &\le \frac{T}{4\vartheta}\kappa^2\gamma^2 + \vartheta \int_0^T \int_0^1 u_n^2 \mathrm{d}x\mathrm{d}s \\ &\le \frac{T}{4\vartheta}\kappa^2\gamma^2 + \vartheta T \max_{0\le t\le T} \int_0^1 u_n^2 \mathrm{d}x.\end{aligned} \tag{2.85}$$

Thus, taking the supremum of (2.84) over $t \in [0,T]$ and using this inequality with $\vartheta T = 1/4$ in the result, we have

$$\begin{aligned}\mathbb{E}\Big[\frac{1}{4}\max_{0\le t\le T}\int_0^1 u_n^2(x,t)\mathrm{d}x + \int_0^T\int_0^1 |\partial_x u_n|^2 \mathrm{d}x\mathrm{d}t + \kappa\int_0^T\int_0^1 u_n^2 \mathrm{d}x\mathrm{d}t \\ + \int_0^T\int_0^1 u_n|\partial_x u_n|^2 \mathrm{d}x\mathrm{d}t\Big] \\ = \mathbb{E}\Big[\frac{T}{4\vartheta}\kappa^2\gamma^2 + \underbrace{\int_0^T\int_0^1 u_n^{3/2}\partial_{xt}W^n}_{\mathrm{II}} + \frac{1}{2}\int_0^1 u_0^2 \mathrm{d}x\Big].\end{aligned} \tag{2.86}$$

Finally, making use of Lemma 1.1.20 we can bound the stochastic term as follows,

$$\begin{aligned}\mathrm{II} &= \mathbb{E}\Big[\int_0^T\int_0^1 u_n^{3/2}\partial_{xt}W^n\Big] \\ &\le \mathbb{E}\Big[\int_0^T \|u_n^{3/2}\|_{H^1(0,1)}\|\partial_{xt}W^n\|_{H^1(0,1)'}\Big] \\ &\le \frac{3\vartheta}{2}\mathbb{E}\int_0^T \|u_n^{1/2}|\partial_x u_n|\|^2_{L^2(0,1)}\mathrm{d}t + \frac{1}{2\vartheta}\mathbb{E}\int_0^T \|\partial_{xt}W^n\|^2_{H^1(0,1)'},\end{aligned} \tag{2.87}$$

where $\int_0^T\int_0^1 |u_n||\partial_x u_n|^2 \mathrm{d}x\mathrm{d}t$ can be observed in the right hand side for a small ϑ. It follows that we have an a priori energy estimate of the form

$$\|u_n\|^2_{L^\infty(0,T;L^2(0,1))} + \|u_n\|^2_{L^2(0,T;H^1(0,1))} \le c_{(2.82)} + \|u_0\|^2_{L^2(0,1)}, \tag{2.88}$$

where $c_{(2.82)} := \frac{T}{4\vartheta}\kappa^2\gamma^2 + \frac{3\vartheta}{2} + \frac{Tc_{(1.47)}}{2\vartheta}$ which is uniform in n, hence we can pass the limit as $k \to \infty$. To estimate $\partial_t u_n$, we note that since $\partial_t u_n(t) \in E_n$

$$\|\partial_t u_n(t)\|_{H^1(0,1)'} = \sup_{v\in E_n\setminus 0} \frac{\langle \partial_t u_n(t)|v\rangle_{H^1(0,1)'\times H^1(0,1)}}{\|v\|_{H^1(0,1)}}. \tag{2.89}$$

From (2.114), (A.32) and Sobolev embedding theorem 1.1.7 we have

$$\begin{aligned}\langle \partial_t u_n(t), v\rangle &\le |a(u_n,v)| + |\langle |\partial_x u_n|^2, v\rangle| + |\langle \sqrt{u_n}\partial_{xt}W^n, v\rangle| + |\langle \kappa\gamma, v\rangle| \\ &\le c_{(2.90)}(\|u_n(t)\|_{H^1(0,1)} + \|\sqrt{u_n}\partial_{xt}W^n\|_{H^1(0,1)'})\|v\|_{H^1(0,1)},\end{aligned} \tag{2.90}$$

for every $v \in H^1(0,1)$, where we used Sobolev embedding theorem in one spacial dimension to obtain a uniform bound in n for the quadratic nonlinearity term of u_n based on the following lemma

2.3.6 Lemma. *For u_n satisfies the energy estimate (2.82) for each $n \in \mathbb{N}$ we have*

$$\int_0^1 |\partial_x u_n|^2 v \leq c_{(2.91)} \|v\|_{H^1(0,1)}. \tag{2.91}$$

Proof. The energy estimate (2.82) implies the fact that

$$\int_0^T \int_0^1 |\partial_x u_n|^2 u_n \leq c_{(2.92)}, \tag{2.92}$$

where $c_{(2.92)}$ is uniform in n. Theorem 1.1.7 states that $W_p^1(\mathbb{R}^n)$ is embedded into the Hölder space $C^0_{1-n/p}(\mathbb{R}^n)$. In our case $n = 1$ and $p = 2$ which implies that $H^1(0,1)$ is continuously embedded into $C^0_{1/2}(0,1)$, which is turn continuously embedded into $L^\infty(0,1)$. Hence there exists a constant $c_{(2.93)}$ such that

$$\|v\|_{L^\infty(0,1)} \leq c_{(2.93)} \|v\|_{H^1(0,1)}. \tag{2.93}$$

Thus we have

$$\int_0^1 |\partial_x u_n|^2 v \leq c_{(2.91)} \|v\|_{H^1(0,1)}, \tag{2.94}$$

for each $v \in H^1(0,1)$, where $c_{(2.91)} := c_{(2.92)} c_{(2.93)}$. □

Now we integrate equation (2.90) with respect to time and using (2.82) in the result, we obtain

$$\|\partial_t u_n\|^2_{L^2(0,T;H^1(0,1)')} \leq c_{(2.90)} (\|u_0\|^2_{L^2(0,1)} + \|\sqrt{u_n}\partial_{xt} W^n\|_{L^2(0,T;H^1(0,1)')}). \tag{2.95}$$

Equations (2.82) and (2.95) complete the proof of Theorem 2.3.5. □

2.3.7 Convergence of the approximate solution

We obtain compactness from the Aubin–Lions theorem because $H^1(0,1) \hookrightarrow L^2(0,1) \hookrightarrow H^1(0,1)'$, with the functions $u_n : (0,T) \to H^1(0,1)$ such that $\{u_n\}$ is uniformly bounded in $L^2(0,T;H^1(0,1))$ and $\{\partial_t u_n\}$ is uniformly bounded in $L^2(0,T;H^1(0,1)')$, then there is a subsequence which we call again, without loss of generality, u_n that converges strongly in $L^2(0,T;L^2(0,1))$. Moreover, u_n converges weakly in $L^2(0,T;H^1(0,1))$ and $\partial_t u_n$ converges weakly in $L^2(0,T;H^1(0,1)')$. We define the limit function as a generalised solution of the original equation. Like any other property of equation (1.10a), showing that the limit function is a mild solution is not straight-forward. Let us "claim" that the limit function is a mild solution in the following notation

$$\begin{aligned} u_n(t) = \mathrm{e}^{-tA}\,\mathcal{P}_n u_0 + \int_0^t \mathrm{e}^{-(t-s)A}\,\mathcal{P}_n |\partial_x u_n|^2 \mathrm{d}s + \int_0^t \mathrm{e}^{-(t-s)A}\,\mathcal{P}_n \sqrt{u_n}\partial_{xs} W^n \\ + \kappa\gamma \int_0^t \mathrm{e}^{-(t-s)A}\,\mathcal{P}_n \mathbb{1}_{(0,1)} \mathrm{d}s. \end{aligned} \tag{2.96}$$

In this section we are pointing out the difficulty behind this claim. The uniform convergence that we obtained form Aubin–Lions theorem implies that for each $\epsilon > 0$ we have

$$\|u_n - u\|_{L^2(\Omega,T;H^1(0,1))} < \epsilon. \tag{2.97}$$

Then we can take the limit as n goes to zero over the sides of (2.96). However, this is NOT the case, where the quadratic nonlinearity term is not easy to bound in the given norm i.e.

$$\Big\| \int_0^T \int_0^t \mathrm{e}^{-(t-s)A}(\mathcal{P}_n|\partial_x u_n|^2 - |\partial_x u|^2)\mathrm{d}s\mathrm{d}t \Big\|_{L^2(\Omega,T;H^1(0,1))} \not< \infty. \tag{2.98}$$

Therefore we introduce the equivalence of weak and mild solutions in the Fourier–Galerkin space E_n in the next remark.

2.3.8 Remark (Equivalence of weak and mild solutions in the Fourier–Galerkin space E_n). For $v \in E_n \subset L^2(0,1)$ we have $\langle v|\chi_j\rangle = \langle v, \chi_j\rangle = 0$ for $v \neq 0$ and $j > n$, where $E_n = span\{\chi_0, \dots, \chi_n\}$. Then we can redefine the semigroup operator that generated by $-A$ to be in the following form

$$\mathrm{e}^{-At}\, v = \sum_{j=0}^{n} \mathrm{e}^{\lambda_j t}\langle v|\chi_j\rangle\chi_j \in E_n, \tag{2.99}$$

Thus if $u_n \in E_n$ we have

$$\begin{aligned}\langle u_n(t), v\rangle - \langle u_0, v\rangle + \int_0^t a(u_n, v)\mathrm{d}s + \int_0^t \langle|\partial_x u_n|^2, v\rangle\mathrm{d}s = \int_0^t \langle\sqrt{u_n}\partial_{xs}W^n, v\rangle \\ + \int_0^t \langle\kappa\gamma, v\rangle\mathrm{d}s,\end{aligned} \tag{2.100}$$

for each $n \in \mathbb{N}$ and $v \in E_n$ if and only if u_n satisfies the mild form

$$\begin{aligned}u_n(t) = \mathrm{e}^{-tA}\,\mathcal{P}_n u_0 + \int_0^t \mathrm{e}^{-(t-s)A}\,\mathcal{P}_n|\partial_x u_n|^2\mathrm{d}s + \int_0^t \mathrm{e}^{-(t-s)A}\,\mathcal{P}_n\sqrt{u_n}\partial_{xs}W^n \\ + \kappa\gamma\int_0^t \mathrm{e}^{-(t-s)A}\,\mathcal{P}_n\mathbb{1}_{(0,1)}\mathrm{d}s.\end{aligned} \tag{2.101}$$

2.4 Positivity of the approximate solution

In this section we show that the solution is positive by using a comparison-principle, where we build a subsolution. The subsolution we build is positive thanks to Feller's condition which prevents the solution to be negative. Simply if two solutions one above the other, then they can not a cross hence the solution stays positive.

2.4.1 Lemma (Comparison). *Suppose three functions $\phi, \psi, f : [0,T] \to \mathbb{R}$, satisfy $\phi_0 \geq \psi_0 > 0$ where,*

$$\phi'(t) + \kappa(\phi(t) - \gamma) - \eta\sqrt{\phi(t)}\dot{W}(t) = f(t), \tag{2.102a}$$

and

$$\psi'(t) + \kappa(\psi(t) - \gamma) - \eta\sqrt{\psi(t)}\dot{W}(t) = 0, \tag{2.102b}$$

it follows that $\phi(t) \geq \psi(t) > 0$ *for all* t.

Proof. The proof of the existence to (2.102a) in $[0, T]$ is similar to the existence proof for $\psi(t)$ in Lemma 2.1.1. We want to show that $\phi(t) \geq \psi(t)$ holds where

$$\phi(t) = \phi_0 - \int_0^t \kappa(\phi(s) - \gamma)\mathrm{d}s + \int_0^t \eta\sqrt{\phi(s)}\mathrm{d}W(s) + \int_0^t f(s)\mathrm{d}s, \tag{2.103}$$

and

$$\psi(t) = \psi_0 - \int_0^t \kappa(\psi_(s) - \gamma)\mathrm{d}s + \int_0^t \eta\sqrt{\psi(s)}\mathrm{d}W(s). \tag{2.104}$$

by the assumption we have $\phi_0 \geq \psi_0$ and let $w(t) := \psi(t) - \phi(t)$. Then

$$w(t) = w_0 - \int_0^t \kappa w(s)\mathrm{d}s + \int_0^t \eta(\sqrt{\psi(s)} - \sqrt{\phi(s)})\mathrm{d}W(s) - \int_0^t f(s)\mathrm{d}s, \tag{2.105}$$

with initial value $w_0 = \psi_0 - \phi_0$. We recall the Yamada–Watanabe result [Karatzas and Shreve, 1988], [Mytnik et al., 2006]. Taking $h := \sqrt{x}$ we know this

$$\int_{(0,\varepsilon)} h^{-2}(x)\mathrm{d}x = \infty, \ \forall \varepsilon > 0. \tag{2.106}$$

Then we have

$$\left|\sqrt{(x,t)} - \sqrt{(y,t)}\right| \leq h(|x - y|), \quad \text{for every } 0 \leq t < \infty. \tag{2.107}$$

Because of the conditions imposed on the function h (i.e. $h : [0, \infty) \to [0, \infty)$ strictly increasing function with $h(0) = 0$ and (2.106)), there exists a strictly decreasing sequence $\{a_n\}_{n=0}^{\infty} \subseteq (0, 1]$ with $a_0 = 1$, $\lim_{n\to\infty} a_n = 0$ and $\int_{a_n}^{a_{n-1}} h^{-2}(x)\mathrm{d}x = n$, for every $n \geq 1$. For each $n \geq 1$, there exists a continuous function ρ_n on $\mathbb{R}$ with support in (a_n, a_{n-1}) so that $0 \leq \rho_n(x) \leq (2/nh^2(x))$ holds for every $x > 0$, and $\int_{a_n}^{a_{n-1}} \rho_n(x)\mathrm{d}x = 1$. Then the function

$$M_n(x) = \int_0^{[x]_+} \int_0^y \rho_n(\xi)\mathrm{d}\xi\mathrm{d}y\mathbb{1}_{(0,\infty)}(x), \ \text{for } x \in \mathbb{R}, \ n \geq 1. \tag{2.108}$$

is even and twice continuously differentiable, with $|M_n'(x)| \leq 1$ and $\lim_{n\to\infty} M_n(x) = [x]_+$. Furthermore, the sequence $\{M_n\}_{n=1}^{\infty}$ is nondecreasing. Then we apply by Itô lemma for $w(t)$ under the fact

$$\mathbb{E}\int_0^t |\sqrt{\psi(s)} - \sqrt{\phi(s)}|^2\mathrm{d}s < \infty, \quad 0 \leq t < \infty. \tag{2.109}$$

as follows

$$M_n(w(t)) = \int_0^t M_n'(w(s))\big(\kappa w(s) - f(s)\big)\mathrm{d}s \tag{2.110}$$
$$+ \frac{1}{2}\int_0^t M_n''(w(s))\Big(\sqrt{\psi(s)} - \sqrt{\phi(s)}\Big)^2\mathrm{d}s$$
$$+ \int_0^t M_n'(w(s))\Big(\sqrt{\psi(s)} - \sqrt{\phi(s)}\Big)\mathrm{d}W(s).$$

The expectation of the stochastic integral in (2.110) is zero thanks to (2.109) we know it is a martingale, i.e., $\mathbb{E}\Big[\int g(s)\mathrm{d}W(s)\Big] = 0$, whereas the expectation of the second integral in (2.110) is bounded above by

$$\mathbb{E}\int_0^t M_n''(w(s))|w(s)|\mathrm{d}s \le \frac{2t}{n}. \tag{2.111}$$

Owing to the positivity of f,we have $\mathbb{E}\big[\int f(s)\mathrm{d}s\big] \ge 0$. Thus we conclude that

$$\mathbb{E}M_n(w(t)) \le \mathbb{E}\int_0^t M_n'(w(s))\big(\kappa w(s)\big)\mathrm{d}s + \frac{t}{n} \tag{2.112}$$
$$\le \kappa\int_0^t \mathbb{E}[w(s)_+]\mathrm{d}s + \frac{t}{n}, \quad t \ge 0, n \ge 1.$$

Now we can let $n \to \infty$ to obtain $\mathbb{E}[w(s)_+] \le \kappa\int_0^t \mathbb{E}[w(s)_+]\mathrm{d}s$, for $0 \le t < \infty$ and by the Gronwall inequality, we have $\mathbb{E}[w(s)_+] = 0$ namely $\phi(t) \ge \psi(t)$ a.s.

