

University of Trento

Department of mathematics
Ph.D. in mathematics
XXIX cycle



Delayed Forward-Backward stochastic PDE's
driven by non Gaussian Lévy noise
with application in finance

FINAL THESIS

Advisor:
Prof. Luca Di Persio

Ph.D. student:
Francesco Giuseppe Cordoni

Committee members:

Prof. S. Albeverio (University of Bonn);
Prof. L. Campi (London school of economics);
Prof. L. Di Persio (University of Verona);
Prof. L. Maticiuc (Gheorghe Asachi Technical University).

Department of mathematics, University of Trento,
November 28, 2016.

To Marta

*"There are 10^{11} stars in the galaxy.
That used to be a huge number.
But it's only a hundred billion.
It's less than the national deficit!
We used to call them astronomical numbers.
Now we should call them economical numbers."*

Richard P. Feynman

Acknowledgement

First, I would like to deeply thank my advisor, Prof. LuCa Di Persio, not only for his support and guidance through the mathematical world, but mostly for the help he has given me, and I am sure he will, in the everyday life. He helped me grow both as a mathematician and more important as a person. You cannot mention LuCa with mentioning his 3/4, so that my sincere thanks go also to Francesca for her fundamental contribution "behind the scene".

I have certainly to thank Prof. S. Albeverio, Prof. L. Campi and Prof. L. Maticiuc, for having accepted to be part of the thesis committee; thanks for sharing you mathematical knowledge and for the fundamental contributions to this thesis.

My sincere thanks go also to Prof. V. Barbu, Prof G. Di Nunno and Prof. A. Zalescu for their expert advices through the last three years.

I would like also to thank Prof. R. Serapioni for the guidance to teaching, which has been a significant part of my Ph.D. experience.

Contents

Overview of the thesis	ix
I Stochastic calculus for delay equations	1
1 Path-dependent calculus and BSDE with time-delayed generator	3
1.1 Introduction	4
1.2 BSDE with delayed generator	9
1.2.1 Preliminaries	9
1.2.2 The forward-backward delayed system	13
1.2.3 Path-dependent PDE – proof of the existence theorem	25
1.3 Financial applications	43
1.3.1 The large investor problem	43
1.3.2 Risk measures via g-expectations	45
1.3.3 Pricing under counterparty risk	48
1.3.4 Conclusion and further development	54
2 Stochastic calculus for functional delay differential equations with jumps	55
2.1 Stochastic systems with memory and jumps	56
2.1.1 Stochastic functional differential equations with jumps	59
2.1.2 \mathcal{D} framework	66
2.1.3 M^P framework	70
2.1.4 Robustness SFDDs	73
2.2 Itô's formula	79
2.3 A nonlinear Kolmogorov equation for stochastic delay differential equations with jumps	91
2.3.1 Notations and preliminaries	95
2.3.2 Forward stochastic functional differential equation with delay	96
2.3.3 Joint quadratic variation	104
2.3.4 Existence of mild solutions of Kolmogorov equation	111
2.4 Financial applications	116
2.4.1 Path-dependent derivative pricing in stochastic volatility models with jumps	117
2.4.2 Application to option pricing in market with memory	123
2.5 Connection with path-dependent calculus	128

2.5.1	Pathwise derivatives and functional Itô's formula	132
2.5.2	The nonlinear Feynman-Kac formula and the connection with PPIDE	133
2.5.3	Comparison with the representation given in [CDPO16]	135
II Infinite dimensional analysis and network models		137
3	Infinite dimensional analysis	139
3.1	Invariant measure for the Vasicek model in the Heath–Jarrow–Morton–Musielà framework	139
3.1.1	Preliminary results	141
3.1.2	The Heath-Jarrow-Morton-Musiela SPDE	145
3.1.3	An approximation problem	151
3.1.4	Conclusions	160
3.2	Optimal control of stochastic FitzHugh-Nagumo equation	160
3.2.1	Existence for equation (3.102)	162
3.2.2	The optimal control of stochastic FitzHugh-Nagumo equation	165
3.2.3	An example	175
3.2.4	The linear multiplicative noise perturbation	176
3.3	Optimal control for the stochastic FitzHugh-Nagumo equation with recovery variable	177
3.3.1	The abstract setting	177
3.3.2	The optimal control problem	179
4	Infinite dimensional analysis on networks	189
4.1	Gaussian estimate on networks with dynamic stochastic boundary conditions	189
4.1.1	General framework	190
4.1.2	The perturbed stochastic problem	200
4.1.3	Application to stochastic optimal control	204
4.1.4	Conclusions	207
4.2	Stochastic reaction-diffusion equations on networks with dynamic time-delayed boundary conditions	208
4.2.1	General framework	209
4.2.2	On the infinitesimal generator	212
4.2.3	The perturbed stochastic problem	216
4.2.4	Application to stochastic optimal control	221
III Rigorous asymptotic expansions		227
5	Small noise asymptotic expansion	229
5.1	Asymptotic expansion for some local volatility models arising in finance	229
5.2	The asymptotic expansion	230
5.2.1	The general setting	230
5.2.2	The asymptotic character of the expansion of the solution X_t^ϵ of the SDE in powers of ϵ	234
5.3	Corrections around the Black-Scholes price (with Brownian, resp. Brownian plus jumps)	237

5.3.1	A correction given by an exponential function	239
5.3.2	A correction given by an exponential function and jumps	246
5.3.3	A correction given by a polynomial function	249
5.3.4	A correction given by a polynomial function and jumps	255
5.3.5	The Black-Scholes model with stochastic interest rate	258
5.4	Small noise expansion for the Lévy perturbed Vasicek model	265
5.4.1	The Vasicek model: a first order correction	266
5.4.2	The Vasicek model: a first order correction with jumps	268
5.4.3	Application to pricing	268

Overview of the thesis

The birth of modern financial mathematics dates back to the very beginning of the last century, when a French doctoral student, Louis Bachelier¹ defended the thesis titled *Théorie de la spéculation* (The theory of speculation), see [Bac00]. The seminal work of Bachelier is not only the first known attempt to concrete modelling financial markets, but it appears also to be the first mathematical definition of the *Brownian motion* (although the first definition of *Brownian motion* is to be attributed to the botanist Robert Brown²).

Bachelier's thesis was published five years earlier than Einstein's³ (independent) work on Brownian motion, when during the so-called *annus mirabilis*, he published the famous paper connecting diffusion theory and molecular physics, see [Ein05]. It is however, Norbert Wiener⁴, the first to rigorously define and construct the *Brownian motion*, which is also called *Wiener process* in his honour. In particular his main results consist in proving the existence of Brownian motion and in constructing the so-called *Wiener measure* which gives the probability distribution of underlying Brownian motion.

Probability, and the theory of stochastic processes, made thus several steps forward thanks to the fundamental works of Andrey Kolmogorov⁵. In the seminal book [Kol56] Kolmogorov, with a methodology similar to Euclid's work in geometry, built a set of fundamental axioms, based mainly on Émile Borel⁶ results on measure theory, on which the modern theory of probability is founded. He also introduced the notion of conditional expectation, that will be the brick on which derivative pricing will be based several years later.

The last key ingredient to the full development of a comprehensive theory of stochastic processes was made by the Japanese mathematician Kiyosi Itô⁷. In fact the Brownian motion is a continuous everywhere function but nowhere differentiable and, being thus an infinite variation process, standard rules of calculus does not apply to the Brownian motion, reason for which an ad hoc calculus was to be properly derived. The major drawbacks are that the theory of Riemann–Stieltjes integration and differential calculus cannot be used,

¹Louis Jean-Baptiste Alphonse Bachelier (Le Havre, France, 11 March 1870–Saint-Servan-sur-Mer, France, 28 April 1946)

²Robert Brown (Montrose, Scotland, 21 December 1773–London, England, 10 June 1858)

³Albert Einstein (Ulm, Kingdom of Württemberg, German Empire, 14 March 1879–Princeton, New Jersey, U.S., 18 April 1955)

⁴Norbert Wiener (Columbia, Missouri, U.S., November 26, 1894–Stockholm, Sweden, March 18, 1964)

⁵Andrey Nikolaevich Kolmogorov (Tambov, Russian Empire, 25 April 1903–Moscow, Soviet Union, 20 October 1987)

⁶Félix Édouard Justin Émile Borel (Saint-Affrique, France, 7 January 1871–Paris, France, 3 February 1956)

⁷Kiyosi Itô (Hokusei, Mie, Honshū, Japan, September 7, 1915–Kyōto, Japan, November 10, 2008)

so that the two main mathematical objects on which all the modern mathematics is founded were thus missing. In [Itô51] Itô derived what is now called *Itô lemma* (or Itô formula), which represents the stochastic counterpart to the chain rule in standard calculus. The impact of Itô lemma to stochastic calculus cannot be overstated, and quoting Steven Shreve "*Stochastic calculus is little more than repeated use of Itô formula in a variety of situations*".

One last honourable mention has to be made; being this thesis concerned with jumps processes, we cannot avoid to mention the French mathematician Paul Lévy⁸. Lévy made important breakthrough in the theory of stochastic processes, introducing among others *martingales*, see [Lév25]; also *Lévy processes* are clearly named in his honour.

While Itô's work was changing the world of stochastic calculus, the economist Harry Markowitz⁹ published his pioneering works in modern portfolio theory, see [Mar52, Mar68]. Markowitz's Ph.D. thesis was perhaps the work that gave the way to the development of the mathematical theory of finance.

Some years later Paul Samuelson¹⁰ realized that Itô's stochastic calculus was tailor-made for modelling financial markets, so that he started to use the newly born stochastic calculus to study prices in financial markets see [Sam73]. Meanwhile two other economists, Fisher Black¹¹ and Myron Scholes¹², published what is nowadays considered the most influential work in financial mathematics, see [BS73], where they introduced the well-known *Black-Scholes model* for pricing financial derivatives. In [BS73] Black and Scholes solved the problem of finding the fair value for a European call option, for this work they were awarded the 1997 Nobel prize in Economics. One year later, Robert Merton¹³, who is perhaps Samuelson's most notable student, published another fundamental work on pricing corporate bonds, see [Mer74] (in his honour the *Black-Scholes model* is some times referred to as *Black-Scholes-Merton model*).

From the very first results, the mathematical theory of financial markets has undergone several changes, mostly due to financial crises who forced the mathematical-economical community to change the basic assumptions on which the whole theory is founded. Consequently a new mathematical foundation were needed. In particular the 2007/2008 credit crunch showed the word that a new financial theoretical framework was necessary, since several empirical evidences emerged that aspects that were neglected prior to these years were in fact fundamental if one has to deal with financial markets.

The goal of the present thesis goes in this direction; we aim at developing rigorous mathematical instruments that allows to treat fundamental problem in modern financial mathematics. In order to do so, the current project is divided into three main parts, which focus on three different topics of modern financial mathematics. The first part is concerned with *delay equations*, the second part deal with *infinite dimensional analysis and network models* whereas the last part treats the topic of *rigorous asymptotic expansions*.

⁸Paul Pierre Lévy (Paris, France, 15 September 1886 – Paris, France, 15 December 1971)

⁹Harry Markowitz (Chicago, Illinois, 24 August 1927)

¹⁰Paul Anthony Samuelson (Gary, Indiana, US, 15 May 1915 – Belmont, Massachusetts, US, 13 December 2009)

¹¹Fischer Sheffey Black (Washington, D.C., U.S., 11 January 1938 – New York, U.S., 30 August 1995)

¹²Myron Samuel Scholes (Timmins, Ontario, Canada, 1 July 1941)

¹³Robert Cox Merton (New York City, New York, USA, 31 July 1944)

Part I: Stochastic calculus for delay equations

Part *I* is devoted to the study of delay differential equations. The study of stochastic delay differential equations has been first done in the 80's, mostly in the case where the driving noise is a standard Brownian motion. From the very first results the importance of delay equations in mathematical finance was clear: in fact on one side they allow to study market imperfections, such as delays due to some physical constraints, i.e. time necessary to transport some commodities, on the other side delay equations provide a fundamental tool in studying path-dependent options.

In the first part, we will carry out a comprehensive study of delay equations, from the existence and uniqueness, to *Feynman–Kac* type results. Thus, we will provide several examples in financial mathematics where the results just derived play a crucial role. The mathematical ingredient, besides *delay equations*, will be backward stochastic differential equations (BSDE), which are by now one of the most intensive area of research in stochastic processes, mainly linked to optimal stochastic control problem and mathematical finance.

This part is divided into two chapters, which consists of, respectively, two and four sections. Chapter 1 will focus on forward–backward system with delay, where the delay may also enter the backward component. This last point being the main novelty of chapter 1, which leads to non-trivial mathematical results. In particular the chapter is based on the new *path-dependent calculus* recently developed in [CF10, CF13]. We will prove the existence of a viscosity solution, in a suitable sense, to an associated (*path-dependent*) PDE.

Thus, several concrete example of possible applications which are relevant in modern mathematical finance are treated. In particular in section 1.3.4 we will show how previously derived results are particularly suitable to model *counterparty risk* and *collateralization*, which the recent credit crunch shows to be fundamental and cannot be neglected in the valuation of over-the-counter derivatives.

Chapter 2 is also concerned with delay differential equations. We will in this chapter consider instead the setting of *stochastic functional delay differential equations* (SFDDE) first introduced in [Moh84]. We will generalize this setting allowing also for a jump noise giving thus a complete characterization of delay equations with jumps, we refer to section 2.1 to an introduction to delay equations. We will first show in 2.1 the existence and uniqueness of a solution to a SFDDE with jumps, both in the space of càdlàg functions and Lebesgue integrable functions. Then using the tools of *calculus via regularization* we will derive in section 2.2 an ad hoc Itô formula for this type of delay equations. Thus, in section 2.3 we will prove a *Feynman–Kac type* result under mild assumptions on differentiability for the coefficients. Using then all the above mentioned results we will provide applications to *path-dependent options* and market with memory. Eventually, we will show how the present setting is in fact connected to the *path-dependent setting* used in section 1.2.3.

Part II: Infinite dimensional analysis and network models

Part *II* is concerned with *infinite dimensional analysis and network models*. The modern theory of partial differential equations relies on weak concepts of solutions, using tools of functional analysis and function spaces. As in the deterministic case, in the stochastic one, *infinite dimensional analysis* allows to deal with a wide variety of problem when considering a given differential equation. In particular, when one consider a problem where the underlying domain is a network, the underlying domain can be incorporated in a suitable choice of a space of functions. This allow to reformulate the problem in an abstract setting, so that one

can rely on standard result in operator theory.

As previous part, part *II*, consists of two chapters. Chapter 3, which is divided into 3 sections, deals with two different models, i.e. the *Vasicek model* in interest rate modelling and the *FitzHugh–Nagumo model* (FHN) in neurobiology. In section 5.4.1 we will show how to write the Vasicek model in the *Heath–Jarrow–Morton–Musielá* framework as an infinite dimensional stochastic PDE, and exploiting the property of the leading semigroup we will characterize the associated invariant measure. The following two sections, namely sections 3.2–3.3, are concerned with an optimal control problem for an infinite dimensional version of the FHN equation; section 3.2 considers the FHN without recovery variable whereas in section 3.3 the recovery variable is also included into the model.

The second chapter of the current section, namely chapter 4, is devoted to *infinite dimensional analysis on networks*. In particular it is of recent interest to consider diffusion problems where the underlying domain is a graph, where possible applications range from quantum mechanics to system of interconnected bank in financial mathematics, this last application being of particular interest in modelling contagion risk of default. Introducing a suitable product function space, the original problem on a network can be rewritten as an infinite dimensional (stochastic) PDE. More recently the attention has been also put on what type of boundary conditions one can endow the problem with, that is what type of dynamics has to be prescribed at the intersection of two adjacent edges of the network. This last point will be the main object of investigation of the chapter. In fact we will generalize the standard boundary condition in network model, that is the so-called *Kirchhoff conditions*. In section 4.1 we will use aforementioned techniques in order to prove a Gaussian estimate for the leading semigroup of a (stochastic) PDE on a network with both static and dynamic *non-local* dynamic conditions; eventually, an application to optimal control is proposed. The second section deals with a similar problem to the one above, where now the boundary condition are assumed to be *non-local* in time exhibiting a delay. In doing so we exploit another different possible technique in dealing with delay equations, that is *semigroup theory*; this also hints that even the first chapter 1 was in fact based on *infinite dimensional* techniques.

Part III: Rigorous asymptotic expansions

Eventually Part *III* deals with asymptotic expansion. Small noise expansion is widely used in physics in order to obtain perturbed solution of *Schrödinger equation*. In fact when considering particularly complex potentials V one cannot hope to obtain an analytical solution for the *Schrödinger equation*. A possible solution is to expand the potential as a Taylor summation and to obtain thus approximated solutions.

Recent evidence has shown how the Black–Scholes model is highly unrealistic, where instead local (or stochastic) volatility models should be used if one aim at capturing several factors in the markets. Clearly the main drawback in considering more general models is that one cannot hope of having anymore explicit analytic solutions, leading in some particular case to problem which are hard to treat also numerically. Borrowing the above idea from physics we apply a perturbation technique to the pricing of a contingent claim in a financial market. The main difficulty in doing so lies in the fact that the underlying function is stochastic; one has thus to derive a rigorous (small noise) asymptotic expansion for a stochastic differential equation (SDE). Therefore we derive in Section 5.2 the rigorous asymptotic expansion for general SDE, allowing also to consider general jump noises. Then

in Section 5.4 we provided several example of approximation prices for contingent claims centred around the Black–Scholes price. We also provide some numerical examples that validate our expansion.

Part I

Stochastic calculus for delay
equations

1 | Path-dependent calculus and BSDE with time-delayed generator

The present chapter is taken from [CDPMZ16, CDP16a].

Abstract

We prove the existence of a viscosity solution of the following path dependent nonlinear Kolmogorov equation:

$$\begin{cases} \partial_t u(t, \phi) + \mathcal{L}u(t, \phi) + f(t, \phi, u(t, \phi), \partial_x u(t, \phi)\sigma(t, \phi), (u(\cdot, \phi))_t) = 0, & t \in [0, T], \phi \in \mathcal{C}, \\ u(T, \phi) = h(\phi), & \phi \in \mathcal{C}, \end{cases}$$

where $\mathcal{C} = \mathcal{C}([0, T]; \mathbb{R}^d)$, $(u(\cdot, \phi))_t := (u(t + \theta, \phi))_{\theta \in [-\delta, 0]}$ and

$$\mathcal{L}u(t, \phi) := \langle b(t, \phi), \partial_x u(t, \phi) \rangle + \frac{1}{2} \text{Tr} [\sigma(t, \phi)\sigma^*(t, \phi)\partial_{xx}^2 u(t, \phi)].$$

The result is obtained by a stochastic approach. In particular we prove a new type of nonlinear Feynman–Kac representation formula associated to a backward stochastic differential equation with time-delayed generator which is of non-Markovian type.

The second part is concerned with financial applications to the large investor problem and risk measures via g -expectations. Also we consider a non-linear pricing problem that takes into account credit risk and funding issues. The aforementioned problem is formulated as a stochastic forward–backward system with delay, both in the forward and in the backward component, whose solution is characterized in terms of viscosity solution to a suitable type of *path-dependent PDE*.

1.1 Introduction

We aim at providing a probabilistic representation of a viscosity solution to the following path-dependent nonlinear Kolmogorov equation (PDKE)

$$\begin{cases} -\partial_t u(t, \phi) - \mathcal{L}u(t, \phi) - f(t, \phi, u(t, \phi), \partial_x u(t, \phi), \sigma(t, \phi), (u(\cdot, \phi))_t) = 0, \\ u(T, \phi) = h(\phi), \end{cases} \quad (1.1)$$

for $t \in [0, T]$, $\phi \in \mathcal{C} := \mathcal{C}([0, T]; \mathbb{R}^d)$ being the space of continuous \mathbb{R}^d -valued functions defined on the interval $[0, T]$, being $T < \infty$ a fixed time horizon. Also, for a fixed delay $\delta > 0$, we have set $(u(\cdot, \phi))_t := (u(t + \theta, \phi))_{\theta \in [-\delta, 0]}$. In equation (1.1) we have denoted by \mathcal{L} the second order differential operator given by

$$\mathcal{L}u(t, \phi) := \frac{1}{2} \text{Tr}[\sigma(t, \phi) \sigma^*(t, \phi) \partial_{xx}^2 u(t, \phi)] + \langle b(t, \phi), \partial_x u(t, \phi) \rangle,$$

with $b : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}^d$ and $\sigma : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}^{d \times d'}$ being two non-anticipative functionals to be better introduced in subsequent section.

In particular we will prove that, under appropriate assumptions on the coefficients, being

$$(X^{t, \phi}(s), Y^{t, \phi}(s), Z^{t, \phi}(s))_{s \in [t, T]},$$

the unique solution to the decoupled forward-backward stochastic differential system

$$\begin{cases} X^{t, \phi}(s) = \phi(t) + \int_t^s b(r, X^{t, \phi}) dr + \int_t^s \sigma(r, X^{t, \phi}) dW(r), & s \in [t, T], \\ Y^{t, \phi}(s) = h(X^{t, \phi}) + \int_s^T f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) dr \\ \quad - \int_s^T Z^{t, \phi}(r) dW(r), & s \in [t, T], \end{cases} \quad (1.2)$$

with $(t, \phi) \in [0, T] \times \mathcal{C}$ and W a standard Brownian motion, then the deterministic non-anticipative functional $u : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$ given by the representation formula $u(t, \phi) := Y^{t, \phi}(t)$ is a viscosity solution, in the sense of [EKT⁺14], to equation (1.1). Above, the notation $Y_r^{t, \phi}$ appearing in the generator f of the backward component in system (1.2) stands for the path of the process $Y^{t, \phi}$ restricted to $[r - \delta, r]$, namely

$$Y_r^{t, \phi} := (Y^{t, \phi}(r + \theta))_{\theta \in [-\delta, 0]}.$$

In particular the forward equation is a functional stochastic differential equation, while the backward equation has time-delayed generator, that is the generator f can depend, unlike the classical backward stochastic differential equations, on the past values of $Y^{t, \phi}$.

Let us stress that if we do not consider delay neither in the forward nor in the backward component, we retrieve standard results of Markovian forward-backward system, so that in this case we obtain $u(t, \phi) = u(t, \phi(t))$, and equation (1.1) becomes

$$\begin{cases} -\partial_t u(t, x) - \mathcal{L}u(t, x) - f(t, x, u(t, x), \partial_x u(t, x), \sigma(t, x)) = 0, & t \in [0, T], x \in \mathbb{R}^d, \\ u(T, x) = h(x), & x \in \mathbb{R}^d, \end{cases}$$

with

$$\mathcal{L}u(t, x) := \frac{1}{2} \text{Tr}[\sigma(t, x)\sigma^*(t, x)\partial_{xx}^2 u(t, x)] + \langle b(t, x), \partial_x u(t, x) \rangle.$$

Let us recall that BSDE's were first introduced, in the linear case, by Bismut [Bis73], whereas the nonlinear case was considered by Pardoux and Peng in [PP90]. Later, in [PP92, Pen91], the connection between BSDEs and semilinear parabolic partial differential equations (PDE's) was established, proving the nonlinear Feynman–Kac formula for Markovian equations stated above. Also, a similar deterministic representation associated with a suitable PDE, can be proved taking into account different types of BSDE's, such as BSDE's with random terminal time, see, e.g. [DP97], reflected BSDE's, see, e.g. [EKKP⁺97], or also backward stochastic variational inequalities, see, e.g. [MR10, MR⁺15b].

When one is to consider the non–Markovian case, the associated PDE becomes path–dependent. In particular in [Pen11a] the author shows for the first time that a non–Markovian BSDE can be linked with a path–dependent PDE. Subsequently in [PW11] the authors proved, in the case of smooth coefficients, the existence and uniqueness of a classical solution for the path–dependent Kolmogorov equation

$$\begin{cases} -\partial_t u(t, \phi) - \frac{1}{2} \partial_{xx}^2 u(t, \phi) - f(t, \phi, u(t, \phi), \partial_x u(t, \phi)) = 0, & t \in [0, T], \phi \in \mathcal{C}, \\ u(T, \phi) = h(\phi), \phi \in \mathcal{C}. \end{cases} \quad (1.3)$$

In particular the authors appealed to a representation formula using the standard non–Markovian BSDE:

$$Y^{t, \phi}(s) = h(W^{t, \phi}) + \int_s^T f(W^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r)) dr - \int_s^T Z^{t, \phi}(r) dW(r), \quad s \in [t, T], \quad (1.4)$$

with the generator and the final condition depending on the Brownian paths: $W^{t, \phi}(s) = \phi(t) + W(s) - W(t)$, if $s \in [t, T]$ and $W^{t, \phi}(s) = \phi(s)$, if $s \in [0, t]$. Then in [Pen11b] a new type of viscosity solution is introduced.

Eventually in [EKT⁺14, ETZ12a, ETZ12b] the authors introduced a new notion of viscosity solutions, which is the definition we will consider in the present work, for semilinear and fully non linear path-dependent PDE, using the framework of functional Itô calculus first set by Dupire [Dup09] and Cont & Fournié [CF13].

We will, in the present work, generalized the results in [EKT⁺14] along two directions. First we will consider a BSDE whose generator depends not only on past valued assumed by a standard Brownian motion W , but the BSDE may depends on a general diffusion process X . Second, and most important generalization, we will prove the connection between path–dependent PDEs and BSDEs with time–delayed generators. We recall that time–delayed BSDE were first introduced in [DI⁺10b] and [DI10a]. More precisely the authors obtained the existence and uniqueness of the solution the the time–delayed BSDE

$$Y(t) = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z(s) dW(s), \quad 0 \leq t \leq T, \quad (1.5)$$

where $Y_s := (Y(r))_{r \in [0, s]}$ and $Z_s := (Z(r))_{r \in [0, s]}$. In particular, the aforementioned existence and uniqueness result holds true if the time horizon T or the Lipschitz constant for the generator f are sufficiently small. To our best knowledge, the link between time–delayed BSDE's and path–dependent PDE's has never been addressed in literature.

We also emphasize that our framework, since our BSDE is time-delayed, requires that the backward equation contains a supplementary initial condition to be satisfied, namely $Y^{t,\phi}(s) = Y^{s,\phi}(s)$, if $s \in [0, t)$. Let us further stress that the Feynman–Kac formula would fail with standard prolongation $Y^{t,\phi}(s) = Y^{t,\phi}(t)$, for $s \in [0, t)$. Although the existence results for equation (1.5) has already been treated in [DI⁺10b, DI10a], this new initial condition imposes a more elaborated proof.

The last part of the paper presents some financial models based on our theoretical results. In recent years delay equations have been of growing interest, mainly motivated by many concrete applications where the effect of delay cannot be neglected, see, e.g. [Moh98, KP07]. On the contrary, BSDEs with time delayed generator have been first introduced as a pure mathematical tool, with no application of interest. Only later in [Del10, Del12] the author proposed some financial applications to pricing, hedging and investment portfolio management, where backward equations with delayed generator provide a fundamental tool.

Based on the recently introduced path-dependent calculus, together with the mild assumptions of differentiability required, the probabilistic representation for a viscosity solution of a non-linear parabolic equations proved in the present paper, finds perhaps its best application in finance. In fact, a wide variety of financial derivatives can be formally treated under the theory developed in what follows, from the more standard European options, to the more exotic path-dependent options, such as Asian options or look-back options.

We propose here different possible applications of forward-backward stochastic differential system (1.2), where the delay in the backward component arises from two different motivations. The first example we will deal with is a generalization of a well-known model in finance, where we will consider the case of a non standard investor acting on a financial market. We will assume, following [CM96, EKPQ01], that a so called *large investor* wishes to invest on a given market, buying or selling a stock. This investor has the peculiarity that his actions on the market can affect the stock price. In particular, we will assume that the stock price S and the bond B are a function of the large investor's portfolio (X, π) , X being the value of the portfolio and π the number of share of the asset S .

This case has been already treated in financial literature, see, e.g. [EKPQ01]. We further generalize the aforementioned results assuming a second market imperfection, that is we assume that it might be a small time delay between the action of the large investor and the reaction of the market, so that we are led to consider the financial system with the presence of the past of X in the coefficients r, μ and σ :

$$\begin{cases} \frac{dB(t)}{B(t)} = r(t, X(t), \pi(t), X_t)dt, & B(0) = 1, \\ \frac{dS(t)}{S(t)} = \mu(t, X(t), \pi(t), X_t)dt + \sigma(t, X(t), X_t)dW(t), & S(0) = s_0 > 0, \end{cases}$$

where the notation X_t stands for the path $(X(t + \theta))_{\theta \in [-\delta, 0]}$, being δ a small enough delay.

The second example we deal with arises from a different situation. Recent literature in financial mathematics has been focused in how to measure the riskiness of a given financial investment. To this extent *dynamic risk measures* have been introduced in [ADEH99]. In particular, BSDEs have been shown to be perhaps the best mathematical tool for modelling dynamic risk measures, via the so called *g-expectations*. In [Del12] the author proposed a risk measure that takes also into account the past values assumed by the investment, that is we will assume that, in making his future choices, the investor will consider not only

the present value of the investment, but also the values assumed in a sufficiently small past interval. This has been motivated by empirical studies that show how the memory effect has a fundamental importance in an investor's choices, see, e.g. [Del12] and references therein for financial studies on the memory effect in financial investment. We therefore consider an investor that tries to quantify the riskiness of a given investment, with Y being his investment, we will assume that the investor looks at the average value of his investment in a sufficiently recent past, that is we consider a generator of the form $\frac{1}{\delta} \int_{-\delta}^0 Y(t + \theta) d\theta$, with $\delta > 0$ being a sufficiently small delay.

Eventually, the last part of the chapter, that is starting from Section 2.5.1 is concerned with a topic of great importance in modern finance, that is the pricing of derivatives in OTC markets taking into account credit/debit value adjustment, collateral and funding constraints.

Starting from the spreading of the credit crunch in 2007, empirical evidences have shown how some aspects of financial markets neglected up to that point by theoretical models, are instead fundamental in concrete economical frameworks. In particular let us mention the violation of standard non-arbitrage relation between forward rates and zero-coupon bonds. Even if we will not address latter problem in the present paper, we would like to underline how it is very connected to the topic we will treat, as witnessed by the recent, wide and growing literature linked to the so called *multi-curve modelling*, see, e.g., [Hen14, MP14, PB13, PT10, GR15], and references therein.

In what follows we will focus on a different issue which emerged after the last financial crisis, namely the problem of pricing derivatives contracts, including the possibility of the *counterparty default*, i.e., the event in which a borrower fails to make the required payments to his lender. Such an event is treated in the framework of the *credit risk* which, according to [oBSfIS00], is defined *as the potential that a bank borrower or counterparty will fail to meet its obligations in accordance with agreed terms*. Even if the number of type of related financial losses is rather huge, it is interesting to note that they may be complete, as in the case of default, or even partial, and can happen in a number of different cases, such, e.g., if a consumer fails to make a payment related to a line of credit, or if an insolvent insurance company does not pay previously stipulated policy obligations, or bank that, because of its insolvency, does not return funds to a depositor, etc. We would like to underline that the *Credit default risk* has a great impact on almost all the credit-sensitive transactions, also including mortgages, loans, securities and derivatives. Hence, its careful determination and forecasting, are crucial tasks, especially in the modern theory of financial markets, see, e.g., [EJY00, JR02], which are widely characterized by sophisticated contracts of the aforementioned type. In particular, the wrong estimation of credit default risk that, at different levels, has been experienced at the end of the last decade, is intrinsically linked to the inadequacy of classical models in describing real financial markets, mainly because of the unrealistic hypotheses of the existence of a unique risk-free rate, i.e. the theoretical rate of return of an investment with zero risk, or the possibility to have unlimited access to funding. Our aim is to derive a mathematical formulation of such problems, while we refer the interested reader to, e.g., [BMP13, BP13] and references therein, for a deep study of related financial implications, see also [LXYW14], where the credit risk is studied in connection with the so called *Catastrophe Bonds*, [WYZW14] where the default probability problem for credit risk is considered, [LMS11] to what concerns a *large deviation* approach.

Recently have appeared several works that try to include counterparty risk, i.e. the risk to each party of a contract that the counterparty will be unable to meet contractual

obligations, as well as funding issues in pricing financial contracts, leading to a systematic treatment of both of them. In particular, to our knowledge, the first attempts in the direction of developing a concrete framework able to treat both the counterparty risk and the funding constraints, can be found in [Cré11, Cré15a, Cré15b]. However, we will mainly refer to a slightly different and yet closely related approach, namely the one developed in [BMP13, BP14]. Let us mention that both approaches identify backward stochastic differential equations (BSDEs) as a fundamental mathematical object to consider a financial setting characterized by counterparty risk, see, e.g., [CDP14b, EKPQ97].

In what follows, we also exploit the approach developed in [BFP15, BP14], where the authors firstly consider the present value of a contract as the discounted present value of future payoffs, and then include margin variations and counterparty risk in their valuation BSDE, which turns out to be risk-free rate independent. In fact, the latter depends only on different funding rates which are directly observable on the markets

The main contribution of the present work is thus to give a rigorous and general mathematical foundation of the previously introduced setting. In particular, following [BFP15, BP14], we will consider the so-called *master pricing equation* generalising it in several direction. Firstly, we will consider possibly path-dependent hedging strategy exploiting the so-called *path-dependent calculus* developed in [CF13, CF10]. Secondly, we will not assume any differentiability assumptions in order to consider viscosity solutions for the related pricing PDE. Finally, as major generalization, we consider a margining procedure that can be path-dependent with respect to the portfolio. In fact, as pointed out in [Cré11, Remark. 5.5], in real world the margining scheme often depends upon its past values. It is worth to mention that the latter is highly non trivial, since, from a mathematical point of view, it implies that the related BSDE generator depends its past values as well.

Let us recall that the first rigorous treatment of delay differential equations dates back to the monograph [Moh84], while more recently, see [Dup09] and [CF13, CF10], new notions of ad hoc derivatives have been introduced to study the stochastic calculus for path dependent stochastic differential equations. Since then, such results have been then generalised in several directions, see, e.g., [BCDNR16, CDPO16, CRb, CDGR, FZ, FMT10, SX11] ,and references therein. We would like to mention that the *path-dependent calculus* has revealed itself since his born, as a powerful tool to model financial markets exhibiting delay, and also path-dependent options.

Analogously, in [EKT⁺14, EKT⁺14, EKT⁺14], the authors proposed an ad hoc notion of viscosity solutions to path-dependent PDE which, similarly to the relation established by the *Feynman-Kac* theorem between a stochastic differential equation (SDE) and its deterministic counterpart, relates a *path-dependent* SDE to a corresponding *path-dependent* PDE, by exploiting the theory of BSDE, hence by using of the notion of *non-linear expectations*, see, e.g., [Pen04].

Recently, the development of the theory of delayed stochastic differential equations, has made one step further to include, besides the delay in the forward SDE, also a delay component in the backward equation. In particular in [DI⁺10b], the authors proposed a new type of stochastic delay equation, whose generator may depends on the past values of the BSDE itself. As mentioned above this peculiarity is highly non trivial, as witnessed by several examples reported in [DI⁺10b], where the authors show how the uniqueness property for the solution fails to be true.

To overcome latter problem some additional assumptions have to be taken into account, as in [CDPMZ16], where, exploiting the notion of viscosity solution proposed in [EKT⁺14],

the connection between forward–backward SDE with delay both in forward and in the backward component and a new type of *path-dependent PDE* has been proved.

In the present paper we exploit the aforementioned results obtained in [CDPMZ16], to generalize the financial setting developed in [BFP15, BP14], allowing for *path-dependent* hedging strategies, i.e. plans to reduce the financial risk associated to adverse price movements of, e.g., assets in which one has invested, and collateralization scheme with delay. We recall that the collateralization represents the situation in which a borrower pledges an asset as recourse to the lender to hedge the case of the borrower’s default. In particular, the collateralization of assets gives to, e.g., banks a sufficient level of reassurance against the default risk. The latter banking practice allows loans to be issued to individuals or companies which do not belong to the set of the ones having optimal credit history or good debt rating. We underline that our approach is particularly suitable to treat financial frameworks characterized by delays in the default procedure, see, e.g., [BP14, Sec. 3, Sec. 4], the latter being the object of our future works.