□

2.4.2 Theorem (Positivity of the solution)**.** *The limit function u defined as the limit of the approximate solution u_n in section 2.3.7 $L^2(0,T;L^2(0,1))$ is positive.*

Since the L^2 convergence of a positive sequence implies the positivity of the limit, it is enough to show that the sequence of approximate solution u_n is positive.

2.4.3 Lemma (Positivity of solution)**.** *A function u_n that satisfies equation (2.116) for a fixed n in the finite element space E_n defined in (2.69) is strictly positive.*

Proof. Let $\xi(t)$ be a global minimum of $u_n(\cdot,t)$ which means $-\partial_{xx}u_n(\xi(t),t) \le 0$, and define

$$\rho(t) := u_n(\xi(t),t), \tag{2.113}$$

where u_n solves,

$$\langle u_n(t),v\rangle - \langle u_0,v\rangle + \int_0^t a(u_n,v) + \int_0^t \langle|\partial_x u_n|^2,v\rangle = \int_0^t \langle\eta\sqrt{u_n}\mathrm{d}W^n,v\rangle$$
$$+ \int_0^t \langle\kappa\gamma,v\rangle \quad \text{for every } v \in E_n. \tag{2.114}$$

where W^n defined as a finite space–time white noise as follows

$$W^n(\cdot,t)=\sum_{i=1}^{n}\chi_i(\cdot)\beta_i(t). \tag{2.115}$$

Rewriting (2.114) in pointwise form we get

$$\mathrm{d}u_n=\big[-A_nu_n+\mathcal{P}_n(|\partial_xu_n|^2)+\gamma\kappa\big]\mathrm{d}t+\mathcal{P}_n(\eta\sqrt{u_n}\mathrm{d}W^n),\quad u_n(0)=\mathcal{P}_nu_0 \tag{2.116}$$

almost everywhere in $[0,T]$. In this notation $A_n:E_n\to E_n$ defined by

$$\langle A_nw,v\rangle=a(w,v),\ \forall w,v\in E_n, \tag{2.117}$$

and $\mathcal{P}_n$ given in the notation of (2.71). Assume the process $\xi(t)$ defined above is an Itô drift-diffusion process that satisfies the stochastic differential equation

$$\mathrm{d}\xi(t)=\mu(t)\mathrm{d}t+\sigma(t)\mathrm{d}B(t), \tag{2.118}$$

where $B(t)$ stands for Brownian motion and hence form Itô lemma we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\rho(t)&=\partial_tu_n(\xi(t),t)+\mu(t)\partial_xu_n(\xi(t),t)+\frac{\sigma^2(t)}{2}\partial_{xx}u_n(\xi(t),t)\\&\quad+\sigma(t)\partial_xu_n(\xi(t),t)\mathrm{d}B(t),\end{aligned} \tag{2.119}$$

since $\xi(t)$ minimises $u_n(\cdot,t)$, $\partial_xu_n(\xi(t),t)=0$ thus

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t)=\partial_tu_n(\xi(t),t)+\frac{\sigma^2(t)}{2}\partial_{xx}u_n(\xi(t),t), \tag{2.120}$$

which implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t)=\partial_{xx}u_n(\xi(t),t)\underbrace{(1+\frac{\sigma^2(t)}{2})}_{>0}-\mathcal{P}_n\kappa(\rho(t)-\gamma)-\mathcal{P}_n\underbrace{|\partial_xu_n(\xi(t),t)|^2}_{=0}+Q(u_n)\mathrm{d}\widetilde{B}^n, \tag{2.121}$$

where $Q(u_n):=\eta\mathcal{P}_n\sqrt{u_n(\xi(t),t)}$, and $\widetilde{B}^n$ is the linear combination given by

$$\widetilde{B}^n(t):=\sum_{i=1}^{n}\chi_i(\xi(t),t)\mathrm{d}\beta_i(t), \tag{2.122}$$

note that $\widetilde{B}$ is a Brownian motion because by Lévy's Characterization Theorem [Mao, 2008], it is a continuous local martingale. Thus,

$$\partial_{xx}u_n(\xi(t),t)=\mathrm{d}\rho(t)+\kappa(\rho(t)-\gamma)-Q(u_n)\mathrm{d}\widetilde{B}^n, \tag{2.123}$$

$$\mathrm{d}\rho(t)+\kappa(\rho(t)-\gamma)-Q(u_n)\mathrm{d}\widetilde{B}^n\geq 0, \tag{2.124}$$

which implies that $\rho(t)$ is positive thanks to Theorem 2.4.1. □

To conclude, as the Yamada and Watanabe theorem allows to considerably relax the Lipschitz condition on the dispersion coefficient in the one-dimensional case. We managed to show the positivity of solution by comparison lemma with some weak minimum principle theorem inspiration, thanks to Feller's condition.

Chapter 3

Numerical experiments

In this chapter we numerically approximate the weak solution to equations (1.10a)-(1.10d) with a finite element method in space and Euler–Maruyama method in time. We start by approximating the linear deterministic version of (1.10a), then we follow that by approximating the non-linear deterministic version of (1.10a). We then add time dependent white noise and then close the chapter by adding a space–time white noise.

3.1 Numerical approximation of the heat equation with a mean reverting term

Considering the deterministic version of (1.10a) without the quadratic nonlinearity term

$$\partial_t u(x,t) = -\kappa(u(x,t)-\gamma) + \partial_{xx}u(x,t) + f(x,t), \text{ on } (0,1)\times\mathbb{R}^+, \tag{3.1a}$$

$$\partial_x u(0,t) = -u(0,t) + g, \ t \geq 0, \tag{3.1b}$$

$$\partial_x u(1,t) = u(1,t) + g, \ t \geq 0, \tag{3.1c}$$

$$u(0) = u_0, \tag{3.1d}$$

where $f(x,t)$ is a source function. We start by finding the weak form of (3.1a)-(3.1d). To do this we multiply (3.1a) by a test function $\varphi \in H^1(0,1)$, use integration by parts and the boundary conditions (3.1b)-(3.1c) to obtain the weak form

$$\langle \partial_t u, \varphi\rangle + \langle \partial_x u, \partial_x \varphi\rangle + \kappa\langle u-\gamma, \varphi\rangle = \langle f, \varphi\rangle + \Big[\partial_x u(x)\varphi(x)\Big]_0^1 \ \text{ for all } \ \varphi \in H^1(0,1), \tag{3.2}$$

A straightforward approach to solve a time-dependent PDEs is to first discretise the time derivative by a finite difference approximation, which yields a sequence of stationary problems, and then apply the finite element method to each. For a uniform time-step $\tau := t^{n+1} - t^n > 0$, for all $n \geq 0$ and initial condition $u_0 \in \mathbb{R}$, the one-step-theta finite difference scheme is defined by

$$\frac{\langle u^{n+1},\varphi\rangle - \langle u^n,\varphi\rangle}{\tau} = \Big[\theta\Big(- \langle \partial_x u^{n+1}, \partial_x\varphi\rangle - \kappa\langle u^{n+1}-\gamma,\varphi\rangle + \langle f^{n+1},\varphi\rangle \tag{3.3}$$
$$+ \Big[\partial_x u^{n+1}(x)\varphi(x)\Big]_0^1\Big) + \Big(1-\theta\Big)\Big(- \langle \partial_x u^{n}, \partial_x\varphi\rangle - \kappa\langle u^{n}-\gamma,\varphi\rangle + \langle f^{n},\varphi\rangle + \Big[\partial_x u^{n}(x)\varphi(x)\Big]_0^1\Big)\Big] \quad \text{for all } \varphi\in H^1(0,1), \tag{3.4}$$

where $\theta\in[0,1]$ is implicitness parameters. Some typical choices of θ from the literature define the following methods

$\theta = 0$	Forward Euler scheme	explicit,
$\theta = \frac{1}{2}$	Crank-Nicolson scheme	implicit,
$\theta = 1$	Backward Euler scheme	implicit.

In the heat equation case we know that the scheme is stable if $\theta \geq 1/2$, for $\theta < 1/2$, it is conditionally stable, requiring a condition on the length of the time step [Thomée, 2006]. Hence, for stability purposes, we take $\theta = 1$ and use the implicit backward Euler scheme. The result is a sequence of spatial (stationary) problems for u^{n+1}, assuming u^n is known from the previous time step:

$$\langle u^{n+1},\varphi\rangle - \langle u^n,\varphi\rangle = -\tau\langle \partial_x u^{n+1},\partial_x\varphi\rangle - \tau\kappa\langle u^{n+1}-\gamma,\varphi\rangle + \tau\langle f^{n+1},\varphi\rangle + \tau\Big[\partial_x u^{n+1}(x)\varphi(x)\Big]_0^1, \quad \text{for all } \varphi\in H^1(0,1). \tag{3.5}$$

$$u^0 = u_0, \tag{3.6}$$

for $n = 0,1,2,\ldots$. Given u^0, we can recursively solve for u_n. The resulting weak form arising from formulation (3.5) can be conveniently written in the standard notation:

$$a(u^{n+1},\varphi) = L_n(\varphi), \tag{3.7}$$

where,

$$a(u^{n+1},\varphi) = \int_0^1 \Big[u^{n+1} + \tau\kappa(u^{n+1}-\gamma)\Big]\varphi \mathrm{d}x + \tau\int_0^1 \partial_x u^{n+1}\cdot\partial_x\varphi \mathrm{d}x, \tag{3.8}$$

$$L_n(\varphi) = \int_0^1 \Big[u^n + \tau f^{n+1}\Big]\varphi \mathrm{d}x + \tau\Big[\partial_x u^{n+1}(x)\varphi(x)\Big]_0^1. \tag{3.9}$$

In addition to the variational problem to be solved in each time step, we also need to approximate the initial condition (3.6). This equation can also be turned into a variational problem:

$$a_0(u,\varphi) = \int_0^1 u\varphi \mathrm{d}x, \tag{3.10}$$

$$L_0(\varphi) = \int_0^1 u_0\varphi \mathrm{d}x. \tag{3.11}$$

When solving this variational problem, u_0 becomes the L^2 projection of the given initial value u_0 into the finite element space. The alternative is to construct u_0 by just interpolating the initial value u_0; that is, if $u_0 = \sum_{j=1}^{n} U_j^0 \phi_j$, we simply set $U_j = u_0(x_j, y_j)$, where (x_j, y_j) are the coordinates of node number j [Langtangen and Logg, 2017].

In summary, we need to finding $u_0 \in H^1(0,1)$ such that $a_0(u_0, \varphi) = L_0(\varphi)$ holds for all $\varphi \in H^1(0,1)$, and then find $u^{n+1} \in H^1(0,1)$ such that $a(u^{n+1}, \varphi) = L^{n+1}(\varphi)$ for all $\varphi \in H^1(0,1)$, for $n = 0, 1, 2, \ldots$

Finally, we discretise space so to have finitely many disjoint points and solve the approximation over the mesh that is created from the discretisation. Since we have one space dimension, the idea is to partition $[0,1]$ into subintervals or elements and choose the finite element spaces to be a set of piecewise linear functions that are defined on the elements. Let $I = [0,1]$, the sequence of nodes of h such that

$$0 = x_0 < x_1 < \cdots < x_J = 1. \tag{3.12}$$

Define the partitions of I as $\varsigma_j = [x_{j-1}, x_j]$, for $j = 1, \ldots, J$. We set $h := \max_{j=1,\ldots,J} |\varsigma_j|$. We perform experiments with decreasing value of the mesh size $h_0 > h_1 > h_2 > \ldots$ and monitor the following error norms

$$\mathcal{E}_{L^2} := \|E\|_{L^2(0,1)}, \quad \mathcal{E}_{L^\infty} := \|E\|_{L^\infty(0,1)}, \tag{3.13}$$

where E stands for the error function given by $E := I^h u - u^h$ and $I^h u$ is the piecewise linear interpolant of u. We quantify the error norms using the estimated order of convergence (eoc)

$$\text{eoc}_j := \frac{\ln(\mathcal{E}_{i,j}/\mathcal{E}_{i,j-1})}{\ln(h_j/h_{j-1})}, \tag{3.14}$$

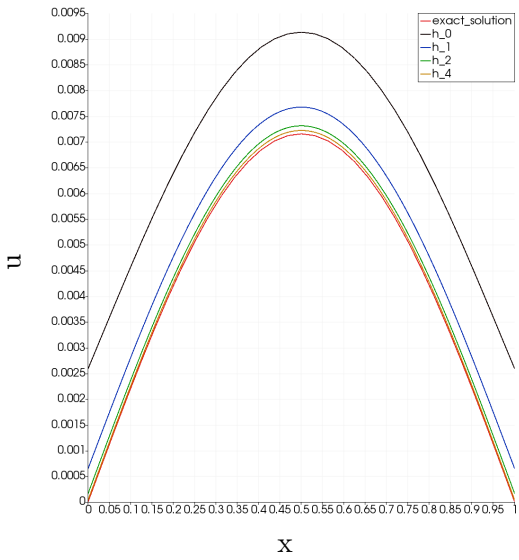
where i corresponds to the error norm, j corresponds to the relative mesh size h_j and $\mathcal{E}_{i,j}$ corresponds to the error of the error norm i at the level j.

3.1.1 Test problem

Considering the initial data $u_0(x) = \sin(\pi x)$ and the forcing $f(x,t) = \kappa(\mathrm{e}^{-\pi^2 t}\sin(\pi x) - \gamma)$ then we see that the solution to (3.1a)-(3.1d) is given by

$$u = \mathrm{e}^{-\pi^2 t}\sin(\pi x), \tag{3.15}$$

with $u(x,t) = -\pi^2 \mathrm{e}^{-\pi^2 t}\sin(\pi x)$ on the boundaries. Comparing the exact and numeric solutions for different levels of mesh sizes namely $h_0 = 1/32, h_1 = 1/64, h_2 = 1/128$, and $h_3 = 1/256$ suggests some type of convergence at the final time in the following figure



Tables 3.1-3.2 below are showing the error norms for (3.1a)-(3.1d) and (3.15). The error obtained by taking the uniform time-step τ to be equal to h^2, $\kappa = 0.25$ and $\gamma = 0.25$ on time interval $[0, 0.5]$ for 1024 time steps is shown in Table 3.1. The error obtained by taking $\tau = h^2$, $\kappa = 1.0$ and $\gamma = 0.5$ on the time interval $[0, 0.5]$ for 1024 time steps is shown in Table 3.2. Finally in Table 3.3 we obtain the error norms for (3.1a)-(3.1d) and (3.15) by taking $\tau = h$, $\kappa = 1.0$ and $\gamma = 0.5$ on the time interval $[0, 0.5]$ for 1024 time steps. All the results exhibit the order of convergence expected from the theory for parabolic-type partial differential equations.

J	$\mathcal{E}_{L^2}$	eoc_1	$\mathcal{E}_{L^\infty}$	eoc_2
32	2.359×10^{-3}	-	2.797×10^{-3}	-
64	5.897×10^{-4}	2.000 514 29	1.735×10^{-3}	2.001 834 65
128	1.474×10^{-4}	2.000 128 75	4.337×10^{-4}	2.000 458 76
256	3.685×10^{-5}	2.000 032 61	1.084×10^{-4}	2.000 114 86
512	9.213×10^{-6}	2.000 008 16	2.710×10^{-5}	2.000 027 92

Table 3.1: Error norms for $\tau = h^2$, $\kappa = 0.25$ and $\gamma = 0.25$.

J	$\mathcal{E}_{L^2}$	eoc_1	$\mathcal{E}_{L^\infty}$	eoc_2
32	1.791×10^{-3}	-	2.123×10^{-3}	-
64	4.479×10^{-4}	1.999 608 83	5.310×10^{-4}	1.999 401 61
128	1.119×10^{-4}	1.999 902 52	1.327×10^{-4}	1.999 850 71
256	2.799×10^{-5}	1.999 975 62	3.319×10^{-5}	1.999 962 67
512	6.999×10^{-6}	2.000 000 31	8.298×10^{-6}	1.999 997 07

Table 3.2: Error norms for $\tau = h^2$, $\kappa = 1.0$ and $\gamma = 0.5$.

J	$\mathcal{E}_{L^2}$	eoc_1	$\mathcal{E}_{L^\infty}$	eoc_2
32	1.791×10^{-3}	-	2.123×10^{-3}	-
64	1.087×10^{-3}	7.198×10^{-1}	1.289×10^{-3}	7.196×10^{-1}
128	5.917×10^{-4}	8.781×10^{-1}	7.015×10^{-4}	8.781×10^{-1}
256	3.078×10^{-4}	9.427×10^{-1}	3.649×10^{-4}	9.427×10^{-1}
512	1.569×10^{-4}	9.722×10^{-1}	1.860×10^{-4}	9.722×10^{-1}

Table 3.3: Error norms for $\tau = h$, $\kappa = 1.0$ and $\gamma = 0.5$.

3.1.2 Numerical approximation of PDE with quadratic nonlinearity

In this section we find the numerical solution of the deterministic part of (1.10a) with quadratic nonlinearity i.e.

$$\partial_t u(x,t) = -\kappa(u(x,t) - \gamma) + \partial_{xx} u(x,t) - |\partial_x u(x,t)|^2 + f(x,t), \text{ on } (0,1) \times \mathbb{R}^+, \tag{3.16a}$$

$$\partial_x u(0,t) = -u(0,t) + g, \ t \geq 0, \tag{3.16b}$$

$$\partial_x u(1,t) = u(1,t) + g, \ t \geq 0, \tag{3.16c}$$

$$u(0) = u_0. \tag{3.16d}$$

Using the discretisation process outlined in the previous section, noting the introduction of the non-linear gradient term, we use the following Implicit-Explicit scheme

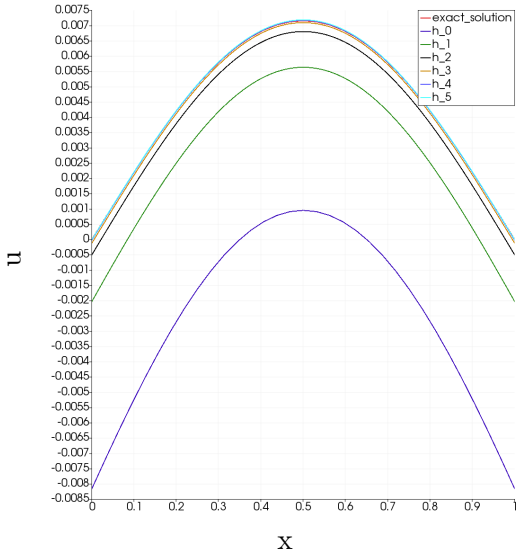
$$a(u^{n+1}, \varphi) = \int_0^1 \Big[u^{n+1} + \tau\kappa(u^{n+1} - \gamma)\Big]\varphi \mathrm{d}x + \tau \int_0^1 \partial_x u^{n+1} \cdot \partial_x \varphi \mathrm{d}x, \tag{3.17}$$

$$L_n(\varphi) = \int_0^1 \Big[u^n + \tau(\partial_x u^n \cdot \partial_x u^n) + \tau f^{n+1}\Big]\varphi \mathrm{d}x + \tau \Big[\partial_x u^{n+1}(x)\varphi(x)\Big]_0^1, \tag{3.18}$$

where we deal with the quadratic nonlinearity by using the known solution u^n from the previous time step , i.e., $(\partial_x u^n \cdot \partial_x u^n)$. Using the same exact solution (3.15) with righthand side given by:

$$f(x,t) = \kappa(e^{-\pi^2 t}\sin(\pi x) - \gamma) + |\pi e^{-\pi^2 t}\cos(\pi x)|^2, \tag{3.19}$$

and Robin boundary with $g(t) = \pi e^{-\pi^2 t}$ to calculate the error norms in the following tables. yields the following graph of the exact and the numeric solutions for different levels of mesh sizes namely $h_0 = 1/32, h_1 = 1/64, h_2 = 1/128, h_3 = 1/256, h_4 = 1/512$ and $h_5 = 1024$



As we can see from the figure above our scheme is generating negative values for the solution even if the Feller condition $2\kappa\gamma > 0$ is met [1]. This is because the Feller condition alone is not enough to guarantee positivity of the discrete time processes [Rouah, 2013]. For our purpose the negative values are floored at zero i.e. we replace $u(t)$ by $u_+(t) = \max(0, u(t))$ everywhere in the discretisation.

Tables 3.4-3.5 below are showing the error norms for (3.16a)-(3.16d) and (3.15). The error obtained by taking $\tau = h^2$, $\kappa = 0.25$ and $\gamma = 0.25$ on time interval $[0, 0.5]$ for 1024 time steps is shown in Table 3.4. The error obtained by taking $\tau = h^2$, $\kappa = 1.0$ and $\gamma = 0.5$ on the time interval $[0, 0.5]$ for 1024 time steps is shown in Table 3.5. Finally in Table 3.6 we obtain the error norms for (3.16a)-(3.16d) and (3.15) by taking $\tau = h$, $\kappa = 1.0$ and $\gamma = 0.5$ on the time interval $[0, 0.5]$ for 1024 time steps. All the results exhibit the order of convergence expected from the theory for parabolic-type partial differential equations.

[1] Note that Feller's condition is given by $2\kappa\gamma > \eta$, however as we are working in a deterministic case we are taking $\eta = 0$

J	$\mathcal{E}_{L^2}$	eoc_1	$\mathcal{E}_{L^\infty}$	eoc_2
32	1.358×10^{-2}	-	1.613×10^{-2}	-
64	3.383×10^{-3}	2.005 201 57	4.020×10^{-3}	2.005 154 37
128	8.451×10^{-4}	2.001 308 17	1.00×10^{-3}	2.001 296 55
256	2.112×10^{-4}	2.000 327 46	2.509×10^{-4}	2.000 324 57
512	5.280×10^{-5}	2.000 082 24	6.274×10^{-5}	2.000 081 51

Table 3.4: Error norms for $\tau = h^2$, $\kappa = 0.25$ and $\gamma = 0.25$.

J	$\mathcal{E}_{L^2}$	eoc_1	$\mathcal{E}_{L^\infty}$	eoc_2
32	4.622×10^{-3}	-	5.491×10^{-3}	-
64	1.154×10^{-3}	2.001 071 28	1.372×10^{-3}	2.000 915 75
128	2.886×10^{-4}	2.000 271 08	3.429×10^{-4}	2.000 232 27
256	7.215×10^{-5}	2.000 067 98	8.573×10^{-5}	2.000 058 29
512	1.803×10^{-5}	2.000 014 55	2.143×10^{-5}	2.000 012 13

Table 3.5: Error norms for $\tau = h^2$, $\kappa = 1.0$ and $\gamma = 0.5$.

J	$\mathcal{E}_{L^2}$	eoc_1	$\mathcal{E}_{L^\infty}$	eoc_2
32	5.491×10^{-3}	-	4.622×10^{-3}	-
64	2.788×10^{-3}	9.779×10^{-1}	2.346×10^{-3}	9.782×10^{-1}
128	1.404×10^{-3}	9.893×10^{-1}	1.181×10^{-3}	9.894×10^{-1}
256	7.047×10^{-4}	9.948×10^{-1}	5.930×10^{-4}	9.948×10^{-1}
512	3.529×10^{-4}	9.974×10^{-1}	2.970×10^{-4}	9.974×10^{-1}

Table 3.6: Error norms for $\tau = h$, $\kappa = 1.0$ and $\gamma = 0.5$.

3.2 Numerical approximation of SPDE driven by time white noise

In this section we examine the numerical approximation of the following equation:

$$\partial_t u(x,t) = -\kappa(u(x,t)-\gamma) + \partial_{xx} u(x,t) - |\partial_x u(x,t)|^2 + \eta\sqrt{u(x,t)}\mathrm{d}W(t), \text{ on } (0,1)\times\mathbb{R}^+, \quad (3.20a)$$

$$\partial_x u(0,t) = -u(0,t), \quad t \geq 0, \quad (3.20b)$$

$$\partial_x u(1,t) = u(1,t), \quad t \geq 0, \tag{3.20c}$$

$$u(0) = u_0, \tag{3.20d}$$

where $\mathrm{d}W(t)$ is Wiener process depend on time with zero initial value. We use the Implicit-Explicit scheme method to discretise in time as follows:

$$u^{n+1} - u^n = -\kappa\tau(u^{n+1} - \gamma) + \tau\partial_{xx}u^{n+1} - \tau(\partial_x u^n \cdot \partial_x u^n) + \tau\eta\sqrt{u^n}\Delta W^n, \tag{3.21}$$

where $\Delta W^n := W(t_{n+1}) - W(t_n)$ (see [Lord et al., 2014] for SDEs Euler–Maruyama). The result is a sequence of spatial (stationary) problems for u^{n+1}, assuming u^n is known from the previous time step:

$$u^0 = u_0, \tag{3.22}$$

$$u^{n+1} = u^n - \tau\kappa(u^{n+1} - \gamma) + \tau\partial_{xx}u^{n+1} - \tau(\partial_x u^n \cdot \partial_x u^n) + \eta\sqrt{u^n}\Delta W^n, \tag{3.23}$$

for $n = 0, 1, 2, \ldots$. Given u^0, we can solve for u^0, u^1, u^2, and so on. We use a finite element method to solve (3.22) and (3.23). This requires turning the equations into weak forms. We multiply by a test function $\varphi \in H^1(0,1)$ and integrate,

$$\int_0^1 u^{n+1}\varphi \mathrm{d}x = \int_0^1 \Big(u^n - \tau\kappa(u^{n+1} - \gamma) + \tau\partial_{xx}u^{n+1} - \tau(\partial_x u^n \cdot \partial_x u^n) + \eta\sqrt{u^n}\Delta W^n\Big)\varphi \mathrm{d}x. \tag{3.24}$$

The resulting weak form arising from formulation (3.23) can be conveniently written in the standard notation

$$a(u^{n+1}, \varphi) = L_n(\varphi), \tag{3.25}$$

where,

$$a(u^{n+1}, \varphi) = \int_0^1 \Big[u^{n+1} + \tau\kappa(u^{n+1} - \gamma)\Big]\varphi \mathrm{d}x + \tau\int_0^1 \partial_x u^{n+1} \cdot \partial_x \varphi \mathrm{d}x \tag{3.26}$$

$$L_n(\varphi) = \int_0^1 (u^n - \tau(\partial_x u^n \cdot \partial_x u^n) + \eta\sqrt{u^n}\Delta W^n)\varphi \mathrm{d}x + \tau\Big[\partial_x u^{n+1}(x)\varphi(x)\Big]_0^1. \tag{3.27}$$

In addition to the variational problem to be solved in each time step, we also need to approximate the initial condition (3.22). This equation can also be turned into a variational problem:

$$a_0(u, \varphi) = \int_0^1 u\varphi \mathrm{d}x, \tag{3.28}$$

$$L_0(\varphi) = \int_0^1 u_0\varphi \mathrm{d}x, \tag{3.29}$$

we need to solve the following sequence of variational problems, to compute the finite element solution to our equation: find $u_0 \in H^1(0,1)$ such that $a_0(u_0, \varphi) = L_0(\varphi)$ holds for all $\varphi \in$

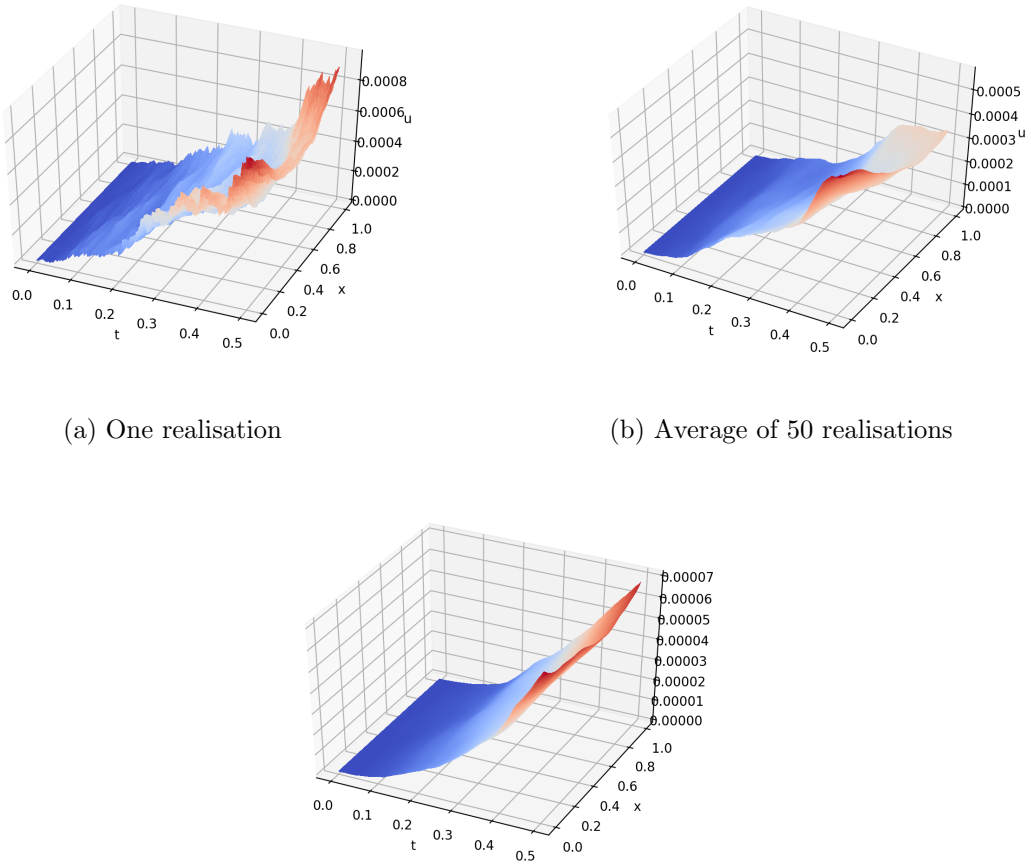
$H^1(0,1)$, and then find $u^{n+1} \in H^1(0,1)$ such that $a(u^{n+1}, \varphi) = L_n(\varphi)$ for all $\varphi \in H^1(0,1)$, $n = 0, 1, 2, \ldots$.