The paper is organized as follows: in Section 1.2.1 we introduce needed notion based on functional Itô’s calculus and the notion of viscosity solution for path-dependent PDE’s. In Section 1.2.2 we prove the existence and uniqueness of a solution for the time-delayed BSDE, whereas Section 1.2.3 is devoted to the main results of the present work, that is the proof of the continuity of the function $u(t, \phi) := Y^{t, \phi}(t)$ as well as the generalization of the Feynman–Kac formula with the core of the present work, that is Theorem 1.2.14. In Section 1.3 we present the financial applications, in particular the large investor application and g -expectations. Therefore in Sec. 2.5.1 we apply above results to the problem of obtaining a portfolio under credit risk and funding issues. Eventually, in Sec. 1.3.3, we will derive the main *path-dependent* pricing PDE.

1.2 BSDE with delayed generator

1.2.1 Preliminaries

Pathwise derivatives and functional Itô’s formula

Let us first introduce the framework on which we shall construct the solutions of PDKE (1.1). In particular for a deep treatment of functional Itô calculus we refer the reader to Dupire [Dup09] and Cont & Fournié [CF13].

Let $\mathcal{D} := \mathbb{D}([0, T]; \mathbb{R}^d)$ be the set of càdlàg \mathbb{R}^d -valued functions, i.e. right continuous, with finite left-hand limits, \hat{B} the canonical process on \mathcal{D} , i.e. $\hat{B}(t, \hat{\phi}) := \hat{\phi}(t)$ and $\hat{\mathbb{F}} := (\hat{\mathcal{F}}_s)_{s \in [0, T]}$ the filtration generated by \hat{B} . On \mathcal{D} , resp. $[0, T] \times \mathcal{D}$, we introduce the following norm, resp. pseudometric, with respect to whom it becomes a Banach space, resp. a complete pseudometric space. Let us thus define, for any $(t, \hat{\phi}), (t', \hat{\phi}') \in [0, T] \times \mathcal{D}$,

$$\begin{aligned} \|\hat{\phi}\|_T &:= \sup_{r \in [0, T]} |\hat{\phi}(r)|, \\ d((t, \hat{\phi}), (t', \hat{\phi}')) &:= |t - t'| + \sup_{r \in [0, T]} |\hat{\phi}(r \wedge t) - \hat{\phi}'(r \wedge t')|. \end{aligned}$$

Let $\hat{u} : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ be an $\hat{\mathbb{F}}$ -progressively measurable non-anticipative process, that is we assume $\hat{u}(t, \hat{\phi})$ depends only on the restriction of $\hat{\phi}$ on $[0, t]$, i.e. $\hat{u}(t, \hat{\phi}) = \hat{u}(t, \hat{\phi}(\cdot \wedge t))$, for any $(t, \hat{\phi}) \in [0, T] \times \mathcal{D}$. We say that \hat{u} is vertically differentiable at $(t, \hat{\phi}) \in [0, T] \times \mathcal{D}$ if

there exist

$$\partial_{x_i} \hat{u}(t, \hat{\phi}) := \lim_{h \rightarrow 0} \frac{\hat{u}(t, \hat{\phi} + h \mathbb{1}_{[[t, T]]} e_i) - \hat{u}(t, \hat{\phi})}{h}$$

for any $i = \overline{1, d}$, where we have denoted by $\{e_i\}_{i=\overline{1, d}}$ the canonical basis of \mathbb{R}^d . The second order derivatives, when they exist, are denoted by $\partial_{x_i x_j}^2 \hat{u}(t, \hat{\phi}) := \partial_{x_i}(\partial_{x_j} \hat{u})$, for any $i, j = \overline{1, d}$. Let us further denote by $\partial_x \hat{u}(t, \hat{\phi})$ the gradient vector, that is we have

$$\partial_x \hat{u}(t, \hat{\phi}) = \left(\partial_{x_1} \hat{u}(t, \hat{\phi}), \dots, \partial_{x_d} \hat{u}(t, \hat{\phi}) \right),$$

and by $\partial_{xx}^2 \hat{u}(t, \hat{\phi})$ the $d \times d$ -Hessian matrix, that is

$$\partial_{xx}^2 \hat{u}(t, \hat{\phi}) = \left(\partial_{x_i x_j}^2 \hat{u}(t, \hat{\phi}) \right)_{i, j = \overline{1, d}}.$$

Let $t \in [0, T]$ and a path $\phi \in \mathcal{D}$, we denote

$$\phi_{(t)} := \phi(\cdot \wedge t) \in \mathcal{D}. \quad (1.6)$$

We say that \hat{u} is horizontally differentiable at $(t, \hat{\phi}) \in [0, T] \times \mathcal{D}$ if there exist

$$\partial_t \hat{u}(t, \hat{\phi}) := \lim_{h \rightarrow 0^+} \frac{\hat{u}(t+h, \hat{\phi}_{(t)}) - \hat{u}(t, \hat{\phi})}{h},$$

for $t \in [0, T)$ and $\partial_t \hat{u}(T, \hat{\phi}) := \lim_{t \rightarrow T^-} \partial_t \hat{u}(t, \hat{\phi})$.

Let $\hat{u} : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ be non-anticipative, we say that $\hat{u} \in \mathcal{C}([0, T] \times \mathcal{D})$ if \hat{u} is continuous on $[0, T] \times \mathcal{D}$ under the pseudometric d ; we write that $\hat{u} \in \mathcal{C}_b([0, T] \times \mathcal{D})$ if $\hat{u} \in \mathcal{C}([0, T] \times \mathcal{D})$ and \hat{u} is bounded on $[0, T] \times \mathcal{D}$. Eventually we say that $\hat{u} \in \mathcal{C}_b^{1,2}([0, T] \times \mathcal{D})$ if $\hat{u} \in \mathcal{C}([0, T] \times \mathcal{D})$ and the derivatives $\partial_x \hat{u}$, $\partial_{xx}^2 \hat{u}$, $\partial_t \hat{u}$ exist and they are continuous and bounded.

Having introduced the needed notations, following [Dup09], we will now work with processes u defined on $[0, T] \times \mathcal{C} \rightarrow \mathbb{R}$, being \mathcal{C} the space of continuous paths, $\mathcal{C}([0, T]; \mathbb{R}^d)$. Let B be the canonical process on Λ , i.e. $B(t, \phi) := \phi(t)$ and $\mathbb{F} := (\mathcal{F}_s)_{s \in [0, T]}$ the filtration generated by B .

From the fact that \mathcal{C} is a closed subspace of \mathcal{D} , we have that $(\Lambda, \|\cdot\|_t)$ is also a Banach space; with an analogous reasoning we claim that $([0, T] \times \mathcal{C}, d)$ is a complete pseudometric space. As done above, we have that if $u : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$ is a non-anticipative process, we write that $u \in \mathcal{C}([0, T] \times \mathcal{C})$ if u is continuous on $[0, T] \times \mathcal{C}$ under the pseudometric d ; if, moreover, u is continuous and bounded on $[0, T] \times \mathcal{C}$, we write $u \in \mathcal{C}_b([0, T] \times \mathcal{C})$. Eventually, following [EKT⁺14], we write that $u \in \mathcal{C}_b^{1,2}([0, T] \times \mathcal{C})$ if there exists $\hat{u} \in \mathcal{C}_b^{1,2}([0, T] \times \mathcal{D})$ such that $\hat{u}|_{[0, T] \times \mathcal{C}} = u$ and by definition we take $\partial_{t_i} u := \partial_t \hat{u}$, $\partial_x u := \partial_x \hat{u}$, $\partial_{xx}^2 u := \partial_{xx}^2 \hat{u}$, notice that definitions are independent of the choice of \hat{u} .

We are now to introduce the shifted spaces of càdlàg and continuous paths. If $t \in [0, T]$, \hat{B}^t is the shifted canonical process on $\mathcal{D}^t := \mathbb{D}([t, T]; \mathbb{R}^d)$, $\hat{\mathbb{F}}^t := (\hat{\mathcal{F}}_s^t)_{s \in [t, T]}$ is the shifted filtration generated by \hat{B}^t ,

$$\begin{aligned} \|\hat{\phi}\|_T^t &:= \sup_{r \in [t, T]} |\hat{\phi}(r)|, \\ d^t((s, \hat{\phi}), (s', \hat{\phi}')) &:= |s - s'| + \sup_{r \in [t, T]} |\hat{\phi}(r \wedge s) - \hat{\phi}'(r \wedge s')|, \end{aligned}$$

for any $(s, \hat{\phi}), (s', \hat{\phi}') \in [t, T] \times \mathcal{D}^t$. Analogously we define the spaces $\mathcal{C}([t, T] \times \mathcal{D}^t)$, $\mathcal{C}_b([t, T] \times \mathcal{D}^t)$ and $\mathcal{C}_b^{1,2}([t, T] \times \mathcal{D}^t)$. Similarly, we denote $\mathcal{C}^t := \mathcal{C}([t, T]; \mathbb{R}^d)$, B^t the shifted canonical process on \mathcal{C}^t , $\mathbb{F}^t := (\mathcal{F}_s^t)_{s \in [t, T]}$ the shifted filtration generated by B^t and we introduce the spaces $\mathcal{C}([t, T] \times \mathcal{C}^t)$, $\mathcal{C}_b([t, T] \times \mathcal{C}^t)$ and $\mathcal{C}_b^{1,2}([t, T] \times \mathcal{C}^t)$.

Let us denote by \mathcal{T} the set of all \mathbb{F} -stopping times τ such that for all $t \in [0, T]$, then we have that the set $\{\phi \in \mathcal{C} : \tau(\phi) > t\}$ is an open subset of $(\mathcal{C}, \|\cdot\|_T)$ and \mathcal{T}^t the be the set of all \mathbb{F} -stopping times τ such that for all $s \in [t, T]$, the set $\{\phi \in \mathcal{C}^t : \tau(\phi) > s\}$ is an open subset of $(\mathcal{C}^t, \|\cdot\|_T^t)$.

For a càdlàg function $\phi \in \mathbb{D}([-\delta, T]; \mathbb{R}^d)$, we denote

$$\phi_t := (\phi(t + \theta))_{\theta \in [-\delta, 0]}. \quad (1.7)$$

We conclude this subsection by recalling the functional version of the Itô's formula (see Cont & Fournié [CF13, Theorem 4.1]).

Theorem 1.2.1 (Functional Itô's formula). *Let A be a d -dimensional Itô process, i.e. $A : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}^d$ is a continuous \mathbb{R}^d -valued semimartingale defined on the probability space $(\mathcal{C}, \mathbb{F}, \mathbb{P})$ which admits the representation*

$$A(t) = A(0) + \int_0^t b(r) dr + \int_0^t \sigma(r) dB(r), \quad \text{for all } t \in [0, T].$$

If $F \in \mathcal{C}_b^{1,2}([0, T] \times \hat{\mathcal{X}})$ then, for any $t \in [0, T]$, the following change of variable formula holds true:

$$\begin{aligned} F(t, A_{(t)}) &= F(0, A_{(0)}) + \int_0^t \partial_t F(r, A_{(r)}) dr + \int_0^t \langle \partial_x F(r, A_{(r)}), b(r) \rangle dr \\ &\quad + \frac{1}{2} \int_0^t \text{Tr}[\sigma(r) \sigma^*(r) \partial_{xx}^2 F(r, A_{(r)})] dr + \int_0^t \langle \partial_x F(r, A_{(r)}), \sigma(r) dB(r) \rangle. \end{aligned}$$

Path-dependent PDEs

We are now to introduced the notion of viscosity solution to equation (1.1), in particular we will use the notion of viscosity solution first introduced in [EKT⁺14], see also [ETZ12a, ETZ12b].

Let $(t, \phi) \in [0, T] \times \mathcal{C}$ be fixed and $(W(t))_{t \geq 0}$ be a d' -dimensional standard Brownian motion defined on some complete probability space $(\Omega, \mathcal{G}, \mathbb{P})$. We denote by $\mathbb{G}^t = (\mathcal{G}_s^t)_{s \in [0, T]}$ the natural filtration generated by $((W(s) - W(t)) \mathbb{1}_{\{s \geq t\}})_{s \in [0, T]}$ and augmented by the set of \mathbb{P} -null events of \mathcal{G} .

Let us take $L \geq 0$ and $t < T$. We denote by \mathcal{U}_t^L the space of \mathbb{G}^t -progressively measurable \mathbb{R}^d -valued processes λ such that $|\lambda| \leq L$. We define a new probability measure $\mathbb{P}^{t, \lambda}$ by $d\mathbb{P}^{t, \lambda} := M^{t, \lambda}(T) d\mathbb{P}$, where

$$M^{t, \lambda}(s) := \exp\left(\int_t^s \lambda(r) dW(r) - \frac{1}{2} \int_t^s |\lambda(r)|^2 dr\right), \quad \mathbb{P}\text{-a.s.}$$

Under some suitable assumptions on the coefficients, to be better specified later on, see also Theorem 1.2.5 in what follows, the existence and uniqueness of a continuous and

adapted stochastic process $(X^{t,\phi}(s))_{s \in [0,T]}$ such that

$$\begin{cases} X^{t,\phi}(s) = \phi(t) + \int_t^s b(r, X^{t,\phi})dr + \int_t^s \sigma(r, X^{t,\phi})dW(r), & s \in [t, T], \\ X^{t,\phi}(s) = \phi(s), & s \in [0, t), \end{cases}$$

where $(t, \phi) \in [0, T] \times \mathcal{C}$ is given.

We are now ready to define the space of the test functions,

$$\begin{aligned} \underline{\mathcal{A}}^L u(t, \phi) &:= \left\{ \varphi \in \mathcal{C}_b^{1,2}([0, T] \times \mathcal{C}) : \exists \tau_0 \in \mathcal{T}_+^t, \varphi(t, \phi) - u(t, \phi) \right. \\ &= \left. \min_{\tau \in \mathcal{T}^t} \underline{\mathcal{E}}_t^L [(\varphi - u)(\tau \wedge \tau_0, X^{t,\phi})] \right\} \end{aligned}$$

and

$$\begin{aligned} \overline{\mathcal{A}}^L u(t, \phi) &:= \left\{ \varphi \in \mathcal{C}_b^{1,2}([0, T] \times \mathcal{C}) : \exists \tau_0 \in \mathcal{T}_+^t, \varphi(t, \phi) - u(t, \phi) \right. \\ &= \left. \max_{\tau \in \mathcal{T}^t} \overline{\mathcal{E}}_t^L [(\varphi - u)(\tau \wedge \tau_0, X^{t,\phi})] \right\}, \end{aligned}$$

where $\mathcal{T}_+^t := \{\tau \in \mathcal{T}^t : \tau > t\}$, if $t < T$ and $\mathcal{T}_+^T := \{T\}$. Also, for any $\xi \in L^2(\mathcal{F}_T^t; \mathbb{P})$, $\underline{\mathcal{E}}_t^L(\xi) := \inf_{\lambda \in \mathcal{U}_t^L} \mathbb{E}^{\mathbb{P}^{t,\lambda}}(\xi)$ and $\overline{\mathcal{E}}_t^L(\xi) := \sup_{\lambda \in \mathcal{U}_t^L} \mathbb{E}^{\mathbb{P}^{t,\lambda}}(\xi)$ are nonlinear expectations.

We are now able to give the definition of a viscosity solution of the functional PDE (1.1), see, e.g. [EKT⁺14, Def. 3.3].

Definition 1.2.2. Let $u \in \mathcal{C}_b([0, T] \times \mathcal{C})$ such that $u(T, \phi) = h(\phi)$, for all $\phi \in \mathcal{C}$.

(a) For any $L \geq 0$, we say that u is a viscosity L -subsolution of (1.1) if at any point $(t, \phi) \in [0, T] \times \mathcal{C}$, for any $\varphi \in \underline{\mathcal{A}}^L u(t, \phi)$, it holds

$$-\partial_t \varphi(t, \phi) - \mathcal{L}\varphi(t, \phi) - f(t, \phi, \varphi(t, \phi), \partial_x \varphi(t, \phi) \sigma(t, \phi), (\varphi(\cdot, \phi))_t) \leq 0.$$

(b) For any $L \geq 0$, we say that u is a viscosity L -supersolution of (1.1) if at any point $(t, \phi) \in [0, T] \times \mathcal{C}$, for any $\varphi \in \overline{\mathcal{A}}^L u(t, \phi)$, we have

$$-\partial_t \varphi(t, \phi) - \mathcal{L}\varphi(t, \phi) - f(t, \phi, \varphi(t, \phi), \partial_x \varphi(t, \phi) \sigma(t, \phi), (\varphi(\cdot, \phi))_t) \geq 0.$$

(c) We say that u is a viscosity subsolution (respectively, supersolution) of (1.1) if u is a viscosity L -subsolution (respectively, L -supersolution) of (1.1) for some $L \geq 0$.

(d) We say that u is a viscosity solution of (1.1) if u is a viscosity subsolution and supersolution of (1.1).

Remark 1.2.3. It is easy to obtain that this definition is equivalent to the classical one in the Markovian framework, see, e.g. [EKT⁺14].

Let us stress that if u is a function from $\mathcal{C}_b^{1,2}([0, T] \times \mathcal{C})$, then it is easy to see that u is a viscosity solution of (1.1) if and only if u is a classical solution for (1.1). Indeed, if u is a viscosity solution then $u \in \underline{\mathcal{A}}^L u(t, \phi) \cap \overline{\mathcal{A}}^L u(t, \phi)$ and therefore u satisfies (1.1). For the reverse statement one can use the nonlinear Feynman-Kac formula proved in this new framework, see Theorem 1.2.17 below, together with functional Itô's formula in order to compute $u(s, X^{t,\phi})$ and the existence and uniqueness result for the corresponding stochastic system (1.2).

Let us also mention that, in accord with standard theory of viscosity solutions, the viscosity property introduced above is a local property, i.e. to check that u is a viscosity solution in (t, ϕ) it is sufficient to know the value of u on the interval $[t, \tau_\epsilon]$, where $\epsilon > 0$ is arbitrarily fixed and $\tau_\epsilon \in \mathcal{T}_+^t$ is given by $\tau_\epsilon := \inf \{s > t : |\phi(s)| \geq \epsilon\} \wedge (t + \epsilon)$.

Eventually, let us notice that since b and σ are Lipschitz, we have uniqueness in law for $X^{t,\phi}$; also, since the filtration on $(\Omega, \mathcal{G}, \mathbb{P})$ is generated by W , every progressively measurable processes λ is a functional of W . Therefore, the spaces of test functions and the above definition are independent on the choice of $(\Omega, \mathcal{G}, \mathbb{P})$ and W .

1.2.2 The forward–backward delayed system

We are now to state the existence and uniqueness results for a delayed forward-backward system, where both the forward and the backward component exhibit a delayed behaviour, that is we will assume that the generator of the backward equation may depend on past values assumed by its solution (Y, Z) . In complete generality, since we will need these results in next sections, we will allow the solution to depend on a general initial time and initial values. Also we remark that in order to ensure the existence and uniqueness result, we need to equip the backward with a suitable condition in the time interval $[0, t)$, being t the initial time, fact that implies a different proof than the one provided in [DI⁺10b].

The main goal is to find a family

$$(X^{t,\phi}, Y^{t,\phi}, Z^{t,\phi})_{(t,\phi) \in [0,T] \times \mathcal{C}},$$

of stochastic processes such that the following decoupled forward–backward system holds

$$\left\{ \begin{array}{l} X^{t,\phi}(s) = \phi(t) + \int_t^s b(r, X^{t,\phi}) dr + \int_t^s \sigma(r, X^{t,\phi}) dW(r), \quad s \in [t, T], \\ X^{t,\phi}(s) = \phi(s), \quad s \in [0, t), \\ Y^{t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), Y_r^{t,\phi}, Z_r^{t,\phi}) dr \\ \quad - \int_s^T Z^{t,\phi}(r) dW(r), \quad s \in [t, T], \\ Y^{t,\phi}(s) = Y^{s,\phi}(s), \quad Z^{t,\phi}(s) = 0, \quad s \in [0, t). \end{array} \right. \quad (1.8)$$

Let us stress once more, that in both the forward and backward equation, the values of $X^{t,\phi}$, resp. $(Y^{t,\phi}, Z^{t,\phi})$, in the time interval $[0, t]$, resp. $[t - \delta, t]$, need to be known; this is one reason we have to impose such initial conditions. The above initial condition for Y is absolutely necessary in view of the Feynman–Kac formula, which will be proven later. We also prolong by convention, $Y^{t,\phi}$ by $Y^{t,\phi}(0)$ on the negative real axis (this is needed in the case that $t < \delta$). For the sake of simplicity, we will take $Z^{t,\phi}(s) := 0$ and $f(s, \cdot, \cdot, \cdot, \cdot, \cdot) := 0$ whenever s become negative.

The forward path-dependent SDE

Let us first focus on the forward component X appearing in the FBDSDE system (1.8), next theorem states the existence and the uniqueness, as well as accurate estimates, for the

process

$$(X^{t,\phi}(r))_{r \in [0,T]}.$$

The existence result is a classical one (see, e.g. [Moh98] or [Moh84]) and the estimates can be obtained by applying Itô's formula together with assumptions (A₁)–(A₂), see, e.g. [Zäl12], and for these reasons we will not state the proof.

In what follows we will assume the following to hold.

Hypothesis 1.2.4. Let us consider two non-anticipative functionals $b : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}^d$ and $\sigma : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}^{d \times d'}$ such that

(A₁) b and σ are continuous;

(A₂) there exists $\ell > 0$ such that for any $t \in [0, T]$, $\phi, \phi' \in \mathcal{C}$,

$$|b(t, \phi) - b(t, \phi')| + |\sigma(t, \phi) - \sigma(t, \phi')| \leq \ell \|\phi - \phi'\|_T.$$

Theorem 1.2.5. *Let b, σ satisfying assumptions 1.2.4 (A₁)–(A₂). Let $(t, \phi), (t', \phi') \in [0, T] \times \mathcal{C}$ be given. Then there exists a unique continuous and adapted stochastic process $(X^{t,\phi}(s))_{s \in [0,T]}$ such that*

$$\begin{cases} X^{t,\phi}(s) = \phi(t) + \int_t^s b(r, X^{t,\phi}) dr + \int_t^s \sigma(r, X^{t,\phi}) dW(r), & s \in [t, T], \\ X^{t,\phi}(s) = \phi(s), & s \in [0, t]. \end{cases} \quad (1.9)$$

Moreover, for any $q \geq 1$, there exists $C = C(q, T, \ell) > 0$ such that

$$\begin{aligned} \mathbb{E}(\|X^{t,\phi}\|_T^{2q}) &\leq C(1 + \|\phi\|_T^{2q}), \\ \mathbb{E}(\|X^{t,\phi} - X^{t',\phi'}\|_T^{2q}) &\leq C\left(\|\phi - \phi'\|_T^{2q} + (1 + \|\phi\|_T^{2q} + \|\phi'\|_T^{2q}) \cdot |t - t'|^q \right. \\ &\quad \left. + \sup_{r \in [t \wedge t', t \vee t']} |\phi(t) - \phi(r)|^{2q}\right), \\ \mathbb{E}\left(\sup_{\substack{s, r \in [t, T] \\ |s-r| \leq \epsilon}} |X^{t,\phi}(s) - X^{t,\phi}(r)|^{2q}\right) &\leq C(1 + \|\phi\|_T^{2q}) \epsilon^{q-1}, \quad \text{for all } \epsilon > 0. \end{aligned} \quad (1.10)$$

The backward delayed SDE

Let us now consider delayed backward SDE appearing in (1.8), in particular in what follows we have d and d' are the fixed constants defined above, whereas $m \in \mathbb{N}^*$ is a new fixed constant. Let us then introduce the main reference spaces we will consider.

Definition 1.2.6. (i) let $\mathcal{H}_t^{2, m \times d'}$ denote the space of \mathbb{G}^t -progressively measurable processes $Z : \Omega \times [t, T] \rightarrow \mathbb{R}^{m \times d'}$ such that

$$\mathbb{E} \left[\int_t^T |Z(s)|^2 ds \right] < \infty;$$

(ii) let $\mathcal{S}_t^{2,m}$ the space of continuous \mathbb{G}^t -progressively measurable processes $Y : \Omega \times [t, T] \rightarrow \mathbb{R}^m$ such that

$$\mathbb{E} \left[\sup_{t \leq s \leq T} |Y(s)|^2 \right] < \infty.$$

Also we will equip the spaces $\mathcal{H}_t^{2,m \times d'}$ and $\mathcal{S}_t^{2,m}$ with the following norms

$$\begin{aligned} \|Z\|_{\mathcal{H}_t^{2,m \times d'}}^2 &= \mathbb{E} \left[\int_t^T e^{\beta s} |Z(s)|^2 ds \right], \\ \|Y\|_{\mathcal{S}_t^{2,m}}^2 &= \mathbb{E} \left[\sup_{t \leq s \leq T} e^{\beta s} |Y(s)|^2 \right], \end{aligned}$$

for a given constant $\beta > 0$.

In what follows, concerning the delayed backward SDE appearing in (1.8), we will assume the following to hold.

Hypothesis 1.2.7. Let

$$f : [0, T] \times \mathcal{C} \times \mathbb{R}^m \times \mathbb{R}^{m \times d'} \times L^2([-\delta, 0]; \mathbb{R}^m) \times L^2([-\delta, 0]; \mathbb{R}^{m \times d'}) \rightarrow \mathbb{R}^m,$$

and

$$h : \mathcal{C} \rightarrow \mathbb{R}^m,$$

such that the following holds:

(A₅) There exist $L, K, M > 0$, $p \geq 1$ and a probability measure α on $([-\delta, 0], \mathcal{B}([-\delta, 0]))$ such that, for any $t \in [0, T]$, $\phi \in \mathcal{C}$, $(y, z), (y', z') \in \mathbb{R}^m \times \mathbb{R}^{m \times d'}$, $\hat{y}, \hat{y}' \in L^2([-\delta, 0]; \mathbb{R}^m)$ and $\hat{z}, \hat{z}' \in L^2([-\delta, 0]; \mathbb{R}^{m \times d'})$, we have

$$\begin{aligned} (i) \quad & \phi \mapsto f(t, \phi, y, z, \hat{y}, \hat{z}) \text{ is continuous,} \\ (ii) \quad & |f(t, \phi, y, z, \hat{y}, \hat{z}) - f(t, \phi, y', z', \hat{y}, \hat{z})| \leq L(|y - y'| + |z - z'|), \\ (iii) \quad & |f(t, \phi, y, z, \hat{y}, \hat{z}) - f(t, \phi, y, z, \hat{y}', \hat{z}')|^2 \\ & \leq K \int_{-\delta}^0 \left(|\hat{y}(\theta) - \hat{y}'(\theta)|^2 + |\hat{z}(\theta) - \hat{z}'(\theta)|^2 \right) \alpha(d\theta), \\ (iv) \quad & |f(t, \phi, 0, 0, 0, 0)| < M(1 + \|\phi\|_T^p). \end{aligned} \tag{1.11}$$

(A₆) The function $f(\cdot, \cdot, y, z, \hat{y}, \hat{z})$ is \mathbb{F} -progressively measurable, for any $(y, z, \hat{y}, \hat{z}) \in \mathbb{R}^m \times \mathbb{R}^{m \times d'} \times L^2([-\delta, 0]; \mathbb{R}^m) \times L^2([-\delta, 0]; \mathbb{R}^{m \times d'})$.

(A₇) The function h is continuous and, for all $\phi \in \mathcal{C}$,

$$|h(\phi)| \leq M(1 + \|\phi\|_T^p).$$

Remark 1.2.8. In order to show the existence and uniqueness of a solution to the backward part of system (1.8), we will use a standard Banach's fixed point argument. For that we are obliged to impose that K or δ should be small enough, see, e.g. restriction (1.30). Also, in

order to obtain the continuity of $Y^{t,\phi}$ with respect to ϕ we are obliged to impose restriction (1.37).

Hence we will assume the following condition to holds, such that both restrictions hold true, in particular we will assume there exists a constant $\gamma \in (0, 1)$ such that

$$K \frac{\gamma e^{\left(\gamma + \frac{6L^2}{\gamma}\right)\delta}}{(1-\gamma)L^2} \max\{1, T\} < \frac{1}{290}. \quad (1.12)$$

For K fixed, it seems at first sight that the expression above cannot be made true by letting δ to 0; however, condition (1.12) will still be verified if we allow L to grow, so we can regard L as a parameter, too.

We are now ready to state the main result of the present section.

Theorem 1.2.9. *Let us assume that assumptions 1.2.7 (A₅)–(A₇) hold true. If K or δ are small enough, that is they satisfy condition (1.12), then there exists a unique solution $(Y^{t,\phi}, Z^{t,\phi})_{(t,\phi) \in [0,T] \times \mathcal{C}}$ for the backward stochastic differential system from (1.8), such that $(Y^{t,\phi}, Z^{t,\phi}) \in \mathcal{S}_t^{2,m} \times \mathcal{H}_t^{2,m \times d'}$, for all $t \in [0, T]$ and $t \mapsto (Y^{t,\phi}, Z^{t,\phi})$ is continuous from $[0, T]$ into $\mathcal{S}_0^{2,m} \times \mathcal{H}_0^{2,m \times d'}$.*

Remark 1.2.10. The main difference between the proof of our result and that of Theorem 2.1 from [DI⁺10b] is due to the supplementary structure condition $Y^{t,\phi}(s) = Y^{s,\phi}(s)$, for $s \in [0, t)$ which should be satisfied by the unknown process $Y^{t,\phi}$.

We also allow T to be arbitrary and we consider that the time horizon is different from the delay $\delta \in [0, T]$; moreover, we separate the Lipschitz constant L with respect to (y, z) by the Lipschitz constant K with respect to \hat{y} , and hence the restriction (1.12) can avoid the constant L .

Proof. The existence and uniqueness will be obtained by the Banach fixed point theorem.

Let $\phi \in \mathcal{C}$ be arbitrarily fixed and let us consider the map Γ defined on $\mathcal{A} \times \mathcal{B}$, with $\mathcal{A} := \mathcal{C}([0, T]; \mathcal{S}_0^{2,m})$ and $\mathcal{B} := \mathcal{C}([0, T]; L^2(\Omega; \mathcal{H}_0^{2,m \times d'}))$, in the following way: for $(U, V) \in \mathcal{A} \times \mathcal{B}$, $\Gamma(U, V) = (Y, Z)$, where for $t \in [0, T]$, the couple of adapted processes (Y^t, Z^t) is solution to the equation

$$\begin{cases} Y^t(s) = h(X^{t,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) dr \\ \quad - \int_s^T Z^t(r) dW(r), \quad s \in [t, T], \\ Y^t(s) := Y^s(s), \quad Z^t(s) := 0, \quad s \in [0, t). \end{cases} \quad (1.13)$$

Step I.

Let us first show that Γ takes values in the Banach space $\mathcal{A} \times \mathcal{B}$. For that, let us take $(U, V) \in \mathcal{A} \times \mathcal{B}$; we will prove that $(Y, Z) := \Gamma(U, V) \in \mathcal{A} \times \mathcal{B}$, i.e. for every $t \in [0, T]$ we have

$$Y^t \in \mathcal{S}_t^{2,m} \subseteq \mathcal{S}_0^{2,m}, \quad Z^t \in \mathcal{H}_t^{2,m \times d'} \subseteq \mathcal{H}_0^{2,m \times d'} \quad (1.14)$$

and the applications

$$\begin{aligned} [0, T] \ni t &\mapsto Y^t \in \mathcal{S}_0^{2,m}, \\ [0, T] \ni t &\mapsto Z^t \in \mathcal{H}_0^{2,m \times d'} \end{aligned} \quad (1.15)$$

are continuous.

Let $t \in [0, T]$ and $t' \in [0, T]$, also, with no loss of generality, we will suppose that $t < t'$ and $t' - t < \delta$.

We have, using (1.13) that

$$\begin{aligned} & \mathbb{E}(\sup_{s \in [0, T]} |Y^t(s) - Y^{t'}(s)|^2) \\ & \leq \mathbb{E}(\sup_{s \in [0, t']} |Y^t(s) - Y^{t'}(s)|^2) + \mathbb{E}(\sup_{s \in [t', T]} |Y^t(s) - Y^{t'}(s)|^2) \\ & \leq 2\mathbb{E}(\sup_{s \in [t, t']} |Y^t(s) - Y^t(t)|^2) + 2\mathbb{E}(\sup_{s \in [t, t']} |Y^t(t) - Y^s(s)|^2) \\ & \quad + \mathbb{E}(\sup_{s \in [t', T]} |Y^t(s) - Y^{t'}(s)|^2). \end{aligned}$$

From the continuity of the solution of equation (1.13) with respect to time, we have that

$$\mathbb{E}(\sup_{s \in [t, t']} |Y^t(s) - Y^t(t)|^2) \rightarrow 0,$$

as $t' \rightarrow t$.