Finally, we discretise space so to have finitely many disjoint points and solve the approximation over the mesh that is created from the discretisation. Since we have one space dimension, the idea is to partition $[0,1]$ into subintervals or elements and choose the finite element spaces to be a set of piecewise linear functions that are defined on the elements. Let $I = [0,1]$, the sequence of nodes of h such that

$$0 = x_0 < x_1 < \cdots < x_J = 1. \tag{3.30}$$

Define the partitions of I as $\varsigma_j = [x_{j-1}, x_j]$, for $j = 1, \ldots, J$. We set $h := \max_{j=1,\ldots,J} |\varsigma_j|$. We perform experiments with decreasing value of the mesh size $h_0 > h_1 > h_2 > \ldots$. We apply Monte Carlo simulations to capture the expectation of the stochastic process u, we have to repeat the algorithm as many times as necessary.

We run a Monte Carlo simulation with number of samples $N = 0, 50$ and 1000 with the equation parameters given by $\kappa = 1.0$, $\gamma = 0.5$, $\eta = 0.05$ and mesh refinement level $J = 512$ on time interval $[0, 0.5]$.



(a) One realisation

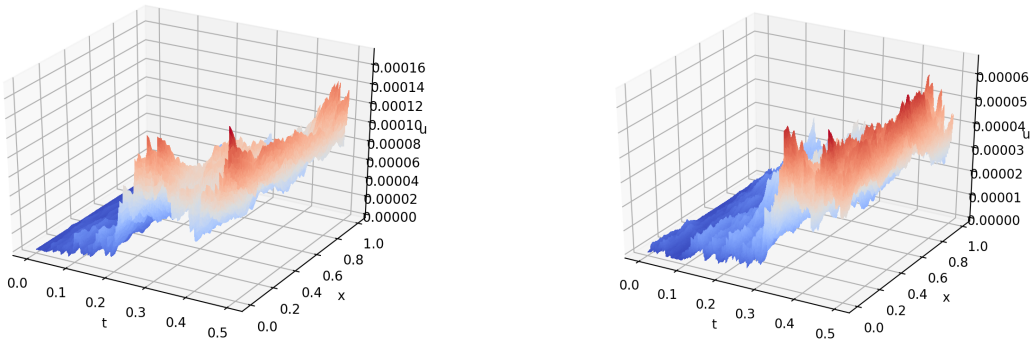
(b) Average of 50 realisations

(c) Average of 1000 realisations

Figure 3.1: Different realisations of an approximated solution to the nonlinear SPDE driven by time white noise

We can see from the graphs above how almost each path of the solution u represent a convex shape or in other word a 'smile'. The solution start from the initial value zero, going up very fast as the long term mean controls overall level of skew. As we increase the Monte Carlo simulation numbers the noise smoothing out.

Repeating the same experiment for a higher mesh refinement level we see that from Figure 3.2 the noise becomes more rough, however the convex shape for almost each path still present.



(a) mesh refinement level $J = 2048$ (b) mesh refinement level $J = 8192$

Figure 3.2: Different mesh refinement levels of an approximated solution to the nonlinear SPDE driven by time white noise

In order to see how big is influence of the equation parameters κ, γ and η, we run a new simulation of one sample with $\kappa = 10, \gamma = 1$ and $\eta = 4$ for the same time interval $t \in [0, 0.5]$ and mesh refinement level $J = 512$. We can see from Figure 3.3 that the volatility η controls convexity of the skew , i.e., high volatility generates more convexity.

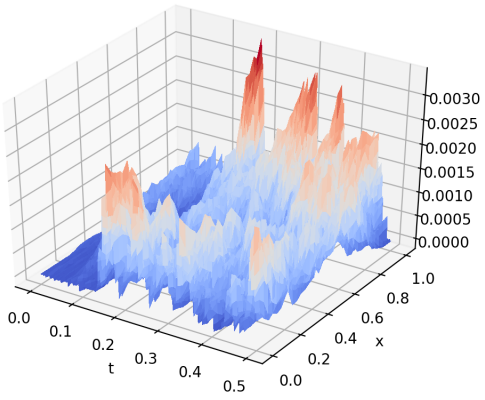


Figure 3.3: Equation parameters $\kappa = 10, \gamma = 1$ and $\eta = 4$.

3.3 Numerical approximation of SPDE driven by space–time white noise

Repeating the previous stochastic experiment for space–time white noise

$$\partial_t u(x,t) = -\kappa(u(x,t)-\gamma)+\partial_{xx}u(x,t)-|\partial_x u(x,t)|^2+\eta\sqrt{u(x,t)}\partial_{x,t}W(t), \text{ on } (0,1)\times\mathbb{R}^+, \quad (3.31a)$$

$$\partial_x u(0,t) = -u(0,t), \quad t \geq 0, \tag{3.31b}$$

$$\partial_x u(1,t) = u(1,t), \quad t \geq 0, \tag{3.31c}$$

$$u(0) = u_0, \tag{3.31d}$$

We use the Implicit-Explicit scheme method to discretise in time as follows:

$$u^{n+1} - u^n = -\kappa\tau(u^{n+1} - \gamma) + \tau\partial_{xx}u^{n+1} - \tau(\partial_x u^n \cdot \partial_x u^n) + \eta\sqrt{u^n}\Delta W^n, \tag{3.32}$$

where $\Delta W_n \sim N(0, (\tau/h)I)$ i.i.d. The result is a sequence of spatial (stationary) problems for u^{n+1}, assuming u^n is known from the previous time step:

$$u^0 = u_0, \tag{3.33}$$

$$u^{n+1} = u^n - \tau\kappa(u^{n+1} - \gamma) + \tau\partial_{xx}u^{n+1} - \tau(\partial_x u^n \cdot \partial_x u^n) + \eta\sqrt{u^n}\Delta W^n, \tag{3.34}$$

for $n = 0, 1, 2, \ldots$. Given u^0, we can solve for u^0, u^1, u^2, and so on. We use a finite element method to solve (3.33) and (3.34). This requires turning the equations into weak forms. We multiply by a test function $\varphi \in H^1(0,1)$ and integrate,

$$\int_0^1 u^{n+1}\varphi \mathrm{d}x = \int_0^1 \Big(u^n - \tau\kappa(u^{n+1} - \gamma) + \tau\partial_{xx}u^{n+1} - \tau(\partial_x u^n \cdot \partial_x u^n) + \eta\sqrt{u^n}\Delta W^n\Big)\varphi \mathrm{d}x. \tag{3.35}$$

The resulting weak form arising from formulation (3.34) can be conveniently written in the standard notation

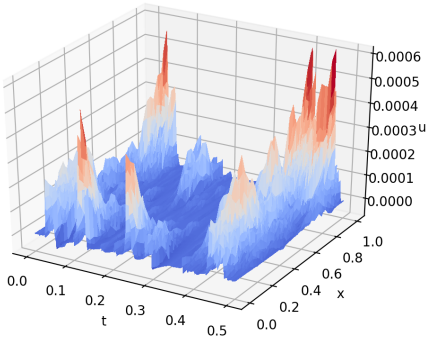
$$a(u^{n+1}, \varphi) = L_n(\varphi), \tag{3.36}$$

where,

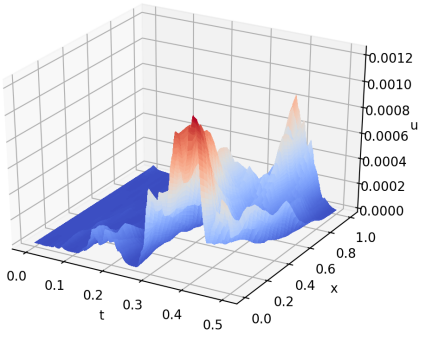
$$a(u^{n+1}, \varphi) = \int_0^1 \Big[u^{n+1} + \tau\kappa(u^{n+1} - \gamma)\Big]\varphi \mathrm{d}x + \tau\int_0^1 \partial_x u^{n+1} \cdot \partial_x \varphi \mathrm{d}x \tag{3.37}$$

$$L_n(\varphi) = \int_0^1 (u^n - \tau(\partial_x u^n \cdot \partial_x u^n) + \eta\sqrt{u^n}\Delta W^n)\varphi \mathrm{d}x + \tau\Big[\partial_x u^{n+1}(x)\varphi(x)\Big]_0^1. \tag{3.38}$$

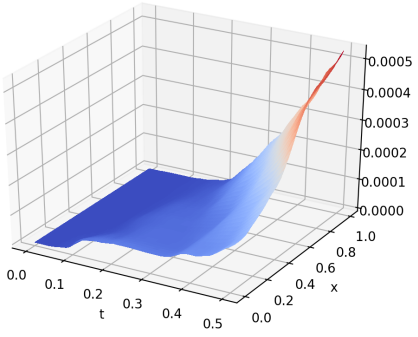
We run a Monte Carlo simulation with number of samples $N = 0, 10, 1000$ with the equation parameters given by $\kappa = 1.0$, $\gamma = 0.5$, $\eta = 0.05$ and mesh refinement level $J = 512$ on time interval $[0, 0.5]$.



(a) One realisation



(b) Average of 10 realisations



(c) Average of 1000 realisations

Figure 3.4: Different realisations of an approximated solution to the nonlinear SPDE driven by space–time white noise

3.4 Conclusions

In this chapter we introduced a numerical approximations based on finite element method in space and Euler–Maruyama in time. We obtained order of convergence expected from the theory for parabolic-type partial differential equations for the deterministic cases. One obstacle we faced in this chapter is the fact that our scheme may break down due to negative values being supplied to the square root function even if the Feller condition is met. This would be come a complex number and this would not make sense in modelling an asset price. A natural fix, is to take the absolute value however this reflects a large negative variance to a large positive variance which it its turn transforms realisations of low volatility into high volatility. Therefore using a full truncation scheme, where the negative values are floored at zero is more safer. In the further research we would like to understand how we can show convergence to the exact solution in the

stochastic cases.

Chapter 4

Applications to financial options

4.1 Option contracts

A financial options contract is an agreement between a buyer and seller, written on underlying asset. The contract gives the purchaser of the option the right to buy or sell the asset on a specified price (exercise or strike price) on or before specified date (expiration date). Options contracts are often used in securities, commodities, and real estate transactions. A call option gives the holder the right to buy an asset at a specified strike price on or before the option's expiration date. The seller of the call, called the writer, is obligated to deliver the assets at the strike price if the option is exercised. A put option gives the holder the right to sell an asset at a specified strike price on or before the option's expiration date. The writer of a put option is obligated to receive those assets and to deliver the required cash. An option is said to be in, at or out of-the-money according to the following table.

Options Classifications	Call Options	Put Options
At-the-money	Asset price equal to Strike price	Asset price equal to Strike price
In-the-money	Asset price greater than Strike price	Asset price less than Strike price
Out-of-the-money	Asset price less than Strike price	Asset price greater than Strike price

Table 4.1: Option Classifications

An option is European if it can be exercised only on the expiration date. It is American if it can be exercised at any time on or before the expiration date.

4.1.1 Option price models

"If options are correctly priced in the market, it should not be possible to make sure profits by creating portfolios of long and short positions in options and their underlying stocks. Using this principle, a theoretical valuation formula for options is derived" as stated by [Black and Scholes, 1973]. Based on this principle, academic researchers used several models to calculate the option prices theoretically such as the binomial options pricing model (BOPM) which was developed by Cox, Ross and Rubinstein in 1979 and the continuous PDE Black–Scholes pricing model. Where these models are dependent on certain values specified and assumed to be known at the present. Most of these values represent the option contract terms like underlying price, strike price and time to maturity along with quantities derived from the price movement themselves such as volatility. Among these, the Black–Scholes pricing model is considered as the most widely used model for option pricing because of its simplicity and explicitly calculable formulas. It is used to calculate the theoretical value of European options. This raises the next question, in the real financial world: do people rely on Black–Scholes formula for pricing options and; if so, how do they adjust parameters involved in the calculation [Wilmott et al., 2010], [Capiński and Zastawniak, 2011].

4.1.2 Black–Scholes pricing model

To derive Black–Scholes PDE for option values on a non-dividend paying assets we assume that:

1. The asset price has a log–normal dynamics (its log is Gaussian). Denote the asset price process by $S(t)$. The proportional change of the asset over small amount of time $\mathrm{d}t$ has a drift μ and stochastic component satisfying the following stochastic differential equation

$$\mathrm{d}S(t) = \mu S(t)\mathrm{d}t + \sigma S(t)\mathrm{d}W(t), \tag{4.1}$$

   where,

   - $W(t)$ is a Wiener process.
   - μ is the constant drift of the asset returns.
   - σ is the constant volatility of the asset price.

2. The interest rate r is known as a constant over time.

3. The option is European style, that is it can only be exercised at the expiration date.

4. Trading the asset can be in continuous portions and short selling is permitted.

Under these conditions with the absence of taxes and the transaction fees, the value of the European option depends only on the stock price $S(t)$ and time t. Suppose $V(S,t)$ the option value at time t. By Itô 's Lemma we obtain:

$$\mathrm{d}V(S,t) = \Big(\mu S\frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2}\Big)\mathrm{d}t + \sigma S\frac{\partial V}{\partial S}\mathrm{d}W_t, \quad \text{where,} \quad 0 < S < \infty, 0 < t < T. \tag{4.2}$$

considering a self-financing trading strategy where at each time t we hold x_t units of the cash account and y_t units of the stock. Then P_t, the time t value of this strategy satisfies

$$P_t = x_t B_t + y_t S_t \tag{4.3}$$

We choose x_t and y_t in such a way that the strategy replicates the value of the option. The self-financing assumption implies that

$$\begin{aligned}\mathrm{d}P_t &= x_t\mathrm{d}B_t + y_t\mathrm{d}S_t \\ &= rx_tB_t\mathrm{d}t + y_t(\mu S_t\mathrm{d}t + \sigma S_t\mathrm{d}B_t) \\ &= (rx_tB_t + y_t\mu S_t)\mathrm{d}t + y_t\sigma S_t\mathrm{d}W_t \end{aligned} \tag{4.4}$$

any gains or losses on the portfolio are due entirely to gains or losses in the underlying securities, i.e. the cash-account and stock, and not due to changes in the holdings x_t and y_t. Returning to our derivation, we can equate terms in (4.2) with the corresponding terms in (4.4) to obtain

$$y_t = \frac{\partial V}{\partial S} \tag{4.5}$$

$$rx_tB_t = \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{S^2} \tag{4.6}$$

If we set $V_0 = P_0$, the initial value of our self-financing strategy, then it must be the case that $V_t = P_t$ for all t since V and P have the same dynamics. This is true by construction after we equated terms in (4.2) with the corresponding terms in (4.4). Substituting (4.5) and (4.6) into (4.3) we obtain

$$\frac{\partial V(S,t)}{\partial t} + rS\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 V}{\partial S^2} - rV = 0. \tag{4.7}$$

The partial differential equation (4.7) is known as the Black–Scholes PDE. Notice that Black–Scholes PDE does not contain the growth rate μ of the underlying share. Hence the price of the options will be independent of how rapidly or slowly an asset grows. Solving the Black–Scholes PDE (4.7) for V with final and boundary conditions we obtain the Black–Scholes formula for call or put European options:

$$C(S(t),t) = SN(d_1) - K\,\mathrm{e}^{-r(T-t)}\,N(d_2), \tag{4.8}$$

$$P(S(t),t) = K\,\mathrm{e}^{-r(T-t)}\,N(d_2) - SN(d_1), \tag{4.9}$$

where

$$d_1 = \frac{\log(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}},$$
$$d_2 = \frac{\log(S/K) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}.$$

In this formula $N(d)$ is the probability that a standard normal random variable is less than d. $N(d_1)$ and $N(d_2)$, both positive but less than one, represent the number of shares and the amount of debt in a portfolio that exactly replicate the price of the option. the Black?Scholes call option price is denoted by C and the put option price by P, K stands for the option's strike price and T stands for the option's time to maturity.

4.1.3 Data analysis

In order to study options volatility from real world data, we started by creating a MySQL database `Option_Historical_Data`. It contains all the options information that we need to compute the Black–Scholes implied volatility (the volatility input to the Black–Scholes formula that generates the market price) and examine the volatility surface that is the collection of all such implied volatilities (UnderlyingSymbol, Underlying Price, Type , Expiration, DataDate, Strike, Last, Bid, Ask, Volume, OpenInterest, IV, Delta, Gamma, Theta, Vega, ..etc) from 2002 until 2016. These financial informations were collected from trusted sources as CSV files and have been converted to tables into MySQL database using Python scripts. In this project S&P 500 (The Standard & Poor's 500) has been selected as the object of study for several reasons. Firstly, S&P 500 is the most important indicators of US stock market performance measurement. It is based on the market capitalisation and share prices of 500 large companies. Moreover, the S&P 500 (SPY) exchange-traded fund's contracts are the most widely traded, seeing more than 1.2 million shares change hands each session. In addition, S&P 500 index SPX can be traded likewise SPY as the underlying asset for financial options. SPX options trade in a very active market while options on some other indices are much less traded. The following table and figure show the trading volume (number of shares that traded during a given day) comparison between SPY and SPX.

Period	SPX trading volume	SPY trading volume	volume ratio SPY/SPX
Pre-Crisis (2005 - Bear Stearns Feb 2008)	330,932,748	282,248,736	0.85
During Crisis (Bear Stearns march 2008 - 2009)	313,390,032	618,345,723	1.97
After Crisis 2010-2015	1,148,879,589	3,431,624,193	2.98
Overall	1,793,202,369	4,332,218,652	2.41

Table 4.2: Volume comparison between SPY and SPX

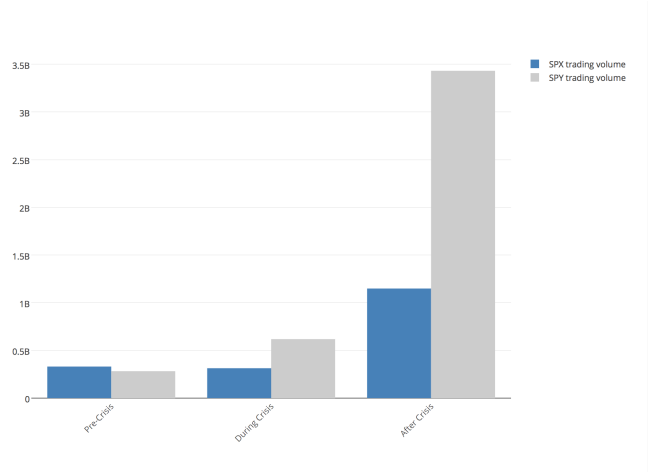


Figure 4.1: Trading volume of SPX, SPY

As shown in Table 4.2 numbers are impressive where the SPY contracts are traded far more than the SPX's. This is evident in the after Bear Stearns period 2010-2015. Thus we are going to choose SPY and its derived options for our investigation.

4.1.4 Black–Scholes implied volatility

The Black–Scholes formulas given by (4.8) are used to determine the option call and put prices for known asset price, strike price, option time to maturity, risk-free rate and some value of σ, where σ is the volatility (the degree of variation of a trading price series over time as measured by the standard deviation of logarithmic returns). The value of σ that makes the price, determined by the Black–Scholes formula, equal to the current market price is called the implied volatility. In other words, σ is the volatility that, when substituted into the Black-Scholes formula with

the other parameters held fixed, gives the market price C as a function of σ.

$$V(S, K, T) = V_{BS}(S, K, \sigma_{BS}(S, K, T), T), \tag{4.10}$$

where, $V(S, K, T)$ represents the current market price and V_{BS} is the corresponding Black–Scholes pricing model for call or put options. In order to determine the implied volatility, we can use interpolation to invert the Black–Scholes formula to obtain σ as a function of the market price. Quite simply, we set the Black–Scholes pricing formulas equal to the market observed price and we use a root finding algorithm to find the volatility parameter which sets the difference (between model and market price) to zero. In our data, C, P, S, K, T represents a list of the Black–Scholes parameters as we have amount of options traded every day. We computed the implied volatility by applying Brent's algorithm[1] in order to find the roots of the Black–Scholes model. Next, we give an example on the fact that different strike prices can generate different implied volatility when the asset price and time to maturity are the same. In this example the market price is given by the Mid price which is the average between the demand and offer prices

Strike price	Mid price	Implied volatility
150.0	51.29	1.4461199
160.0	41.3	1.1764081
182.0	19.335	0.6079909
184.0	17.33	0.5524254
186.0	15.34	0.5024264
193.0	8.49	0.3454955
194.0	7.5500	0.3272877
195.0	6.635	0.3113043

Table 4.3: Computing the Black–Scholes implied volatility on the 4th of January 2016

As we can see from the table above although the asset price is the same for all options $S = 201.01992$ and same time to maturity 4 days to expiration with constant rate $r = 0.0034$ we obtain a different implied volatilities

[1] root-finding algorithm combining the bisection method, the secant method and inverse quadratic interpolation

4.1.5 Uniqueness of the implied volatility

One may ask, is there always a unique solution $\sigma(K,T)$ satisfies equation (4.10)? To answer this question we would like to look at the behaviour of $F(\sigma)$, where

$$F(\sigma) = V_{BS}(S,K,\sigma_{BS}(S,K,T),T) - V(S,K,T). \tag{4.11}$$

As volatility represents the price movements, it can reach values from 0 to ∞ only. In particular, if an option has a constant price, then its volatility would be zero, where the lowest possible volatility option is zero volatility. Hence, the volatility domain can be given from 0 to ∞. Taking for example, a call option with underlying asset price $S = 201.0192$, strike price $K = 150$, interest rate $r = 0.0034$, time to maturity $T = 0.01095$, $C = 51.29$ and plot its $F(\sigma)$ when the volatility is changing from 0 to 100.

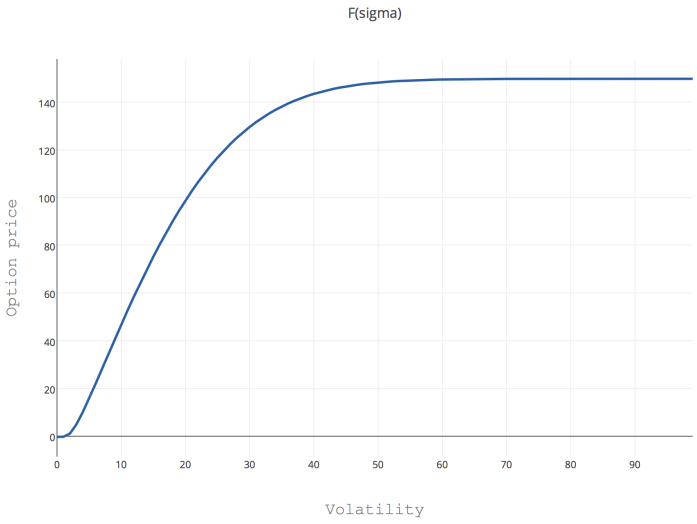


Figure 4.2: Relationship between volatility and option price

Moreover, we know that the option price is going to increase because its derivative with respect to σ given by $\frac{\partial V}{\partial \sigma} = SN(d1)\sqrt{T-t}$ is strictly positive so, V is an increasing function of σ. Since the Black–Scholes formula is continuous and increasing in σ, there will a be a unique solution of σ.

4.1.6 Negative implied volatility

Using real world data calculations of the Black–Scholes implied volatility for call and put options we observed, some undefined values of implied volatility (NaN)[2] or what we call during this

[2]We get NaN when the results of implied volatilities become less than Brent's method chosen tolerance, , i.e., positive values smaller than $1.0e^{-6}$

project the negative volatilities. This happens for options with a wide spread between bid and ask prices, in other words, this happens for underpriced options. For example, on 24 June 2016 which is the day that the vote of the referendum on Britain exiting from the European Union was announced, we get a negative Black–Scholes implied volatility for 251 out of 1385 call options.

Underlying price	Strike price	Mid price	Days to expiration	Interest Rate	Type	Implied volatility
203.2425	130.0	72.95	6	0.0038	call	NAN
203.2425	150.0	52.95	7	0.0038	call	NAN
203.2425	130.0	73.2	98	0.0038	call	NAN
203.2425	120.0	83.15	119	0.0038	call	NAN
203.2425	105.0	98.0500	175	0.0038	call	NAN
⋮	⋮	⋮	⋮	⋮	⋮	⋮
203.2425	130.0	73.9	539	0.0038	call	NAN

Table 4.4: Undefined implied volatilities phenomenon

Since the European call option can not exceed the stock price and no options can have a negative price, we have an upper and lower bound of European call option prices given by:

$$\max(S(t) - Ke^{-r(T-t)}, 0) \leq C(S,t) \leq S(t), \tag{4.12}$$

Let us now discuss these bounds on both call and put option classifications (in/out of the money). For in-the-money call options, when the asset price is greater than strike price, the lower bound of the call option price is greater than zero. For example, taking in-the-money call option with underlying asset price $S = 100$, strike price $K = 90$, interest rate $r = 0.2$, time to maturity $T = 0.2$, and computeing its Black–Scholes call price when the volatility is changing from 0 to 10, we obtain the following graph:

Figure 4.3: Relationship between volatility and in-the-money call options

As we can observe, the Black–Scholes option price is bounded from below by the zero volatility price, which is a positive number and bounded from above by the asset price. However, if we use real world data, it is often observed that the prices will fall below the theoretical 'zero volatility value'. In this case, we cannot define the volatility, where the volatility is a monotonically increasing function of option price. Therefore, we have to use a formula to estimate the volatility for contracts, whose option price is less than the theoretical 'zero volatility value'.

Let us denote V as the Black–Scholes option price, V_0 as the 'zero volatility' Black–Scholes option price, V_C as the real world recorded option price, σ_V as the implied volatility of the option when the price is V and σ_N as the negative implied volatility. Then we can define a formula to obtain the negative implied volatilities for the undervalued options as follows:

$$\begin{aligned} V &= V_0 + |V_C - V_0|, \\ \sigma_N &= -\sigma_V. \end{aligned} \tag{4.13}$$

However, for out-of-the-money call options where the strike price is greater than the asset price, this implies that $\max(S_t - Ke^{-r(T)}, 0) = 0$. That means the range of option price is from zero to the asset price, so there is no possible negative implied volatility for out-of-the-money call options. We can clarify this by the out-of-the-money option with underlying asset price $S = 100$, strike price $K = 110$, interest rate $r = 0.2$, time to maturity $T = 0.2$. We compute its Black–Scholes call price when the volatility is changing from 0 to 10:

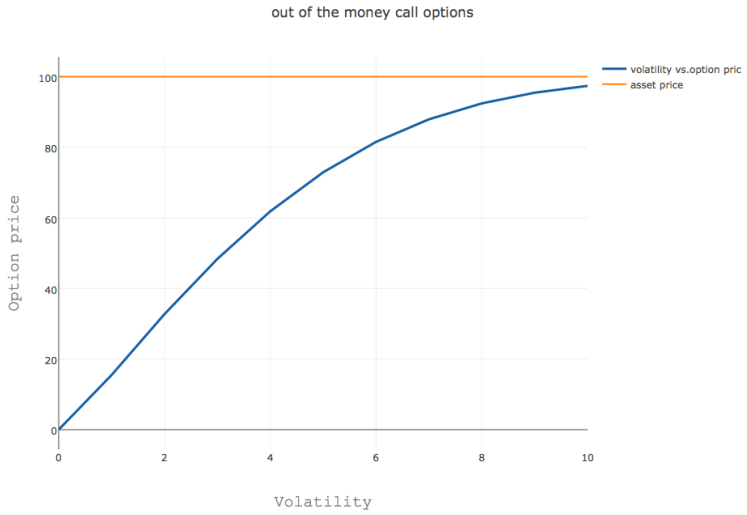


Figure 4.4: Relationship between volatility and out-of-the-money call options

Hence, we can use the formula that is represented by (4.13) to redefine the implied volatilities of Table 4.4 to become:

Asset price	Strike	Time	Rate	V_0	V	σ_v	σ_N
203.2425	130.0	6	0.0038	73.2506	73.5512	1.7372781	-1.7372781
203.2425	150.0	7	0.0038	53.2534	53.5568	1.15191658	- 1.15191658
203.2425	165.0	14	0.0038	38.2665	38.4830	0.55825998	- 0.55825998
203.2425	105.0	21	0.0038	98.2654	98.5309	1.29197354	-1.29197354
203.2425	150.0	28	0.0038	53.2862	53.4724	0.53317069	-0.53317069
203.2425	150.0	35	0.0038	53.2971	53.4942	0.48113754	-0.48113754
203.2425	105.0	56	0.0038	98.303	98.6073	0.80731072	-0.80731072
203.2425	105.0	84	0.0038	98.3342	98.6185	0.65304074	-0.65304074
203.2425	130.0	98	0.0038	73.3750	73.5501	0.39751150	-0.39751150
203.2425	120.0	119	0.0038	83.3910	83.6321	0.437532761	-0.437532761
203.2425	105.0	175	0.0038	98.4336	98.8172	0.473997975	-0.473997975
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
203.2425	130.0	539	0.0038	73.9699	74.0399	0.15203624	-0.15203624

Table 4.5: Negative implied volatilities phenomenon

The amount of underpriced options which generate a negative implied volatilities is exhibit in the following figures

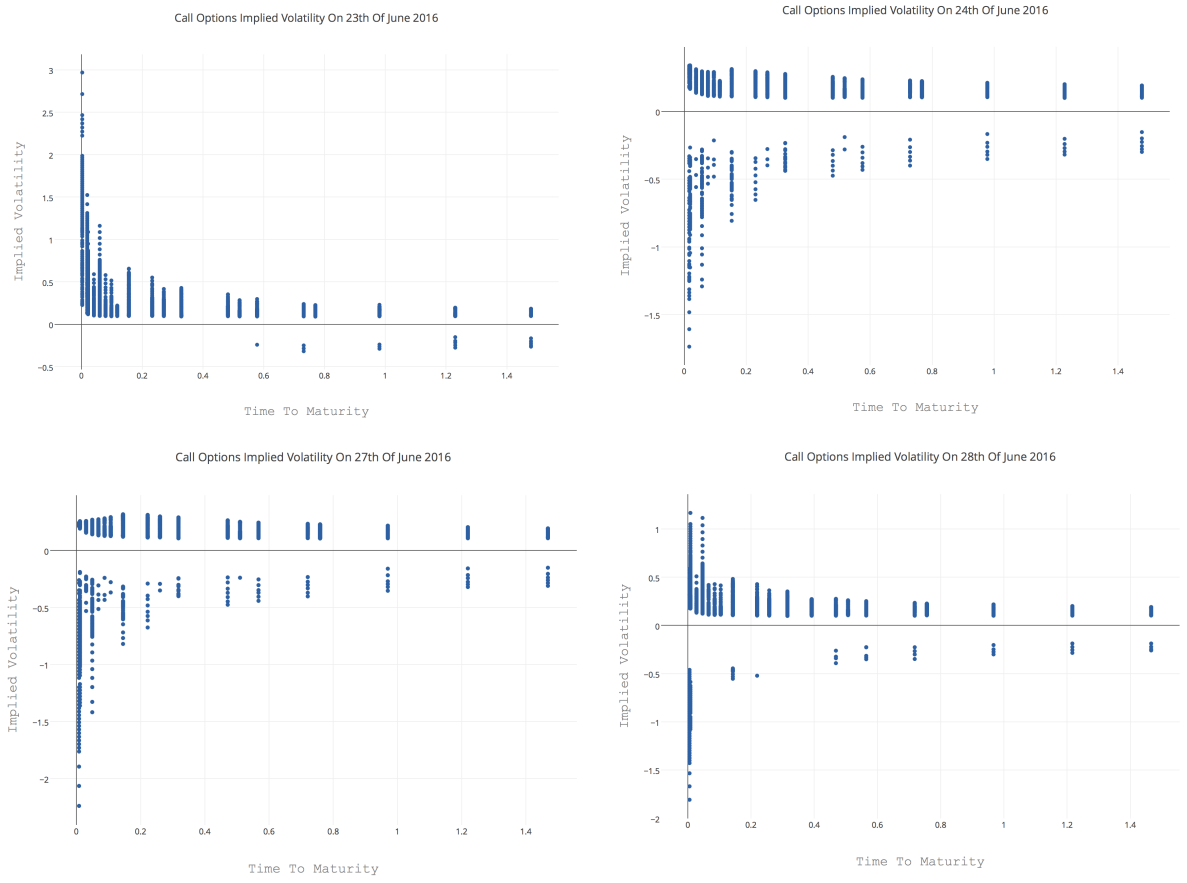


Figure 4.5: Black–Scholes implied volatility of call options on 24th-28th June 2016

The following figures exhibit that the majority of underpriced options are in In-the-money side (i.e. asset price greater than Strike price).