Concerning the term $\mathbb{E}(\sup_{s \in [t', T]} |Y^t(s) - Y^{t'}(s)|^2)$ let us denote for short, only throughout this step,

$$\begin{aligned} \Delta Y(r) &:= Y^t(r) - Y^{t'}(r), & \Delta Z(r) &:= Z^t(r) - Z^{t'}(r) \\ \Delta U(r) &:= U^t(r) - U^{t'}(r), & \Delta V(r) &:= V^t(r) - V^{t'}(r) \end{aligned}$$

and

$$\begin{aligned} \Delta h &:= h(X^{t, \phi}) - h(X^{t', \phi}), \\ \Delta f(r) &:= f(r, X^{t, \phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) - f(r, X^{t', \phi}, Y^{t'}(r), Z^{t'}(r), U_r^{t'}, V_r^{t'}). \end{aligned}$$

Exploiting Itô's formula we have, for any $\beta > 0$ and any $s \in [t, T]$,

$$\begin{aligned} & e^{\beta s} |\Delta Y(s)|^2 + \beta \int_s^T e^{\beta r} |\Delta Y(r)|^2 dr + \int_s^T e^{\beta r} |\Delta Z(r)|^2 dr \\ & = e^{\beta T} |\Delta Y(T)|^2 - 2 \int_s^T e^{\beta r} \langle \Delta Y(r), \Delta Z(r) \rangle dW(r) \\ & \quad + 2 \int_s^T e^{\beta r} \langle \Delta Y(r), f(r, X^{t, \phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) \\ & \quad \quad - f(r, X^{t', \phi}, Y^{t'}(r), Z^{t'}(r), U_r^{t'}, V_r^{t'}) \rangle dr, \end{aligned}$$

so that, from assumptions 1.2.7, and noting that it holds

$$\begin{aligned}
& \int_s^T e^{\beta r} \left(\int_{-\delta}^0 |\Delta U(r+\theta)|^2 + |\Delta V(r+\theta)|^2 \alpha(d\theta) \right) dr \\
&= \int_{-\delta}^0 \int_s^T e^{\beta r} (|\Delta U(r+\theta)|^2 + |\Delta V(r+\theta)|^2) dr \alpha(d\theta) \\
&= \int_{-\delta}^0 \left(\int_{s+\theta}^{T+\theta} e^{\beta(r-\theta)} |\Delta U(r)|^2 + |\Delta V(r)|^2 dr \right) \alpha(d\theta) \\
&\leq e^{\beta\delta} \cdot \int_{-\delta}^0 \alpha(d\theta) \cdot \int_0^T e^{\beta r} (|\Delta U(r)|^2 + |\Delta V(r)|^2) dr \\
&\leq T e^{\beta\delta} \sup_{r \in [0, T]} (e^{\beta r} |\Delta U(r)|^2) + e^{\beta\delta} \int_0^T e^{\beta r} |\Delta V(r)|^2 dr,
\end{aligned}$$

we have for any $a > 0$,

$$\begin{aligned}
& 2 \int_s^T e^{\beta r} \langle \Delta Y(r), f(r, X^{t, \phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) \\
& \quad - f(r, X^{t, \phi'}, Y^{t'}(r), Z^{t'}(r), U_r^{t'}, V_r^{t'}) \rangle dr \\
&\leq a \int_s^T e^{\beta r} |\Delta Y(r)|^2 + \frac{3}{a} \int_s^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{6L^2}{a} \int_s^T e^{\beta r} (|\Delta Y(r)|^2 + |\Delta Z(r)|^2) dr \\
& \quad + \frac{3TK e^{\beta\delta}}{a} \sup_{r \in [0, T]} (e^{\beta r} |\Delta U(r)|^2) + \frac{3K e^{\beta\delta}}{a} \int_0^T e^{\beta r} |\Delta V(r)|^2 dr.
\end{aligned}$$

Therefore we have

$$\begin{aligned}
& e^{\beta s} |\Delta Y(s)|^2 + \left(\beta - a - \frac{6L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Y(r)|^2 dr + \left(1 - \frac{6L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Z(r)|^2 dr \\
&\leq e^{\beta T} |\Delta Y(T)|^2 + \frac{3}{a} \int_s^T e^{\beta r} |\Delta f(r)|^2 dr - 2 \int_s^T e^{\beta r} \langle \Delta Y(r), \Delta Z(r) \rangle dW(r) \\
& \quad + \frac{3TK e^{\beta\delta}}{a} \sup_{r \in [0, T]} e^{\beta r} |\Delta U(r)|^2 + \frac{3K e^{\beta\delta}}{a} \int_0^T e^{\beta r} |\Delta V(r)|^2 dr.
\end{aligned}$$

We now choose $\beta, a > 0$ such that

$$a + \frac{6L^2}{a} < \beta \quad \text{and} \quad \frac{6L^2}{a} < 1. \quad (1.16)$$

so that we obtain

$$\begin{aligned}
& \left(1 - \frac{6L^2}{a} \right) \mathbb{E} \int_s^T e^{\beta r} |\Delta Z(r)|^2 dr \leq \mathbb{E} (e^{\beta T} |\Delta h|^2) + \frac{3}{a} \mathbb{E} \int_s^T e^{\beta r} |\Delta f(r)|^2 dr \\
& \quad + \frac{3TK e^{\beta\delta}}{a} \mathbb{E} \left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U(r)|^2 \right) + \frac{3K e^{\beta\delta}}{a} \mathbb{E} \int_0^T e^{\beta r} |\Delta V(r)|^2 dr.
\end{aligned} \quad (1.17)$$

and, exploiting Burkholder–Davis–Gundy’s inequality, we have that

$$\begin{aligned} & 2\mathbb{E}\left(\sup_{s \in [t', T]} \left| \int_s^T e^{\beta r} \langle \Delta Y(r), \Delta Z(r) \rangle dW(r) \right| \right) \\ & \leq \frac{1}{4}\mathbb{E}\left(\sup_{s \in [t', T]} e^{\beta s} |\Delta Y(s)|^2\right) + 144\mathbb{E}\int_{t'}^T e^{\beta r} |\Delta Z(r)|^2 dr. \end{aligned}$$

which immediately implies

$$\begin{aligned} & \frac{3}{4}\mathbb{E}\left(\sup_{s \in [t', T]} e^{\beta s} |\Delta Y(s)|^2\right) \leq \mathbb{E}(e^{\beta T} |\Delta h|^2) + \frac{3}{a}\mathbb{E}\int_{t'}^T e^{\beta r} |\Delta f(r)|^2 dr \\ & + \frac{3TK e^{\beta \delta}}{a}\mathbb{E}\left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U(r)|^2\right) + \frac{3K e^{\beta \delta}}{a}\mathbb{E}\int_0^T e^{\beta r} |\Delta V(r)|^2 dr \\ & + 144\mathbb{E}\int_{t'}^T |\Delta Z(r)|^2 dr. \end{aligned}$$

Hence, we have that

$$\begin{aligned} & \frac{3}{4}\mathbb{E}\left(\sup_{s \in [t', T]} e^{\beta s} |\Delta Y(s)|^2\right) \leq C_1\mathbb{E}(e^{\beta T} |\Delta h|^2) + \frac{3}{a}C_1\mathbb{E}\int_{t'}^T e^{\beta r} |\Delta f(r)|^2 dr \\ & + \frac{3TK e^{\beta \delta}}{a}\left(1 + \frac{144}{1 - 6L^2/a}\right)\mathbb{E}\left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U(r)|^2\right) \\ & + \frac{3K e^{\beta \delta}}{a}\left(1 + \frac{144}{1 - 6L^2/a}\right)\mathbb{E}\int_0^T e^{\beta r} |\Delta V(r)|^2 dr. \end{aligned} \tag{1.18}$$

Exploiting thus assumptions (A₅) and (A₇) together with the fact that $X^{\cdot\phi}$ is continuous and bounded, we have that

$$C_1\mathbb{E}(e^{\beta T} |\Delta h|^2) + \frac{3}{a}C_1\mathbb{E}\int_{t'}^T e^{\beta r} |\Delta f(r)|^2 dr \rightarrow 0 \quad \text{as } t' \rightarrow t.$$

Since $(U, V) \in \mathcal{A} \times \mathcal{B}$, and therefore we have that

$$\mathbb{E}\left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U(r)|^2\right) \rightarrow 0,$$

and

$$\mathbb{E}\int_0^T e^{\beta r} |\Delta V(r)|^2 dr \rightarrow 0,$$

as $t' \rightarrow t$, we have that

$$\mathbb{E}\left(\sup_{s \in [t', T]} e^{\beta s} |\Delta Y(s)|^2\right) \rightarrow 0 \quad \text{and} \quad \mathbb{E}\int_{t'}^T e^{\beta r} |\Delta Z(r)|^2 dr \rightarrow 0, \quad \text{as } t' \rightarrow t. \tag{1.19}$$

We are now left to show that the term $\mathbb{E}(\sup_{s \in [t, t']} |Y^t(t) - Y^s(s)|^2)$ is also converging to 0 as $t' \rightarrow t$.

Since the map $t \mapsto Y^t(t)$ is deterministic, we have from equation (1.13),

$$\begin{aligned}
Y^t(t) - Y^s(s) &= \mathbb{E}[Y^t(t) - Y^s(s)] \\
&= \mathbb{E}[h(X^{t,\phi}) - h(X^{s,\phi})] + \mathbb{E} \int_t^T f(r, X^{t,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) dr \\
&\quad - \mathbb{E} \int_s^T f(r, X^{s,\phi}, Y^s(r), Z^s(r), U_r^s, V_r^s) dr \\
&= \mathbb{E}[h(X^{t,\phi}) - h(X^{s,\phi})] + \mathbb{E} \int_t^s f(r, X^{t,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) dr \\
&\quad + \mathbb{E} \int_s^T [f(r, X^{t,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) - f(r, X^{s,\phi}, Y^s(r), Z^s(r), U_r^s, V_r^s)] dr.
\end{aligned}$$

Using then assumption (1.11) we have

$$\begin{aligned}
&|Y^{t,\phi}(t) - Y^{s,\phi}(s)| \\
&\leq \mathbb{E}|h(X^{t,\phi}) - h(X^{s,\phi})| + \mathbb{E} \int_t^s L(|Y^t(r)| + |Z^t(r)|) dr \\
&+ \sqrt{K \int_t^s \mathbb{E} \left(\int_{-\delta}^0 |U^t(r+\theta)|^2 + |V^t(r+\theta)|^2 \alpha(d\theta) \right) dr} \cdot \sqrt{s-t} \\
&\quad + \mathbb{E} \int_t^s |f(r, X^{t,\phi}, 0, 0, 0, 0)| dr \\
&+ \mathbb{E} \int_s^T |f(r, X^{t,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) - f(r, X^{s,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t)| dr \\
&+ \mathbb{E} \int_s^T L(|Y^t(r) - Y^s(r)| + |Z^t(r) - Z^s(r)|) dr \\
&+ \sqrt{K(T-s) \int_s^T \mathbb{E} \left(\int_{-\delta}^0 |U^t(r+\theta) - U^s(r+\theta)|^2 + |V^t(r+\theta) - V^s(r+\theta)|^2 \alpha(d\theta) \right) dr}
\end{aligned}$$

and therefore we obtain

$$\begin{aligned}
& |Y^t(t) - Y^s(s)| \\
& \leq \mathbb{E}|h(X^{t,\phi}) - h(X^{s,\phi})| + L\sqrt{s-t}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|Y^t(r)|^2 + \mathbb{E}\int_0^T|Z^t(r)|^2dr} \\
& \quad + \sqrt{K}\sqrt{s-t}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|U^t(r)|^2 + \mathbb{E}\int_0^T|V^t(r)|^2dr} \\
& \quad + (s-t)M(1 + \mathbb{E}\|X^{t,\phi}\|_T^p) \\
& \quad + \mathbb{E}\int_s^T|f(r, X^{t,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t) - f(r, X^{s,\phi}, Y^t(r), Z^t(r), U_r^t, V_r^t)|dr \\
& \quad + L\sqrt{T-s}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|Y^t(r) - Y^s(r)|^2 + \mathbb{E}\int_0^T|Z^t(r) - Z^s(r)|^2dr} \\
& \quad + \sqrt{K}\sqrt{T-s}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|U^t(r) - U^s(r)|^2 + \mathbb{E}\int_0^T|V^t(r) - V^s(r)|^2dr}.
\end{aligned}$$

Taking again into account the fact that $(U, V) \in \mathcal{A} \times \mathcal{B}$, properties (1.10) and assumptions (A_5) and (A_7) , we infer that

$$\mathbb{E}\left(\sup_{s\in[t,t']}|Y^t(t) - Y^s(s)|\right) \rightarrow 0, \quad \text{as } t' \rightarrow t. \quad (1.20)$$

Concerning the term $\mathbb{E}\int_0^T|Z^t(r) - Z^{t'}(r)|^2dr$, we see that

$$\begin{aligned}
& \mathbb{E}\int_0^T|Z^t(r) - Z^{t'}(r)|^2dr \\
& = \mathbb{E}\int_0^{t'}|Z^t(r) - Z^{t'}(r)|^2dr + \mathbb{E}\int_{t'}^T|Z^t(r) - Z^{t'}(r)|^2dr \\
& = \mathbb{E}\int_t^{t'}|Z^t(r)|^2dr + \mathbb{E}\int_{t'}^T|Z^t(r) - Z^{t'}(r)|^2dr,
\end{aligned}$$

hence, by (1.17),

$$\mathbb{E}\int_0^T|Z^t(r) - Z^{t'}(r)|^2dr \rightarrow 0, \quad \text{as } t' \rightarrow t. \quad (1.21)$$

Step II.

We are now to prove that Γ is a contraction on the space $\mathcal{A} \times \mathcal{B}$ with respect to the norms

$$\| \| (Y, Z) \| \|_{\mathcal{A} \times \mathcal{B}} := (\| \| Y \| \|_1^2 + \| \| Z \| \|_2^2)^{1/2},$$

where

$$\| \| Y \| \|_1^2 := \sup_{t\in[0,T]} \mathbb{E}\left(\sup_{r\in[0,T]} e^{\beta r} |Y^t(r)|^2\right)$$

$$\| \| Z \| \|_2^2 := \sup_{t\in[0,T]} \mathbb{E}\int_0^T e^{\beta r} |Z^t(r)|^2dr.$$

Let us recall that $\Gamma : \mathcal{A} \times \mathcal{B} \rightarrow \mathcal{A} \times \mathcal{B}$ is defined by $\Gamma(U, V) = (Y, Z)$, where (Y, Z) is the solution of the BSDE (1.13).

Let us consider $(U^1, V^1), (U^2, V^2) \in \mathcal{A} \times \mathcal{B}$ and $(Y^1, Z^1) := \Gamma(U^1, V^1)$, $(Y^2, Z^2) := \Gamma(U^2, V^2)$. For the sake of brevity, we will denote in what follows

$$\begin{aligned} \Delta f^t(r) &:= f(r, X^{t,\phi}, Y^{1,t}(r), Z^{1,t}(r), U_r^{1,t}, V_r^{1,t}) \\ &\quad - f(r, X^{t,\phi}, Y^{2,t}(r), Z^{2,t}(r), U_r^{2,t}, V_r^{2,t}), \\ \Delta U^t(r) &:= U^{1,t}(r) - U^{2,t}(r), \quad \Delta V^t(r) := V^{1,t}(r) - V^{2,t}(r), \\ \Delta Y^t(r) &:= Y^{1,t}(r) - Y^{2,t}(r), \quad \Delta Z^t(r) := Z^{1,t}(r) - Z^{2,t}(r). \end{aligned}$$

Proceeding as in **Step I**, we have from Itô's formula, for any $s \in [t, T]$ and $\beta > 0$,

$$\begin{aligned} &e^{\beta s} |\Delta^t Y(s)|^2 + \beta \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 dr + \int_s^T e^{\beta r} |\Delta Z^t(r)|^2 dr \\ &= 2 \int_s^T e^{\beta r} \langle \Delta Y^t(r), \Delta f^t(r) \rangle dr - 2 \int_s^T e^{\beta r} \langle \Delta Y^t(r), \Delta Z^t(r) \rangle dW(r). \end{aligned} \tag{1.22}$$

Noticing that it holds

$$\begin{aligned} &\frac{2K}{a} \int_s^T e^{\beta r} \left(\int_{-\delta}^0 (|\Delta U^t(r+\theta)|^2 + |\Delta V^t(r+\theta)|^2) \alpha(d\theta) \right) dr \\ &\leq \frac{2K}{a} \int_{-\delta}^0 \left(\int_s^T e^{\beta r} (|\Delta U^t(r+\theta)|^2 + |\Delta V^t(r+\theta)|^2) dr \right) \alpha(d\theta) \\ &\leq \frac{2K}{a} \int_{-\delta}^0 \left(\int_{s+r}^{T+r} e^{\beta(r'-\theta)} (|\Delta U^t(r')|^2 + |\Delta V^t(r')|^2) dr' \right) \alpha(d\theta) \\ &\leq \frac{2K}{a} \int_{-\delta}^0 e^{-\beta\theta} \alpha(d\theta) \cdot \int_{s-\delta}^T e^{\beta r} (|\Delta U^t(r)|^2 + |\Delta V^t(r)|^2) dr \\ &\leq \frac{2K e^{\beta\delta}}{a} \int_{s-\delta}^T e^{\beta r} (|\Delta U^t(r)|^2 + |\Delta V^t(r)|^2) dr. \end{aligned}$$

we immediately have, from assumptions 1.2.7, that for any $a > 0$,

$$\begin{aligned}
& 2 \left| \int_s^T e^{\beta r} \langle \Delta Y^t(r), \Delta F^t(r) \rangle dr \right| \leq 2 \int_s^T e^{\beta r} |\langle \Delta Y^t(r), \Delta f^t(r) \rangle| dr \\
& \leq a \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 + \frac{1}{a} \int_s^T e^{\beta r} |\Delta f^t(r)|^2 dr \\
& \leq a \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 + \frac{2}{a} \int_s^T e^{\beta r} L^2 (|\Delta Y^t(r)| + |\Delta Z^t(r)|)^2 dr \\
& \quad + \frac{2}{a} \int_s^T e^{\beta r} \left(K \int_{-\delta}^0 (|\Delta U^t(r+\theta)|^2 + |\Delta V^t(r+\theta)|^2) \alpha(d\theta) \right) dr \\
& \leq a \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 + \frac{4L^2}{a} \int_s^T e^{\beta r} (|\Delta Y^t(r)|^2 + |\Delta Z^t(r)|^2) dr \\
& \quad + \frac{2Ke^{\beta\delta}}{a} \int_{s-\delta}^T e^{\beta r} (|\Delta U^t(r)|^2 + |\Delta V^t(r)|^2) dr.
\end{aligned} \tag{1.23}$$

Therefore equation (1.22) yields

$$\begin{aligned}
& e^{\beta s} |\Delta Y^t(s)|^2 + \left(\beta - a - \frac{4L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Y^t(r)|^2 dr + \left(1 - \frac{4L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Z^t(r)|^2 dr \\
& \leq \frac{2Ke^{\beta\delta}}{a} T \sup_{r \in [0, T]} e^{\beta r} |\Delta U^t(r)|^2 + \frac{2Ke^{\beta\delta}}{a} \int_0^T e^{\beta r} |\Delta V^t(r)|^2 dr \\
& \quad - 2 \int_s^T e^{\beta r} \langle \Delta Y^t(r), \Delta Z^t(r) \rangle dW(r).
\end{aligned} \tag{1.24}$$

Let now $\beta, a > 0$ satisfying

$$\beta > a + \frac{4L^2}{a} \quad \text{and} \quad 1 > \frac{4L^2}{a}, \tag{1.25}$$

we have that

$$\begin{aligned}
& \left(1 - \frac{4L^2}{a} \right) \mathbb{E} \int_s^T e^{\beta r} |\Delta Z^t(r)|^2 dr \\
& \leq \frac{2TKe^{\beta\delta}}{a} \mathbb{E} \left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U^t(r)|^2 \right) + \frac{2Ke^{\beta\delta}}{a} \mathbb{E} \int_0^T e^{\beta r} |\Delta V^t(r)|^2 dr.
\end{aligned} \tag{1.26}$$

Exploiting now Burkholder–Davis–Gundy’s inequality, we have that

$$\begin{aligned}
& 2\mathbb{E} \left[\sup_{s \in [t, T]} \left| \int_s^T e^{\beta r} \langle \Delta Y^t(r), \Delta Z^t(r) \rangle dW(r) \right| \right] \\
& \leq 4\mathbb{E} \left[\sup_{s \in [t, T]} \left| \int_t^s e^{\beta r} \langle \Delta Y^t(r), \Delta Z^t(r) \rangle dW(r) \right| \right] \\
& \leq \frac{1}{2} \mathbb{E} \left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y^t(s)|^2 \right) + 72\mathbb{E} \int_t^T e^{\beta r} |\Delta Z^t(r)|^2 dr,
\end{aligned}$$

which implies

$$\begin{aligned}
& \mathbb{E}\left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y^t(s)|^2\right) \\
& \leq \frac{2Ke^{\beta\delta}}{a} T \mathbb{E}\left(\sup_{s \in [0, T]} e^{\beta s} |\Delta U^t(s)|^2\right) + \frac{2Ke^{\beta\delta}}{a} \mathbb{E} \int_0^T e^{\beta r} |\Delta V^t(r)|^2 dr \\
& \quad + 2\mathbb{E}\left[\sup_{s \in [t, T]} \left| \int_s^T e^{\beta r} \langle \Delta Y^t(r), \Delta Z^t(r) \rangle dW(r) \right|\right] \\
& \leq \frac{2Ke^{\beta\delta}}{a} T \mathbb{E}\left(\sup_{s \in [0, T]} e^{\beta s} |\Delta U^t(s)|^2\right) + \frac{2Ke^{\beta\delta}}{a} \mathbb{E} \int_0^T e^{\beta r} |\Delta V^t(r)|^2 dr \\
& \quad + \frac{1}{2} \mathbb{E}\left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y^t(s)|^2\right) + 72 \mathbb{E} \int_t^T e^{\beta r} |\Delta Z^t(r)|^2 dr.
\end{aligned}$$

Hence, we have that

$$\begin{aligned}
& \mathbb{E}\left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y^t(s)|^2\right) \\
& \leq \frac{4TKe^{\beta\delta}}{a} C_1 \mathbb{E}\left(\sup_{s \in [0, T]} e^{\beta s} |\Delta U^t(s)|^2\right) + \frac{4Ke^{\beta\delta}}{a} C_1 \mathbb{E} \int_0^T e^{\beta r} |\Delta V^t(r)|^2 dr,
\end{aligned} \tag{1.27}$$

where we have denoted by $C_1 := 1 + \frac{72}{1-4L^2/a}$.

Let us now consider the term $\mathbb{E}(\sup_{s \in [0, t]} e^{\beta s} |\Delta Y(s)|^2)$. From equation (1.13), we see that,

$$\begin{aligned}
& \mathbb{E}\left(\sup_{s \in [0, t]} e^{\beta s} |\Delta Y^t(s)|^2\right) = \mathbb{E}\left(\sup_{s \in [0, t]} e^{\beta s} |Y^{1,t}(s) - Y^{2,t}(s)|^2\right) \\
& = \mathbb{E}\left(\sup_{s \in [0, t]} e^{\beta s} |Y^{1,s}(s) - Y^{2,s}(s)|^2\right) = \sup_{s \in [0, t]} \mathbb{E}(e^{\beta s} |\Delta Y^s(s)|^2) = \sup_{s \in [0, t]} \mathbb{E}(e^{\beta s} |\Delta Y^s(s)|^2)
\end{aligned} \tag{1.28}$$

so that, exploiting Itô's formula and proceeding as above, we obtain that

$$\begin{aligned}
& \mathbb{E}(e^{\beta s} |\Delta Y^s(s)|^2) \\
& \leq \frac{2TKe^{\beta\delta}}{a} \mathbb{E}\left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U^s(r)|^2\right) + \frac{2Ke^{\beta\delta}}{a} \mathbb{E} \int_0^T e^{\beta r} |\Delta V^s(r)|^2 dr.
\end{aligned} \tag{1.29}$$

Thus from inequalities (1.26–1.29) we obtain

$$\begin{aligned}
& \mathbb{E}\left(\sup_{s \in [0, T]} e^{\beta r} |\Delta Y^t(s)|^2\right) + \mathbb{E} \int_0^T e^{\beta r} |\Delta Z^t(r)|^2 dr \\
& \leq \frac{4TKe^{\beta\delta}}{a} C_1 \mathbb{E}\left(\sup_{s \in [0, T]} e^{\beta s} |\Delta U^t(s)|^2\right) + \frac{4Ke^{\beta\delta}}{a} C_1 \mathbb{E} \int_0^T e^{\beta r} |\Delta V^t(r)|^2 dr \\
& \quad + \frac{2TKe^{\beta\delta}}{a(1-4L^2/a)} \mathbb{E}\left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U^t(r)|^2\right) + \frac{2Ke^{\beta\delta}}{a(1-4L^2/a)} \mathbb{E} \int_0^T e^{\beta r} |\Delta V^t(r)|^2 dr \\
& \quad + \frac{2TKe^{\beta\delta}}{a} \cdot \sup_{s \in [0, t]} \mathbb{E}\left(\sup_{r \in [0, T]} e^{\beta r} |\Delta U^s(r)|^2\right) + \frac{2Ke^{\beta\delta}}{a} \cdot \sup_{s \in [0, t]} \mathbb{E} \int_0^T e^{\beta r} |\Delta V^s(r)|^2 dr.
\end{aligned}$$

Passing then to the supremum for $t \in [0, T]$ we get

$$\begin{aligned} & \| \| Y^1 - Y^2 \| \|_1^2 + \| \| Z^1 - Z^2 \| \|_2^2 \\ & \leq \frac{2Ke^{\beta\delta}}{a} \left(3 + \frac{145}{1 - 4L^2/a} \right) \max\{1, T\} \left[\| \| U^1 - U^2 \| \|_1^2 + \| \| V^1 - V^2 \| \|_2^2 \right]. \end{aligned}$$

By choosing now $a := \frac{4L^2}{\gamma}$ and β slightly bigger than $\gamma + \frac{4L^2}{\gamma}$, condition (1.25) is satisfied and, by (1.12) we have that

$$\frac{2Ke^{\beta\delta}}{a} \left(3 + \frac{145}{1 - 4L^2/a} \right) \max\{1, T\} < 1. \quad (1.30)$$

Eventually, since U and V were chosen arbitrarily, it follows that the application Γ is a contraction on the space $\mathcal{A} \times \mathcal{B}$. Therefore there exists a unique fixed point $\Gamma(Y, Z) = (Y, Z) \in \mathcal{A} \times \mathcal{B}$ and this finishes the proof of the existence and a uniqueness of a solution to equation (1.8). \square

Remark 1.2.11. Using Itô's formula and proceeding as in the proof of Theorem 1.2.9, we can easily show that the solution $(Y^{t,\phi}, Z^{t,\phi})$ to equation (1.8) satisfies the following inequality. For any $q \geq 1$, there exists $C > 0$ such that

$$\begin{aligned} & \mathbb{E} \left(\sup_{r \in [0, T]} |Y^{t,\phi}(r)|^{2q} \right) + \mathbb{E} \left(\int_0^T |Z^{t,\phi}(r)|^2 dr \right)^q \\ & \leq C \left[\mathbb{E} |h(X^{t,\phi})|^{2q} + \mathbb{E} \left(\int_0^T |f(r, X^{t,\phi}, 0, 0, 0, 0)| dr \right)^{2q} \right] \\ & \leq C(1 + \|\phi\|_T^{2q}). \end{aligned} \quad (1.31)$$

1.2.3 Path-dependent PDE – proof of the existence theorem

The current section is devoted to the study of viscosity solution to the path-dependent equation (1.1). In particular, in order to obtain existence of a viscosity solution, we will impose some additional assumptions on the generator f and on the terminal condition h in equation (1.8), in particular we will assume that the generator f depends only on past values assumed by Y and not by past values assumed by the control Z . In what follows we will assume the following to hold.

Hypothesis 1.2.12. Let

$$f : [0, T] \times \mathcal{C} \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{C}([-\delta, 0]; \mathbb{R}) \rightarrow \mathbb{R},$$

and

$$h : \mathcal{C} \rightarrow \mathbb{R},$$

such that the following holds.

(A₈) the functions f and h are continuous; also $f(\cdot, \cdot, y, z, \hat{y})$ is non-anticipative;

(A₉) there exist $L, K, M > 0$ and $p \geq 1$ such that for any $(t, \phi) \in [0, T] \times \mathcal{C}$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$ and $\hat{y}, \hat{y}' \in \mathcal{C}([-\delta, 0]; \mathbb{R})$:

$$(i) \quad |f(t, \phi, y, z, \hat{y}) - f(t, \phi, y', z', \hat{y})| \leq L(|y - y'| + |z - z'|),$$

$$(ii) \quad |f(t, \phi, y, z, \hat{y}) - f(t, \phi, y, z, \hat{y}')|^2 \leq K \int_{-\delta}^0 |\hat{y}(\theta) - \hat{y}'(\theta)|^2 \alpha(d\theta),$$

$$(iii) \quad |f(t, \phi, 0, 0, 0)| \leq M(1 + \|\phi\|_T^p),$$

$$(iv) \quad |h(\phi)| \leq M(1 + \|\phi\|_T^p),$$

being α a probability measure on $([-\delta, 0], \mathcal{B}([-\delta, 0]))$.

Remark 1.2.13. Generators f that satisfy assumptions 1.2.7 are of the following form:

$$f_1(t, \phi, y, z, \hat{y}) := K \int_{-\delta}^0 \hat{y}(s) ds,$$

$$f_2(t, \phi, y, z, \hat{y}) := K \hat{y}(t - \delta).$$

In general, being $g : [0, T] \rightarrow \mathbb{R}$ a measurable and bounded function with $g(t) = 0$ for $t < 0$, the following linear time delayed generator

$$f(t, \phi, y, z, \hat{y}) = \int_{-\delta}^0 g(t + \theta) \hat{y}(\theta) \alpha(d\theta),$$

satisfies assumptions 1.2.12.

Next is the main result of the present paper.

Theorem 1.2.14 (Existence). *Let us assume assumptions 1.2.4-1.2.12 hold. If the delay δ or the Lipschitz constant K are sufficiently small, i.e. condition (1.12) is verified, then the path-dependent PDE (1.1) admits at least one viscosity solution.*

The proof of this result uses in an essential way the nonlinear representation Feynman-Kac type formula, which links the functional SDE (1.9) to a suitable BSDE with time-delayed generators.

Under assumptions (A₁)-(A₂)-(A₈)-(A₉), it follows from Theorem 1.2.9, in the case $m = 1$, that, for each $(t, \phi) \in [0, T] \times \mathcal{C}$, there exists a unique solution $(X^{t, \phi}, Y^{t, \phi}, Z^{t, \phi})$ of \mathbb{G}^t -progressively measurable processes such that $(Y^{t, \phi}, Z^{t, \phi}) \in \mathcal{S}_t^{2,1} \times \mathcal{H}_t^{2,1 \times d'}$, with $Y^{t, \phi}(s) = Y^{s, \phi}(s)$, for any $s \in [0, T]$, solution to the BSDE:

$$Y^{t, \phi}(s) = h(X^{t, \phi}) + \int_s^T f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) dr - \int_s^T Z^{t, \phi}(r) dW(r),$$

for all $s \in [t, T]$.

(1.32)

Let us further observe that the generator f depends on ω only via the the forward process $X^{t, \phi}$.

Before proving Theorem 1.2.14, we need to prove some results. In particular, let us define the function $u : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$ by

$$u(t, \phi) := Y^{t, \phi}(t), \quad (t, \phi) \in [0, T] \times \mathcal{C},$$

(1.33)

notice that $u(t, \phi)$ is a deterministic function since $Y^{t, \phi}(t)$ is $\mathcal{G}_t^t \equiv \mathcal{N}$ -measurable.

Theorem 1.2.15. *Under the assumptions of Theorem 1.2.14, the function u is continuous.*

Proof. Let us first prove the continuity of $\mathcal{C} \ni \phi \mapsto u(t, \phi)$, uniformly with respect to $t \in [0, T]$.

Let us thus take $t \in [0, T]$, $\phi, \phi' \in \mathcal{C}$, and let us denote

$$\Delta Y(r) := Y^{t, \phi}(r) - Y^{t, \phi'}(r), \quad \Delta Z(r) := Z^{t, \phi}(r) - Z^{t, \phi'}(r)$$

and

$$\Delta f(r) := f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) - f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}).$$

By Itô's formula we have, for any $\beta > 0$ and any $s \in [t, T]$,

$$\begin{aligned} e^{\beta s} |\Delta Y(s)|^2 + \beta \int_s^T e^{\beta r} |\Delta Y(r)|^2 dr + \int_s^T e^{\beta r} |\Delta Z(r)|^2 dr &= e^{\beta T} |\Delta Y(T)|^2 \\ + 2 \int_s^T e^{\beta r} \langle \Delta Y(r), f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) & \\ - f(r, X^{t, \phi'}, Y^{t, \phi'}(r), Z^{t, \phi'}(r), Y_r^{t, \phi'}) \rangle dr & \\ - 2 \int_s^T e^{\beta r} \langle \Delta Y(r), \Delta Z(r) \rangle dW(r). & \end{aligned}$$

Exploiting thus assumptions 1.2.4–1.2.12, we have for any $a > 0$,

$$\begin{aligned} 2 \left| \int_s^T e^{\beta r} \langle \Delta Y(r), f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) \right. & \\ \left. - f(r, X^{t, \phi'}, Y^{t, \phi'}(r), Z^{t, \phi'}(r), Y_r^{t, \phi'}) \rangle dr \right| &\leq \\ \leq a \int_s^T e^{\beta r} |\Delta Y(r)|^2 + \frac{3}{a} \int_s^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{3}{a} \int_s^T e^{\beta r} L^2 (|\Delta Y(r)| + |\Delta Z(r)|)^2 dr & \\ + \frac{3K e^{\beta \delta}}{a} \int_{s-\delta}^T e^{\beta r} |\Delta Y(r)|^2 dr, & \end{aligned}$$

so that it holds

$$\begin{aligned} e^{\beta s} |\Delta Y(s)|^2 + \left(\beta - a - \frac{6L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Y(r)|^2 dr + \left(1 - \frac{6L^2}{a} \right) \int_s^T e^{\beta r} |\Delta Z(r)|^2 dr & \\ \leq e^{\beta T} |\Delta Y(T)|^2 + \frac{3}{a} \int_s^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{3K e^{\beta \delta}}{a} \int_{s-\delta}^T e^{\beta r} |\Delta Y(r)|^2 dr & \\ - 2 \int_s^T e^{\beta r} \langle \Delta Y(r), \Delta Z(r) \rangle dW(r). & \end{aligned}$$

Let now $\beta, a > 0$ such that

$$\beta > a + \frac{6L^2}{a} \quad \text{and} \quad 1 > \frac{6L^2}{a}, \quad (1.34)$$

then

$$\begin{aligned} & \left(1 - \frac{6L^2}{a}\right) \mathbb{E} \int_t^T e^{\beta r} |\Delta Z(r)|^2 dr \\ & \leq \mathbb{E}(e^{\beta T} |\Delta Y(T)|^2) + \frac{3}{a} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{3Ke^{\beta\delta}}{a} \mathbb{E} \int_{t-\delta}^T e^{\beta r} |\Delta Y(r)|^2 dr \end{aligned} \quad (1.35)$$

and therefore, exploiting Burkholder–Davis–Gundy’s inequality, we have that

$$\begin{aligned} & 2\mathbb{E} \left[\sup_{s \in [t, T]} \left| \int_s^T e^{\beta r} \langle \Delta Y(r), \Delta Z(r) \rangle dW(r) \right| \right] \\ & \leq \frac{1}{4} \mathbb{E} \left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y(s)|^2 \right) + 144 \mathbb{E} \int_t^T e^{\beta r} |\Delta Z(r)|^2 dr, \end{aligned}$$

which immediately implies

$$\begin{aligned} & \frac{3}{4} \mathbb{E} \left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y(s)|^2 \right) \\ & \leq \mathbb{E}(e^{\beta T} |\Delta Y(T)|^2) + \frac{3}{a} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{3Ke^{\beta\delta}}{a} \mathbb{E} \int_{t-\delta}^T e^{\beta r} |\Delta Y(r)|^2 dr \\ & \quad + 144 \mathbb{E} \int_t^T e^{\beta r} |\Delta Z(r)|^2 dr \\ & \leq \mathbb{E}(e^{\beta T} |\Delta Y(T)|^2) + \frac{3}{a} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{3Ke^{\beta\delta}}{a} \mathbb{E} \int_{t-\delta}^T e^{\beta r} |\Delta Y(r)|^2 dr \\ & \quad + 144 \left(1 - \frac{6L^2}{a}\right)^{-1} \left[\mathbb{E}(e^{\beta T} |\Delta Y(T)|^2) + \frac{3}{a} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr \right] \\ & \quad + 144 \left(1 - \frac{6L^2}{a}\right)^{-1} \left[\frac{3Ke^{\beta\delta}}{a} \mathbb{E} \int_{t-\delta}^T e^{\beta r} |\Delta Y(r)|^2 dr \right]. \end{aligned}$$

Hence,

$$\begin{aligned} & \frac{3}{4} \mathbb{E} \left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y(s)|^2 \right) \\ & \leq C_1 \mathbb{E}(e^{\beta T} |\Delta Y(T)|^2) + \frac{3}{a} C_1 \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{3Ke^{\beta\delta}}{a} C_1 \mathbb{E} \int_{t-\delta}^T e^{\beta r} |\Delta Y(r)|^2 dr, \end{aligned}$$

where we have denoted

$$C_1 := 1 + \frac{144}{1 - 6L^2/a}. \quad (1.36)$$

Choosing now $a := \frac{6L^2}{\gamma}$ and β slightly bigger than $\gamma + \frac{6L^2}{\gamma}$, condition (1.34) is satisfied and

$$T \frac{3Ke^{\beta\delta}}{a} C_1 < \frac{1}{4}, \quad (1.37)$$

by (1.12). We then have

$$\begin{aligned} & \frac{1}{2} \mathbb{E} \left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y(s)|^2 \right) \\ & \leq C_1 \mathbb{E} (e^{\beta T} |\Delta Y(T)|^2) + \frac{3C_1}{a} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr + \frac{3K\delta e^{\beta\delta}}{a} C_1 \mathbb{E} \left(\sup_{s \in [t-\delta, t]} e^{\beta s} |\Delta Y(s)|^2 \right). \end{aligned} \quad (1.38)$$

Exploiting the initial conditions satisfied by the $Y^{t, \phi}$, we can rewrite equation (1.38) as

$$\begin{aligned} & \mathbb{E} \left(\sup_{s \in [t-\delta, t]} e^{\beta s} |\Delta Y(s)|^2 \right) = \mathbb{E} \left(\sup_{s \in [t-\delta, t]} e^{\beta s} |Y^{t, \phi}(s) - Y^{t, \phi'}(s)|^2 \right) \\ & = \sup_{s \in [t-\delta, t]} \mathbb{E} (e^{\beta s} |Y^{s, \phi}(s) - Y^{s, \phi'}(s)|^2) \leq \sup_{s \in [t-\delta, t]} \mathbb{E} \left(\sup_{r \in [s, T]} e^{\beta(s-r)} e^{\beta r} |Y^{s, \phi}(r) - Y^{s, \phi'}(r)|^2 \right), \end{aligned}$$

and therefore we obtain that

$$\begin{aligned} & \mathbb{E} \left(\sup_{s \in [t, T]} e^{\beta s} |\Delta Y(s)|^2 \right) \leq 2C_1 \mathbb{E} (e^{\beta T} |\Delta Y(T)|^2) \\ & \quad + \frac{6C_1}{a} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr \\ & \quad + \frac{6K\delta e^{\beta\delta}}{a} C_1 \sup_{s \in [t-\delta, t]} \mathbb{E} \left(\sup_{r \in [s, T]} e^{\beta(s-r)} e^{\beta r} |Y^{s, \phi}(r) - Y^{s, \phi'}(r)|^2 \right). \end{aligned}$$

Passing to the supremum for $t \in [0, T]$ we have that,

$$\begin{aligned} & \sup_{t \in [0, T]} \mathbb{E} \left(\sup_{s \in [t, T]} e^{\beta s} |Y^{t, \phi}(s) - Y^{t, \phi'}(s)|^2 \right) \\ & \leq 2C_1 \sup_{t \in [0, T]} \mathbb{E} (e^{\beta T} |h(X^{t, \phi}) - h(X^{t, \phi'})|^2) + \frac{6C_1}{a} \sup_{t \in [0, T]} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr \\ & \quad + \frac{6K\delta e^{\beta\delta}}{a} C_1 \sup_{s \in [0, T]} \mathbb{E} \left(\sup_{r \in [s, T]} e^{\beta r} |Y^{s, \phi}(r) - Y^{s, \phi'}(r)|^2 \right). \end{aligned} \quad (1.39)$$

We can now see that

$$\begin{aligned} & \mathbb{E} \left(\sup_{s \in [0, t]} e^{\beta s} |Y^{t, \phi}(s) - Y^{t, \phi'}(s)|^2 \right) = \sup_{s \in [0, t]} \mathbb{E} (e^{\beta s} |Y^{s, \phi}(s) - Y^{s, \phi'}(s)|^2) \\ & \leq \sup_{s \in [0, T]} \mathbb{E} \left(\sup_{r \in [s, T]} e^{\beta r} |Y^{s, \phi}(r) - Y^{s, \phi'}(r)|^2 \right), \end{aligned} \quad (1.40)$$

so that we can apply again inequality (1.39).