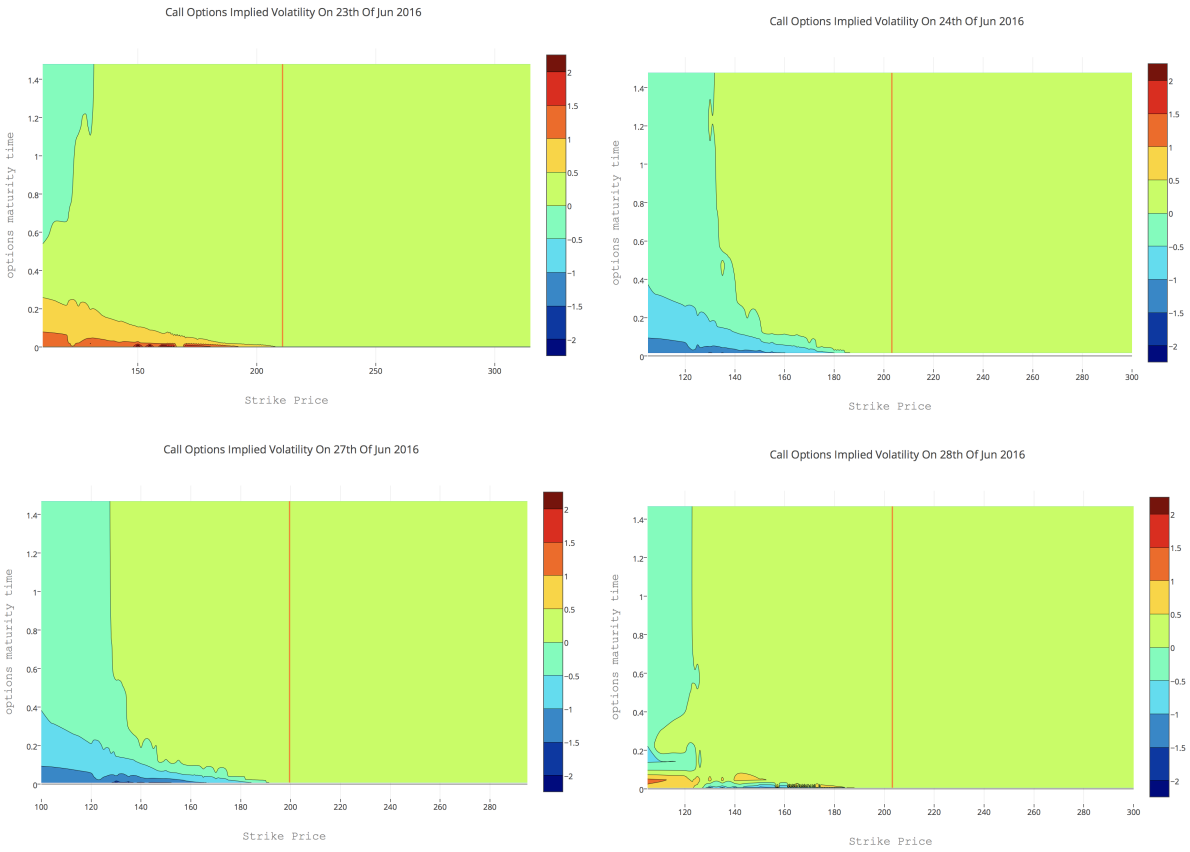


Figure 4.6: Contours of Black–Scholes implied volatility of call options on 24th-28th June 2016

On the other hand, the European put options can not exceed the present value of the strike price K and cannot be higher than discounted value of the strike price. We can give, the upper and lower bounds of European put option prices as,

$$\max(Ke^{-r(T-t)} - S_t, 0) \leq P(S,t) \leq Ke^{-r(T-t)}. \tag{4.14}$$

For out-of-the-money put options, when the asset price is greater than strike price, this implies that $\max(Ke^{-r(T-t)} - S_t, 0) = 0$. That means the range of put option prices is from zero to the asset price, so there is no possible negative implied volatility for out-of-the-money put options. We can give an example by taking the same values of Black–Scholes variables given previously, and computing its Black–Scholes put price when the volatility is changing from 0 to 10.

Figure 4.7: Relationship between volatility and out-of-the-money put options

In contrast, for in-the-money put options where the asset prices is less than the strike price we have the following figure:

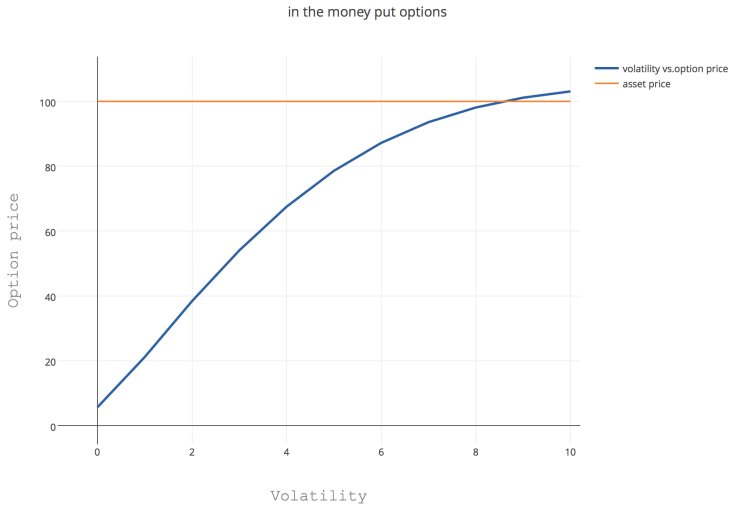


Figure 4.8: R elationship between volatility and in-the-money put options

As we can observe, the Black–Scholes option price is bounded from below by the zero volatility price which is a positive number and bounded from above by the asset price. However, if we use real world data, it is often observed that the prices will fall below the theoretical 'zero volatility value'. In this case, we need to use the same argument of undervalued call options to define the implied volatility of undervalued put options. The amount of underpriced options which generate a negative implied volatilities is exhibit in the following figures

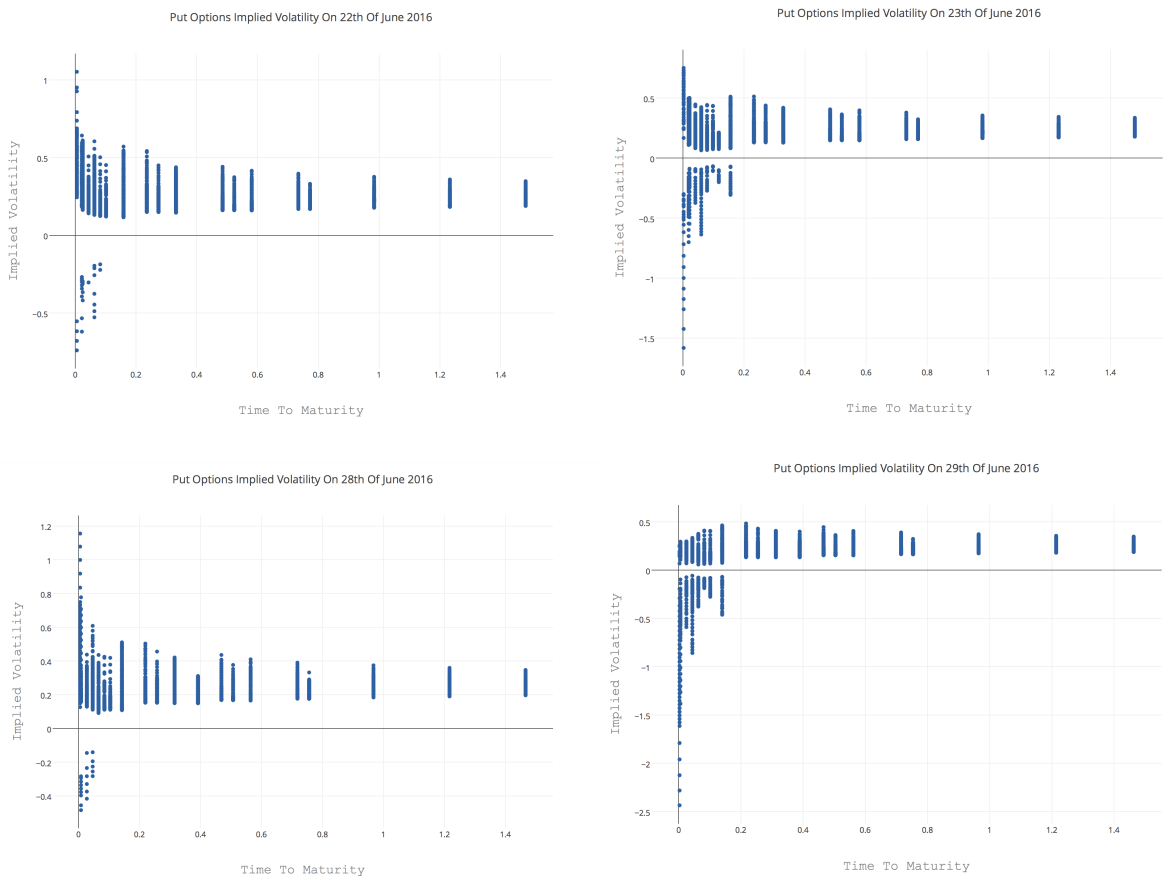


Figure 4.9: Black–Scholes implied volatility of put options on 22th-29th June 2016

The following figures exhibit that the majority of underpriced options are in out-of-the-money side (i.e. asset price less than Strike price).

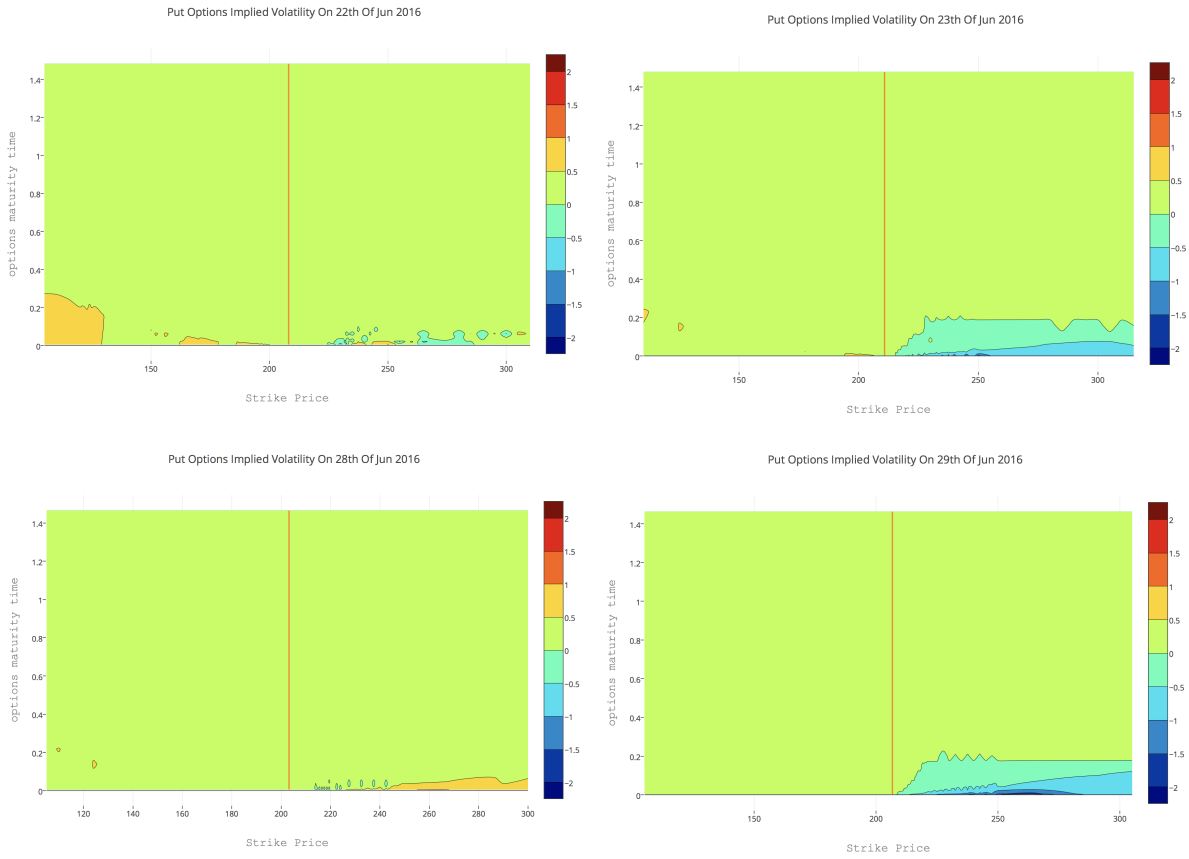


Figure 4.10: Contours of Black–Scholes implied volatility of put options on 22th-29th June 2016

To sum up, from the Black–Scholes implied volatility calculation using a real world data, we notice that the Black–Scholes formula tends to undervalue in-the-money call and put options which gives us a negative implied volatility. The negative implied volatility phenomena has also been observed and commented by [Gatheral and Taleb, 2011] as well as a number of other authors.

4.1.7 The volatility surface

The fundamental concept of the Black–Scholes pricing model is the assumption that the volatility is known and the option's implied volatility should differ from the true volatility, only because the random events. In other words, for given time to maturity T, the volatility is constant with respect to the strike price, which leads to a flat volatility surface with $\sigma(K,T) = \sigma$ for all K. In practice, however, not only is the volatility surface not flat but it actually varies, often significantly, across various variables such as time [Haugh, 2009]. Meanwhile examining the volatility surface for call options, we observed that:

- options with lower strikes and shorter time to maturity tend to have higher implied volat-

ilities.

- options with lowest and highest strike prices have the highest implied volatility with time to maturity held fixed. This phenomenon is known as the volatility smile.
- For a given strike K, the implied volatility can be either increasing or decreasing with time-to-maturity. In general, however, σ tends to converge to a constant as T increses.

4.1.8 Volatility smile in real world data

A volatility smile is a common phenomena that results from plotting Black–Scholes implied volatility for a series of options with the same expiration date against their strike prices. The shape of the curve is symmetrical around the asset price. However, in real world data, the curve is really more a skew or a smirk. The next figures illustrate the implied volatility surface for the SPY as it was around December 2016. Smiles for call options with various expiry dates are shown below.

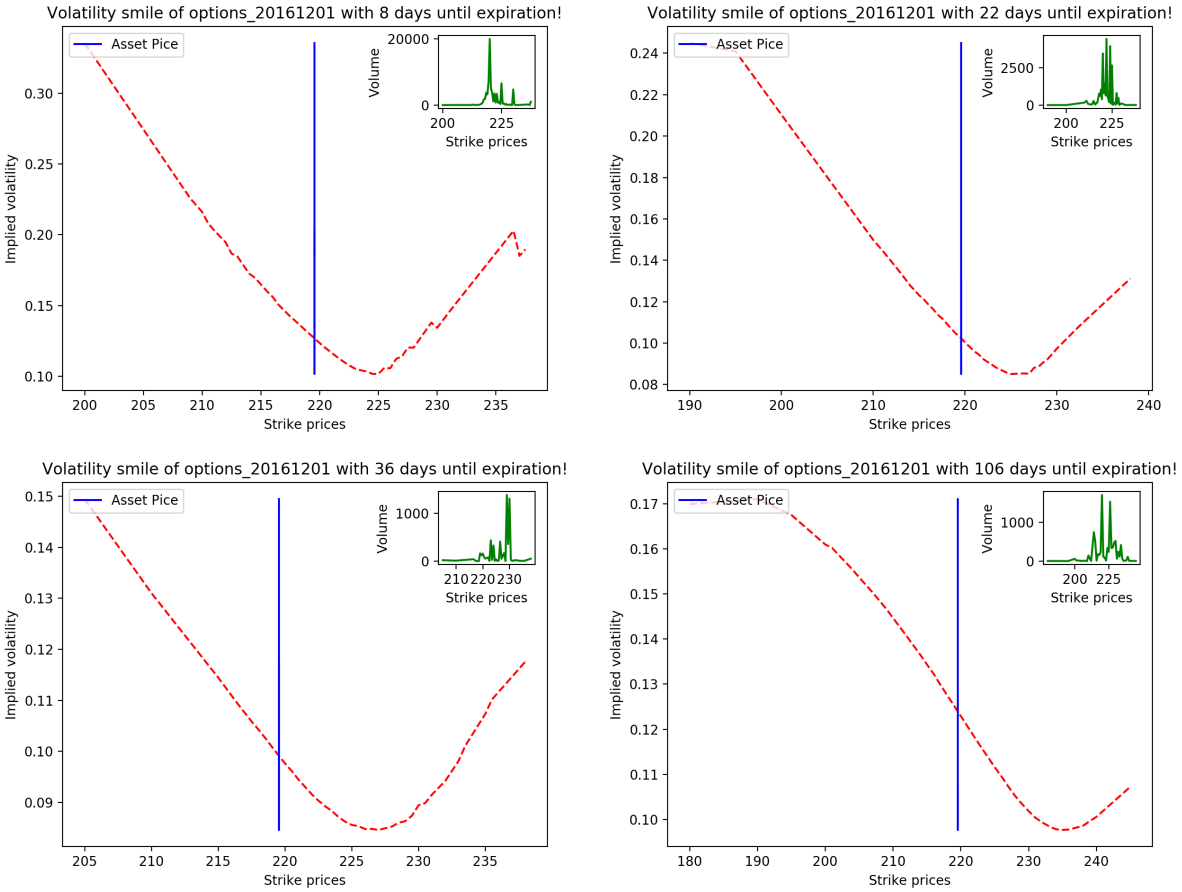


Figure 4.11: Volatility smile

Figure 4.11 shows that, in addition to expiration time T dependence, the implied volatility

also depends on its strike prices x. Hence, the smile shape is generated from the fact that implied volatility can be written as a function of two variables, time and strike $\sigma_{\text{imp}}(x, T)$.

4.1.9 Heston's model

Heston's model assumes that the interest rate is constant, there is no dividend payment and the stock price $S(t)$ and its variance $u(t)$ at (the real) time $t \geq 0$ satisfys the following SDEs:

$$\mathrm{d}S(t) = \mu(t)S(t)\mathrm{d}t + \sqrt{u(t)}S(t)\mathrm{d}W_1(t), \tag{4.15}$$

$$\mathrm{d}u(t) = -\kappa(u(t) - \gamma)\mathrm{d}t + \eta\sqrt{u(t)}\mathrm{d}W_2(t) \tag{4.16}$$

with

$$\langle \mathrm{d}W_1 \mathrm{d}W_2 \rangle = \rho \mathrm{d}t$$

where κ, γ, $\eta > 0$ and $|\rho| < 1$ are the model parameters, with $2\kappa\gamma > \eta^2$, which ensures that zero is an unattainable boundary for the process u.

- $\mu(t)$: is the (deterministic) instantaneous drift of stock price returns.
- κ: is the speed of reversion of $u(t)$ to its long-term mean γ, as t tends to infinity, the expected value of $u(t)$ tends to γ.
- η: is the volatility of variance.
- ρ: is the correlation's coefficient between random stock price returns and changes in $u(t)$.
- $\mathrm{d}W_1$ and $\mathrm{d}W_2$: are Wiener processes.

The stochastic differential equation (4.15) models the behaviour of the asset price $S(t)$ as a stochastic process driven by the Wiener process W_1, as it is assumed in the Black–Scholes model. However, the volatility σ, is the square root of the variance $u(t)$ which in turn, described as a stochastic process driven by a second Wiener process W_2. The first part of the right hand side in equation (4.16) models a mean-reverting, with reversion speed κ and mean γ. Note that the initial variance $u(t) > 0$ is assumed to be non-random at time zero. If the stochastic variance u moves far from the reversion level γ during strophic events in the financial market, it will be pulled back towards a central location γ by the mean-reverting term $-\kappa(u(t) - \gamma)\mathrm{d}t$. The parameter η is known as the volatility of the volatility as it directly affects the stock price volatility. Increasing η increases the convexity of the smile generated by the model. There is a correlation between the asset and variance processes, represented by the coefficient ρ, that

can be used to generate a skewed volatility smile. The process u is a Cox–Ingersoll–Ross (CIR) diffusion (also known as the square root process), and the stochastic differential equation for the CIR diffusion satisfies the Yamada-Watanabe condition, so it admits a unique strong solution.

To sum up we introduced this chapter to explain the financial concepts and to show the shortcomes of Black–Scholes model and the stochastic volatility models such as Heston's model. This is what motivate us to suggests a new model which give a better explanation of the volatility smile. although our model is driven by a space time white noise while Heston's model presumed that the asset price (4.15) and the variance (4.16) are driven by time only dependent noises, with correlation coefficient. Therefore, in order to restrict the forcing term in (1.10a), to consider stochasticity in time only, we project the noise's space factor. Of course time noise is a special case of space–time white noise. As the theory of space–time white noise in stochastic PDE's is more general than noises with respect to time, we can project the space variable

4.1.10 Example. Space–time white noise process W defined by

$$W(t) = \sum_{i=0}^{\infty} \chi_i(x)\beta_i(t), \tag{4.17}$$

where χ_i is any an orthonormal basis of L^2 and $\beta_i, i = 0, 1, \dots$ are iid $\mathcal{F}_t$-Brownian motions. For our purpose, we will rewrite the space-time white noise with respect to a sequence of real coefficients α_i as follows;

$$W^{\alpha}(t) = \sum_{i=0}^{\infty} \alpha_i\chi_i(x)\beta_i(t), \tag{4.18}$$

assuming $\alpha = (1, 0, 0, \dots)$ and $\chi_i(x) = c_i \cos(2i\pi x)$ as the orthonormal basis of $L^2(0, 1)$. This implies

$$W^{\alpha}(t) = \beta_0(t), \quad \text{for} \quad \chi_0(x) = 1, \tag{4.19}$$

which shows that, time noise $\mathrm{d}W(t)$ is the first component of space–time white noise $\mathrm{d}W(x, t)$ with orthonormal basis being one, which is what we need.

Then, we are able to define the projection operator $\mathcal{P}$ acting on the orthonormal basis χ_i such that

$$\chi_i \to P\chi_i = \begin{cases} \chi_0 & \text{if } i = 0, \\ 0 & \text{otherwise .} \end{cases} \tag{4.20}$$

Conclusion

We propose an extension to the Heston and CIR models that model the implied volatility with an extra independent variable representing the "current strike price" as a more accurate model. The model consists of an SPDE which is semilinear in both the deterministic part and the noise coefficient. Furthermore, the deterministic nonlinearity involves the unknown's gradient squared, reminiscent of the KPZ equation but with an opposite sign. Thanks to this feature we can have a more "concrete" approach avoiding "infinite constants" introduced by [Hairer, 2014] to study similar equations via regularity structures.

We have shown that there exists a subsequence of approximate solution of the nonlinear SPDE in finite dimensional subspace by applying Fourier–Galerkin approximations. The nonlinear first order negative sign makes it possible to show that the approximate solution satisfies an energy estimate which is uniform and thus provides us convergence to a limit. We define the limit u to be a generalised solution of (1.10a). The nonlinear first order negative sign appears as an obstacle but turns out to be the crucial tool to solve the problem. We have good reasons to believe and gained good intuition on how to show that the limit is a mild solution.

A challenge encountered in proved is establishing the positivity of solution analytically and the related issue of how to preserve such positivity in the computations. Our proving to show the solution positive is based on a comparison-principle, where we build a subsolution which is positive by Feller's condition that prevents the solution to be negative. Simply if two solutions one above the other then they can not cross hence the solution stays positive. A numerical method such as the Euler–Maruyama applied to (1.10a) may break down due to negative values being supplied to the square root function even if the Feller condition is met. A natural fix, is to take the absolute value or by using a full truncation scheme, where the negative values are floored at zero.

In Chapter 4 we analysed the financial data that we selected form the real world life in order to study options volatility. By using $S\&P500$ (SPY) data as an example to chapter the volatility surface implied by Black–Scholes model. We observed, some undefined values of implied volatility or what we called during this project the negative volatilities. This happens

for options with a wide spread between bid and ask prices, in other words, this happens for underpriced options. The mean feature that we capture is options with lowest and highest strike prices have the highest implied volatility with time to maturity held fixed, which is known as volatility smile. We also noted that none of the Black–Scholes model assumptions satisfied perfectly in real world observed data.

Studying (1.10a)-(1.10d) raised other questions I wish to consider for further research. For instance form the analytic point of view, in finance, according to [Capiński and Zastawniak, 2011] and [Neftci and Hirsa, 2000] Girsanov theorem is used to derive an asset's or rate's dynamics under a new probability measure. In the Black–Scholes model the probability measure is moves from the historic measure P to the risk neutral measure Q. This makes us curious to investigate what happen the SPDE (1.10a) if we apply Girsanov's theorem to meet the financial requirements.

Moreover from financial point of view, the financial applications of the model (1.10a)–(1.10d) have the potential to be explored. One very important question is how to apply (1.10a) in the financial field. If the volatility really depends on the strike price as the graphs show, does it mean the asset price depend -in one way or another- on the strike price as well. It is known from the logic in options trading the strike price is the set price at which a derivative contract can be bought or sold when it is exercised, which mean it determines after writing the option. Hence, it is not possible for the asset price to depend on the strike price. However, it might depend on the average of all the strike prices. Another interesting approach for abetter fit of the volatility smile is to consider the fractional diffusion. As it appears from the volatility surface, the smile not always smiling (convex) but it can be more pointy in some cases. Tracking the 'turning point' which is the lowest implied volatility value that a set of implied volatilities with the same time to maturity and asset price can take is also motivating. Real world data are showing that the volume of options trading around the turning point the very higher than any other spots.

Appendix A

Useful lemmas and inequalities

Here we collect some lemmas and inequalities that we have used during the project

A.0.1 Definition (Lipschitz continuity)**.** A real-valued function $f : \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous if there exists a positive real constant $c_{(A.1)}$ such that, for all real x and y we have,

$$|f(x) - f(y)| \leq c_{(A.1)}|x - y|, \quad \forall x, y, \in \mathbb{R}. \tag{A.1}$$

A.0.2 Definition (Linear growth)**.** A real-valued function $f : \mathbb{R} \to \mathbb{R}$ has a linear growth condition if there exists a positive real constant $c_{(A.2)}$ such that

$$|f(x)| \leq c_{(A.2)}(1 + |x|), \quad \forall x \in \mathbb{R}. \tag{A.2}$$

A.0.3 Lemma. *For $a, b \in \mathbb{R}$ we can bound the difference between $\sqrt{a}$ and $\sqrt{b}$ by the following inequality;*

$$|\sqrt{a} - \sqrt{b}| \leq c_{(A.3)}|a - b|, \tag{A.3}$$

where $c_{(A.3)} := \frac{1}{2\sqrt{a \wedge b}} \vee 1$.

A.0.4 Lemma. *(Gronwall) Suppose that $z(t)$ satisfies*

$$0 \leq z(t) \leq a + \int_0^t b(s)z(s)\mathrm{d}s, \ t \geq 0, \tag{A.4}$$

for $a \geq 0$ and a non-negative and integrable function $b(s)$. Then,

$$z(t) \leq a\, \mathrm{e}^{B(t)}, \quad \textit{where} \quad B(t) := \int_0^t b(s)\mathrm{d}s. \tag{A.5}$$

A.0.5 Lemma. *[Operator norm property] The set $\mathcal{L}(X, Y)$ is a Banach space with the operator norm*

$$\|L\|_{\mathcal{L}(X,Y)} := \sup_{u \neq 0} \frac{\|Lu\|_Y}{\|u\|_X}. \tag{A.6}$$

If $L \in \mathcal{L}(X, Y)$, then

$$\|Lu\|_Y \leq \|L\|_{\mathcal{L}(X,Y)} \|u\|_X, \tag{A.7}$$

for any $u \in X$.

A.0.6 Lemma. *For a Hilbert space H with inner product $\langle \cdot, \cdot \rangle$ and that the linear operator $-A : D(A) \subset H \to H$ has a complete orthonormal set of eigenfunctions $\phi_j : j \in \mathbb{N}$ and eigenvalues $\lambda_j > 0$, ordered so that $\lambda_{j+1} \geq \lambda_j$. Then*

i. For $u \in D(A^\alpha)$ and $\alpha \in \mathbb{R}$, we have $A^\alpha \, \mathrm{e}^{-tA} \, u = \mathrm{e}^{-tA} \, A^\alpha u$ for $t \geq 0$.

ii. For each $\alpha \geq 0$, there exists a constant $c_{\text{A.8}}$ such that

$$\|A^\alpha \, \mathrm{e}^{-tA} \, \|_{\mathcal{L}(H)} \leq c_{\text{A.8}} t^{-\alpha}, \qquad t > 0. \tag{A.8}$$

iii. For $\alpha \in [0, 1]$, there is a constant $c_{\text{A.9}}$ such that

$$\|A^{-\alpha}(I - \mathrm{e}^{-tA})\|_{\mathcal{L}(H)} \leq c_{\text{A.9}} t^\alpha, \qquad t \geq 0. \tag{A.9}$$

Proof. See [Kruse, 2014]. □

A.0.7 Theorem. *[Existence and uniqueness [Mao, 2008]] Assume that for every real number $T > 0$ and integer $n > 1$, there exists a positive constant $c_{\text{A.10}}$ such that for all $t \in [0, T]$ and all $\phi, \psi \in \mathbb{R}^d$ with $|\phi| \vee |\psi| \leq n$,*

$$|f(\phi, t) - f(\psi, t)|^2 \vee |g(\phi, t) - g(\psi, t)|^2 \leq c_{\text{A.10}} |\phi - \psi|^2. \tag{A.10}$$

Assume also that for every $T > 0$, there exists a positive constant $c_{\text{A.11}}$ such that for all $(\phi, t) \in \mathbb{R}^d \times [0, T]$,

$$\phi^T f(\phi, t) + \frac{1}{2} |g(\phi, t)|^2 \leq c_{\text{A.11}} (1 + |\phi|^2). \tag{A.11}$$

Then there exists a unique global solution $\phi(t)$ to equation (2.75).