From inequalities (1.39) and (1.40) we can conclude that

$$\begin{aligned} & \sup_{t \in [0, T]} \mathbb{E} \left(\sup_{s \in [0, T]} e^{\beta s} |Y^{t, \phi}(s) - Y^{t, \phi'}(s)|^2 \right) \\ & \leq 4C_1 \sup_{t \in [0, T]} \mathbb{E} (e^{\beta T} |h(X^{t, \phi}) - h(X^{t, \phi'})|^2) + \frac{12C_1}{a} \sup_{t \in [0, T]} \mathbb{E} \int_t^T e^{\beta r} |\Delta f(r)|^2 dr \\ & \quad + \frac{12K\delta e^{\beta\delta}}{a} C_1 \sup_{s \in [0, T]} \mathbb{E} \left(\sup_{r \in [s, T]} e^{\beta r} |Y^{s, \phi}(r) - Y^{s, \phi'}(r)|^2 \right). \end{aligned}$$

Since $\delta \leq T$, by (1.37) we also have

$$\frac{12K\delta e^{\beta\delta}}{a}C_1 < 1$$

and so

$$\begin{aligned} & \left(1 - \frac{12K\delta e^{\beta\delta}}{a}C_1\right) \sup_{t \in [0, T]} \mathbb{E} \left(\sup_{s \in [0, T]} e^{\beta s} |Y^{t, \phi}(s) - Y^{t, \phi'}(s)|^2 \right) \\ & \leq 4C_1 \sup_{t \in [0, T]} \mathbb{E} (e^{\beta T} |h(X^{t, \phi}) - h(X^{t, \phi'})|^2) \\ & \quad + \frac{12C_1}{a} \sup_{t \in [0, T]} \mathbb{E} \int_t^T e^{\beta r} |f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) + \\ & \quad \quad \quad - f(r, X^{t, \phi'}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi})|^2 dr. \end{aligned}$$

Let us now fix $\phi \in \mathcal{C}$. In order to prove that u is continuous in ϕ , uniformly with respect to $t \in [0, T]$, it is enough to show that

$$\begin{aligned} & \mathbb{E} (|h(X^{t, \phi}) - h(X^{t, \phi'})|^2) \\ & \quad + \mathbb{E} \int_0^T |f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) - f(r, X^{t, \phi'}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi})|^2 dr \end{aligned}$$

converge to 0 as $\phi' \rightarrow \phi$, uniformly in $t \in [0, T]$.

Since we have no guarantee that the family $\{|Z^{t, \phi}|^2\}_{t \in [0, T]}$ is uniformly integrable, we will use the Lipschitz property of f in the argument (y, z, u) in order to replace $[0, T]$ with a finite subset. By Theorem 1.2.9, the mapping $t \mapsto (Y^{t, \phi}, Z^{t, \phi})$ is continuous from $[0, T]$ into $\mathcal{S}_0^{2,1} \times \mathcal{H}_0^{2,d'}$ and therefore uniformly continuous. Consequently, as $n \rightarrow \infty$, we have that

$$\sup_{|t-t'| \leq \frac{1}{n}} \mathbb{E} \left[\sup_{s \in [0, T]} (Y^{t, \phi}(s) - Y^{t', \phi}(s))^2 + \int_0^T (Z^{t, \phi}(s) - Z^{t', \phi}(s))^2 ds \right] \rightarrow 0.$$

Let, for $n \in \mathbb{N}^*$, $\pi_n := \{0, \frac{T}{n}, \dots, \frac{(n-1)T}{n}, T\}$, then, by (A_g), we see that

$$\begin{aligned} & \sup_{t \in [0, T]} \sup_{t' \in \pi_n} \mathbb{E} \int_0^T |f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) + \\ & \quad \quad \quad - f(r, X^{t', \phi}, Y^{t', \phi}(r), Z^{t', \phi}(r), Y_r^{t', \phi})|^2 dr, \end{aligned}$$

converges to

$$\sup_{t \in [0, T]} \mathbb{E} \int_0^T |f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) - f(r, X^{t, \phi'}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi})|^2 dr,$$

uniformly in ϕ' . We are thus left to prove that

$$\begin{aligned} & \mathbb{E} (|h(X^{t, \phi}) - h(X^{t, \phi'})|^2) \\ & \quad + \mathbb{E} \int_0^T |f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi}) - f(r, X^{t, \phi'}, Y^{t, \phi}(r), Z^{t, \phi}(r), Y_r^{t, \phi})|^2 dr \end{aligned}$$

converge to 0 as $\phi' \rightarrow \phi$, uniformly in $t \in [0, T]$, for fixed $n \in \mathbb{N}^*$ and $t' \in \pi_n$.

Let us thus introduce the modulus of continuity of the functions h and f :

$$\begin{aligned} m_{h,f}(\epsilon, \mathcal{K}, \mathcal{U}, \kappa) &= \\ &:= \sup_{\substack{\phi', \phi'' \in \mathcal{K}, t \in [0, T], (y, u) \in \mathcal{U} \\ |z| \leq \kappa, \|\phi - \phi'\|_T \leq \epsilon}} (|h(\phi') - h(\phi'')| + |f(t, \phi', y, z, u) - f(t, \phi'', y, z, u)|), \end{aligned}$$

where $\epsilon > 0$, \mathcal{K} is a compact in \mathcal{C} , \mathcal{U} is a compact in $\mathbb{R} \times L^2([-\delta, 0]; \mathbb{R})$ and $\kappa \in \mathbb{R}_+$.

Let $\epsilon > 0$ be fixed, but arbitrary; with no loss generality, we can suppose that the function ϕ' lies in a compact $\mathcal{K} \subseteq \mathcal{C}$.

By Theorem 1.2.5, we know that the family $\left\{ (X^{t, \phi'}, Y^{t', \phi}) \right\}_{(t, \phi') \in [0, T] \times \mathcal{K}}$ is tight with respect to the product topology on $\mathcal{C} \times \mathcal{C}([0, T])$, and therefore, for every $\epsilon > 0$, there exist compact subsets $\mathcal{K}_\epsilon \subseteq \mathcal{C}$ and $\mathcal{K}'_\epsilon \subseteq \mathcal{C}([0, T])$ such that

$$\mathbb{P} \left(X^{t, \phi'} \in \mathcal{K}_\epsilon, Y^{t', \phi} \in \mathcal{K}'_\epsilon \right) \geq 1 - \epsilon, \quad \text{for all } (t, \phi') \in [0, T] \times \mathcal{K}.$$

For ease of the notation, let us define $\Phi : \mathcal{C} \times \mathcal{C} \times \mathcal{C}([0, T]) \times \mathbb{R}^d \rightarrow \mathbb{R}$ by

$$\Phi(r, \phi', \phi'', y, z) := \frac{1}{T} |h(\phi') - h(\phi'')|^2 + |f(r, \phi', y(r), z, y_r) - f(r, \phi'', y(r), z, y_r)|^2.$$

We can see by (A₈), that it holds

$$\Phi(\phi', \phi'', y, z) \leq C \left(1 + \|\phi'\|_T^{2p} + \|\phi''\|_T^{2p} + \|y\|_T^2 + |z|^2 \right),$$

where in what follows we will denote by C several possibly different constants depending only on K, L, M and T . Then, for all $t \in [0, T]$, $\phi', \phi'' \in \mathcal{C}$, we have from the a priori estimate (1.31) on the processes $Y^{t, \phi}$ and $Z^{t, \phi}$, we have that,

$$\mathbb{E} \left[\int_0^T \Phi \left(r, X^{t, \phi'}, X^{t, \phi''}, Y^{t', \phi}, Z^{t', \phi}(r) \right) dr \right]^{p'} \leq C \left(1 + \|\phi'\|_T^{pp'} + \|\phi''\|_T^{pp'} \right).$$

Let now \mathcal{U}_ϵ be the image of $[0, T] \times \mathcal{K}'_\epsilon$ through the continuous application

$$(r, y) \mapsto (y(r), y_r),$$

and we also have that, \mathcal{U} is compact in $\mathbb{R} \times L^2([-\delta, 0]; \mathbb{R})$.

For arbitrary $\epsilon', \kappa > 0$, we see that,

$$\begin{aligned}
& \mathbb{E} \int_0^T \Phi \left(r, X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi}(r) \right) dr \\
& \leq \mathbb{E} \int_0^T \Phi \left(X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi} \right) \\
& \quad \cdot \mathbb{1}_{\left\{ \left\{ (X^{t,\phi}, Y^{t',\phi}), (X^{t,\phi'}, Y^{t',\phi}) \in \mathcal{K}_\epsilon \times \mathcal{K}'_\epsilon, |Z^{t',\phi}| \leq \kappa, \|X^{t,\phi} - X^{t,\phi'}\|_T \leq \epsilon' \right\}} \right) dr \\
& \quad + \mathbb{E} \int_0^T \Phi \left(X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi} \right) \mathbb{1}_{\left\{ (X^{t,\phi}, Y^{t',\phi}, Z^{t',\phi}(r)) \notin \mathcal{K}_\epsilon \times \mathcal{K}'_\epsilon \right\}} dr \\
& \quad + \mathbb{E} \int_0^T \Phi \left(X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi} \right) \mathbb{1}_{\left\{ (X^{t,\phi}, Y^{t',\phi}, Z^{t',\phi}(r)) \notin \mathcal{K}_\epsilon \times \mathcal{K}'_\epsilon \right\}} dr \\
& \quad + \mathbb{E} \int_0^T \Phi \left(X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi} \right) \mathbb{1}_{\left\{ |Z^{t',\phi}(r)| > \kappa \right\}} dr \\
& \quad + \mathbb{E} \int_0^T \Phi \left(X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi} \right) \mathbb{1}_{\left\{ \|X^{t,\phi} - X^{t,\phi'}\|_T > \epsilon' \right\}} dr
\end{aligned}$$

and therefore

$$\begin{aligned}
& \mathbb{E} \int_0^T \Phi \left(r, X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi}(r) \right) dr \\
& \leq T m_{h,f}(\epsilon', \mathcal{K}_\epsilon, \mathcal{U}_\epsilon, \kappa) + 2 \left\{ \mathbb{E} \left[\int_0^T \Phi \left(r, X^{t,\phi}, X^{t,\phi'}, Y^{t',\phi}, Z^{t',\phi}(r) \right) dr \right]^{p'} \right\}^{1/p'} \epsilon^{1-\frac{1}{p'}} \\
& \quad + C \mathbb{E} \left[\left(1 + \|X^{t,\phi}\|_T^{2p} + \|X^{t,\phi'}\|_T^{2p} + \|Y^{t',\phi}\|_T^2 \right) \int_0^T \mathbb{1}_{\{|Z^{t',\phi}(r)| > \kappa\}} dr \right] \\
& \quad + C \mathbb{E} \int_0^T |Z^{t',\phi}(r)|^2 \mathbb{1}_{\{|Z^{t',\phi}(r)| > \kappa\}} dr \\
& \quad + \left\{ \mathbb{E} \left[\int_0^T \Phi \left(r, X^{t,\phi'}, X^{t,\phi''}, Y^{t',\phi}, Z^{t',\phi}(r) \right) dr \right]^{p'} \right\}^{1/p'} \\
& \quad \quad \cdot \left[\mathbb{P} \left(\|X^{t,\phi} - X^{t,\phi'}\|_T > \epsilon' \right) \right]^{1-\frac{1}{p'}} \\
& \leq T m_{h,f}(\epsilon', \mathcal{K}_\epsilon, \mathcal{U}, \kappa) \\
& \quad + C \left(1 + \|\phi\|_T^p + \|\phi'\|_T^p \right) \left[\epsilon^{1-\frac{1}{p'}} + \frac{\mathbb{E} \|X^{t,\phi} - X^{t,\phi'}\|_T^{p'-1}}{(\epsilon')^{p'-1}} + \frac{(\mathbb{E} \int_0^T |Z^{t',\phi}(r)|^2 dr)^{1/2}}{\kappa^2} \right] \\
& \quad + C \mathbb{E} \int_0^T |Z^{t',\phi}(r)|^2 \mathbb{1}_{\{|Z^{t',\phi}(r)| > \kappa\}} dr.
\end{aligned}$$

Eventually, by Theorem 1.2.5, we have

$$\begin{aligned}
& \sup_{t \in [0, T]} \mathbb{E} \int_0^T \Phi \left(r, X^{t, \phi}, X^{t, \phi'}, Y^{t', \phi}, Z^{t', \phi}(r) \right) dr \\
& \leq Tm_{h, f}(\epsilon', \mathcal{K}_\epsilon, \mathcal{U}_\epsilon, \kappa) \\
& \quad + C \left(1 + \|\phi\|_T^p + \|\phi'\|_T^p \right) \left[\epsilon^{1 - \frac{1}{p'}} + \frac{\left(1 + \|\phi\|_T^{p-1} + \|\phi'\|_T^{p-1} \right) \mathbb{E} \|\phi - \phi'\|_T^{p'-1}}{(\epsilon')^{p'-1}} + \frac{1}{\kappa^2} \right] \\
& \quad + C \mathbb{E} \int_0^T |Z^{t', \phi}(r)|^2 \mathbb{1}_{\{|Z^{t', \phi}(r)| > \kappa\}} dr.
\end{aligned}$$

Passing now to the limit as $\phi' \rightarrow \phi$, $\epsilon' \rightarrow 0$, $(\epsilon, \kappa) \rightarrow (0, +\infty)$, we obtain the claim.

Concerning the continuity of $[0, T] \ni t \rightarrow u(t, \phi)$, this is an immediate consequence of the continuity of the stochastic process $Y^{t, \phi}$, together with the continuity of the mapping $t \mapsto Y^{t, \phi}$ from $[0, T]$ into $\mathcal{S}_0^{2,1}$. \square

In order to prove the generalized Feynman–Kac formula suitable for our framework we first consider the particular case when the generator f is independent of the past values of Y and Z , namely $(Y^{t, \phi}, Z^{t, \phi})$ is the solution of the standard BSDE with Lipschitz coefficients

$$Y^{t, \phi}(s) = h(X^{t, \phi}) + \int_s^T f(r, X^{t, \phi}, Y^{t, \phi}(r), Z^{t, \phi}(r)) dr - \int_s^T Z^{t, \phi}(r) dW(r), \quad s \in [t, T]. \tag{1.41}$$

Theorem 1.2.16. *Let us assume that assumptions 1.2.4–1.2.12 hold. Then there exists a continuous non-anticipative functional $u : [0, T] \times \mathcal{Z} \rightarrow \mathbb{R}$ such that*

$$Y^{t, \phi}(s) = u(s, X^{t, \phi}), \quad \text{for all } s \in [t, T], \quad \text{a.s.},$$

for every $(t, \phi) \in [0, T] \times \mathcal{C}$.

Proof. Again, for the sake of readability, we split the proof into several steps.

Step I.

Let $0 = t_0 < t_1 < \dots < t_n = T$ and suppose that

$$\begin{aligned}
b(t, \phi) &= b_1(t, \phi(t)) \mathbb{1}_{[[0, t_1]]}(t) + b_2(t, \phi(t_1), \phi(t) - \phi(t_1)) \mathbb{1}_{[[t_1, t_2]]}(t) + \dots \\
& \quad + b_n(t, \phi(t_1), \phi(t_2) - \phi(t_1), \dots, \phi(t) - \phi(t_{n-1})) \mathbb{1}_{[[t_{n-1}, T]]}(t), \\
\sigma(t, \phi) &= \sigma_1(t, \phi(t)) \mathbb{1}_{[[0, t_1]]}(t) + \sigma_2(t, \phi(t_1), \phi(t) - \phi(t_1)) \mathbb{1}_{[[t_1, t_2]]}(t) + \dots \\
& \quad + \sigma_n(t, \phi(t_1), \phi(t_2) - \phi(t_1), \dots, \phi(t) - \phi(t_{n-1})) \mathbb{1}_{[[t_{n-1}, T]]}(t)
\end{aligned}$$

and

$$\begin{aligned}
f(t, \phi, y, z) &= f_1(t, \phi(t), y, z) \mathbb{1}_{[[0, t_1]]}(t) + f_2(t, \phi(t_1), \phi(t) - \phi(t_1), y, z) \mathbb{1}_{[[t_1, t_2]]}(t) + \dots \\
& \quad + f_n(t, \phi(t_1), \phi(t_2) - \phi(t_1), \dots, \phi(t) - \phi(t_{n-1})) \mathbb{1}_{[[t_{n-1}, T]]}(t),
\end{aligned}$$

$$h(\phi) = \varphi(\phi(t_1), \phi(t_2) - \phi(t_1), \dots, \phi(T) - \phi(t_{n-1})),$$

for every $\phi \in \Lambda$.

Let us first show that the terms $(X^{t,\phi}(t_1), \dots, X^{t,\phi}(r) - X^{t,\phi}(t_k))$, $0 \leq k \leq n-1$ are related to the solution of a SDE equation of Itô type in $\mathbb{R}^{n \times d}$. Let

$$\tilde{b}(t, x_1, \dots, x_n) := \begin{pmatrix} b_1(t, x_1) \mathbb{1}_{[[0, t_1]]}(t) \\ b_2(t, x_1, x_2) \mathbb{1}_{[[t_1, t_2]]}(t) \\ \vdots \\ b_n(t, x_1, \dots, x_n) \mathbb{1}_{[[t_{n-1}, t_n]]}(t) \end{pmatrix},$$

$$\tilde{\sigma}(t, x_1, \dots, x_n) := \begin{pmatrix} \sigma_1(t, x_1) \mathbb{1}_{[[0, t_1]]}(t) \\ \sigma_2(t, x_1, x_2) \mathbb{1}_{[[t_1, t_2]]}(t) \\ \vdots \\ \sigma_n(t, x_1, \dots, x_n) \mathbb{1}_{[[t_{n-1}, t_n]]}(t) \end{pmatrix}.$$

Let then $\tilde{X}^{t,\mathbf{x}}$, with $t \in [0, T]$ and $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^{d \times n}$ be the unique solution of the following stochastic differential equation:

$$\tilde{X}^{t,\mathbf{x}}(s) = \mathbf{x} + \int_t^s \tilde{b}(r, \tilde{X}^{t,\mathbf{x}}) dr + \int_t^s \tilde{\sigma}(r, \tilde{X}^{t,\mathbf{x}}) dW(r), \quad s \in [t, T].$$

We assert that, for $t \in [t_{k_0}, t_{k_0+1})$, $s \in [t_k, t_{k+1}]$, $s \geq t$, with $0 \leq k_0 \leq k \leq n-1$, we have

$$(X^{t,\phi}(t_1), \dots, X^{t,\phi}(s) - X^{t,\phi}(t_k)) = \left(\tilde{X}^{i,t,(\phi(t_1), \dots, \phi(t) - \phi(t_{k_0}), 0, \dots, 0)}(s) \right)_{i=\overline{1, k+1}}.$$

Let us stress that for $k=0$ this reads $X^{t,\phi}(s) = \tilde{X}^{1,t,\phi(t)}(s)$; so that for $k > k_0 = 0$, it is interpreted as $(X^{t,\phi}(t_1), \dots, X^{t,\phi}(s) - X^{t,\phi}(t_k)) = \left(\tilde{X}^{i,t,(\phi(t_1), \dots, \phi(t) - \phi(t_{k_0}), 0, \dots, 0)}(s) \right)_{i=\overline{1, k+1}}$.

We will prove this statement by induction on k . If $k=0$, then $k_0=0$ and we obviously have

$$X^{t,\phi}(s) = \tilde{X}^{1,t,\phi(t)}(s).$$

Let us suppose that the statement holds true for $k-1$, for the sake of brevity, in what follows we will denote by $\mathbf{x} := (\phi(t_1), \dots, \phi(t) - \phi(t_{k_0}), 0, \dots, 0)$.

If $k_0 \leq k-1$ then, from the induction hypothesis, we have that

$$(X^{t,\phi}(t_1), \dots, X^{t,\phi}(r) - X^{t,\phi}(t_{k-1})) = \left(\tilde{X}^{i,t,(\phi(t_1), \dots, \phi(t) - \phi(t_{k_0}), 0, \dots, 0)}(r) \right)_{i=\overline{1, k}}$$

for every $r \in [t_{k-1}, t_k]$, so that, for $s \in [t_k, t_{k+1}]$ we have,

$$\tilde{X}^{j,t,\mathbf{x}}(s) = \tilde{X}^{j,t,\mathbf{x}}(t_k) = X^{t,\phi}(t_j) - X^{t,\phi}(t_{j-1}), \quad 1 \leq j \leq k,$$

with the convention $X^{t,\phi}(t_0) = 0$.

In the case $k_0 = k$, for $s \in [t, t_{k+1}]$ we also have:

$$\tilde{X}^{j,t,\mathbf{x}}(s) = \mathbf{x}^j = \phi(t_j) - \phi(t_{j-1}) = X^{t,\phi}(t_j) - X^{t,\phi}(t_{j-1}), \quad 1 \leq j \leq k,$$

again with the convention $\phi(t_0) = 0$.

Consequently, on $[t \vee t_k, t_{k+1}]$, it holds,

$$\begin{aligned} \tilde{X}^{k+1,t,\mathbf{x}}(s) &= \mathbf{x}^{k+1} + \int_{t \vee t_k}^s b_{k+1}(r, \tilde{X}^{1,t,\mathbf{x}}(r), \dots, \tilde{X}^{k+1,t,\mathbf{x}}(r)) dr \\ &\quad + \int_{t \vee t_k}^s \sigma_{k+1}(r, \tilde{X}^{1,t,\mathbf{x}}(r), \dots, \tilde{X}^{k+1,t,\mathbf{x}}(r)) dW(r) \\ &= \mathbf{x}^{k+1} + \int_{t \vee t_k}^s b_{k+1}(r, X^{t,\phi}(t_1), \dots, X^{t,\phi}(t_k) - X^{t,\phi}(t_{k-1}), \tilde{X}^{k+1,t,\mathbf{x}}(r)) dr \\ &\quad + \int_{t \vee t_k}^s \sigma_{k+1}(r, X^{t,\phi}(t_1), \dots, X^{t,\phi}(t_k) - X^{t,\phi}(t_{k-1}), \tilde{X}^{k+1,t,\mathbf{x}}(r)) dW(r). \end{aligned}$$

If $k_0 \leq k-1$ then $\mathbf{x}^{k+1} = 0$; if $k_0 = k$, then $\mathbf{x}^{k+1} = \phi(t) - \phi(t_{k_0}) = X^{t,\phi}(t) - X^{t,\phi}(t_k)$. By uniqueness, since $X^{t,\phi}$ satisfies

$$\begin{aligned} X^{t,\phi}(s) &= X^{t,\phi}(t \vee t_k) + \int_{t \vee t_k}^s b_{k+1}(r, X^{t,\phi}(t_1), \dots, X^{t,\phi}(r) - X^{t,\phi}(t_k)) dr \\ &\quad + \int_{t \vee t_k}^s \sigma_{k+1}(r, X^{t,\phi}(t_1), \dots, X^{t,\phi}(r) - X^{t,\phi}(t_k)) dW(r), \quad s \in [t \vee t_k, t_{k+1}]. \end{aligned}$$

we obtain $\tilde{X}^{k+1,t,\mathbf{x}}(s) = X^{t,\phi}(s) - X^{t,\phi}(t_k)$, for all $s \in [t \vee t_k, t_{k+1}]$. We thus have proved that the statement holds true for k .

The next step is to derive a Feynman–Kac type formula linking $X^{t,\phi}$ and $Y^{t,\phi}$. For that, let us consider the following BSDE:

$$\tilde{Y}^{t,\mathbf{x}}(s) = \varphi(\tilde{X}^{t,\mathbf{x}}(T)) + \int_t^T \tilde{f}(r, \tilde{X}^{t,\mathbf{x}}(r), \tilde{Y}^{t,\mathbf{x}}(r), \tilde{Z}^{t,\mathbf{x}}(r)) dr + \int_t^s \tilde{Z}^{t,\mathbf{x}} dW(r), \quad s \in [t, T],$$

where \tilde{f} is defined by

$$\begin{aligned} \tilde{f}(t, x_1, \dots, x_n, y, z) \\ = f_1(t, x_1, y, z) \mathbb{1}_{[[0, t_1]]}(t) + f_2(t, x_1, x_2, y, z) \mathbb{1}_{[[t_1, t_2]]}(t) + \dots + f_n(t, x_1, x_2, \dots, x_n) \mathbb{1}_{[[t_{n-1}, T]]}(t). \end{aligned}$$

From [EKPQ97], there exist some measurable functions \tilde{u}, \tilde{d} such that for all $(t, \mathbf{x}) \in [0, T] \times \mathbb{R}^{d \times n}$ it holds,

$$\begin{aligned} \tilde{Y}^{t,\mathbf{x}}(s) &= \tilde{u}(s, \tilde{X}^{t,\mathbf{x}}(s)), \quad \text{for all } s \in [t, T]; \\ \tilde{Z}^{t,\mathbf{x}}(s) &= \tilde{d}(s, \tilde{X}^{t,\mathbf{x}}(s)) \tilde{\sigma}(s, \tilde{X}^{t,\mathbf{x}}), \quad ds\text{-a.e. on } [t, T]. \end{aligned}$$

On the other hand, let $t \in [t_{k_0}, t_{k_0+1})$, with $0 \leq k_0 \leq n-1$ and denote, for simplicity, $\mathbf{x} = (\phi(t_1), \dots, \phi(t) - \phi(t_{k_0}), 0, \dots, 0)$.

If $s \in [t_k, t_{k+1}]$, $s \geq t$, and therefore $k \geq k_0$, we have

$$(X^{t,\phi}(t_1), \dots, X^{t,\phi}(s) - X^{t,\phi}(t_k)) = \left(\tilde{X}^{i,t,\mathbf{x}}(s) \right)_{i=\overline{1, k+1}}$$

Thus, on $[t \vee t_k, t_{k+1}]$

$$\begin{aligned} \tilde{f}(s, \tilde{X}^{t,\mathbf{x}}(s), \tilde{Y}^{t,\mathbf{x}}(s), \tilde{Z}^{t,\mathbf{x}}(s)) &= f_{k+1}(s, X^{t,\phi}(t_1), \dots, X^{t,\phi}(s) - X^{t,\phi}(t_k), \tilde{Y}^{t,\mathbf{x}}(s), \tilde{Z}^{t,\mathbf{x}}(s)) \\ &= f(s, X^{t,\phi}, \tilde{Y}^{t,\mathbf{x}}(s), \tilde{Z}^{t,\mathbf{x}}(s)). \end{aligned}$$

Allowing k to vary, we obtain the equality

$$\tilde{f}(s, \tilde{X}^{t,\mathbf{x}}(s), \tilde{Y}^{t,\mathbf{x}}(s), \tilde{Z}^{t,\mathbf{x}}(s)) = f(s, X^{t,\phi}, \tilde{Y}^{t,\mathbf{x}}(s), \tilde{Z}^{t,\mathbf{x}}(s)), \quad \text{for all } s \in [t, T].$$

Also,

$$\varphi(\tilde{X}^{t,\mathbf{x}}(T)) = \varphi(X^{t,\phi}(t_1), \dots, X^{t,\phi}(T) - X^{t,\phi}(t_{n-1})) = h(X^{t,\phi}),$$

and by the uniqueness of the solution of the BSDE, we get that

$$(\tilde{Y}^{t,\mathbf{x}}, \tilde{Z}^{t,\mathbf{x}}) = (Y^{t,\phi}, Z^{t,\phi})$$

and, consequently,

$$\begin{aligned} Y^{t,\phi}(s) &= \tilde{u}(s, X^{t,\phi}(t_1), \dots, X^{t,\phi}(s) - X^{t,\phi}(t_k), 0, \dots, 0), \quad \text{for all } s \in [t \vee t_k, t_{k+1}], \\ Z^{t,\phi}(s) &= \tilde{d}_{k+1}(s, X^{t,\phi}(t_1), \dots, X^{t,\phi}(s) - X^{t,\phi}(t_k), 0, \dots, 0) \\ &\quad \cdot \sigma_{k+1}(s, X^{t,\phi}(t_1), \dots, X^{t,\phi}(s) - X^{t,\phi}(t_k)), \quad ds\text{-a.e. on } [t \vee t_k, t_{k+1}]. \end{aligned}$$

By setting, for $0 \leq k \leq n-1$, $t \in [t_k, t_{k+1}]$ and $\phi \in \Lambda$,

$$\begin{aligned} u(t, \phi) &:= \tilde{u}(t, \phi(t_1), \dots, \phi(t) - \phi(t_k), 0, \dots, 0); \\ d(t, \phi) &:= \tilde{d}_{k+1}(t, \phi(t_1), \dots, \phi(t) - \phi(t_k), 0, \dots, 0), \end{aligned}$$

we get that, for every $t \in [0, T]$ and $\phi \in \Lambda$

$$\begin{aligned} Y^{t,\phi}(s) &= u(s, X^{t,\phi}), \quad \text{for all } s \in [t, T]; \\ Z^{t,\phi}(s) &= d(s, X^{t,\phi})\sigma(t, X^{t,\phi}), \quad ds\text{-a.e. on } [t, T]. \end{aligned}$$

Step II.

Let us notice that the same conclusion holds for b , σ and f of the form

$$\begin{aligned} b(t, \phi) &= b_1(t, \phi(t))\mathbf{1}_{[0, t_1)}(t) + b_2(t, \phi(t_1), \phi(t))\mathbf{1}_{[t_1, t_2)}(t) + \dots \\ &\quad + b_n(t, \phi(t_1), \dots, \phi(t_{n-1}), \phi(t))\mathbf{1}_{[t_{n-1}, T]}(t), \\ \sigma(t, \phi) &= \sigma_1(t, \phi(t))\mathbf{1}_{[0, t_1)}(t) + \sigma_2(t, \phi(t_1), \phi(t))\mathbf{1}_{[t_1, t_2)}(t) + \dots \\ &\quad + \sigma_n(t, \phi(t_1), \dots, \phi(t_{n-1}), \phi(t))\mathbf{1}_{[t_{n-1}, T]}(t) \end{aligned}$$

and

$$\begin{aligned} f(t, \phi, y, z) &= f_1(t, \phi(t), y, z)\mathbf{1}_{[0, t_1)}(t) + f_2(t, \phi(t_1), \phi(t), y, z)\mathbf{1}_{[t_1, t_2)}(t) + \dots \\ &\quad + f_n(t, \phi(t_1), \dots, \phi(t_{n-1}), \phi(t))\mathbf{1}_{[t_{n-1}, T]}(t), \\ h(\phi) &= \varphi(\phi(t_1), \dots, \phi(t_{n-1}), \phi(T)), \end{aligned}$$

for every $\phi \in \Lambda$.

Step III.

For $0 = t_0 \leq t_1 < \dots < t_k \leq T$ and $x_1, \dots, x_k \in \mathbb{R}^d$, let $\Phi_{t_1, \dots, t_k}^{x_1, \dots, x_k} : [0, T] \rightarrow \mathbb{R}^d$ be such that

$$\Phi_{t_1, \dots, t_k}^{x_1, \dots, x_k}(t_i) = x_i, \quad i = \overline{1, k}; \quad \Phi_{t_1, \dots, t_k}^{x_1, \dots, x_k}(T) = x_k, \quad \Phi_{t_1, \dots, t_k}^{x_1, \dots, x_k}(0) = x_1$$

and is prolonged to $[0, T]$ by linear interpolation.

Let us consider partitions of $[0, T]$, $0 = t_0^n < t_1^n < \dots < t_n^n = T$, $t_k^n := \frac{kT}{n}$. For $k \in \{1, \dots, n\}$, $t \in [0, T]$ and $x_1, \dots, x_k \in \mathbb{R}^d$, we define

$$b_k^n(t, x_1, \dots, x_k) := b(t, \Phi_{t_1^n, \dots, t_k^n}^{x_1, \dots, x_k}).$$

Notice that $b_k^n(t, \cdot)$ are continuous functions.

Finally, for $t \in [0, T]$ and $x_1, \dots, x_n \in \mathbb{R}^d$ we set

$$\bar{b}_n(t, x_1, \dots, x_n) := b_1^n(t, x_1) \mathbb{1}_{[0, t_1^n]} + \dots + b_n^n(t, x_1, \dots, x_n) \mathbb{1}_{[t_{n-1}^n, t_n^n]}$$

and, for $(t, \phi) \in [0, T] \times \Lambda$,

$$b_n(t, \phi) := \bar{b}_n(t, \phi(t \wedge t_1^n), \dots, \phi(t \wedge t_n^n)).$$

If $t \in [t_{k-1}^n, t_k^n)$ with $k \in \{1, \dots, n\}$, or if $t = T$ and $k = n$, then

$$b_n(t, \phi) = b_k^n(t, \phi(t_1^n), \dots, \phi(t_{k-1}^n), \phi(t)),$$

so b_n has the form described in **Step II**. Also in this case it holds

$$\begin{aligned} |b_n(t, \phi) - b(t, \phi)| &= |b(t, \Phi_{t_1^n, \dots, t_{k-1}^n, t_k^n}^{\phi(t_1^n), \dots, \phi(t_{k-1}^n), \phi(t)}) - b(t, \phi)| \\ &\leq \sup_{\|\psi - \phi\|_T \leq \omega(\phi, T/n)} |b(t, \psi) - b(t, \phi)| \leq \ell\omega(\phi, T/n), \end{aligned}$$

where $\omega(\phi, \epsilon) := \sup_{|s-r| \leq \epsilon} |\phi(s) - \phi(r)|$.