A.0.8 The Robin Laplacian

Consider the linear operator $Au := -\partial_{xx} u(x) + \kappa u(x)$. We study the ecliptic equation $Au = f$ subject to Robin boundary conditions as follows:

$$Au = f(x), \tag{A.12}$$

$$-\alpha_0 \partial_x u(0) = u(0) \tag{A.13}$$

$$\alpha_1 \partial_x u(1) = u(1) \tag{A.14}$$

where $\kappa > 0$. Multiply the equation by a test function $v \in H^1(0,1)$ in order to obtain the variational formulation of the equation as follows:

$$-\int_0^1 \partial_{xx} uv + \kappa \int_0^1 uv = \int_0^1 fv, \tag{A.15}$$

perform integration by parts:

$$\int_0^1 \partial_x u \partial_x v - \Big[\frac{1}{\alpha_1} u(1)v(1) + \frac{1}{\alpha_0} u(0)v(0)\Big] + \kappa \int_0^1 uv = \int_0^1 fv, \tag{A.16}$$

rearrange

$$\underbrace{\int_0^1 \partial_x u \partial_x v - \sum_{x=0,1} \frac{1}{\alpha_x} u(x)v(x) + \kappa \int_0^1 uv}_{a[u,v]} = \int_0^1 fv. \tag{A.17}$$

The first term of the left hand side can be bounded by using Cauchy–Schwarz and Young inequalities,

$$\int_0^1 \partial_x u \partial_x v + \kappa \int_0^1 uv \le \frac{1}{2} \|\partial_x u\|_{L^2(0,1)} \|\partial_x v\|_{L^2(0,1)} + \frac{\kappa^2}{2} \|u\|_{L^2(0,1)} \|v\|_{L^2(0,1)}, \tag{A.18}$$

$$\le \kappa \|u\|_{H^1(0,1)} \|v\|_{H^1(0,1)}, \tag{A.19}$$

and the trace inequality for the second term,

$$\sum_{x=0,1} -\frac{1}{\alpha_x} u(x)v(x) \le c \|u\|_{L^2(\partial D)} \|v\|_{L^2(\partial D)}. \tag{A.20}$$

Hence,

$$a[u,v] \le \kappa \|u\|_{H^1(0,1)} \|v\|_{H^1(0,1)} + c \|u\|_{L^2(\partial D)} \|v\|_{L^2(\partial D)}. \tag{A.21}$$

In order to satisfy the coercivity we have

$$a[u,u] = \int_0^1 \partial_x u \partial_x u - \sum_{x=0}^{1} \frac{1}{\alpha_x} u^2(x) + \kappa \int_0^1 u^2, \tag{A.22}$$

for $x_* \in (0,1)$, where $x_* = \arg\min_{(0,1)} |u|^2$ and $0 \le u \in C^1$,

$$u(0) = \int_0^{x_*} \partial_x u + u(x_*), \tag{A.23}$$

$$u(1) = \int_{x_*}^1 \partial_x u + u(x_*). \tag{A.24}$$

Considering the L^2 norm, yields

$$|u(0)|^2 \le 2\Big(\int_0^{x_*} |\partial_x u|^2 \Big) x_* + 2|u(x_*)|^2, \tag{A.25}$$

$$|u(1)|^2 \le 2\Big(\int_{x_*}^1 |\partial_x u|^2 \Big) (1 - x_*) + 2|u(x_*)|^2. \tag{A.26}$$

Since $x_* = \arg\min_{(0,1)} |u|^2$ then we can say that $|u(x_*)| \leq |u(x)|$, hence

$$\int_0^1 |u(x_*)|^2 \mathrm{d}x \leq \int_0^1 |u(x)|^2 \mathrm{d}x, \tag{A.27}$$

yields

$$|u(x_*)|^2 \leq \|u\|^2_{L^2(0,1)}. \tag{A.28}$$

Plugging in (A.25) and (A.26) respectively,

$$|u(0)|^2 \leq 2x_* \Big(\int_0^{x_*} |\partial_x u|^2 \Big) + 2\|u\|^2_{L^2(0,1)}, \tag{A.29}$$

$$|u(1)|^2 \leq 2(1-x_*) \Big(\int_{x_*}^1 |\partial_x u|^2 \Big) + 2\|u\|^2_{L^2(0,1)}. \tag{A.30}$$

We can conclude that for $\bar{\alpha} = \min_{x=0,1} \alpha_x$,

$$\frac{1}{\bar{\alpha}} \Big(|u(0)|^2 + |u(1)|^2 \Big) \leq 4 \int_0^1 |\partial_x u|^2 + 4\|u\|^2_{L^2(0,1)}. \tag{A.31}$$

Thus, for $\bar{\alpha}$ big enough ($-\frac{1}{\bar{\alpha}}$ small enough), we have

$$a[u,u] \geq c\|u\|^2_{H^1(0,1)}. \tag{A.32}$$

Then by the Lax–Milgram theorem, there is a unique solution $u \in H^1(0,1)$ to the equation $a(u,v) = f(v)$.

A.0.9 Heat kernel estimates

Here we collect all estimates on heat kernel that we use. Consider the following basis

$$\phi_n(x) = a_n(-\sqrt{\lambda_n} \cos(\sqrt{\lambda_n} x) + \sin(\sqrt{\lambda_n} x)), \tag{A.33}$$

where a_n and $n \in \mathbb{N}$. Then the heat kernel for the Laplace operator on $(0,1)$ with separated boundary conditions is given by

$$\mathcal{K}_t(x,y) = \sum_{n=0}^{\infty} \phi_n(x)\phi_n(y)\, \mathrm{e}^{-\lambda_n^2 t}, \tag{A.34}$$

where, λ_n solves the following

$$\tan\sqrt{\lambda_n} = \frac{2\sqrt{\lambda_n}}{1-\lambda_n}. \tag{A.35}$$

A.0.10 Lemma. *(Heat kernel estimates). There are constant $c_{A.36}, c_{A.37}, c_{A.38}, c_{A.39} > 0$ such that for every $x, x_1, x_2 \in (0,1)$ and $t, t_1, t_2 \geq 0$,*

$$\|\mathcal{K}_t(x,\cdot)\|_{L^2(0,1)} \leq c_{A.36}(1 + t^{-1/4}), \tag{A.36}$$

$$\|\mathcal{K}_t(x_1,\cdot)-\mathcal{K}_t(x_2,\cdot)\|_{L^2(0,1)} \le c_{A.37} t^{-1/4}(1\wedge|x_1-x_2|t^{-1/2}), \tag{A.37}$$

$$\|\partial\mathcal{K}_t(x,\cdot)\|_{L^2(0,1)} \le c_{A.38} t^{-3/4}, \tag{A.38}$$

$$\|\partial_t\mathcal{K}_t(x,y)u(y)\|_{H^{-1}(0,1)} \le c_{A.39} t^{-5/4}. \tag{A.39}$$

Proof. First we show inequality (A.36). Since $(\S_h)_{n>0}$ is an orthonormal basis in $L^2(0,1)$,

$$\begin{aligned}
\|\mathcal{K}_t(x,y)\|_{L^2(0,1)} &= \Big(\int_0^1 |\mathcal{K}_t(x,y)|^2\mathrm{d}y\Big)^{1/2} \\
&= \Big(\int_0^1 |\sum_{n=0}^{\infty}\phi_n(x)\phi_n(y)\,\mathrm{e}^{-\lambda_n^2 t}\,|^2\mathrm{d}y\Big)^{1/2} \\
&= \Big(|\sum_{n=0}^{\infty}\phi_n(x)\,\mathrm{e}^{-\lambda_n^2 t}\,|^2\Big)^{1/2} \\
&= \Big(\sum_{n=0}^{\infty}\mathrm{e}^{-2\lambda_n^2 t}\Big)^{1/2} \\
&\le (1+ct^{-1/2})^{1/2} \\
&\le c(1+t^{-1/4}).
\end{aligned} \tag{A.40}$$

Inequality (A.37) can be proved as follows,

$$\|\mathcal{K}_t(x_1,\cdot)-\mathcal{K}_t(x_2,\cdot)\|_{L^2(0,1)} \tag{A.41}$$

For (A.38)

$$\begin{aligned}
\|\partial_y\mathcal{K}_t(x,y)\|_{L^2(0,1)} &= \Big(\sum_{n=0}^{\infty} n^2\pi^2\phi_n(x)^2\,\mathrm{e}^{-2\lambda_n^2 t}\Big)^{1/2} \\
&= \Big(\sum_{n=0}^{\infty} n^2\pi^2\,\mathrm{e}^{-2\lambda_n^2 t}\Big)^{1/2} \\
&\le ct^{-3/4}.
\end{aligned} \tag{A.42}$$

Estimating (A.39)

$$\|\partial_t\mathcal{K}_t(x,y)u(y)\|_{H^{-1}(0,1)} = \sup_{\|\varphi\|_{H_0^1(0,1)}=1} \underbrace{\int_0^1\int_0^1 \partial_t\mathcal{K}_t(x,y)u_0(y)\varphi(x)\mathrm{d}y\mathrm{d}x}_{i} \tag{A.43}$$

$$\begin{aligned}
i &= \int_0^1\int_0^1 \partial_t\mathcal{K}_t(x,y)u_0(y)\varphi(x)\mathrm{d}y\mathrm{d}x \\
&= \Big|\int_0^1\int_0^1\sum_{n=0}^{\infty}\lambda_n^2\phi_n(x)\phi_n(y)\,\mathrm{e}^{-\lambda_n^2 t}\,u_0(y)\varphi(x)\mathrm{d}y\mathrm{d}x\Big| \\
&= \Big|\sum_{n=0}^{\infty}\lambda_n^2\,\mathrm{e}^{-\lambda_n^2 t}\underbrace{\int_0^1\phi_n(x)\varphi(x)\mathrm{d}x}_{i_a}\underbrace{\int_0^1\phi_n(y)u(y)\mathrm{d}y}_{\hat{u}_n}\Big|
\end{aligned} \tag{A.44}$$

$$
\begin{aligned}
i_a &= \int_0^1 \phi_n(x)\varphi(x)\mathrm{d}x \\
&= \int_0^1 \partial_x S_h(x)\varphi(x)\mathrm{d}x \\
&= \underbrace{\Big[\varphi(x)S_h(x)\Big]_0^1}_{=0} - \int_0^1 S_h(x)\partial_x\varphi(x)\mathrm{d}x \\
&= -\int_0^1 S_h(x)\partial_x\varphi(x)\mathrm{d}x,
\end{aligned}
\tag{A.45}
$$

where $S_h(x) = \int_0^x \phi_n(\varsigma)\mathrm{d}\varsigma$ and $\hat{u}_n$ are Fourier coefficient of u. Hence,

$$
\begin{aligned}
&\Big|\sum_{n=0}^{\infty} \lambda_n^2\, \mathrm{e}^{-\lambda_n^2 t}\, \hat{u}_n \int_0^1\int_0^x \phi_n(\varsigma)\mathrm{d}\varsigma\partial_x\varphi(x)\mathrm{d}x\Big| \\
&\le \Big(\sum_{n=0}^{\infty} \lambda_n^4\, \mathrm{e}^{-2\lambda_n^2 t}(\hat{u}_n)^2\Big)^{1/2} \Big(\underbrace{\sum_{n=0}^{\infty}\Big(\int_0^1\int_0^x S_h(\varsigma)\mathrm{d}\varsigma\partial_x\varphi(x)\mathrm{d}x\Big)^2}_{\le c}\Big)^{1/2} \\
&\le \sqrt{c}\Big(\underbrace{\sum_{n=0}^{\infty} \lambda_n^4\, \mathrm{e}^{-2\lambda_n^2 t}(\hat{u}_n)^2}_{\le 3\sqrt{\pi}/2^{11/2}t^{5/2}}\Big)^{1/2} \\
&ct^{-5/4}.
\end{aligned}
\tag{A.46}
$$

□

Bibliography

Austing, P. (2014). *Smile pricing explained*. Financial Engineering Explained. Palgrave Macmillan.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Political Economy*, (81):637–654.

Bréhier, C.-E. (2014). A short introduction to Stochastic PDEs. The content is based on the lectures delivered at CERMICS in March 2014.

Bruned, Y. (2016). *Singular KPZ Type Equations*. PhD thesis, Mathematics.

Capiński, M. and Zastawniak, T. (2011). *Mathematics for finance: an introduction to financial engineering*. Springer. OCLC: 670041078.

Da Prato, G. and Zabczyk, J. (2014). *Stochastic equations in infinite dimensions*. Cambridge : Cambridge University Press.

Dalang, R. C. and Lluis, Q.-S. (2011). Stochastic integrals for spde's: a comparison. *Expositiones Mathematicae*, 29(1):67–109.

Evans, L. (2013). *An Introduction to Stochastic Differential Equations*. Providence, Rhode Island American Mathematical Society.

Funaki, T. and Quastel, J. (2015). KPZ equation, its renormalization and invariant measures. *Stochastic Partial Differential Equations: Analysis and Computations*, 3(2):159–220.

Gatheral, J. and Taleb, N. N. (2011). *The Volatility surface A Practitioner's Guide*. John Wiley & Sons, New York, NY. OCLC: 899182374.

Gikhman, I. I. (2011). A short remark on feller's square root condition. *SSRN Electronic Journal*.

Hairer, M. (2013). Solving the KPZ equation. *Annals of Mathematics*.

Hairer, M. (2014). A theory of regularity structures. *Analysis of PDEs (math.AP); Mathematical Physics (math-ph); Probability (math.PR).*

Haugh, M. (2009). *IEOR E4707: Financial Engineering: Continuous-Time Models.*

Karatzas, I. and Shreve, S. E. (1988). *Brownian motion and stochastic calculus.* Springer-Verlag, New York.

Kardar, M., Parisi, G., and Zhang, Y.-C. (1986). Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889–892.

Kelley, J. L. (1975). *General topology.* Springer-Verlag.

Kruse, R. (2014). *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations.* LectureNotesinMathematics 2093. Springer.

Langtangen, H. P. and Logg, A. (2017). *Solving PDEs in Python.* Springer.

Lord, G., Powell, C., and Shardlow, T. (2014). *An Introduction to Computational Stochastic PDEs.* Cambridge University Press.

Mao, X. (2008). *Stochastic Differential Equations and Applications.* Woodhead Publishing, 2nd edition edition.

Mytnik, L., Perkins, E., and Sturm, A. (2006). On pathwise uniqueness for stochastic heat equations with non-lipschitz coefficients. *The Annals of Probability*, 34(5):19910–1959.

Neftci, S. and Hirsa, A. (2000). *An Introduction to the Mathematics of Financial Derivatives.* Academic Press.

Øksendal, B. (1985). *Stochastic Differential Equations, An Introduction with Applications.* Universitext. Springer-Verlag Berlin Heidelberg, 1 edition.

Ross, S. (2009). *Probability and Statistics For Engineers and Scientists.* Elsevier Academic Press.

Rouah, F. (2013). *Heston Model and its Extensions in Matlab and C.* Hoboken : Wiley.

Thomée, V. (2006). *Galerkin Finite Element Methods for Parabolic Problems.* Springer.

Walsh, J. B. (1986). Ecole d'été de probabilités de Saint-Flour XIV – 1984. In *An introduction to stochastic partial differential equations,Lecture Notes in Mathematics*, Berlin Heidelberg New York Tokyo. Springer-Verlage.

Wilmott, P., Howison, S., and Dewynne, J. (2010). *The mathematics of financial derivatives: a student introduction*. Cambridge University Press.