In a similar way, one introduces σ_n , f_n and h_n and we have, for every $(t, \phi, y, z) \in [0, T] \times \Lambda \times \mathbb{R} \times \mathbb{R}^d$,

$$\begin{aligned} |\sigma_n(t, \phi) - \sigma(t, \phi)| &\leq \ell\omega(\phi, T/n); \\ |f_n(t, \phi, y, z) - f(t, \phi, y, z)| &\leq \sup_{\|\psi - \phi\|_T \leq \omega(\phi, T/n)} |f(t, \psi, y, z) - f(t, \phi, y, z)|; \\ |h_n(\phi) - h(\phi)| &\leq \sup_{\|\psi - \phi\|_T \leq \omega(\phi, T/n)} |h(\psi) - h(\phi)|. \end{aligned}$$

We also have that b_n , σ_n , f_n and h_n satisfy assumptions (A₁)–(A₂)–(A₈)–(A₉) with the same constants. Corresponding to these coefficients, we define the solution of the associated FBSDE,

$$(X^{n,t,\phi}, Y^{n,t,\phi}, Z^{n,t,\phi}).$$

By the Feynman–Kac formula, already proven in this case, we have the existence of some non-anticipative functions u_n and d_n , such that, for every $(t, \phi) \in [0, T] \times \Lambda$, we have:

$$\begin{aligned} Y^{n,t,\phi}(s) &= u_n(s, X^{n,t,\phi}), \quad \text{for all } s \in [t, T]; \\ Z^{n,t,\phi}(s) &= d_n(s, X^{n,t,\phi})\sigma(t, X^{n,t,\phi}), \quad ds\text{-a.e. on } [t, T]. \end{aligned}$$

In order to pass to the limit in the first formula above, we need to show that $X^{n,t,\phi} \rightarrow X^{t,\phi}$ in probability in Λ and that u_n converges to u on compact subsets of Λ .

Let $t \in [0, T]$ and $\phi, \phi' \in \Lambda$. By Itô's formula we have that,

$$\begin{aligned} & |X^{n,t,\phi'}(s) - X^{t,\phi}(s)|^2 = |\phi'(t) - \phi(t)|^2 \\ & + 2 \int_t^s \langle b_n(r, X^{n,t,\phi'}) - b(r, X^{t,\phi}), X^{n,t,\phi'}(r) - X^{t,\phi}(r) \rangle dr \\ & + \int_t^s |\sigma_n(r, X^{n,t,\phi'}) - \sigma(r, X^{t,\phi})|^2 dr \\ & + 2 \int_t^s \langle \sigma_n(r, X^{n,t,\phi'}) - \sigma(r, X^{t,\phi}), (X^{n,t,\phi'}(r) - X^{t,\phi}(r)) dW(r) \rangle. \end{aligned}$$

By standard calculations using Gronwall's lemma, Burkholder-Davis-Gundy inequalities and Theorem 1.2.5, we get, for some arbitrary $p > 1$,

$$\begin{aligned} & \mathbb{E} \sup_{s \in [0, T]} |X^{n,t,\phi'}(s) - X^{t,\phi}(s)|^{2p} \leq C \left(\|\phi' - \phi\|_T^{2p} + \mathbb{E} \omega(X^{t,\phi}, T/n)^{2p} \right) \\ & \leq C \left(\|\phi' - \phi\|_T^{2p} + \frac{1 + \|\phi\|_T^{2p}}{n^{p-1}} + \omega(\phi, T/n)^{2p} \right), \end{aligned} \quad (1.42)$$

where the positive constant C is depending only on p and the parameters of our FBSDE, also as above the constant C is allowed to change value from line to line during this proof.

Applying now Itô's formula to $e^{\beta s} |Y^{n,t,\phi'}(s) - Y^{t,\phi}(s)|^2$, we obtain for $\beta > 0$,

$$\begin{aligned} & e^{\beta s} |\Delta_Y^{n,t,\phi,\phi'}(s)|^2 + \beta \int_s^T e^{\beta r} |\Delta_Z^{n,t,\phi,\phi'}(r)|^2 dr + \int_s^T e^{\beta r} |\Delta_Z^{n,t,\phi,\phi'}(r)|^2 dr \\ & = e^{\beta T} |\Delta_h^{n,t,\phi,\phi'}|^2 + 2 \int_s^T \langle \Delta_f^{n,t,\phi,\phi'}(r), e^{\beta r} \Delta_Y^{n,t,\phi,\phi'}(r) \rangle dr \\ & + 2 \int_s^T \langle f_n(s, X^{n,t,\phi'}, Y^{n,t,\phi'}(s), Z^{n,t,\phi'}(s)) \\ & \quad - f_n(s, X^{n,t,\phi'}, Y^{t,\phi}(s), Z^{t,\phi}(s)), e^{\beta r} \Delta_Y^{n,t,\phi,\phi'}(r) \rangle dr \\ & - 2 \int_s^T \langle e^{\beta r} \Delta_Y^{n,t,\phi,\phi'}(r), \Delta_Z^{n,t,\phi,\phi'}(r) dW(r) \rangle, \end{aligned}$$

where, for the sake of simplicity, we have denoted

$$\begin{aligned} \Delta_Y^{n,t,\phi,\phi'}(s) & := Y^{n,t,\phi'}(s) - Y^{t,\phi}(s); \\ \Delta_Z^{n,t,\phi,\phi'}(s) & := Z^{n,t,\phi'}(s) - Z^{t,\phi}(s); \\ \Delta_h^{n,t,\phi,\phi'} & := h_n(X^{n,t,\phi'}) - h(X^{t,\phi}); \\ \Delta_f^{n,t,\phi,\phi'}(s) & := f_n(s, X^{n,t,\phi'}, Y^{t,\phi}(s), Z^{t,\phi}(s)) - f(s, X^{t,\phi}, Y^{t,\phi}(s), Z^{t,\phi}(s)). \end{aligned}$$

Again, for β sufficiently large, exploiting Burkholder-Davis-Gundy inequalities and the Lipschitz property in $(y, z) \in \mathbb{R} \times \mathbb{R}^{d'}$ of f_n , we infer that

$$\mathbb{E} \sup_{s \in [t, T]} |\Delta_Y^{n,t,\phi,\phi'}(s)|^2 + \mathbb{E} \int_t^T |\Delta_Z^{n,t,\phi,\phi'}(r)|^2 dr \leq C \mathbb{E} \left[|\Delta_h^{n,t,\phi,\phi'}|^2 + \int_t^T |\Delta_f^{n,t,\phi,\phi'}(r)|^2 dr \right]. \quad (1.43)$$

We now have

$$\begin{aligned}
& \left| \Delta_h^{n,t,\phi,\phi'} \right| \leq \left| h_n(X^{n,t,\phi'}) - h(X^{n,t,\phi'}) \right| + \left| h(X^{n,t,\phi'}) - h(X^{t,\phi}) \right| \\
& \leq \sup_{\|\psi - X^{n,t,\phi'}\|_T \leq \omega(X^{n,t,\phi'}, T/n)} \left| h(\psi) - h(X^{n,t,\phi'}) \right| + \left| h(X^{n,t,\phi'}) - h(X^{t,\phi}) \right| \\
& \leq \sup_{\|\psi - X^{t,\phi}\|_T \leq \omega(X^{t,\phi}, T/n) + 3\|X^{n,t,\phi'} - X^{t,\phi}\|_T} \left| h(\psi) - h(X^{t,\phi}) \right| + 2\left| h(X^{n,t,\phi'}) - h(X^{t,\phi}) \right| \\
& \leq 3 \sup_{\|\psi - X^{t,\phi}\|_T \leq \omega(X^{t,\phi}, T/n) + 3\|X^{n,t,\phi'} - X^{t,\phi}\|_T} \left| h(\psi) - h(X^{t,\phi}) \right|,
\end{aligned} \tag{1.44}$$

since

$$\omega(X^{n,t,\phi'}, T/n) \leq \omega(X^{t,\phi}, T/n) + 2\|X^{n,t,\phi'} - X^{t,\phi}\|_T.$$

Similarly, we obtain,

$$\begin{aligned}
\left| \Delta_f^{n,t,\phi,\phi'}(s) \right| & \leq 3 \sup_{\|\psi - X^{t,\phi}\|_T \leq \omega(X^{t,\phi}, T/n) + 3\|X^{n,t,\phi'} - X^{t,\phi}\|_T} \left| f(s, \psi, Y^{t,\phi}(s), Z^{t,\phi}(s)) \right. \\
& \quad \left. - f(s, X^{t,\phi}, Y^{t,\phi}(s), Z^{t,\phi}(s)) \right|.
\end{aligned} \tag{1.45}$$

Let now $(t, \phi) \in [0, T] \times \Lambda$ and let (ϕ_n) be a sequence converging to ϕ in Λ . It is clear from relation (1.42) that (X^{n,t,ϕ_n}) converges in $L^{2p}(\Omega; \Lambda)$ to $X^{t,\phi}$, therefore there exists a subsequence converging a.s. in Λ to $X^{t,\phi}$. Without restricting the generality, we will still denote (X^{n,t,ϕ_n}) this subsequence. Since $\omega(X^{t,\phi}, T/n) + 3\|X^{n,t,\phi'} - X^{t,\phi}\|_T$ converges to 0 a.s., it is clear by relations (1.44) and (1.42) that $\Delta_h^{n,t,\phi,\phi'}$ and $\Delta_f^{n,t,\phi,\phi'}(s)$ converge to 0, a.s., respectively $ds\mathbb{P}$ -a.e.

Then, by the dominated convergence theorem, we obtain

$$\mathbb{E} \left[\left| \Delta_h^{n,t,\phi,\phi_n} \right|^2 + \int_t^T \left| \Delta_f^{n,t,\phi,\phi_n}(r) \right|^2 dr \right] \rightarrow 0,$$

which, combined with estimate (1.43), gives that

$$\mathbb{E} \sup_{s \in [t, T]} |Y^{n,t,\phi_n}(s) - Y^{t,\phi}(s)|^2 \rightarrow 0,$$

which implies, assuming $\phi_n \equiv \phi$

$$\mathbb{E} \sup_{s \in [t, T]} |Y^{n,t,\phi}(s) - Y^{t,\phi}(s)|^2 \rightarrow 0,$$

and, letting $s = t$, we obtain,

$$u_n(t, \phi_n) \rightarrow u(t, \phi).$$

Therefore we can pass to the limit in the relation

$$Y^{n,t,\phi}(s) = u_n(s, X^{n,t,\phi}), \quad \text{for all } s \in [t, T], \text{ a.s.},$$

replacing in the right term ϕ_n by $X^{n,t,\phi}$ and ϕ by $X^{n,t,\phi}$, and find that

$$Y^{t,\phi}(s) = u(s, X^{t,\phi}), \quad \text{for all } s \in [t, T], \text{ a.s.}$$

□

The following result represents the Feynman–Kac formula adapted to our framework.

Theorem 1.2.17. *Let us assume that assumptions 1.2.4–1.2.12 hold and condition (1.12) is verified. Then there exists a continuous non-anticipative functional $u : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$ such that*

$$Y^{t,\phi}(s) = u(s, X^{t,\phi}), \quad \text{for all } s \in [0, T], \quad \text{a.s.,}$$

for any $(t, \phi) \in [0, T] \times \mathcal{C}$.

Proof. The continuity of u was already asserted in Theorem 1.2.15, so that the proof follows with the same arguments.

Let now consider BSDE with delayed generator (1.52)

$$Y^{t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), Y_r^{t,\phi}) dr - \int_s^T Z^{t,\phi}(r) dW(r), \quad (1.46)$$

and the corresponding iterative equations

$$\begin{aligned} Y^{n+1,t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^{n+1,t,\phi}(r), Z^{n+1,t,\phi}(r), Y_r^{n,t,\phi}) dr \\ - \int_s^T Z^{n+1,t,\phi}(r) dW(r), \quad s \in [t, T], \end{aligned} \quad (1.47)$$

with $Y^{0,t,\phi} \equiv 0$ and $Z^{0,t,\phi} \equiv 0$.

Let us suppose that there exists $u_n : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$ a \mathbb{F} -progressively measurable functional such that u_n is continuous and

$$Y^{n,t,\phi}(s) = u_n(s, X^{t,\phi}),$$

for every $t, s \in [0, T]$ and $\phi \in \mathcal{C}$.

Let us now consider the term

$$Y_r^{n,t,\phi} = (Y^{n,t,\phi}(r + \theta))_{\theta \in [-\delta, 0]},$$

in particular we have that, if $r + \theta \geq 0$, then $Y^{n,t,\phi}(r + \theta) = u_n(r + \theta, X^{t,\phi})$ and if $r + \theta < 0$, then $Y^{n,t,\phi}(r + \theta) = Y^{n,t,\phi}(0) = u_n(0, X^{t,\phi})$. By defining then

$$\tilde{u}_n(t, \phi) := (u_n(\mathbb{1}_{[0, T]}(t + \theta), \phi))_{\theta \in [-\delta, 0]},$$

we have

$$Y_r^{n,t,\phi} = \tilde{u}_n(r, X^{t,\phi}).$$

Then we can apply Theorem 1.2.16 in order to infer that

$$Y^{n+1,t,\phi}(s) = u_{n+1}(s, X^{t,\phi}),$$

for a continuous non-anticipative functional $u_{n+1} : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$.

Notice that $(Y^{n,t,\phi}, Z^{n,t,\phi})$ is the Picard iteration sequence used for constructing the solution $(Y^{t,\phi}, Z^{t,\phi})$, recall that

$$(Y^{n+1,\cdot,\phi}, Z^{n+1,\cdot,\phi}) = \Gamma(Y^{n,\cdot,\phi}, Z^{n,\cdot,\phi}),$$

being Γ the contraction defined in the proof of Theorem 1.2.9, we have that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\sup_{s \in [0, T]} |Y^{n, t, \phi}(s) - Y^{t, \phi}(s)|^2 \right) = 0.$$

Of course $u_n(t, \phi)$ converges to

$$u(t, \phi) := \mathbb{E} Y^{t, \phi}(t),$$

for every $t \in [0, T]$ and $\phi \in \Lambda$. This implies that the nonlinear Feynman–Kac formula

$$Y^{t, \phi}(s) = u(s, X^{t, \phi})$$

holds. □

We are now able to prove Theorem 1.2.14, which shows the existence of a viscosity solution to the PDKE (1.1).

Proof of Theorem 1.2.14. We will prove that function u defined by (1.33) is a viscosity solution to (1.1). In particular we will only show that u is a viscosity subsolution, the supersolution case is similar.

Suppose, by contrary, that u is not a viscosity subsolution. Then, for any $L_0 \geq 0$, there exists $L \geq L_0$ such that u is not a viscosity L -subsolution in the sense of Definition 1.2.2. Therefore, there exists $(t, \phi) \in [0, T] \times \mathcal{C}$ and there exists $\varphi \in \underline{\mathcal{A}}^L u(t, \phi)$ such that

$$\partial_t \varphi(t, \phi) + \mathcal{L} \varphi(t, \phi) + f(t, \phi, \varphi(t, \phi), \partial_x \varphi(t, \phi) \sigma(t, \phi), (\varphi(\cdot, \phi))_t) \leq -c < 0$$

for some $c > 0$.

Using the definition of $\underline{\mathcal{A}}^L u(t, \phi)$ we see that there exists $\tau_0 \in \mathcal{T}_+^t$ such that

$$\varphi(t, \phi) - u(t, \phi) = \min_{\tau \in \mathcal{T}^t} \underline{\mathcal{E}}_t^L [(\varphi - u)(\tau \wedge \tau_0, X^{t, \phi})],$$

and by the continuity of the coefficients, we deduce that there exists $\tilde{\tau} \in \mathcal{T}_+^t$ such that

$$\begin{aligned} & \partial_t \varphi(s, X^{t, \phi}) + \mathcal{L} \varphi(s, X^{t, \phi}) + \\ & + f(s, X^{t, \phi}, \varphi(s, X^{t, \phi}), \partial_x \varphi(s, X^{t, \phi}) \sigma(s, X^{t, \phi}), (\varphi(\cdot, X^{t, \phi}))_s) \leq -c/2 < 0, \end{aligned}$$

for any $s \in [t, \tilde{\tau}]$. Let us now take

$$\begin{aligned} \tilde{\tau} = T \wedge \tau_0 \wedge \inf \{ s > t : & \partial_t \varphi(s, X^{t, \phi}) + \mathcal{L} \varphi(s, X^{t, \phi}) \\ & + f(s, X^{t, \phi}, \varphi(s, X^{t, \phi}), \partial_x \varphi(s, X^{t, \phi}) \sigma(s, X^{t, \phi}), (\varphi(\cdot, X^{t, \phi}))_s) > c/2 \}. \end{aligned}$$

Let us denote

$$\begin{aligned} (Y^1(s), Z^1(s)) & := (\varphi(s, X^{t, \phi}), \partial_x \varphi(s, X^{t, \phi}) \sigma(s, X^{t, \phi})), \quad \text{for any } s \in [t, T], \\ (Y^2(s), Z^2(s)) & := (Y^{t, \phi}(s), Z^{t, \phi}(s)), \quad \text{for any } s \in [0, T] \end{aligned}$$

and

$$\Delta Y(s) := Y^1(s) - Y^2(s), \quad \Delta Z(s) := Z^1(s) - Z^2(s), \quad \text{for any } s \in [t, T].$$

Using Itô's formula we deduce, for any $s \in [t, T]$,

$$\varphi(s, X^{t,\phi}) = \varphi(t, \phi) + \int_t^s (\partial_t \varphi(r, X^{t,\phi}) + \mathcal{L}\varphi(r, X^{t,\phi})) dr + \int_t^s \langle \partial_x \varphi(r, X^{t,\phi}), \sigma(r, X^{t,\phi}) dW(r) \rangle$$

so that we obtain

$$\begin{aligned} & \Delta Y(\tilde{\tau}) - \Delta Y(t) \\ &= \int_t^{\tilde{\tau}} (\partial_t \varphi(r, X^{t,\phi}) + \mathcal{L}\varphi(r, X^{t,\phi}) + f(r, X^{t,\phi}, Y^2(r), Z^2(r), Y_r^2)) dr + \int_t^{\tilde{\tau}} \Delta Z(r) dW(r). \end{aligned}$$

Since for any $r \in [t, \tilde{\tau}]$ we have that

$$\begin{aligned} & \partial_t \varphi(r, X^{t,\phi}) + \mathcal{L}\varphi(r, X^{t,\phi}) + \partial_t \varphi(r, X^{t,\phi}) + \mathcal{L}\varphi(r, X^{t,\phi}) + f(r, X^{t,\phi}, Y^2(r), Z^2(r), Y_r^2) \\ & \leq -\frac{c}{2} + f(r, X^{t,\phi}, Y^2(r), Z^2(r), Y_r^2) \\ & \quad - f(r, X^{t,\phi}, \varphi(r, X^{t,\phi}), \partial_x \varphi(r, X^{t,\phi}) \sigma(r, X^{t,\phi}), (\varphi(\cdot, X^{t,\phi}))_r), \end{aligned}$$

we deduce, using Feynman-Kac formula, namely $Y^2(r) = u(r, X^{t,\phi})$, that,

$$\begin{aligned} \Delta Y(\tilde{\tau}) - \Delta Y(t) & \leq -\frac{c}{2} (\tilde{\tau} - t) + \int_t^{\tilde{\tau}} \Delta Z(r) dW(r) \\ & \quad + \int_t^{\tilde{\tau}} (f(r, X^{t,\phi}, \varphi(r, X^{t,\phi}), Z^2(r), (\varphi(\cdot, X^{t,\phi}))_r) \\ & \quad - f(r, X^{t,\phi}, \varphi(r, X^{t,\phi}), Z^1(r), (\varphi(\cdot, X^{t,\phi}))_r)) dr. \end{aligned}$$

We thus have that there exists $\lambda \in \mathcal{U}_t^L$ such that

$$\begin{aligned} & f(r, X^{t,\phi}, \varphi(r, X^{t,\phi}), Z^1(r), (\varphi(\cdot, X^{t,\phi}))_r) - f(r, X^{t,\phi}, \varphi(r, X^{t,\phi}), Z^2(r), (\varphi(\cdot, X^{t,\phi}))_r) \\ & = \langle \lambda(r), \Delta Z(r) \rangle, \end{aligned}$$

and therefore

$$\Delta Y(\tilde{\tau}) - \Delta Y(t) \leq -\frac{c}{2} (\tilde{\tau} - t) + \int_t^{\tilde{\tau}} \Delta Z(r) (dW(r) - \lambda(r) dr).$$

Noticing now that $W(s) - \int_t^s \lambda(r) dr$ is $\mathbb{P}^{t,\lambda}$ -martingale, we obtain that

$$\begin{aligned} & \Delta Y(t) = \mathbb{E}^{\mathbb{P}^{t,\lambda}}(\Delta Y(t)) \\ & \geq \mathbb{E}^{\mathbb{P}^{t,\lambda}}(\Delta Y(\tilde{\tau})) + \frac{c}{2} \mathbb{E}^{\mathbb{P}^{t,\lambda}}(\tilde{\tau} - t) + \mathbb{E}^{\mathbb{P}^{t,\lambda}} \int_t^{\tilde{\tau}} \Delta Z(r) (dW(r) - \lambda(r) dr) \\ & > \mathbb{E}^{\mathbb{P}^{t,\lambda}}(\Delta Y(\tilde{\tau})) = \mathbb{E}^{\mathbb{P}^{t,\lambda}}(\varphi(\tilde{\tau}, X^{t,\phi}) - Y^{t,\phi}(\tilde{\tau})) = \mathbb{E}^{\mathbb{P}^{t,\lambda}}[(\varphi - u)(\tilde{\tau}, X^{t,\phi})] \\ & \geq \underline{\mathcal{E}}_t^L[(\varphi - u)(\tilde{\tau}, X^{t,\phi})] = \underline{\mathcal{E}}_t^L[(\varphi - u)(\tau_0 \wedge \tilde{\tau}, X^{t,\phi})] \\ & \geq \min_{\tau \in \mathcal{T}^t} \underline{\mathcal{E}}_t^L[(\varphi - u)(\tau_0 \wedge \tilde{\tau}, X^{t,\phi})] = \varphi(t, X^{t,\phi}) - Y^{t,\phi}(t), \end{aligned}$$

which is a contradiction and the proof is thus complete. \square

1.3 Financial applications

In what follows we shall apply theoretical results developed in previous sections to analyse some particular models of great interest in modern finance. Financial literature that shows how delay naturally arise when dealing with asset price evolution or in general with certain financial instruments, is nowadays wide and developed, see, for instance, [AHMP07, CY99, KSW05a, KSW07a] and reference therein. On the other hand, not much is done when the delay enters the backward component. We aim here to give some financial applications where also the backward equation exhibits a delayed behaviour. We remark that since the goal of the present work is purely theoretical, the examples provided will not be stated in complete generality. In fact in this section we will show how the study of BSDEs with delayed generator, and the associated path-dependent Kolmogorov equation, may lead to the study of a completely new class of financial problems that have not been studied before. Nevertheless, we intend to address the study of these problems in a complete generality in a future work.

BSDEs with delay have been introduced in [DI⁺10b] as a pure mathematical tool with no financial application of interest. Later on, some works have appeared showing that the delay in the backward component arise naturally in several applications, see, e.g. [Del10, Del12].

In what follows we will provide two extensions of forward-backward models that have been proposed in past literature where the backward components can exhibit a short-time delay.

1.3.1 The large investor problem

Following the model studied in [CM96], see also, e.g. [EKPQ01], we will consider in the present example a non standard investor acting on a financial market. We assume this investor, usually referred to in literature as *the large investor*, has superior information about the stock prices and/or he is willing to invest a large amount of money in the stock. This fact implies that the large investor may influence the behaviour of the stock price with his actions. It is further natural to assume that there is a short time delay between the investor's actions and the reaction of the market to the large investor's actions. In particular we assume that the drift coefficient of the underlying S at time t depends on how the large investor acts on the market in the interval $(t - \delta, t)$.

Let us then consider a risky asset S and a riskless bond B evolving according to

$$\begin{cases} \frac{dB(t)}{B(t)} = r(t, X(t), \pi(t), X_t) dt, & B(0) = 1, \\ \frac{dS(t)}{S(t)} = \mu(t, X(t), \pi(t), X_t) dt + \sigma(t, X(t), X_t) dW(t), & S(0) = s_0 > 0. \end{cases} \quad (1.48)$$

Here we have denoted by X the portfolio of the large investor. Also we used the notations introduced by (1.6) and (1.7). We suppose that the coefficients r, μ and σ satisfy some suitable regularity assumptions.

We have that the portfolio X , composed at any time $t \in [0, T]$ by $\pi(t)$, the amount of shares (held by the large investor) of the risky asset S and by $X(t) - \pi(t)$, the amount shares

of the riskless bond B , evolves according to

$$\begin{aligned} dX(t) &= \frac{\pi(t)}{S(t)} dS(t) + \frac{X(t) - \pi(t)}{B(t)} dB(t) \\ &= \pi(t) \cdot [\mu(t, X(t), \pi(t), X_t) dt + \sigma(t, X(t), X_t) dW(t)] \\ &\quad + [X(t) - \pi(t)] \cdot r(t, X(t), \pi(t), X_t) dt, \end{aligned}$$

with the final condition $X(T) = h(S)$.

Hence, for $t \in [0, T]$,

$$X(t) = h(S) + \int_t^T F(s, X(s), \pi(s), X_s, \pi_s) ds - \int_t^T \pi(s) \sigma(s, X(s), X_s) dW(s), \quad (1.49)$$

where we have denoted for short

$$\begin{aligned} F(s, X(s), \pi(s), X_s, \pi_s) &:= -[X(s) - \pi(s)] \cdot r(s, X(s), \pi(s), X_s) \\ &\quad - \pi(s) \cdot \mu(s, X(s), \pi(s), X_s). \end{aligned} \quad (1.50)$$

Since the forward equations from (1.48) can be solved explicitly by

$$\begin{aligned} S(t) = s_0 \exp \left[\int_0^t \left(\mu(s, X(s), \pi(s), X_s) - \frac{1}{2} \sigma^2(s, X(s), X_s) \right) ds \right. \\ \left. + \int_0^t \sigma(s, X(s), X_s) dW(s) \right], \end{aligned}$$

we deduce that S is a functional of X, π and W , i.e. there exists \tilde{h} such that the final condition becomes $X(T) = \tilde{h}(W, X, \pi)$.

A first remark is that we can impose some suitable assumptions on the functions r, μ and σ such that the function $\bar{F} : [0, T] \times \mathbb{R} \times \mathbb{R} \times L^2([-\delta, 0]; \mathbb{R}) \rightarrow \mathbb{R}$, defined by

$$\bar{F}(s, y, z, \hat{y}) := -(y - z\sigma^{-1}(s, y, \hat{y})) \cdot r(s, y, z\sigma^{-1}(s, y, \hat{y}), \hat{y}) - z \cdot \mu(s, y, z\sigma^{-1}(s, y, \hat{y}), \hat{y}),$$

satisfies assumptions (A₅)–(A₇).

The second remark concerns the fact that Theorem 1.2.9 is still true, with a slight adjustment of the proof, if we consider in the backward equation (1.8) the final condition

$$\bar{h}(W, X, Z) := \tilde{h}(W, X, \sigma^{-1}(\cdot, X, X) Z),$$

instead of a functional of W only, with \bar{h} satisfying a Lipschitz condition:

$$|\bar{h}(x, y, z) - \bar{h}(x, y', z')|^2 \leq L \left[\int_0^T |y(s) - y'(s)|^2 ds + \int_0^T |z(s) - z'(s)|^2 ds \right].$$

Therefore we can rewrite (1.49) as

$$X(t) = \bar{h}(W, X, Z) + \int_t^T \bar{F}(s, X(s), Z(s), X_s) ds - \int_t^T Z(s) dW(s), \quad t \in [0, T] \quad (1.51)$$

and we deduce from Theorem 1.2.9 that, under proper assumptions on the coefficients, there exists a unique solution (X, Z) to equation (1.51).

Hence equation (1.49) admits a unique solution (X, π) , where

$$\pi(s) := Z(s) \sigma^{-1}(s, X(s), X_s).$$

In order to obtain the connection with the PDE we consider first the problem

$$\left\{ \begin{array}{l} \bar{W}^{t,\phi}(s) = \phi(t) + \int_t^s dW(r), \quad s \in [t, T], \\ \bar{W}^{t,\phi}(s) = \phi(s), \quad s \in [0, t], \\ X^{t,\phi}(s) = \bar{h}(\bar{W}^{t,\phi}, X^{t,\phi}, Z^{t,\phi}) + \int_s^T \bar{F}(r, X^{t,\phi}(r), Z^{t,\phi}(r), X_r^{t,\phi}) dr \\ \quad - \int_s^T Z^{t,\phi}(r) dW(r), \quad s \in [t, T], \\ X^{t,\phi}(s) = X^{s,\phi}(s), \quad \pi^{t,\phi}(s) = 0, \quad s \in [0, t]. \end{array} \right. \quad (1.52)$$

Using Theorem 1.2.9 we see that, under suitable assumptions on the coefficients, there exists a unique solution $(X^{t,\phi}, \pi^{t,\phi})_{(t,\phi) \in [0,T] \times \mathcal{C}}$ of the above system.

From the results of the previous sections, in particular from theorem 1.2.17, we have the following representation for the solution (X, Z) of the backward component in system (1.52). In particular we have that, for every $(t, \phi) \in [0, T] \times \mathcal{C}$,

$$X^{t,\phi}(s) = u(s, \bar{W}^{t,\phi}), \quad \text{for all } s \in [t, T],$$

where $u(t, \phi) = X^{t,\phi}(t)$ is a viscosity solution of the following path-dependent PDE:

$$\left\{ \begin{array}{l} -\partial_t u(t, \phi) - \frac{1}{2} \partial_{xx}^2 u(t, \phi) - \bar{F}(t, u(t, \phi), \partial_x u(t, \phi), (u(\cdot, \phi))_t) = 0, \quad \phi \in \mathcal{C}, \quad t \in [0, T], \\ u(T, \phi) = \bar{h}(\phi, (u(\cdot, \phi))), \quad \phi \in \mathcal{C}. \end{array} \right.$$

An example of this type has been developed first in [CM96] and then treated by many authors, see, e.g. [EKPQ01], where they considered a case where the drift and the diffusion component of the price equation depend both on the wealth process X and the underlying process S . This would lead to a fully coupled forward-backward system which does not fit in our setting. On the other hand in our model we have assumed that the drift μ and the interest rate r may be influenced also by past values of the wealth process X , whereas in cited papers no delay in the backward component is assumed.

1.3.2 Risk measures via g-expectations

A key problem in financial mathematics is the risk management of an investment. Such a problem has been widely studied in finance since the introductory paper [ADEH99] where the notion of *risk measure* has been first introduced. Since then, several empirical studies have been done concerning the key task of risk-management, showing in particular that the best way to quantify the risk of a given financial position should be a *dynamic risk measure*, rather than a classic static one. Starting from this fact the notion of *g-expectation* has

been first introduced in [Pen97], as a fundamental mathematical tool if one is to deal with *dynamic risk measure*, we refer also to [Pen04, Gia06] for a comprehensive and exhaustive introduction to *dynamic risk measures*.

The main purpose of a risk measure is to quantify in a single number the riskness of a given financial position. The next one is the mathematical formulation of the notion of *risk measure*, see, e.g. [Del13, Definition 13.1.1].

Definition 1.3.1. A family $(\rho_t)_{t \in [0, T]}$ of mapping $\rho_t : L^2(\Omega, \mathcal{F}_T) \rightarrow L^2(\Omega, \mathcal{F}_t)$, such that $\rho_T(\xi) = -\xi$ is called *dynamic risk measure*.

From a practical point of view, denoting by ξ the terminal value of a given financial position, $\rho_t(\xi)$ quantifies the risk the investor takes in the position ξ at terminal time T . Clearly, in order to have a concrete financial use, a risk measure has to satisfy a set of properties, usually referred to as *axioms of risk measures*, we refer to [Del13] to a complete list of the aforementioned axioms.

From a mathematical point of view, it has been shown that BSDE's, and the related forward-backward system, are a perfect tool to tackle the problem of risk management. In particular, one possible way to define a *dynamic risk measure*, is to specify the generator g of the driving BSDE, from here the name g -expectation, where the generator g determines the properties of the *dynamic risk measure*. A direct approach to g -expectation is therefore to introduce a BSDE of the form

$$Y(t) = \xi + \int_t^T g(s, Y(s), Z(s)) ds - \int_t^T Z(s) dW(s), \quad (1.53)$$

where the generator g is called the generator of a g -expectation. In this sense we have that

$$\rho_t(\xi) = \mathbb{E}^g[\xi | \mathcal{F}_t] = Y(t),$$

where we have used the subscript g to emphasize the role played by the generator g . Heuristically speaking we have the relation

$$\mathbb{E}^g[dY(t) | \mathcal{F}_t] = -g(s, Y(s), Z(s)), \quad 0 \leq t \leq T,$$

so that intuitively the coefficient g reflects the agent's belief on the expected change of risk.

Once we have chosen a risk measure g , such that it is financially reasonable, we then solve the BSDE (1.53) endowed with a suitable final condition which represents the investor's wealth at terminal time T , we refer to [BEK07, Gia06] for a detailed introduction to the usage of BSDE as *dynamic risk measures*.

In literature it has always been considered a generator g that depends only on the present value at time t of the risk measure $Y(t)$ and its variability $Z(t)$, but, as pointed out in [Del12], if we want to model an investor's preferences we cannot leave aside the memory effect, that is it is reasonable to assume that an investor makes his choices based also on what happened on the past. In [Del12] the author proposed to consider a g -expectation which incorporates a disappointment effect through a BSDE of moving average. In fact in the just mentioned work the author suggests that when dealing with investor's preferences, to consider Markovian systems is restrictive since it is natural that an investor takes into account the past history of a given investment when he is to make some choice. We refer to [Del12] for a short but exhaustive review of different economical studies of how memory effect cannot be neglected when dealing with an investor's choice. Regarding the case considered in [Del12], we make

the assumption that the investor has a *short memory*, that is, in making his choices he just consider what has happened in the recent past.

This leads to consider a g -expectation of the form $g(s, y, z) = \beta \bar{y}$, where \bar{y} is the time-average in a sufficiently small time interval and $\beta \in \mathbb{R}$ a given financial parameter, that is we will be dealing with a BSDE with delayed generator of the form

$$Y(t) = \xi + \frac{\beta}{\delta} \int_t^T \int_{-\delta}^0 Y(s+r) dr ds - \int_t^T Z(s) dW(s), \quad (1.54)$$

with ξ the terminal payoff of the investment to be better introduced in a while and δ small enough such that equation (1.12) holds.

Let us now assume that the financial market is composed by one risky asset S and one riskless bond B . The generalization to d risky assets can be easily derived from the present case. We assume, in a complete generality, that both the bond and the asset may exhibit delay. For the case of delay in the forward component a more established theory exists, with existence and uniqueness as well as regularity results, see, e.g. [AHMP07, FT05, FMT10].

We consider in what follows the delayed market model introduced in [CY99]. In what concerns the stock price we assume that S evolves according to the following stochastic delay differential equation:

$$\begin{cases} \frac{dS(t)}{S(t)} = \mu(t, S) dt + \sigma(t, S) dW(t), \\ S_0 = s_0 \in \mathbb{R}, \end{cases} \quad (1.55)$$

where $\mu, \sigma : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$ are some given functions, where the notation is introduced in Section 1.1.

Let us assume that μ and σ satisfy assumptions of type (A₁)–(A₂), so that there exists a unique solution of equation (1.58) satisfying estimates (1.10) (this is a consequence of Theorem 1.2.5).

Also we assume the investor subscribe a claim with terminal payoff $h : \mathcal{C} \rightarrow \mathbb{R}$ so that the BSDE (1.76) becomes

$$Y(t) = h(S_T) + \frac{\beta}{\delta} \int_t^T \int_{-\delta}^0 Y(s+r) dr ds - \int_t^T Z(s) dW(s), \quad (1.56)$$

where we assume h to satisfy assumption 1.2.12 (A₈)–(A₉)(iv), also let us stress that the generator $g : L^2([-\delta, 0]; \mathbb{R}) \rightarrow \mathbb{R}$ defined above satisfies assumptions 1.2.12 (A₈)–(A₉), see, e.g. Remark 1.2.13. We are naturally led to consider the following forward-backward system with delay

$$\begin{cases} S^{t,\phi}(s) = \phi(t) + \int_t^s S^{t,\phi}(r) \mu(t, S^{t,\phi}) dr + \int_t^s S^{t,\phi}(r) \sigma(t, S^{t,\phi}) dW(r), & s \in [t, T], \\ S^{t,\phi}(s) = \phi(s), & s \in [0, t], \\ Y(s) = h(S^{t,\phi}) + \frac{\beta}{\delta} \int_s^T \int_{-\delta}^0 Y^{t,\phi}(r+\theta) d\theta dr - \int_s^T Z^{t,\phi}(r) dW(r), \\ Y^{t,\phi}(s) = Y^{s,\phi}(s), & s \in [0, t], \end{cases} \quad (1.57)$$

and by theorems 1.2.5–1.2.9 we have that the forward-backward system (1.74) admits a unique solution.

Let us also stress that the great majority of possible claim that can be considered in finance satisfies above assumptions on the terminal payoff h , also we allow the option to possibly be *path-dependent*, that is its terminal value at time T depends explicitly on past values assumed by the asset S .

From the results of the previous sections, we thus have the following characterization for the FBSDE (1.74). In particular we obtain that theorem 1.2.17 holds, so that, for every $(t, \phi) \in [0, T] \times \mathcal{C}$,

$$Y^{t, \phi}(s) = u(s, S^{t, \phi}), \quad \text{for all } s \in [t, T],$$

where $u(t, \phi) = Y^{t, \phi}(t)$ is a viscosity solution of the following path-dependent PDE:

$$\begin{cases} \partial_t u(t, \phi) + \frac{1}{2} \sigma^2(t, \phi) \partial_{xx}^2 u(t, \phi) + \phi \mu(t, \phi) \partial_x u(t, \phi) \\ \quad = \frac{\beta}{\delta} \int_{-\delta}^0 u(t+r, \phi) dr, \quad \phi \in \mathcal{C}, \quad t \in [0, T), \\ u(T, \phi) = h(\phi), \quad \phi \in \mathcal{C}. \end{cases}$$

1.3.3 Pricing under counterparty risk

In the present section, which is mainly taken from [CDP16a], we are going to apply previously derived results to the pricing of financial derivatives under counterparty risk and funding issues. In order to derive the pricing equation we closely follow [BFP15, Sec. 2], see also, e.g., [BP13, BPP11] and references therein.

Let us consider a standard filtered probability space $(\Omega, \mathcal{G}, (\mathcal{G}_s)_{s \in [0, T]}, \mathbb{Q})$, being $T < +\infty$ a fixed positive constant, while the filtration \mathcal{G}_s represents all the information available on the market at a given time s . Our goal is to derive a portfolio of financial contracts between two parties, namely the investor, which will be denoted by I , and the *counterparty*, which it will be denoted by C .

In order to work in a realistic and concrete financial framework, we include the risk of default. In particular, we denote by τ_I , resp. τ_C , a \mathcal{G} -measurable stopping time representing the default time of I , resp. the default time of C . Moreover, we prescribe that τ_I , resp. τ_C , has intensity $\lambda_I > 0$, resp. $\lambda_C > 0$, and we indicate by $\tau := \tau_I \wedge \tau_C$, and by $\lambda := \lambda_I + \lambda_C$. Recalling that a *risk-neutral measure* \mathbb{Q} is nothing but a probability measure such that each share price equal the discounted expectation of the share price under \mathbb{Q} itself, in what follows $(\mathcal{F}_s)_{s \in [0, T]}$ indicates the default free filtration generated by the underlying which evolves under the *risk-neutral measure* \mathbb{Q} , and according to the following SDE

$$\begin{cases} \frac{dS(t)}{S(t)} = r_t dt + \sigma(t, S_t) dW(t), \\ S_0 = s_0 > 0 \in \mathbb{R}_+, \end{cases} \quad (1.58)$$

where r is a \mathcal{F} -measurable process indicating the *risk-free rate*. We also assume that there exists a *risk-free account* B whose dynamic is given by

$$\begin{cases} dB(t) = r_t B(t) dt, \\ B_0 = 1, \end{cases} \quad (1.59)$$

where we have used the notation introduced in Section 1.2.3, while the assumptions on the coefficients appearing in equations (1.58)–(1.59) will be specified in a while. We underline that we indicate by $S(t)$ the present \mathbb{R} -value of the process S , whereas S_t denotes the whole path up to time t , so that, in complete generality, we have assumed both the risky asset and the *risk-less rate* to be *path-dependent*.

Remark 1.3.2. Until now, we have worked under the strong assumption that there exists a *risk-free rate* r with a corresponding *risk-free account*. Nevertheless latter assumption turns out to be rather unrealistic in concrete financial markets, and this enlightens a major strength of our approach, since it allows to derive a portfolio that is independent on the *risk-free rate* r . We will treat deeper such a key point in what follows.

Given a rate $\xi(s)$, we will denote the discount factor associated to ξ as

$$D(u, t; \xi) := e^{-\int_u^t \xi_s ds},$$

and we also define $D(u, t) := D(u, t; r)$.

Following [BFP15, Sec. 2.1], we construct a replicating portfolio taking all future cash flows and then discounting them at the *risk-free rate* r . Moreover, to treat the problem in its full generality, we will assume the following processes to be possibly path-dependent, hence stating a difference with respect to what has been done in [BFP15].

In particular we have first to consider the payments due to the contract itself, that is a predictable process π_t and the terminal payoff of the claim $\Phi(S_T)$, so that, at time t , the cumulated discounted flow is given by

$$\mathbb{1}_{\{\tau > T\}} D(0, T) \Phi(S_T) + \int_t^T D(t, s) \pi_s ds. \quad (1.60)$$

We also have a random variable θ_τ representing the cash flow due to the default of one of the parties, hence the resulting cash flow is given by

$$\mathbb{1}_{\{t < \tau < T\}} D(t, \tau) \theta_\tau = \mathbb{1}_{\{t < \tau < T\}} \int_t^T D(t, s) \theta_s d\mathbb{1}_{\{\tau \leq s\}}. \quad (1.61)$$

We consider further a collateral account C_t , namely an asset or some other financial good that a borrower offers to a lender to secure a loan, and we will use the convention that $C_t > 0$, resp. $C_t < 0$, for the investor being the collateral taker, resp. the collateral provider. Moreover, we assume that the collateral account is subjected to an interest rate c_t , which might be different according to what party is the collateral taker, namely we have

$$c_t = \mathbb{1}_{\{C_t > 0\}} c_t^+ + \mathbb{1}_{\{C_t < 0\}} c_t^-. \quad (1.62)$$

Allowing also for rehypotecation, namely allowing banks and brokers to use assets that have been posted as collateral by their clients, we end up with the following cash flow

$$\int_t^T D(t, s) (r_s - c_s) C_s ds. \quad (1.63)$$

The contract will be hedged by a cash position denoted by H_t , resp. by exploiting a position in a risky asset, denoted by F_t . As we have already done, we use the convention that $F_t > 0$,

resp. $F_t < 0$, if the investor is borrowing money, resp. if he is investing money. Again we assume the existence of two different rates

$$f_t = \mathbb{1}_{\{F_t > 0\}} f_t^+ + \mathbb{1}_{\{F_t < 0\}} f_t^-, \quad (1.64)$$

so that the funding component cash flows reads as follow

$$\int_t^\tau D(t, s)(r_s - f_s)F_s ds. \quad (1.65)$$

A similar convention holds for H_t , so that equations (1.64)–(1.65) becomes

$$h_t = \mathbb{1}_{\{H_t > 0\}} h_t^+ + \mathbb{1}_{\{H_t < 0\}} h_t^-, \quad (1.66)$$

and

$$\int_t^\tau D(t, s)(h_s - f_s)H_s ds. \quad (1.67)$$

Eventually, summing up all the aforementioned cash flows (1.60)–(1.61)–(1.63)–(1.65)–(1.67), we have that the value of the portfolio V is given by

$$\begin{aligned} V(t) = & \mathbb{E}_t^{\mathcal{G}} \left[\int_t^\tau D(t, s) (\pi_s + (r_s - c_s)C_s + (r_s - f_s)F_s - (f_s - h_s)H_s) \right] + \\ & + \mathbb{E}_t^{\mathcal{G}} [\mathbb{1}_{\{\tau > T\}} D(t, T)\Phi(S_T) + \mathbb{1}_{\{t < \tau < T\}} D(t, \tau)\theta_\tau], \end{aligned} \quad (1.68)$$

where we have denoted by $\mathbb{E}_t^{\mathcal{G}}[\cdot] := \mathbb{E}[\cdot | \mathcal{G}_t]$. Further, since $V(s) = F_s + H_s + C_s$, we can substitute it in eq. (1.68), obtaining

$$\begin{aligned} V(t) = & \mathbb{E}_t^{\mathcal{G}} \left[\int_t^\tau D(t, s) (\pi_s + (f_s - c_s)C_s + (r_s - f_s)V(s) - (r_s - h_s)H_s) \right] + \\ & + \mathbb{E}_t^{\mathcal{G}} [\mathbb{1}_{\{\tau > T\}} D(t, T)\Phi(S_T) + \mathbb{1}_{\{t < \tau < T\}} D(t, \tau)\theta_\tau]. \end{aligned} \quad (1.69)$$

A concrete example of how equation (1.69) is practically derived, we refer to [BFP15, Sec. 2.2].

Switching to the default free filtration, we can exploit the results stated in [BR13, Sec. 5.1] and [BJPR09, Lemma 3.8.1], see also [BFP15, Lemma 3.1, Lemma 3.3]. In particular, let us recall that \mathcal{G} is the standard filtration, whereas \mathcal{F} denotes the default-free filtration.

Lemma 1.3.3. *For any X \mathcal{G} -measurable random variable and any $t \geq 0$ it holds*

$$\mathbb{E}_t^{\mathcal{G}} [\mathbb{1}_{\{t < \tau \leq s\}} X] = \mathbb{1}_{\{\tau > t\}} \frac{\mathbb{E}_t^{\mathcal{F}} [\mathbb{1}_{\{t < \tau \leq s\}} X]}{\mathbb{E}_t^{\mathcal{F}} [\mathbb{1}_{\{\tau > t\}}]}.$$

In particular we have that for any \mathcal{G}_t -measurable random variable Y , there exists an \mathcal{F}_t -measurable random variable Z such that

$$\mathbb{1}_{\{\tau > t\}} X = \mathbb{1}_{\{\tau > t\}} Z.$$

Lemma 1.3.4. *Let φ_s be a predictable process and τ_I , resp. τ_C , a stopping time with intensity $\lambda_t^I > 0$, resp. $\lambda_t^C > 0$. Assuming that τ_I and τ_C are independent and denoting $\tau := \tau_I \wedge \tau_C$ and $\lambda_t = \lambda_t^I + \lambda_t^C$, then we have*

$$\mathbb{E}_t^{\mathcal{G}} [\mathbb{1}_{\{t < \tau < T\}} \mathbb{1}_{\{\tau_I < \tau_C\}} \varphi_\tau] = \mathbb{1}_{\{\tau > t\}} \mathbb{E}_t^{\mathcal{F}} \left[\int_t^T D(t, s; \lambda) \lambda_t^I \varphi_s ds \right].$$

So that, by an application of Lemmas 1.3.3–1.3.4, and with a slightly abuse of notation, we denote again by $V(t)$ the portfolio evaluated under the default-free filtration \mathcal{F} , namely

$$V(t) = \mathbb{E}_t^{\mathcal{F}} \left[\int_t^T D(t, s; r + \lambda) (\pi_s + (f_s - c_s)C_s + (r_s - f_s)V(s) - (r_s - h_s)H_s) ds \right] + \mathbb{E}_t^{\mathcal{F}} [D(t, T; r + \lambda)\Phi(S_T) + D(t, \tau; r + \lambda)\theta_\tau]. \quad (1.70)$$

By equation (1.70), and proceeding as in [BFP15, Sec. 3], we can immediately obtain the BSDE formulation for the portfolio V , that is we have that V evolves according to

$$\begin{cases} dV(s) = -(\pi_s + (f_s - c_s)C_s - (\lambda_s + f_s)V(s) - (r_s - h_s)H_s) ds - Z(s)dW(s), \\ V(T) = \Phi(S_T). \end{cases} \quad (1.71)$$

The pricing path-dependent PDE

In the previous section we have derived the BSDE that describes the evolution of the financial portfolio, while, in the present section, we are going to better specify the mathematical assumptions regarding equation (1.71). We would like to underline that while some of the following assumptions are mainly taken from [BFP15, Sec. 4], some others are here formalized, to our best knowledge, for the first time.

We assume that the dividend process π depends on time t and on the underlying S , moreover we assume the dependence to be possibly path dependent, namely we have $\pi(t, S_t)$. Further, we assume that π satisfies assumptions 1.2.4.

In what follows, all the rates r, f, c, h, λ are taken to be deterministic and bounded in time and possibly dependent on past values.

We assume that θ has the form

$$\theta_t = \epsilon(t) - \mathbb{1}\{\tau_C < \tau_I\}LGC_C(\epsilon(t) - C_t)^+ + \mathbb{1}\{\tau_I < \tau_C\}LGC_I(\epsilon(t) - C_t)^-, \quad (1.72)$$

where LGD denotes the *loss given default*, commonly defined as the share of an asset that is lost when a borrower defaults, and $(\epsilon(t) - C_t)^+$, resp. $(\epsilon(t) - C_t)^-$, denotes the positive part, resp. the negative part. We will not enter here on financial details regarding θ_t , since it would go beyond the aim of the present work, but we refer the interested reader to [BP13] for a deep treatment of close-out values. We assume then that $\epsilon(t) = V(t)$.

The hedging term H is of the form $H_s = H(s, S_s, V(s), Z(s))$ and it satisfies assumptions 1.2.7, moreover the diffusion term σ , appearing in equation (1.58), satisfies assumptions 1.2.4.

Last but not least, since this is the main novelty of the present approach, we assume that the collateral depends on portfolio past values. As said above this implies that the BSDE is highly irregular, and even the existence and uniqueness of a solution is in general not granted under standard assumptions. However, as pointed out in [Cré11, Rem. 5.5], in practice it often happens that the collateralization scheme is path-dependent in V . We thus assume that the collateral C is of the following particular form

$$C_{(s)} = \alpha_t \bar{V}_{(t)}, \quad \bar{V}_{(t)} := \frac{1}{\delta} \int_{-\delta}^0 V(t+s) ds,$$

where $\alpha_t \in (0, 1]$, that is we assume the collateral to be a fraction of a time average of the portfolio. Notice that in principle also θ in equation (1.72) depends on $\bar{V}_{(t)}$ so that in what

follows we will use the notation $\theta_s(\bar{V}_{(s)})$. Moreover, in order to satisfy condition (1.12), we have that δ is positive and small enough. Alternatively, we can set $\delta = T$, at the cost of assuming the Lipschitz constant in equation (1.58) to satisfies condition (1.12).

We would like to underline that previous choice is just one of the possible collateralization schemes admitted in our mathematical setting, see, e.g., Remark (1.2.13).

The aforementioned assumptions can be better formalized as follows

Hypothesis 1.3.5.

(C1) the parameters r, f, c, h, λ are all bounded elements of \mathcal{C} ;

(C2) $H : [0, T] \times \mathcal{C} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, and there exists positive constants $L_H > 0$ and $M_H > 0$ and $p \geq 1$ such that

$$\begin{aligned} (i) \quad & \phi \mapsto H(t, \phi, y, z) \text{ is continuous,} \\ (ii) \quad & |H(t, \phi, y, z) - H(t, \phi, y', z')| \leq L_H(|y - y'| + |z - z'|), \\ (iii) \quad & |H(t, \phi, 0, 0)| < M_H(1 + \|\phi\|_T^p); \end{aligned} \quad (1.73)$$

(C3) $\Phi : \mathcal{C} \rightarrow \mathbb{R}$ is continuous and for all $\phi \in \mathcal{C}$, there exist $M_\Phi > 0$ and $p \geq 1$, such that

$$|\Phi(\phi)| \leq M_\Phi(1 + \|\phi\|_T^p).$$

(C4) $\sigma : [0, T] \times \mathcal{C} \rightarrow \mathbb{R}$ is continuous and for any $t \in [0, T]$, $\phi, \phi' \in \mathcal{C}$, there exists $\ell_\sigma > 0$ such that

$$|\sigma(t, \phi) - \sigma(t, \phi')| \leq \ell_\sigma \|\phi - \phi'\|_T.$$

In the light of assumptions 1.3.5, let us consider the following forward-backward collateralization scheme

$$\begin{cases} S^{u, \phi}(t) = \phi(t) + \int_u^t r_s ds + \int_u^t \sigma(s, S_s^{u, \phi}) dW(s), & s \in [u, T], \\ S^{u, \phi}(t) = \phi(t), & t \in [0, u], \\ V^{u, \phi}(t) = \Phi(S_T^{u, \phi}) + \int_t^T B\left(s, S_s^{u, \phi}, V^{u, \phi}(s), Z^{u, \phi}(s), V_{(s)}^{u, \phi}\right) ds - \int_t^T Z^{u, \phi}(s) dW(s), \\ V^{u, \phi}(t) = V^{t, \phi}(t), & t \in [0, u], \end{cases} \quad (1.74)$$

with,

$$\begin{aligned} B := & (f_s + \lambda_s) V^{u, \phi}(s) - \pi_s - \theta_s(\bar{V}_{(s)}) + \\ & + (r_s - h_s) H(s, S_s, V(s), Z(s)) - \alpha_s (f_s - c_s) \bar{V}_{(s)}. \end{aligned}$$

Remark 1.3.6. It is worth to mention that in the setting represented by (1.74), we are generalizing the framework considered in [BFP15]. In fact, we have that, besides assuming the collateralization scheme to be dependent on the past values of V , we are also considering both the hedging strategy and the terminal payoff to be possibly *path-dependent*. Moreover, we do not require any differentiability assumption on Φ . The latter being a crucial point when dealing with concrete financial derivatives often characterized by terminal payoffs which are Lipschitz continuous, but fail to be differentiable.

With respect to the forward–backward system (1.74), we have

Theorem 1.3.7. *Let us consider the forward–backward delayed system (1.74) with assumptions 1.3.5. Then for every $(u, \phi) \in [0, T] \times \mathcal{C}$, it holds*

$$V^{u, \phi}(t) = u(t, S^{u, \phi}), \quad \text{for all } t \in [u, T],$$

where $u(t, \phi) = V^{t, \phi}(t)$ is a viscosity solution of the following path–dependent PDE:

$$\begin{cases} \partial_t u(t, \phi) + \frac{1}{2} \sigma^2(t, \phi) \partial_{xx}^2 u(t, \phi) + r_t \partial_x u(t, \phi) \\ \quad = B(t, \phi, u(t, \phi), \partial_x u(t, \phi), \sigma(t, \phi), (u(\cdot, \phi))_t), & \phi \in \mathcal{C}, t \in [0, T), \\ u(T, \phi) = \Phi(\phi), & \phi \in \mathcal{C}, \end{cases} \quad (1.75)$$

with

$$\begin{aligned} B(t, \phi, y, z, \hat{y}) &:= (f_s + \lambda_s)y - \pi_s - \theta_s(\hat{y}) + \\ &\quad + (r_s - h_s)H(s, \phi, y, z) - \alpha_s(f_s - c_s)\hat{y}. \end{aligned}$$

Proof. Because of assumptions 1.3.5, we have that (1.74) satisfies assumptions 1.2.4–1.2.7, hence, exploiting theorems 1.2.5–1.2.9, we have existence and uniqueness of its solution. Thus, by using the results stated in the previous sections, we can derive the characterization of (1.74) given in equation (1.77). In particular, we obtain that Theorem 1.2.17 holds true, and the claim follows. \square

We would like to underline that the scheme described by (1.74) still depends on the non realistic assumption of a *risk-free rate* r . Then, in order to consider a more concrete case, namely without the *risk-free rate*, we will exploit the results stated in [BFP15, Sec. 6]. In particular we will assume the hedging strategy to be a classical *delta-hedging strategy*, namely

$$H(s, S_s, V(s), Z(s)) := S_t \frac{Z(t)}{\sigma(t, S_t)},$$

so that equation (1.74) can be rewritten as follows

$$\begin{cases} S^{u, \phi}(t) = \phi(t) + \int_u^t h_s ds + \int_u^t \sigma(s, S_s^{u, \phi}) dW(s), & s \in [u, T], \\ S^{u, \phi}(t) = \phi(t), & t \in [0, u), \\ V^{u, \phi}(t) = \Phi(S_T^{u, \phi}) + \int_t^T \tilde{B}(s, S_s^{u, \phi}, V^{u, \phi}(s), V_{(s)}^{u, \phi}) ds - \int_t^T Z^{u, \phi}(s) dW(s), \\ V^{u, \phi}(t) = V^{t, \phi}(t), & t \in [0, u), \end{cases} \quad (1.76)$$

with,

$$\tilde{B} := (f_s + \lambda_s)V^{u, \phi}(s) - \pi_s - \theta_s(\bar{V}_{(s)}) - \alpha_s(f_s - c_s)\bar{V}_{(s)}.$$

Thus, we are now dealing with a scheme, i.e. the one represented by system (1.76), which is independent of the *risk-free rate* r , and where the parameter h mimics the role played by the *risk-free rate* r in the classical *Black–Scholes equation*. For the latter setting, the following result holds

Theorem 1.3.8. *Let us consider the forward-backward delayed system (1.76), assuming both assumption 1.3.5 and the existence of positive constants k_σ and K_σ such that, for any $(t, \phi) \in [0, T] \times \mathcal{C}$, the following holds*

$$k_\sigma < \sigma(t, \phi) < K_\sigma.$$

Then for every $(u, \phi) \in [0, T] \times \mathcal{C}$,

$$V^{u, \phi}(t) = u(t, S^{u, \phi}), \quad \text{for all } t \in [u, T],$$

where $u(t, \phi) = V^{t, \phi}(t)$ is a viscosity solution of the following path-dependent PDE:

$$\begin{cases} \partial_t u(t, \phi) + \frac{1}{2} \sigma^2(t, \phi) \partial_{xx}^2 u(t, \phi) + r_t \partial_x u(t, \phi) \\ \quad = \tilde{B}(t, \phi, u(t, \phi), \partial_x u(t, \phi) \sigma(t, \phi), (u(\cdot, \phi))_t), & \phi \in \mathcal{C}, t \in [0, T), \\ u(T, \phi) = \Phi(\phi), & \phi \in \mathcal{C}, \end{cases} \quad (1.77)$$

with

$$\tilde{B}(t, \phi, y, \hat{y}) := (f_s + \lambda_s)y - \pi_s - \theta_s(\hat{y}) - \alpha_s(f_s - c_s)\hat{y}.$$

Proof. The proof is analogous to the one provided for Th. 1.3.7. \square

1.3.4 Conclusion and further development

Inspired by the increasing attention to financial models which take into account credit risk factors, we have generalized results provided in [BFP15, BP13] by exploiting techniques developed to treat backward stochastic differential equations (BSDEs).

In particular, contrary to what happens in [BFP15, BP13], our BSDEs approach allows, for the coefficients in the pricing PDE, to be possibly path-dependent. Moreover, as major novelty of our work, we are also able to treat a collateralization scheme that can depend on the past value of the contract.

We would like to underline that previous approach will be the base of our future works related to the fundamental topic of allowing for close-out rule with delay. Previous situation usually happens when one consider the time-gap between the actual default of a party and the real closure of a contract. In such interval of time it may happen that also the second party could default. Therefore, when one has to price a contract, such a time delay has to be taken into account, see, e.g., [BP13], for a detailed treatment of the topic.

With respect to the aforementioned setting, we believe that both BSDEs techniques and the *path-dependent calculus*, could turn to be useful tools to treat the problem in concrete financial frameworks.

2 | Stochastic calculus for functional delay differential equations with jumps

Sections 2.1–2.2 are taken from [BCDNR16], Section 2.3 is taken from [CDPO16], financial applications exposed in Section 2.4 are taken from [CDP15c] whereas the connection with *path-dependent calculus* developed in Section 2.5 is mainly taken from the appendix in [BCDNR16] and from [CDP15b].

Abstract

Stochastic systems with memory naturally appear in life science, economy, and finance. We take the modelling point of view of stochastic functional delay equations and we study these structures when the driving noises admit jumps. Our results concern existence and uniqueness of strong solutions, estimates for the moments and the fundamental tools of calculus, such as the Itô formula. We study the robustness of the solution to the change of noises. Specifically, we consider the noises with infinite activity jumps versus an adequately corrected Gaussian noise. The study is presented in two different frameworks: we work with random variables in infinite dimensions, where the values are considered either in an appropriate L^p -type space or in the space of càdlàg paths. The choice of the value space is crucial from the modelling point of view as the different settings allow for the treatment of different models of memory or delay. Our techniques involve tools of infinite dimensional calculus and the stochastic calculus via regularisation.

Then we consider a stochastic functional delay differential equation, we lift the problem in the infinite dimensional space of square integrable Lebesgue functions in order to show that its solution is an L^2 -valued Markov process whose uniqueness can be shown under standard assumptions of locally Lipschitzianity and linear growth for the coefficients. Coupling the aforementioned equation with a standard backward differential equation, and deriving some ad hoc results concerning the Malliavin derivative for systems with memory, we are able to derive a *non-linear Feynman-Kac* representation theorem under mild assumptions

of differentiability.

We also consider a particular type of infinite dimensional PIDE which can be used to price a rather general class of stochastic volatility models with jumps arising in finance, deriving results for continuous-type *Asian options*. Our approach relies on recent results obtained by Yang Feng, Salah-Eldin A. Mohammed, and ourselves, for *stochastic functional delay differential equations* (SFDDEs), and on a non-linear *Feynman-Kac theorem* for SFDDEs linked to a forward-backward system with delay. Under mild assumptions of regularity, which is connected to an infinite dimensional PIDE whose solution is only required to be Lipschitz, via the mild notion of gradient first introduced by Marco Fuhrman and Gianmario Tessitore in 2005, we are able to treat the previously cited financial model by deriving an appropriate set of infinite dimensional (pricing) PIDE.

Eventually we provide existence and uniqueness result as well as a *Feynman-Kac*-type representation theorem for a stochastic coupled Delayed Forward Backward System (DFBS) with values in a suitable space of càdlàg functions, hence showing that the aforementioned solution can be retrieved studying the related *path-dependent partial integro-differential equation* (PPIDE). A link between the proposed approach and the one obtained by the calculus via regularization is also shown.

2.1 Stochastic systems with memory and jumps

Delay equations are differential equations whose coefficients depend also on the past history of the solution. Besides being of mathematical interest on their own, delay equations naturally arise in many applications, ranging from mathematical biology to mathematical finance, where often the effect of the memory or delay on the evolution of the system cannot be neglected, we refer to [CY99, CY07, KSW05b, KSW07b, Kua93, KP07, Moh98] and references therein for applications in different areas.

When dealing with a delay differential equation (DDE), one cannot in general rely on standard existence and uniqueness theorems, but ad hoc results have to be proven. In general this is done by lifting the DDE, from having a solution with values in a finite dimensional state space, such as \mathbb{R}^d , to having values in an infinite dimensional path space, which has to be carefully chosen according to the specific problem. For the case of *deterministic* delay differential equations an extensive literature exists, we refer the reader to the monographs [DVGLW95, EN00a] for details.

When considering *stochastic* delay differential equations (SDDE), that is DDE perturbed by a stochastic noise, one encounters problems that did not appear in the deterministic case or in classical stochastic differential equations. In particular the SDDE fails to satisfy the Markov property, hence one cannot rely on the well established setting of Markov processes for the study of the solution. As in the deterministic case, however, one can apply the key idea to lift the SDDE to have values in a suitable infinite dimensional path space. In doing so, one is able to recover the Markov property, nevertheless the main drawback is that now one is dealing with an infinite dimensional stochastic partial differential equation (SPDE). Although a well established theory for SPDE's exists, some fundamental results, known in the finite dimensional case, fail to hold true in the infinite dimension. In particular when considering infinite dimensional SPDE's, the concept of *quadratic variation* is not a

straightforward generalisation of the classical notion of quadratic variation. We recall that this concept is crucial in essential tools of stochastic analysis, such as the Itô formula. Some concrete results around the concept of quadratic variation in infinite dimensions have appeared only recently, see [DGR14].

SDDE's have been first studied in the seminal works [CM78, Moh84] and then extensively studied in [FMT10, Moh98, YM05], though in a different, but yet related setting. Recently there has been a renewed interest in SDDE's motivated by financial applications. In [Dup09] a path dependent stochastic calculus was first suggested and then widely developed in [CF10, CF13].

From the stochastic calculus point of view, the technique of regularisation, recently introduced, proved to be powerful to define the a stochastic integral and to prove a general Itô formula for stochastic differential equation both in finite and infinite dimensions. The first results exploiting the stochastic calculus via regularisation are found in [RV95, RV96a], where a generalisation of the Itô formula was proved. More recently in [DGR12] a new concept of quadratic variation for Banach space-valued processes was proposed and applied to prove a suitable Itô formula for infinite dimensional stochastic processes. This triggered a stream of studies aimed at deriving a suitable Itô's formula for delay equations and at studying deterministic problems that can be tackled by a stochastic approach. Particular attention was given to the Kolmogorov equation. We refer to [CDGR, CRa, CRb, FZ, FMT10]. Eventually in [CRa, FZ] the relationship between the *path-wise calculus* and the *Banach-space* calculus was detailed.

We remark that all the aforementioned results for delay equations are proved in the case when the driving noise is continuous, such as for a standard Brownian motion. Very few results exist when the noise allows for random jumps to happen, see, e.g. [Rei02, RvG06].

The aim of the present section is to extend the theory of SFDDE, studied in [Moh84] for a Brownian driving noise, to include jumps, that is to deal with noises of jump-diffusion type. Specifically we aim at settling the existence and unicity of solutions, and derive the fundamental tools of a stochastic calculus for SFDE's with jumps. We also study the robustness of the solutions of SFDDEs to changes of the driving noise. This is an important analysis in view of the future applications. From a finite dimensional point of view this was studied in e.g. [BDNK11].

We consider an \mathbb{R}^d -valued SDE of the form

$$\begin{aligned} dX(t) = & f(t, X(t), X(t + \cdot))dt + g(t, X(t), X(t + \cdot))dW(t) \\ & + \int_{\mathbb{R}_0} h(t, X(t), X(t + \cdot))(z)\tilde{N}(dt, dz), \quad t \in [0, T], \end{aligned} \quad (2.1)$$

where W is a standard Brownian motion, \tilde{N} is a compensated Poisson random measure and f , g and h are some given suitable functional coefficients. With the notation $X(t + \cdot)$ we mean that the coefficient may depend also on the past values of the solution on the interval $[t - r, t]$ for some fixed delay $r > 0$. It is this dependence on the past values of the evolution that is identified as memory or, equivalently, delay. The formal introduction of the current notation will be carried out in the next section. Notice that at this stage equation (2.66) is a finite dimensional SDE with values in \mathbb{R}^d .

We now lift the process (2.66) to have values in a suitable infinite dimensional path space. The choice of the suitable space is truly a key issue. As illustration, consider the purely diffusive case and denote the maximum delay appearing in (2.66) by $r > 0$. Then,

in [YM05] a product space of the form $M^p := L^p([-r, 0]; \mathbb{R}^d) \times \mathbb{R}^d$, $p \in [2, \infty)$, was chosen, whereas in [Moh98] the space of continuous functions $\mathcal{C} := \mathcal{C}([-r, 0]; \mathbb{R}^d)$ was taken as reference space. With the former choice one can rely on well-established results and techniques for L^p -spaces. Nevertheless this choice may seem artificial when dealing with a past path-dependent memory. For this reason the second choice, the space of continuous functions, is often considered the right space where to study delay equations, though it requires mathematically careful considerations.

The natural extension of SFDDE's to the jump-diffusion case correspondingly leads to two possible choices of setting: the product space M^p and the space of càdlàg (right continuous with finite left limit) functions $\mathcal{D} := \mathcal{D}([-r, 0]; \mathbb{R}^d)$. We decided to carry out our study in both settings in order to give a general comprehensive and critical presentation of when and in what sense it may be more suitable to treat the study in the one or the other setting. In fact, on the one side, we have the inclusion $\mathcal{D} \subset M^p$, with the injection being continuous, so that the M^p -setting appears to be more general, on the other side we see that existence and uniqueness of the solution of an SFDDE cannot be established in general in the space M_p . This in fact depends on the choice of type of delay or memory. In fact the drawback of the M^p approach is that it does not apply to SFDDE's with discrete delay, which are equations of the form

$$X(t) = \int_0^t X(s + \rho) ds + \int_0^t X(s + \rho) dW(s), \quad (2.2)$$

where $\rho \in [-r, 0)$ is a fixed parameter. Certainly, the reason is that $X(t + \rho)$ is actually interpreted as an evaluation of the segment $X_t = \{X(t + s), s \in [-r, 0]\}$ at the point $s = \rho$. Such operation is not well-defined in the M^p -setting. To see this, simply take two elements $(\eta_1, \eta_1(0)), (\eta_2, \eta_2(0)) \in M^p$ such that $\eta_1(s) = \eta_2(s)$ for all $s \in [-r, 0] \setminus \{\rho\}$ and $\eta_1(\rho) \neq \eta_2(\rho)$. Then, clearly $(\eta_1, \eta_1(0))$ and $(\eta_2, \eta_2(0))$ belong to the same class in M^p , but the evaluation at ρ is not uniquely determined. The discrete delay case can be treated in the setting given by \mathcal{D} .

In the sequel, we thus lift equation (2.66) to have values either in M^p , with $p \in [2, \infty)$, or in \mathcal{D} , exploiting the notion of *segment*. We then study a SFDDE of the form,

$$\begin{aligned} dX(t) &= f(t, X_t)dt + g(t, X_t)dW(t) + \int_{\mathbb{R}_0} h(t, X_t)(z)\tilde{N}(dt, dz) \\ X_0 &= \eta \end{aligned} \quad (2.3)$$

where we denote X_t the *segment* of the process X on an interval $[t - r, t]$, that is

$$X_t := \{X(t + \theta) : \theta \in [-r, 0]\},$$

being $r \geq 0$ the maximum delay. By $X(t)$ we denote the present value of the process at time t . Also η is a function on $[-r, 0]$.

In this paper, first we establish existence, uniqueness and moment estimates for the equation (2.3) where the *segment* X_t takes values either in M^p or in \mathcal{D} . Then we look at the robustness of the model to changes of the noise. In particular, we study what happens if we replace the small jumps of the infinite activity Poisson random measure N , by a continuous

Brownian noise B . This is done by comparing a process X with dynamics

$$\begin{aligned} dX(t) &= f(t, X_t)dt + g(t, X_t)dW(t) + \int_{\mathbb{R}_0} h_0(t, X_t)\lambda(z)\tilde{N}(dt, dz) \\ X_0 &= \eta, \end{aligned} \quad (2.4)$$

to the process $X^{(\epsilon)}$ defined by

$$\begin{aligned} dX^{(\epsilon)}(t) &= f(t, X_t^{(\epsilon)})dt + g(t, X_t^{(\epsilon)})dW(t) + h_0(t, X_t^{(\epsilon)}) \int_{|z| < \epsilon} |\lambda(z)|^2 \nu(dz) dB(t) \\ &\quad + \int_{|z| \geq \epsilon} h_0(t, X_t^{(\epsilon)})\lambda(z)\tilde{N}(dt, dz) \\ X_0^{(\epsilon)} &= \eta. \end{aligned} \quad (2.5)$$

We remark that the choice of this approximate guarantees the same so-called total volatility, using a terminology from financial modelling.

Eventually, exploiting the *stochastic calculus via regularisation* we prove an Itô type formula for stochastic delay equations with jumps, showing that the results are in fact coherent with the results obtained in [Moh98, YM05]. We work with forward integrals in the following sense. In general, given the stochastic processes $X = \{X_s, s \in [0, T]\}$, and $Y = \{Y_s, s \in [0, T]\}$, taking values in $L^p([0, T], \mathbb{R}^d)$ and its topological dual, respectively, we define the *forward integral* of Y against X as

$$\int_0^t {}_q \langle Y_s, dX_s \rangle_p := \lim_{\epsilon \downarrow 0} \int_0^t \left\langle Y_s, \frac{X_{s+\epsilon} - X_s}{\epsilon} \right\rangle_p ds, \quad (2.6)$$

where the limit holds in *probability* and we denoted by $\langle \cdot, \cdot \rangle_p$ the pairing between $L^p([-r, 0], \mathbb{R}^d)$ and its dual. Furthermore, if the above limit holds *uniformly on probability on compact sets* (ucp), we immediately have that the process

$$\left(\int_0^t {}_q \langle Y_s, dX_s \rangle_p \right)_{t \in [0, T]},$$

admits a càdlàg version and we say that the *forward integral* exists. When the two processes have values in the space M^p and M^{p*} , we are able to show that the above limit holds in fact ucp, characterizing thus the *forward integral* in terms of the derivative of the process X , which coincides with the operators introduced in [YM05] and in [FZ].

The present work is structured as follows. In Section 2.1.1 we introduce the main notation used throughout the whole paper. In Section 2.1.2 we study existence and uniqueness results for equation (2.3) with values in \mathcal{D} , whereas in Section 2.1.3 we prove the same results in the M^p setting. Then, in Section 2.1.4 we prove the robustness of equation (2.3) to the change of the noise. Eventually in Section 2.2 we prove a suitable Itô-type formula for SFDDE's with values in M^p and in \mathcal{D} .

2.1.1 Stochastic functional differential equations with jumps

Notation

Let $(\Omega, \mathcal{F}, \mathbb{F} := \{\mathcal{F}_t\}_{t \in [0, T]}, P)$ be a complete, filtered probability space satisfying the usual hypotheses for some finite time horizon $T < \infty$. Let $r \geq 0$ be a non-negative constant

denoting the maximum delay of the equations considered. We extend the filtration by letting $\mathcal{F}_s = \mathcal{F}_0$ for all $s \in [-r, 0]$. This will still be denoted by \mathbb{F} .

Let $W = (W^1, \dots, W^m)^\top$ be an m -dimensional \mathbb{F} -adapted Brownian motion and $N = (N^1, \dots, N^n)^\top$ be the jump measures associated with n independent \mathbb{F} -adapted Lévy processes, with Lévy measures $\nu = (\nu_1, \dots, \nu_n)$ respectively. We denote by \tilde{N} , the compensated Poisson random measure

$$\tilde{N}(dt, dz) := (N^1(dt, dz) - \nu_1(dz)dt, \dots, N^n(dt, dz) - \nu_n(dz)dt)^\top.$$

Consider the equation

$$\begin{aligned} dX(t) &= f(t, X_t)dt + g(t, X_t)dW(t) + \int_{\mathbb{R}_0} h(t, X_t)(z)\tilde{N}(dt, dz) \\ X_0 &= \eta, \end{aligned} \tag{2.7}$$

where f, g and h are some given functionals on a space containing the segments X_t , $t \in [0, T]$ of the process X . We will give precise definitions of the segments and the coefficient functionals below. Equations of the form (2.7) will be referred to as *stochastic functional delay differential equation* (SFDDE).

We remark that equation (2.7) is to be interpreted component-wise as a system of SFDDE of the following form:

$$\begin{aligned} dX^i(t) &= f^i(t, X_t)dt + \sum_{j=1}^m g^{i,j}(t, X_t)dW^j(t) + \sum_{j=1}^n \int_{\mathbb{R}_0} h^{i,j}(t, X_t, z)\tilde{N}^j(dt, dz), \\ X_0^i &= \eta^i, \quad i = 1, \dots, d. \end{aligned}$$

With the component-wise interpretation in mind, it is natural to require that the images $f(t, X_t)$ and $g(t, X_t)$ of the coefficient functionals f and g are contained in the spaces $L^2(\Omega, \mathbb{R}^d)$ and $L^2(\Omega, \mathbb{R}^{d \times m})$ respectively. Similarly, we want the image $h(t, X_t)$ of h to be contained in a set of matrices with j 'th column in $L^2(\Omega, L^2(\nu_j, \mathbb{R}^d))$. To express the space of all such matrices in a compact manner, we introduce the following notation. For the \mathbb{R}^n -valued measure $\nu = (\nu_1, \dots, \nu_n)^\top$, we will write $L^p(\nu) = L^p(\nu, \mathbb{R}^{d \times n})$ ($p \geq 2$), to denote the set of measurable functions

$$H : \mathbb{R}_0 \rightarrow \mathbb{R}^{d \times n},$$

such that

$$\|H\|_{L^p(\nu)}^p := \sum_{j=1}^n \|H^{j,\cdot}\|_{L^p(\nu_j, \mathbb{R}^d)}^p < \infty. \tag{2.8}$$

Here we have used the notation $H^{j,\cdot}$ to denote the j 'th column of H . Notice also that the Bochner space

$$L^q(\Omega, L^p(\nu, \mathbb{R}^{d \times n})) \quad (q \geq 2)$$

consists of the measurable functions $\mathcal{H} : \Omega \mapsto L^p(\nu, \mathbb{R}^{d \times n})$ such that

$$\|\mathcal{H}\|_{L^p(\Omega, L^p(\nu, \mathbb{R}^{d \times n}))}^q := E[\|\mathcal{H}\|_{L^p(\nu, \mathbb{R}^{d \times n})}^q] < \infty. \tag{2.9}$$

For convenience, we will sometimes omit to explicitly specify the spaces \mathbb{R}^d , $\mathbb{R}^{d \times m}$ and $\mathbb{R}^{d \times n}$, when it is clear from the context which space to consider and no confusion is possible. In this paper, when no confusion will occur, the standard Euclidean norm $\|\cdot\|_{\mathbb{R}^{k \times l}}$, will be denoted by $|\cdot|$ for any $k, l \in \mathbb{N}$.

Hereafter we introduce the relevant spaces we work with in the sequel. For $0 \leq u \leq T$, let

$$\mathcal{D}_u := \mathcal{D}([-r, u], \mathbb{R}^d), \quad (2.10)$$

denote the space of all càdlàg functions from $[-r, u]$ to \mathbb{R}^d , equipped with the uniform norm

$$\|\eta\|_{\mathcal{D}_u} := \sup_{-r \leq \theta \leq u} \{|\eta(\theta)|\}, \quad \eta \in \mathcal{D}_u. \quad (2.11)$$

Set $\mathcal{D} := \mathcal{D}_0$. For $2 \leq p < \infty$, let $L_u^p := L^p([-r, u], \mathbb{R}^d)$ and

$$M_u^p := L_u^p \times \mathbb{R}^d$$

with norm given by

$$\|(\eta, v)\|_{M_u^p}^p := \|\eta\|_{L_u^p}^p + |v|^p, \quad \eta \in M_u^p$$

Set $L^p := L_0^p$ and $M^p := M_0^p$.

We recall that the M_u^p -spaces are separable Banach spaces and M_u^2 is also a Hilbert space. On the other side \mathcal{D}_u equipped with the topology given by (2.11) is a non-separable Banach space. The space \mathcal{D}_u equipped with the Skorohod topology is separable metric space. Moreover, there exists also a topology on \mathcal{D}_u , equivalent to the Skorohod topology, such that \mathcal{D}_u is a complete separable metric space. See e.g. [Bil68, Par67].

Observe that if $\eta \in \mathcal{D}$, then

$$\|(\eta \mathbb{1}_{[-r, 0]}, \eta(0))\|_{M^p}^p = \|\eta \mathbb{1}_{[-r, 0]}\|_{L^p}^p + |\eta(0)|^p \leq (r+1)\|\eta\|_{\mathcal{D}}^p. \quad (2.12)$$

By (2.12), and since the elements in M^p have at most one càdlàg representative, the linear functional

$$\eta \mapsto (\eta \mathbb{1}_{[-r, 0]}, \eta(0))$$

is a linear continuous embedding of \mathcal{D} into M^p . Note that we will write $\|\eta\|_{M^p}$ in place of $\|(\eta \mathbb{1}_{[-r, 0]}, \eta(0))\|_{M^p}$.

We now introduce the notion of *segment* that will play an important role in this paper.

Definition 2.1.1. For any stochastic process $Y : [-r, T] \times \Omega \rightarrow \mathbb{R}^d$, and each $t \in [0, T]$, we define the *segments*

$$Y_t : [-r, 0] \times \Omega \rightarrow \mathbb{R}^d, \quad \text{by } Y_t(\theta, \omega) := Y(t + \theta, \omega), \quad \theta \in [-r, 0], \quad \omega \in \Omega.$$

In view of the arguments above, for each t , the segment can also be regarded as a function

$$\Omega \ni \omega \mapsto Y_t(\cdot, \omega) \in \mathcal{D}$$

or

$$\Omega \ni \omega \mapsto (Y_t(\cdot, \omega) \mathbb{1}_{[-r, 0]}, Y(t)) \in M^p$$

provided the necessary conditions of càdlàg paths or integrability.

We recall the following definitions. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ -algebra on Ω , containing all the P -null sets. Let \mathcal{D} be equipped with the σ -algebra \mathfrak{D} generated by the Skorohod topology.

Definition 2.1.2. We say that a function $\eta : \Omega \rightarrow \mathcal{D}$ is a (\mathcal{G} -measurable) \mathcal{D} -valued random variable if it is \mathcal{G} -measurable with respect to the σ -algebra \mathfrak{D} , or equivalently, if the \mathbb{R}^d -valued function $\omega \mapsto \eta(\theta, \omega)$ is \mathcal{G} -measurable for each $\theta \in [-r, 0]$.

Recall that $\mathfrak{D} \subsetneq \mathcal{B}(\mathcal{D})$, where $\mathcal{B}(\mathcal{D})$ is the Borel σ -algebra generated by the topology given by the norm (2.11).

Definition 2.1.3. We say that a function $(\eta, v) : \Omega \rightarrow M^p$ is a (\mathcal{G} -measurable) M^p -valued random variable if it is measurable with respect to the σ -algebras \mathcal{G} and $\mathcal{B}(M^p)$, or equivalently if the function

$$\omega \mapsto \int_{-r}^0 \eta(\theta, \omega) \phi(\theta) d\theta + v(\omega) \cdot u$$

is \mathcal{G} -measurable for every $(\phi, u) \in M^{p*} = M^{\frac{p}{p-1}}$.

Notice also that if η is a \mathcal{G} -measurable \mathcal{D} -valued-random variable, then it is \mathcal{G} -measurable as an M^p -valued random variable. Corresponding definitions apply in the cases of the D_u or M_u^p spaces above.

We are now ready to introduce the spaces of measurable \mathcal{D} -valued and M^p -valued random variables.

Recall that \mathcal{D}_u is equipped with the σ -algebra \mathfrak{D}_u generated by the Skorohod topology on \mathcal{D}_u . Let η be a \mathcal{D}_u -valued random variable. For $p \geq 2$, define

$$\|\eta\|_{S^p(\Omega; \mathcal{D}_u)}^p := E \left[\sup_{\theta \in [-r, u]} |\eta(\theta)|^p \right] = E[\|\eta(\theta)\|_{\mathcal{D}}^p],$$

and the equivalence relation $\eta_1 \sim \eta_2 \Leftrightarrow \|\eta_1 - \eta_2\|_{S^p(\Omega; \mathcal{D}_u)} = 0$. Let

$$S^p(\Omega, \mathcal{G}; \mathcal{D}_u)$$

denote the space of equivalence classes of \mathcal{D} -valued random variables $\omega \mapsto \eta(\omega, \cdot)$ such that $\|\eta\|_{S^p(\Omega; \mathcal{D}_u)} < \infty$.

For $p \geq 2$, let

$$L^p(\Omega, \mathcal{G}; M_u^p),$$

denote the Bochner spaces $L^p(\Omega, M_u^p)$ consisting of the M_u^p -valued random variables (η, v) such that the norm given by

$$\|(\eta, v)\|_{L^p(\Omega; M_u^p)}^p := E[\|(\eta, v)\|_{M_u^p}^p]$$

is finite. We recall that both $S^p(\Omega; \mathcal{D}_u)$ and $L^p(\Omega; M_u^p)$ are Banach spaces. Observe that if $\eta \in S^p(\Omega, \mathcal{G}; \mathcal{D})$, then

$$\|(\eta, \eta(0))\|_{L^p(\Omega; M^p)}^p \leq (r+1) \|\eta\|_{S^p(\Omega; \mathcal{D})}^p \quad (2.13)$$

thus, it also holds that

$$S^p(\Omega, \mathcal{G}; \mathcal{D}) \subset L^p(\Omega, \mathcal{G}; M^p),$$

and the embedding is continuous. With the appropriate boundedness and integrability conditions on a càdlàg adapted process Y , then for each t , the segment Y_t can be regarded as an element in the spaces $S^p(\Omega, \mathcal{F}_t; \mathcal{D})$ or $L^p(\Omega, \mathcal{F}_t; M^p)$.

In line with the definitions given above, we also use the following notation for any $u \in [0, T]$ and $2 \leq p < \infty$. Let

$$S_{ad}^p(\Omega, \mathcal{F}_u; \mathcal{D}_u) \subseteq S^p(\Omega, \mathcal{F}_u; \mathcal{D}_u)$$

denote the subspace of elements in $S^p(\Omega, \mathcal{F}_u; \mathcal{D}_u)$ admitting a \mathbb{F} -adapted representative. We remark that if $Z \in S^p(\Omega, \mathcal{F}_T; \mathcal{D}_T)$, then we have that

$$\|Z\|_{S^p(\Omega; \mathcal{D})} \leq \|Z\|_{S^p(\Omega; \mathcal{D}_t)} \leq \|Z\|_{S^p(\Omega; \mathcal{D}_T)}. \quad (2.14)$$

Also, consider the Banach space

$$L^p(\Omega; L_u^p)$$

with the usual norm given by:

$$\|Y\|_{L^p(\Omega; L_u^p)}^p := \mathbb{E}[\|Y\|_{L_u^p}^p] < \infty.$$

Then

$$L_{ad}^p(\Omega; L_u^p) \subseteq L^p(\Omega; L_u^p)$$

denotes the subspace of elements admitting \mathbb{F} -adapted representative.

Suppose now that $Y \in L_{ad}^p(\Omega; L_T^p)$. Since $Y(t)$ is well-defined for a.e. $t \in [-r, T]$, it makes sense to consider the segments Y_t as elements in $L^p(\Omega, \mathcal{F}_t; L_t^p)$ for a.e. t . Then,

$$\begin{aligned} \int_{-r}^u \|(Y_t, Y(t))\|_{L^p(\Omega; M^p)}^p dt &\leq \int_{-r}^u \left(\|Y_t\|_{L^p(\Omega; L_t^p)}^p + \|Y(t)\|_{L^p(\Omega; \mathbb{R}^d)}^p \right) dt \\ &\leq 2(r+u) \|Y_t\|_{L^p(\Omega; L_u^p)}^p. \end{aligned} \quad (2.15)$$

Even though we can not consider $S_{ad}^p(\Omega; \mathcal{D}_t)$ as a subspace of $L_{ad}^p(\Omega; L_t^p)$, since the function

$$S_{ad}^p(\Omega; \mathcal{D}_t) \ni \eta \mapsto \eta \in L_{ad}^p(\Omega; L_t^p)$$

is not injective, this function is continuous, and

$$\|Y_t\|_{L^p(\Omega; L_t^p)}^p \leq (t+r) \|Y_t\|_{S^p(\Omega; \mathcal{D}_t)}^p. \quad (2.16)$$

Remark 2.1.4. In the continuous setting (see e.g. [Moh98]), the segments of an SFFDE are often considered as elements of the Bochner space $L^2(\Omega; \mathcal{C})$, where \mathcal{C} denotes the set of continuous functions from $[-r, 0]$ to \mathbb{R}^d . We remark that the càdlàg counterpart, namely the Bochner space $L^p(\Omega, \mathcal{G}; \mathcal{D})$ of \mathcal{D} -valued functions turns out to be too restrictive to contain a sufficiently large class of càdlàg segments. This can be seen from the following lemma:

Lemma 2.1.5. *Suppose that X is a càdlàg Lévy-Itô process with $X \in L^p(\Omega, \mathcal{G}; \mathcal{D}[a, b])$. Then X is continuous with probability 1.*

To see why this holds, we first recall that by an equivalent definition of Bochner spaces (see [DS88] for more on these spaces), $L^p(\Omega, \mathcal{G}; \mathcal{D})$ consists of equivalence classes of the $(\mathcal{G}, \mathcal{B}(\mathcal{D}))$ -measurable functions $X : \Omega \rightarrow \mathcal{D}$ such that the image $X(\Omega_0)$ is separable for some subset $\Omega_0 \subset \Omega$ with $P(\Omega_0) = 1$, and $E[\|X\|_{\mathcal{D}([a,b])}^p] < \infty$ holds¹. By [JK15, lemma 9.12], we know that $X(\Omega_0)$ is separable if and only if there exist a countable set $\mathbb{T}_0 \in [a, b]$, such that $\Delta X(t, \omega) = 0$ whenever $t \notin \mathbb{T}_0, \omega \in \Omega_0$. In other words, except for a negligible set of sample paths of X , all the jumps of X occur at a countable number of times.

Proof of Lemma 2.1.5. Since $X \in L^p(\Omega, \mathcal{G}; \mathcal{D}([a, b]))$, we can choose Ω_0, \mathbb{T}_0 be as above. Now since X is a càdlàg Lévy-Itô process, it also holds that $P(\omega : \Delta X(t, \omega) \neq 0) = 0$ for every t , and hence

$$\mathcal{N} := \bigcup_{t \in \mathbb{T}_0} \{\omega \in \Omega_0 : \Delta X(t, \omega) \neq 0\}$$

is a null set. But then if $\omega \in \Omega_0 \setminus \mathcal{N}$, it holds that $\Delta X(t, \omega) = 0$ for every t , that is X is continuous on $\Omega_0 \setminus \mathcal{N}$ and $P(\Omega_0 \setminus \mathcal{N}) = 1$. \square

Examples

To illustrate some possible ways to model memory or delay in a stochastic differential equation, we include some examples of delay terms appearing in applications.

Distributed delay : the functional

$$S_t \mapsto \int_{-r}^0 S(t + \theta) \alpha(d\theta). \quad (2.17)$$

where α is a finite Borel measure on $[-r, 0]$, is an example of a *distributed delay*-functional. This is a general type of delay in the sense that examples below, can be regarded as particular cases of this one.

A general financial framework in this setting has been studied in [CY99, CY07] where the authors considered a price evolution for the stock of the form

$$dS(t) = M(S_t)dt + N(S_t)dW(t) = \int_{-r}^0 S(t+s) \alpha_M(ds)dt + \int_{-r}^0 S(t+s) \alpha_N(ds)dW(t),$$

α_M and α_N being suitable functions of bounded variation. See also [Moh98, Sec. V], where α is taken as a *probability measure*.

Absolutely continuous distributed delay: in the particular case $\alpha \ll \mathcal{L}$, where we have denoted by \mathcal{L} the Lebesgue measure, we have that the measure α admits a density $\kappa := \frac{d\alpha}{d\mathcal{L}}$. Therefore the functional (2.17) reads as

$$S_t \mapsto \int_{-r}^0 S(t + \theta) \kappa(\theta) d\theta, .$$

¹in fact this definition is valid for $\mathcal{D}([a, b])$ replaced by any Banach space V

A more advanced example has been provided in [KSW07b] where a functional of the form

$$(t, S_t) \mapsto \int_{-r}^0 \ell(t, S(t+\theta)) h(\theta) d\theta,$$

for some functional ℓ , has been treated.

Discrete delay : if we let $\alpha = \delta_\tau$, in equation (2.17), where δ_τ is the Dirac measure concentrated at $\tau \in [-r, 0]$, then we have a *discrete delay* functional, namely

$$S_t \mapsto \int_{-r}^0 S(t+\theta) \delta_\tau(d\theta) = S(t-\tau). \quad (2.18)$$

A discrete delay model using functionals on the form (2.18), is widely used in concrete applications, spanning from mathematical biology, as in the case of the delayed Lotka-Volterra model, see, e.g. [DVGLW95, Kua93, Moh98], to mathematical finance, as it happens for the delayed Black-Scholes model, see, e.g. [AHMP07, KSW05b], or for commodities markets, see, e.g., [KP07]. In particular, in [AHMP07], the authors give an explicit form for the price a European call option written on an underlying evolving as

$$dS(t) = \mu S(t-a)dt + \sigma(S(t-b))dW(t),$$

for $\mu \in \mathbb{R}$ and a suitable function σ .

A particular case of the discrete delay example is the no delay case, i.e. $\tau = \delta_0$. A multiple delay case, can be defined by letting $\alpha = \sum_{i=1}^N \delta_{\tau_i}$, $\tau_i \in [-r, 0]$, $i = 1, 2, \dots, N$.

Brownian delay: our setting allows also to consider delays with respect to a Brownian motion, namely

$$S_t \mapsto \int_{t-r}^t S(\theta) dW(\theta).$$

Hence this permits to take *noisy memory* models into account. These cases are arising e.g. in the modelling of stochastic volatility see, e.g. [KSW05b, Swi13] and when dealing with stochastic control problems, see e.g. [DMØR].

Lévy delay: similarly to the Brownian delay, we can also consider a delay with respect to a square integrable Lévy process of the form

$$S_t \mapsto \int_{t-r}^t S(\theta) dL(\theta).$$

Such type of delay has been employed in [Swi13] in order to consider some stochastic volatility models related to energy markets.

Mean field delay : we can consider a delay of the form

$$S_t \mapsto \mathbb{E} \left[\int_{-r}^0 S(t+\theta) \alpha(d\theta) \right],$$

where α is as in distributed delay example, see e.g. [AR].

2.1.2 \mathcal{D} framework

Fix $p \in [2, \infty)$. Consider again the equation

$$\begin{aligned} dX(t) &= f(t, X_t)dt + g(t, X_t)dW(t) + \int_{\mathbb{R}_0} h(t, X_t)(z)\tilde{N}(dt, dz) \\ X_0 &= \eta, \end{aligned} \quad (2.19)$$

In this section, we require that $f(t, \cdot)$, $g(t, \cdot)$, $h(t, \cdot)$ are defined on $S^p(\Omega, \mathcal{F}_t; \mathcal{D})$ for each fixed t . Therefore, we introduce the space

$$\mathbf{S}_p^{\mathbb{F}} := \{(t, \psi) \in [0, T] \times S^p(\Omega, \mathcal{F}; \mathcal{D}) \text{ such that } \psi \in S^p(\Omega, \mathcal{F}_t; \mathcal{D})\}, \quad (2.20)$$

as the domain of the coefficient functionals f, g, h in the SFDDE (2.19). In particular, we will require that:

$$\begin{aligned} f &: \mathbf{S}_p^{\mathbb{F}} \rightarrow L^p(\Omega, \mathbb{R}^d) \\ g &: \mathbf{S}_p^{\mathbb{F}} \rightarrow L^p(\Omega, \mathbb{R}^{d \times m}), \\ h &: \mathbf{S}_p^{\mathbb{F}} \rightarrow L^p(\Omega, L^2(\nu, \mathbb{R}^{d \times n})). \end{aligned}$$

Moreover,

$$\eta \in S^p(\Omega, \mathcal{F}_0; \mathcal{D})$$

To ensure that the integrals are well-defined, the following assumptions are imposed on the coefficient functionals f, g and h .

Hypothesis 2.1.6 (\mathcal{P}). Whenever $Y \in S_{ad}^p(\Omega; \mathcal{D}_T)$, the process

$$[0, T] \times \Omega \times \mathbb{R}_0 \ni (t, \omega, z) \mapsto h(t, Y_t)(\omega)(z) \in \mathbb{R}^{d \times n} \quad (2.21)$$

has a predictable version, and

$$\begin{aligned} [0, T] \times \Omega &\ni (t, \omega) \mapsto f(t, Y_t)(\omega) \in \mathbb{R}^d, \\ [0, T] \times \Omega &\ni (t, \omega) \mapsto g(t, Y_t)(\omega) \in \mathbb{R}^d \end{aligned}$$

have progressively measurable versions.

Predictable and progressive should be interpreted in the standard sense for \mathbb{R}^k -valued processes (see e.g. [App09]). We emphasise that the integrals in (2.19) should be interpreted with respect to the predictable and progressive versions of the respective integrands. For a range of SFDE's likely to be encountered in applications, the assumption \mathcal{P} is fairly easy to verify.

Example 2.1.1. Most of the examples presented in Section 2.1.1 satisfy Assumption \mathcal{P} . For instance, the functional displayed in the distributed delay example above, which is more general than the absolutely continuous and discrete delay, is predictable whenever the point zero is not an atom of the measure α , i.e. the discrete delay in (2.18) is not allowed when $\tau = 0$. The mean-field delay is deterministic and hence predictable. The Brownian delay can also be considered, since the process $t \mapsto \int_{t-r}^t S(\theta)dW(\theta)$ is a continuous martingale, in particular it admits a version with left-limits.

Definition 2.1.7. Suppose that the assumption \mathcal{P} holds. We say that $X \in S_{ad}^p(\Omega; \mathcal{D}_T)$ is a *strong solution* to the equation (2.19) if for each $t \in [0, T]$

$$\begin{aligned} X(t) &= \eta(0) + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dW(s) + \int_0^t \int_{\mathbb{R}_0} h(s, X_s)(z) \tilde{N}(ds, dz) \\ X_0 &= \eta. \end{aligned} \quad (2.22)$$

If the solution is unique, we will write ${}^\eta X$ to denote the solution of (2.22) with initial datum ${}^\eta X_0 = \eta$.

To prove existence and uniqueness of the solution of the SFDDE, we rely on the following result.

Lemma 2.1.8 (Kunita's inequality). *Let $q \geq 2$. Suppose that F, G and H are predictable processes taking values in \mathbb{R}^d , $\mathbb{R}^{d \times m}$ and $\mathbb{R}^{d \times n}$ respectively. If*

$$Y(t) = Y_0 + \int_0^t F(s) ds + \int_0^t G(s) dW(s) + \int_0^t \int_{\mathbb{R}_0} H(s, z) \tilde{N}(ds, dz), \quad t \in [0, T],$$

then there exists a constant $C = C(q, d, m, n, T)$, independent of the processes F, G and H and the initial value Y_0 , such that whenever $t \leq T$ the following inequality holds

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq u \leq t} |Y(u)|^q \right] &\leq C \left\{ \|Y_0\|_{L^q(\Omega, \mathbb{R}^d)}^q + \int_0^t \left(\|F(s)\|_{L^q(\Omega, \mathbb{R}^d)}^q + \|G(s)\|_{L^q(\Omega, \mathbb{R}^{d \times m})}^q \right. \right. \\ &\quad \left. \left. + \|H(s)\|_{L^q(\Omega, L^q(\nu))}^q + \|H(s)\|_{L^q(\Omega, L^2(\nu))}^q \right) ds \right\} \end{aligned} \quad (2.23)$$

For $n = 1$ (and arbitrary m and d), this is a rewritten version of Corollary 2.12 in [Kun04]. We have justified the extension to general n in Appendix 2.1.4.

Existence, uniqueness and moment estimates

Before giving sufficient conditions for existence and uniqueness of solutions to the equation (2.19), we will establish a set of hypotheses.

Hypothesis 2.1.9. (D₁) There exists $L > 0$, such that whenever $t \in [0, T]$ and $\eta_1, \eta_2 \in S^p(\Omega, \mathcal{F}_t; \mathcal{D})$, then

$$\begin{aligned} &\|f(t, \eta_1) - f(t, \eta_2)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(t, \eta_1) - g(t, \eta_2)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \\ &\quad + \|h(t, \eta_1) - h(t, \eta_2)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(t, \eta_1) - h(t, \eta_2)\|_{L^p(\Omega, L^2(\nu))}^p \\ &\leq L \|\eta_1 - \eta_2\|_{S^p(\Omega; \mathcal{D})}^p. \end{aligned}$$

(D₂) There exists $K > 0$, such that whenever $t \in [0, T]$ and $\eta \in S^p(\Omega, \mathcal{F}_t; \mathcal{D})$, then

$$\begin{aligned} &\|f(t, \eta)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(t, \eta)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \\ &\quad + \|h(t, \eta)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(t, \eta)\|_{L^p(\Omega, L^2(\nu))}^p \\ &\leq K(1 + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p). \end{aligned} \quad (2.24)$$

Remark 2.1.10. As usual, \mathbf{D}_2 is implied by \mathbf{D}_1 , if we assume that whenever $\eta = 0$, the left-hand-side of inequality (2.24) is bounded by some K' , uniformly in $t \in [0, T]$.

Theorem 2.1.11 (Existence and Uniqueness I). *Consider equation (2.19) with \mathcal{P} satisfied.*

- (i) *Suppose that assumption \mathbf{D}_1 holds. If $X, Y \in S_{ad}^p(\Omega; \mathcal{D}_T)$ are strong solutions to (2.7), then $X = Y$.*
- (ii) *Suppose that assumptions \mathbf{D}_1 and \mathbf{D}_2 hold. Then there exists a strong solution $X \in S_{ad}^p(\Omega; \mathcal{D}_T)$ to the equation (2.7). Moreover, there exists $D = D(K, p, T, d, m, n) > 0$, such that*

$$\|X\|_{S_{ad}^p(\Omega; \mathcal{D}_T)}^p \leq e^{Dt}(Dt + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p) \quad (2.25)$$

whenever $t \leq T$.

Proof. We will use a standard Picard iteration argument to show that a solution exists. First, we define, for each $k \geq 0$, a sequence of processes in $S_{ad}^p(\Omega; \mathcal{D}_T)$ inductively by

$$\begin{aligned} X^1(t) &= \eta(0), & t \in [0, T], \\ X_0^1 &= \eta \\ X^{k+1}(t) &= \eta(0) + \int_0^t f(s, X_s^k) ds + \int_0^t g(s, X_s^k) dB(s) \\ &\quad + \int_0^t h(s, X_s^k)(z) \tilde{N}(ds, dz), & t \in [0, T] \\ X_t^{k+1} &= \eta. \end{aligned}$$

We immediately have that $X^1 \in S_{ad}^p(\Omega; \mathcal{D}_T)$. Also if we assume that $X^k \in S_{ad}^p(\Omega; \mathcal{D}_T)$, then by assumption $f(X^k)$, $g(X^k)$, and $h(X^k)$ admit progressive and predictable versions respectively. Thus by assumption (\mathbf{D}_2) it follows that

$$\begin{aligned} &\int_0^T \left(\|f(t, X_t^k)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(t, X_t^k)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \right. \\ &\quad \left. + \|h(t, X_t^k)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(t, X_t^k)\|_{L^p(\Omega, L^2(\nu))}^p \right) dt \\ &\leq \int_0^T K(1 + \|X_t^k\|_{S^p(\Omega; \mathcal{D})}^p) dt \leq KT(1 + \|X^k\|_{S_{ad}^p(\Omega; \mathcal{D}_T)}^p) < \infty. \end{aligned} \quad (2.26)$$

In particular, the integrands of X^{k+1} are Itô integrable, so that X^{k+1} is càdlàg and adapted, and finally by Kunita's inequality, we have that $X^{k+1} \in S_{ad}^p(\Omega; \mathcal{D}_T)$.

We now claim that for each $k \in \mathbb{N}$ the following estimate holds for every $t \in [0, T]$,

$$\|X^{k+1} - X^k\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}^p \leq \frac{(LCt)^{k-1}}{(k-1)!} \|X^2 - X^1\|_{S_{ad}^p(\Omega; \mathcal{D}_T)}^p. \quad (2.27)$$

This trivially holds when $k = 1$. Now suppose that (2.27) holds for each $t \in [0, T]$. Using the

definition of X^{k+2} , X^{k+1} , Kunita's inequality (2.23), and assumption (\mathbf{D}_2) , we find that

$$\begin{aligned} \|X^{k+2} - X^{k+1}\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}^p &\leq C \int_0^t \left(\|f(s, X_s^{k+1}) - f(s, X_s^k)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(s, X_s^{k+1}) - g(s, X_s^k)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \right. \\ &\quad \left. + \|h(s, X_s^{k+1}) - h(s, X_s^k)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(s, X_s^{k+1}) - h(s, X_s^k)\|_{L^p(\Omega, L^2(\nu))}^p \right) ds \\ &\leq LC \int_0^t \|X_s^{k+1} - X_s^k\|_{S^p(\Omega; \mathcal{D})}^p ds \leq LC \int_0^t \|X^{k+1} - X^k\|_{S^p(\Omega; \mathcal{D}_s)}^p ds \\ &\leq LC \int_0^t \frac{(LCs)^{k-1}}{(k-1)!} \|X^2 - X^1\|_{S_{ad}^p(\Omega; \mathcal{D}_T)}^p ds = \frac{(LCt)^k}{k!} \|X^2 - X^1\|_{S_{ad}^p(\Omega; \mathcal{D}_T)}^p. \end{aligned}$$

Now, by induction, (2.27) holds for each $k \in \mathbb{N}$. In particular

$$\|X^k - X^i\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}^p \leq \|X^2 - X^1\|_{S_{ad}^p(\Omega; \mathcal{D}_T)}^p \sum_{j=\min\{k,i\}}^{\infty} \frac{(LTC)^{j-1}}{(j-1)!} \rightarrow 0, \quad \text{as } k, i \rightarrow \infty,$$

so that $\{X^k\}_{k \geq 0}$ is a Cauchy sequence in $S_{ad}^p(\Omega; \mathcal{D}_T)$. Since $S_{ad}^p(\Omega; \mathcal{D}_T)$ is complete, we have that $\{X^k\}_{k \geq 0}$ converges to some X in $S_{ad}^p(\Omega; \mathcal{D}_T)$. Clearly $X_0 = \eta$ P -a.s.

We will now show that the limit X satisfies (2.22) by showing that

$$\begin{aligned} d := \mathbb{E} \left[\sup_{0 \leq t \leq T} \left| X(t) - \left\{ \eta(0) + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dW(s) \right. \right. \right. \\ \left. \left. \left. + \int_0^t \int_{\mathbb{R}_0} h(s, X_s)(z) \tilde{N}(ds, dz) \right\} \right|^p \right]^{1/p} = 0 \end{aligned} \quad (2.28)$$

For arbitrary k , we subtract X^{k+1} and add its integral representation inside the supremum in (2.28). Then by the triangle inequality, Kunita's inequality, and finally the Lipschitz condition (\mathbf{D}_1) we find that

$$\begin{aligned} d &\leq \|X - X^{k+1}\|_{S_{ad}^p(\Omega; \mathcal{D}_T)} + \left\{ C \int_0^T \left(\|f(t, X_t^k) - f(t, X_t)\|_{L^p(\Omega; \mathbb{R}^d)}^p \right. \right. \\ &\quad \left. + \|g(t, X_t) - g(t, X_t^k)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p + \|h(t, X_t) - h(t, X_t^k)\|_{L^p(\Omega, L^p(\nu))}^p \right. \\ &\quad \left. \left. + \|h(t, X_t) - h(t, X_t^k)\|_{L^p(\Omega, L^2(\nu))}^p \right) dt \right\}^{1/p} \\ &\leq \|X - X^{k+1}\|_{S_{ad}^p(\Omega; \mathcal{D}_T)} + \left\{ CL \int_0^T \|X_t - X_t^k\|_{S^p(\Omega; \mathcal{D})}^p dt \right\}^{1/p} \\ &\leq \|X - X^{k+1}\|_{S_{ad}^p(\Omega; \mathcal{D}_T)} + (CLT)^{1/p} \|X - X^k\|_{S_{ad}^p(\Omega; \mathcal{D}_T)} \rightarrow 0. \end{aligned}$$

Since for any $\epsilon > 0$ we have that $0 \leq d < \epsilon$, it follows that $d = 0$, and hence a solution exists.

Suppose now that X and Y are solutions of (2.19). We will show that $X = Y$. Exploiting the integral representation of X and Y , Kunita's inequality and the Lipschitz condition (\mathbf{D}_1) , we have that, for all $t \in [0, T]$,

$$\begin{aligned} \|X - Y\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}^p &\leq C \int_0^t \left(\|f(s, X_s) - f(s, Y_s)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(s, X_s) - g(s, Y_s)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \right. \\ &\quad \left. + \|h(s, X_s) - h(s, Y_s)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(s, X_s) - h(s, Y_s)\|_{L^p(\Omega, L^2(\nu))}^p \right) ds \\ &\leq CL \int_0^t \|X_s - Y_s\|_{S^p(\Omega; \mathcal{D})}^p ds \leq CL \int_0^t \|X - Y\|_{S^p(\Omega; \mathcal{D}_s)}^p ds. \end{aligned}$$

and thus we have $\|X - Y\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}^p = 0$ for every $t \in [0, T]$ from Grönwall's inequality.

Similarly, if X is a solution to (2.19), from the integral representations, Kunita's inequality and the linear growth condition (D₂) we have that

$$\begin{aligned} \|X\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}^p &\leq C \left\{ \|\eta\|_{S^p(\Omega; \mathcal{D})}^p + \int_0^t \left(\|f(s, X_s)\|_{L^p(\Omega; \mathbb{R}^d)}^p \right. \right. \\ &\quad \left. \left. + \|g(s, X_s)\|_{L^p(\Omega; \mathbb{R}^d \times \mathbb{R}^n)}^p + \|h(s, X_s)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(s, X_s)\|_{L^p(\Omega, L^2(\nu))}^p \right) ds \right\} \\ &\leq C \left(\|\eta\|_{S^p(\Omega; \mathcal{D})}^p + K \left(\int_0^t 1 + \|X_s\|_{S^p(\Omega; \mathcal{D})}^p ds \right) \right) \leq C \|\eta\|_{S^p(\Omega; \mathcal{D})}^p + CKt + CK \int_0^t \|X\|_{S^p(\Omega; \mathcal{D}_s)}^p ds, \end{aligned}$$

so applying Grönwall's inequality we obtain

$$\|X\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}^p \leq \left(C \|\eta\|_{S^p(\Omega; \mathcal{D})}^p + CKt \right) e^{CKt},$$

□

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

Remark 2.1.12 (Path dependent SDEs). Suppose that for each $t \in [0, T]$ and every $\eta \in S_{ad}^p(\Omega, \mathcal{F}_0; \mathcal{D})$ it holds that,

$$\begin{aligned} f(t, \eta)(\omega) &= F(t, \eta(\omega)) \\ g(t, \eta)(\omega) &= G(t, \eta(\omega)) \\ h(t, \eta, \omega, \zeta) &= H(t, \eta(\omega), \zeta), \end{aligned}$$

P -a.s for some deterministic functionals

$$\begin{aligned} F : [0, T] \times \mathcal{D} &\rightarrow \mathbb{R}^d, \\ G : [0, T] \times \mathcal{D} &\rightarrow \mathbb{R}^{d \times m}, \\ H : [0, T] \times \mathcal{D} &\rightarrow L^2(\nu) \cap L^p(\nu). \end{aligned}$$

then the assumptions (D₁) and (D₂) hold whenever F, G are Lipschitz continuous in the second variable, uniformly with respect to the first, and H is Lipschitz continuous in the second variable, uniformly with respect to the first, using both norms $\|\cdot\|_{L^2(\nu)}$ and $\|\cdot\|_{L^p(\nu)}$.

2.1.3 M^p framework

Now, consider equation

$$\begin{aligned} dX(t) &= f(t, X_t, X(t))dt + g(t, X_t, X(t))dW(t) + \int_{\mathbb{R}_0} h(t, X_t, X(t))(z)\tilde{N}(dt, dz) \\ (X_0, X(0)) &= (\eta, x). \end{aligned} \tag{2.29}$$

Here (2.29) we have used the notation $f(\cdot, X_t, X(t))$ to emphasize the structure of the product space of M^p . Now for each $t \in [0, T]$ we will require that $(X_t, X(t))$ belongs to the space $L^p(\Omega, \mathcal{F}_t; M^p)$ for some $p \in [2, \infty)$, that will be fixed throughout the section. Therefore, we introduce

$$\mathbf{L}_p^{\mathbb{F}} := \{(t, (\psi, v)) \in [0, T] \times L^p(\Omega, \mathcal{F}; M^p) \text{ such that } (\psi, v) \in L^p(\Omega, \mathcal{F}_t; M^p)\}, \tag{2.30}$$

In particular, we will require that:

$$\begin{aligned} f &: \mathbf{L}_p^{\mathbb{F}} \rightarrow L^p(\Omega, \mathbb{R}^d) \\ g &: \mathbf{L}_p^{\mathbb{F}} \rightarrow L^p(\Omega, \mathbb{R}^{d \times m}), \\ h &: \mathbf{L}_p^{\mathbb{F}} \rightarrow L^p(\Omega, L^2(\nu, \mathbb{R}^{d \times n})). \end{aligned}$$

Moreover,

$$(\eta, x) \in L^p(\Omega, \mathcal{F}_0; M^p).$$

To ensure that the integrals are well-defined, the following assumptions are imposed on the coefficient functionals f, g and h .

Hypothesis 2.1.13 (Q). For $Y \in L_{ad}^p(\Omega; L_T^p)$, the process

$$[0, T] \times \Omega \times \mathbb{R}_0 \ni, (t, \omega, z) \mapsto h(t, Y_t, Y(t))(\omega)(z) \in \mathbb{R}^{d \times n}$$

has a predictable version, and

$$\begin{aligned} [0, T] \times \Omega \ni, (t, \omega) &\mapsto f(t, Y_t, Y(t))(\omega) \in \mathbb{R}^d, \\ [0, T] \times \Omega \ni, (t, \omega) &\mapsto g(t, Y_t, Y(t))(\omega) \in \mathbb{R}^{d \times m} \end{aligned}$$

have progressively measurable versions.

Definition 2.1.14. We say that $X \in L_{ad}^p(\Omega; L_T^p)$ is a *strong solution* to (2.29) if for each $t \in [0, T]$

$$\begin{aligned} X(t) &= x + \int_0^t f(s, X_s, X(s))ds + \int_0^t g(s, X_s, X(s))dW(s) + \int_0^t \int_{\mathbb{R}_0} h(s, X_s, X(s))(z)\tilde{N}(ds, dz) \\ (X_0, X(0)) &= (\eta, x). \end{aligned} \tag{2.31}$$

If the solution is unique, we will sometimes write $\eta, x X$ to denote the solution of (2.31) with initial data $(X_0, X(0)) = (\eta, x)$.

Proposition 2.1.15. Let $Y : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ be a stochastic process with a.s. càdlàg sample paths. Then the associated M^p -valued segment process

$$[0, T] \times \Omega \ni, (t, \omega) \mapsto (Y_t(\omega), Y(t, \omega)) \in M^p \tag{2.32}$$

is a.s. càdlàg.

Observe that the property that the segment process is càdlàg whenever Y is càdlàg, depends on the topology of the infinite dimensional space M^p . In general, such property does not hold if we replace M^p with \mathcal{D} .

Proof of Proposition 2.1.15. It suffices to show that if $Y(\omega) : [-r, T] \rightarrow \mathbb{R}^d$ is a càdlàg path, then the function

$$[0, T] \ni t \mapsto (Y_t(\omega), Y(t, \omega)) \in M^p \tag{2.33}$$

is also càdlàg. The function (2.33) is right continuous. In fact, for every sequence $r_k, k \in \mathbb{N}$ with $r_k > 0$ and $r_k \rightarrow 0$ as $k \rightarrow \infty$, we have that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \|Y_{t+r_k}(\omega) - Y_t(\omega), Y(t+r_k, \omega) - Y(t, \omega)\|_{M^p}^p \\ &= \lim_{k \rightarrow \infty} \int_{-r}^0 |Y(t+r_k+\beta, \omega) - Y(t+\beta, \omega)|^p d\beta + \lim_{k \rightarrow \infty} |Y(t+r_k, \omega) - Y(t, \omega)|^p = 0. \end{aligned}$$

by the dominated convergence theorem. Now given $t \in [0, T]$, we define $(Y_t^-(\omega), Y^-(t, \omega)) \in M^p$ by

$$Y_t^-(\theta, \omega) = \begin{cases} Y_t(\theta, \omega), & \theta \in [-r, 0) \\ \lim_{u \rightarrow 0^-} Y_t(u, \omega), & \theta = 0. \end{cases}$$

Consider r_k as above, we can use the dominated convergence theorem to observe that

$$\begin{aligned} & \lim_{k \rightarrow \infty} \|(Y_{t-r_k}(\omega) - Y_t^-(\omega), Y(t-r_k, \omega) - Y^-(t, \omega))\|_{M^p}^p \\ &= \lim_{k \rightarrow \infty} \int_{-r}^0 |Y(t-r_k+\beta, \omega) - Y^-(t+\beta, \omega)|^p d\beta + \lim_{k \rightarrow \infty} |Y(t-r_k, \omega) - Y^-(t, \omega)|^p = 0, \end{aligned}$$

and hence the function (2.33) has left limits. \square

Existence and uniqueness

The $L^p(\Omega; M^p)$ -analogue of the hypotheses (\mathbf{D}_1) and (\mathbf{D}_2) , are defined below.

Hypothesis 2.1.16. (\mathbf{L}_1) There exists $L > 0$, such that whenever $t \in [0, T]$ and $(\eta_1, x_1), (\eta_2, x_2) \in L^p(\Omega, \mathcal{F}_t; M^p)$, then

$$\begin{aligned} & \|f(t, \eta_1, x_1) - f(t, \eta_2, x_2)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(t, \eta_1, x_1) - g(t, \eta_2, x_2)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \\ &+ \|h(t, \eta_1, x_1) - h(t, \eta_2, x_2)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(t, \eta_1, x_1) - h(t, \eta_2, x_2)\|_{L^p(\Omega; L^2(\nu))}^p \\ &\leq L \|(\eta_1, x_1) - (\eta_2, x_2)\|_{L^p(\Omega; M^p)}^p. \end{aligned}$$

(\mathbf{L}_2) There exists $K > 0$, such that whenever $t \in [0, T]$ and $(\eta, v) \in L^p(\Omega, \mathcal{F}_t; M^p)$, then

$$\begin{aligned} & \|f(t, \eta, x)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(t, \eta, x)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \\ &+ \|h(t, \eta, x)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(t, \eta, x)\|_{L^p(\Omega; L^2(\nu))}^p \\ &\leq K(1 + \|(\eta, x)\|_{L^p(\Omega; M^p)}^p). \end{aligned}$$

Theorem 2.1.17 (Existence and Uniqueness II). *Consider (2.29) with \mathcal{Q} satisfied.*

(i) *Let $(\eta, x) \in L^p(\Omega, \mathcal{F}_0; M^p)$ such that η is càdlàg P -a.s. and $X \in L_{ad}^p(\Omega; L_T^p)$ be a strong solution to equation (2.29). Then the segment process*

$$\Omega \times [0, T] \ni (t, \omega) \mapsto (X_t(\omega), X(t, \omega)) \in M^p \quad (2.34)$$

has a càdlàg modification.

- (ii) Suppose that assumption (\mathbf{L}_1) holds. If $X, Y \in L_{ad}^p(\Omega; L_T^p)$ are strong solutions to (2.29), then $X = Y$.
- (iii) Suppose that assumptions (\mathbf{L}_1) and (\mathbf{L}_2) hold. Then there exists a strong solution X to equation (2.29). Moreover, there exists $D = D(K, p, T, d, m, n) > 0$, such that

$$\|X\|_{L_{ad}^p(\Omega; L_t^p)}^p \leq e^{Dt} (Dt + \|(\eta, x)\|_{L^p(\Omega; M^p)}^p), \quad (2.35)$$

whenever $t \leq T$.

Proof. (i) Recall that since X is a strong solution of (2.29), it is a semimartingale on $[0, T]$ and hence it admits a modification which is càdlàg on $[0, T]$. Since $X_0 = \eta$ is càdlàg, X is càdlàg, on $[-r, T]$. By Proposition 2.1.15 (i) holds.

(ii,iii) The proof is based on the same argument as for the proof of Theorem 2.1.11. For the sake of brevity we do not write out the details. However, we remark that if one replaces the norms $\|\cdot\|_{S^p(\Omega; \mathcal{D})}$ and $\|\cdot\|_{S_{ad}^p(\Omega; \mathcal{D}_t)}$, with the norms $\|\cdot\|_{L^p(\Omega; M^p)}$ and $\|\cdot\|_{L_{ad}^p(\Omega; L_t^p)}$ respectively, then all the inequalities hold true, except for the choice of constants. As an example, we provide the following M^p analogue of (2.26), namely

$$\begin{aligned} & \int_0^T \left(\|f(t, X_t^k, X^k(t))\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(t, X_t^k, X^k(t))\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \right. \\ & \quad \left. + \|h(t, X_t^k, X^k(t))\|_{L^p(\Omega, L^p(\nu))}^p + \|h(t, X_t^k, X^k(t))\|_{L^p(\Omega, L^2(\nu))}^p \right) dt \\ & \leq \int_0^T K(1 + \|X_t^k, X^k(t)\|_{L^p(\Omega; M^p)}^p) dt \leq KT(1 + (T + R)\|X^k\|_{L_{ad}^p(\Omega; L_t^p)}^p) < \infty. \end{aligned} \quad (2.36)$$

This follows immediately by the assumption \mathbf{L}_2 and inequality (2.15) \square

Let us stress that when the initial value is càdlàg, then the setting of Section (2.1.2) is more general than the one in this section. In fact, the assumptions (\mathbf{L}_1) and (\mathbf{L}_2) imply assumptions (\mathbf{D}_1) and (\mathbf{D}_2) , respectively.

2.1.4 Robustness SFDDEs

In the present section we study robustness of SFDDE to changes of the noise. In particular, we want to approximate the solution of an SFDDE X , with an approximate processes X^ϵ , where X^ϵ are defined by substituting the integrals with respect to the small jumps with integrals with respect to scaled Brownian motions. We follow rather closely, the presentation in [BDNK11] for ordinary SDE's and remark that a related problem is also considered in [KT11]. In this paper we also include a new ingredient, by giving sufficient conditions which ensure that the approximations X^ϵ converge to X in the p 'th mean.

The main motivation for studying such robustness problem is that it is difficult to perform simulations of distributions corresponding to a Lévy process. Indeed, simulation of such distributions are often performed by neglecting the jumps below a certain size ϵ . However, when needed to preserve the variation of the infinite activity Lévy process, a scaled Brownian motion is typically replacing the small jumps. Under some additional assumptions, it is

known that given a square integrable (1-dimensional) Lévy process with Lévy measure μ and compensated Poisson random measure \tilde{M} , the expression

$$\int_{|z|<\epsilon} z^2 \mu(dz)^{-1/2} \int_0^t \int_{|z|<\epsilon} z \tilde{M}(dz, ds),$$

converges in distribution to a standard Brownian motion W , as ϵ tends to 0. We refer to [AR01, CR07] for more details on this topic. We remark that the robustness problem in this paper, does not rely on the above mentioned additional assumptions.

The model

Fix $p \in [2, \infty)$. We want to consider the following dynamical systems with memory and jumps in the setting of Section 2.1.2:

$$\begin{aligned} X(t) &= \eta(0) + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dB(s) + \int_0^t \int_{\mathbb{R}_0} h_0(s, X_s) \lambda(z) \tilde{N}(ds, dz), \quad t \in [0, T] \\ X_0 &= \eta \in S^p(\Omega, \mathcal{D}). \end{aligned} \quad (2.37)$$

Here,

$$h_0 : \mathbf{S}_p^{\mathbb{F}} \rightarrow L^p(\Omega, \mathbb{R}^{d \times k}),$$

and

$$\lambda \in L^2(\nu, \mathbb{R}^{k \times n}) \cap L^p(\nu, \mathbb{R}^{k \times n})$$

for some $k \in \mathbb{N}$. Observe that

$$h := h_0 \lambda : \mathbf{S}_p^{\mathbb{F}} \rightarrow L^p(\Omega, L^2(\nu, \mathbb{R}^{d \times n})).$$

Example 2.1.2. Suppose that $n = k = d$ and that $h_0(t, \eta)$ and λ are diagonal matrices. In particular that,

$$h_0(t, \eta) \lambda(z)$$

is a diagonal matrix with entries $h_0^{i,i}(t, \eta) \lambda_{i,i}(z)$ for $i = 1, \dots, n$. Then the component-wise form of the jump integral in the SFDDE (2.37) is given by

$$\int_0^t \int_{\mathbb{R}_0} h_0^{i,i}(s, X_s) \lambda_{i,i}(z) \tilde{N}_i(ds, dz).$$

If we let the delay parameter r be equal to 0, then this example reduces to the problem of *robustness to model choice* treated in [Khe12].

Now, let us impose the following assumptions on f, g, h_0 and λ :

Hypothesis 2.1.18. **(i)** The coefficient functionals f and g are assumed to satisfy the assumptions (\mathcal{P}) , (\mathbf{D}_1) , (\mathbf{D}_2) of Section 2.1.2.

(ii) Whenever Y is a càdlàg adapted process on $[-r, T]$, then $h_0(t, Y_t)$ is predictable. Moreover, the functional h_0 satisfies the Lipschitz and linear growth conditions:

$$\begin{aligned} \|h_0(t, \eta_1) - h_0(t, \eta_2)\|_{L^p(\Omega, \mathbb{R}^{d \times k})}^p &\leq L \|\eta_1 - \eta_2\|_{S^p(\Omega; \mathcal{D})}^p, \\ \|h_0(t, \eta)\|_{L^p(\Omega, \mathbb{R}^{d \times k})}^p &\leq K(1 + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p). \end{aligned}$$

We claim now that the map $h := h_0\lambda$ satisfies the assumptions (\mathcal{P}) , (\mathbf{D}_1) , (\mathbf{D}_2) from Section 2.1.2. In fact observe that

$$\begin{aligned} &\|h(t, \eta_1) - h(t, \eta_2)\|_{L^p(\Omega, L^p(\nu))}^p + \|h(t, \eta_1) - h(t, \eta_2)\|_{L^p(\Omega, L^2(\nu))}^p \\ &= \sum_{j=1}^n \|(h_0(t, \eta_1) - h_0(t, \eta_2))\lambda^{i,j}\|_{L^p(\Omega, L^p(\nu_j))}^p + \|(h_0(t, \eta_1) - h_0(t, \eta_2))\lambda^{i,j}\|_{L^p(\Omega, L^2(\nu_j))}^p \\ &\leq \|h_0(t, \eta_1) - h_0(t, \eta_2)\|_{L^p(\Omega; \mathbb{R}^{d \times k})}^p \sum_{j=1}^n (\|\lambda^{i,j}\|_{L^p(\nu_j)}^p + \|\lambda^{i,j}\|_{L^2(\nu_j)}^p) \\ &\leq L \|\eta_1 - \eta_2\|_{S^p(\Omega; \mathcal{D})}^2 (\|\lambda\|_{L^p(\nu)}^p + \|\lambda\|_{L^2(\nu)}^p). \end{aligned}$$

Thus h satisfies the Lipschitz assumption (\mathbf{D}_1) . A similar argument yields h satisfies the linear growth assumption (\mathbf{D}_2) . Thus, by the existence and uniqueness Theorem 2.1.11, the following result holds.

Corollary 2.1.19. *The equation (2.37) has a unique solution ${}^n X$ in $S_{ad}^p(\Omega; \mathcal{D}_T)$. Moreover, there exists $D = D(K, \lambda, p, T, d, m, n) > 0$, such that*

$$E \left[\sup_{-r \leq s \leq t} |{}^n X(s)|^p \right] \leq e^{Dt} (Dt + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p), \quad (2.38)$$

for any $t \leq T$.

The approximating model

Let us first introduce some notation. For any $\varepsilon \in (0, 1)$, define $\lambda_\varepsilon(z) \in \mathbb{R}^{k \times n}$ by

$$\lambda_\varepsilon(z) = 1_{\{|z| < \varepsilon\}}(z) \lambda(z),$$

for a.e. z . Now, let B be an n -dimensional \mathbb{F} -adapted Brownian motion, independent of \tilde{N} . Independence of B and W is not required, (see e.g. [BDNK11]). We want to approximate equation (2.37) by replacing the integral with respect to the small jumps with an integral with respect to the Brownian motion B . More specifically, we will replace the integrators

$$\int_{\mathbb{R}_0} \lambda_\varepsilon^{i,j}(z) \tilde{N}^j(dt, dz), \quad (2.39)$$

with the integrators

$$\Lambda^{i,j}(\varepsilon) dB^j(t), \quad (2.40)$$

for $i = 1, \dots, k; j = 1, \dots, n$. Here, $\Lambda(\varepsilon)$ can be any bounded deterministic function with values in $\mathbb{R}^{k \times n}$ converging to 0 as $\varepsilon \rightarrow 0$. We choose to let

$$\Lambda^{i,j}(\varepsilon) = \|\lambda_\varepsilon^{i,j}\|_{L^2(\nu_j)}.$$

This choice corresponds to what has previously been used in the literature, see e.g. [BDNK11]. A justification of this choice is considered in Remark 2.1.20 below. Notice now that

$$|\Lambda(\epsilon)|_2 = \|\lambda_\epsilon\|_{L^2(\nu_j)}.$$

Remark 2.1.20. The choice $\Lambda^{i,j}(\epsilon) = \|\lambda_\epsilon^{i,j}\|_{L^2(\nu_j)}$ above is reasonable in the sense that for a given predictable square integrable process Y , this change of integrator preserves the variance of the integrals, i.e.

$$E\left[\left(\int_0^t \int_{\mathbb{R}_0} Y(s) \lambda_\epsilon^{i,j}(z) \tilde{N}^j(ds, dz)\right)^2\right] = E\left[\int_0^t Y(s)^2 ds\right] \|\lambda_\epsilon^{i,j}\|_{L^2(\nu_j)}^2 = E\left[\left(\int_0^t Y(s) \Lambda^{i,j}(\epsilon) dB^j(s)\right)^2\right],$$

for $i = 1, \dots, k; j = 1, \dots, n$. From a financial terminology perspective where these models can be applied (see e.g. [AHMP07, BDNK11, CY99, CY07, KSW05b, KSW07b, KP07]), this choice of Λ , preserves the *total volatility* of a process, when (2.39) is replaced by (2.40). However, this particular choice of Λ is not necessary for the analysis, as we will see in Remark 2.1.23 below.

Now, we are ready to exploit the dynamics of the *approximated processes* X^ϵ . Consider

$$\begin{aligned} X^{(\epsilon)}(t) &= \eta(0) + \int_0^t f(s, X_s^\epsilon) ds + \int_0^t g(s, X_s^\epsilon) dW(s) \\ &\quad + \int_0^t h_0(s, X_s^\epsilon) \Lambda(\epsilon) dB(s) + \int_0^t \int_{\mathbb{R}_0} h_0(s, X_s^\epsilon) (\lambda(z) - \lambda_\epsilon(s)) \tilde{N}(ds, dz) \\ X_0^{(\epsilon)} &= \eta \end{aligned} \quad (2.41)$$

Before proceeding to the main result of this section, we make the following observations regarding the functionals in the approximated equation (2.41):

- The functionals

$$\begin{aligned} \eta &\xrightarrow{g_1} h_0(t, \eta) \Lambda(\epsilon), \\ \eta &\xrightarrow{h_1} h_0(t, \eta) (\lambda - \lambda_\epsilon), \end{aligned}$$

satisfy the corresponding hypotheses from Section 2.1.2

- The Lipschitz and linear growth constant appearing in assumptions **(D₁)** and **(D₂)** can be chosen independent of ϵ . In particular, we can deduce the following linear growth estimate:

$$\begin{aligned} &\|h_0(t, \eta) \Lambda(\epsilon)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p + \|h_0(t, \eta) (\lambda - \lambda_\epsilon)\|_{L^p(\Omega, L^p(\nu))}^p + \|h_0(t, \eta) (\lambda - \lambda_\epsilon)\|_{L^p(\Omega, L^2(\nu))}^p \\ &\leq \|h_0(t, \eta)\|_{L^p(\Omega; \mathbb{R}^{d \times k})}^p \sup_{\epsilon \in (0, 1)} |\Lambda(\epsilon)|^p + \|h_0(t, \eta)\|_{L^p(\Omega; \mathbb{R}^{d \times k})}^p \|\lambda\|_{L^p(\nu)}^p + \|h_0(t, \eta)\|_{L^p(\Omega; \mathbb{R}^{d \times k})}^p \|\lambda\|_{L^2(\nu)}^p \\ &\leq K'(1 + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p). \end{aligned}$$

A similar estimate holds for the Lipschitz condition **(D₁)**.

The following existence and uniqueness result immediately follows from Theorem 2.1.11.

Corollary 2.1.21. *For each $\epsilon > 0$, there exists a unique strong solution ${}^\eta X^\epsilon$ to the equation (2.41). Moreover, there exists a $D = D(K, \lambda, p, T, d, m, n) > 0$, independent of ϵ , such that*

$$E \left[\sup_{-r \leq s \leq t} |{}^\eta X^\epsilon(s)|^p \right] \leq e^{Dt} (Dt + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p) \quad (2.42)$$

for any $t \leq T$.

Now, we are ready to state the main result of the present section. This result guarantees that, when ϵ tends to 0, X_ϵ converges to X in $S_{ad}^p(\Omega; \mathcal{D}_T)$.

Theorem 2.1.22 (Robustness). *Suppose that X satisfies equation (2.37) and X^ϵ satisfies equation (2.41). Then there exist a constant $A := A(p, T, \eta, K, L, \lambda) > 0$, independent of ϵ , such that*

$$E \left[\sup_{-r \leq s \leq t} |{}^\eta X(s) - {}^\eta X^\epsilon(s)|^p \right] \leq A e^{At} (\|\lambda_\epsilon\|_{L^2(\nu)}^p + \|\lambda_\epsilon\|_{L^p(\nu)}^p). \quad (2.43)$$

Proof. Writing out the integral representation of $X(s)$ and $X^\epsilon(s)$, we have that

$$\begin{aligned} X(s) - X^\epsilon(s) &= \int_0^s f(u, X_u) - f(u, X_u^\epsilon) du + \int_0^s g(u, X_u) - g(u, X_u^\epsilon) dW(u) \\ &\quad + \int_0^s \int_{\mathbb{R}^d} (h_0(u, X_u) - h_0(u, X_u^\epsilon)) \lambda(z) + h_0(u, X_u^\epsilon) \lambda_\epsilon(z) \tilde{N}(du, dz) \\ &\quad - \int_0^s h_0(u, X_u^\epsilon) \Lambda_p(\epsilon) dB(u), \\ X_0 - X_0^\epsilon &= 0. \end{aligned}$$

Let us first consider some estimates for the integrands of \tilde{N} and B . Observe that

$$\begin{aligned} &\|(h_0(u, X_u) - h_0(u, X_u^\epsilon)) \lambda + h_0(u, X_u^\epsilon) \lambda_\epsilon\|_{L^p(\Omega; L^p(\nu))} \\ &\leq \|(h_0(u, X_u) - h_0(u, X_u^\epsilon))\|_{L^p(\Omega; \mathbb{R}^{d \times k})} \|\lambda\|_{L^p(\nu)} + \|h_0(u, X_u^\epsilon)\|_{L^p(\Omega; \mathbb{R}^{d \times k})} \|\lambda_\epsilon\|_{L^p(\nu)} \\ &\leq L^{1/p} \|X_u - X_u^\epsilon\|_{S^p(\Omega; \mathcal{D})} \|\lambda\|_{L^p(\nu)} + K^{1/p} (1 + \|X_u^\epsilon\|_{S^p(\Omega; \mathcal{D})})^{1/p} \|\lambda_\epsilon\|_{L^p(\nu)}, \end{aligned}$$

and hence

$$\begin{aligned} &\|(h_0(u, X_u) - h_0(u, X_u^\epsilon)) \lambda + h_0(u, X_u^\epsilon) \lambda_\epsilon\|_{L^p(\Omega; L^p(\nu))}^p \\ &\leq 2^{p-1} \left(L \|X_u - X_u^\epsilon\|_{S^p(\Omega; \mathcal{D})}^p \|\lambda\|_{L^p(\nu)}^p + K (1 + \|X_u^\epsilon\|_{S^p(\Omega; \mathcal{D})})^p \|\lambda_\epsilon\|_{L^p(\nu)}^p \right). \end{aligned} \quad (2.44)$$

In an analogous manner we have that

$$\begin{aligned} &\|(h_0(u, X_u) - h_0(u, X_u^\epsilon)) \lambda + h_0(u, X_u^\epsilon) \lambda_\epsilon\|_{L^p(\Omega; L^2(\nu))}^p \\ &\leq 2^{p-1} \left(L \|X_u - X_u^\epsilon\|_{S^p(\Omega; \mathcal{D})}^p \|\lambda\|_{L^2(\nu)}^p + K (1 + \|X_u^\epsilon\|_{S^p(\Omega; \mathcal{D})})^p \|\lambda_\epsilon\|_{L^2(\nu)}^p \right), \end{aligned} \quad (2.45)$$

and that

$$\|h_0(u, X_u^\epsilon) \Lambda(\epsilon)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \leq K (1 + \|X_u^\epsilon\|_{S^p(\Omega; \mathcal{D})})^p |\Lambda(\epsilon)|^p$$

Using Lemma 2.1.8, the Lipschitz condition (\mathbf{D}_1) , estimates (2.44), (2.45), and Corollary 2.1.21 we have that there exist a constant $D' := D'(p, K, L, \lambda)$, independent of ε such that

$$\begin{aligned}
\alpha_\varepsilon(t) &:= \mathbb{E} \left[\sup_{-r \leq s \leq t} |{}^\eta X(s) - {}^\eta X^\varepsilon(s)|^p \right] \\
&\leq \int_0^t \left(\|f(u, X_u) - f(u, X_u^\varepsilon)\|_{L^p(\Omega; \mathbb{R}^d)}^p + \|g(u, X_u) - g(u, X_u^\varepsilon)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \right. \\
&\quad + \|(h_0(u, X_u) - h_0(u, X_u^\varepsilon))\lambda + h_0(u, X_u^\varepsilon)\lambda_\varepsilon\|_{L^p(\Omega; L^p(\nu))}^p \\
&\quad + \|(h_0(u, X_u) - h_0(u, X_u^\varepsilon))\lambda + h_0(u, X_u^\varepsilon)\lambda_\varepsilon\|_{L^p(\Omega, L^2(\nu))}^p \\
&\quad \left. + \|h_0(u, X_u^\varepsilon)\Lambda(\varepsilon)\|_{L^p(\Omega; \mathbb{R}^{d \times n})}^p \right) du \\
&\leq D' \int_0^t \left(\|X_u - X_u^\varepsilon\|_{S^p(\Omega; \mathcal{D})}^p + (1 + \|X_u^\varepsilon\|_{S^p(\Omega; \mathcal{D})}^p) (\|\lambda_\varepsilon\|_{L^2(\nu)}^p + \|\lambda_\varepsilon\|_{L^p(\nu)}^p + |\Lambda(\varepsilon)|^p) \right) du \\
&\leq D' \int_0^t \alpha_\varepsilon(u) du + tD'(1 + e^{Dt}(Dt + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p)) (\|\lambda_\varepsilon\|_{L^2(\nu)}^p + \|\lambda_\varepsilon\|_{L^p(\nu)}^p + |\Lambda(\varepsilon)|^p).
\end{aligned}$$

Now, set $B_t := tD'(1 + e^{Dt}(Dt + \|\eta\|_{S^p(\Omega; \mathcal{D})}^p))$ which is a non-decreasing function in t and hence by Grönwall's inequality, it follows that

$$\alpha_\varepsilon(t) \leq B_t e^{D't} (\|\lambda_\varepsilon\|_{L^2(\nu)}^p + \|\lambda_\varepsilon\|_{L^p(\nu)}^p + |\Lambda(\varepsilon)|^p).$$

Since $|\Lambda(\varepsilon)|^p = \|\lambda_\varepsilon\|_{L^2(\nu)}^p$, the result holds with $A := \max\{2B_T, D'\}$. \square

Remark 2.1.23. We have chosen to scale the Brownian motions B^j in the equation (2.41) for X^ε with $\Lambda^{i,j}(\varepsilon) := \|\lambda_\varepsilon^{i,j}\|_{L^2(\nu_j)}$. However, if we return to (2.39)-(2.40), we could let Λ_ε be any $\mathbb{R}^{k \times n}$ -valued function $\Lambda(\varepsilon) \geq 0$, $\varepsilon \geq 0$, bounded from above and converging to 0 as $\varepsilon \rightarrow 0$. Corollary 2.1.21 and Theorem 2.1.22 still hold, with the inequality (2.43) replaced by

$$E \left[\sup_{-r \leq s \leq t} |{}^\eta X(s) - {}^\eta X^\varepsilon(s)|^p \right] \leq A' e^{A't} (\|\lambda_\varepsilon\|_{L^2(\nu)}^p + \|\lambda_\varepsilon\|_{L^p(\nu)}^p + |\Lambda(\varepsilon)|^p). \quad (2.46)$$

This can be easily seen by reexamining the proofs of Corollary 2.1.21 and Theorem 2.1.22.

Kunita's inequality

In Section 2.1.2, we introduced a general version of *Kunita's inequality*, (Corollary 2.12 in [Kun04]). For $n = 1$, this is a rewritten version of Corollary 2.12 in [Kun04]. Below, we explain how to extend the result to general n .

Proof of Lemma 2.1.8. Notice that since norms on \mathbb{R}^n are equivalent, it holds that

$$\begin{aligned}
\sum_{j=1}^n |a_j|^q &\leq C_0 \left(\sum_{j=1}^n |a_j| \right)^q, \text{ and} \\
\left(\sum_{j=1}^n |a_j| \right)^q &\leq C_1 \sum_{j=1}^n |a_j|^q,
\end{aligned}$$

for some constants C_0, C_1 depending only on n and q . We may assume that $C_0 > 1$

$$\begin{aligned} \sum_{j=1}^n \|H^{:j}(s)\|_{L^2(\nu_j, \mathbb{R}^d)}^q &= \sum_{j=1}^n \left(\int_{\mathbb{R}_0} |H^{:j}(s, z)|^2 \nu_j(dz) \right)^{\frac{q}{2}} \\ &\leq C_0 \left(\sum_{j=1}^n \int_{\mathbb{R}_0} |H^{:j}(s, z)|^2 \nu_j(dz)^{\frac{1}{2}} \right)^q = C_0 \|H(s)\|_{L^2(\nu, \mathbb{R}^d)}^q. \end{aligned}$$

Then, if we write out "the columns wise" form of the \tilde{N} -integral, we obtain

$$\begin{aligned} \sup_{0 \leq u \leq t} \left| \sum_{j=1}^n \int_0^u \int_{\mathbb{R}_0} H^{:j}(s, z) \tilde{N}(ds, dz) \right|^q &\leq n^{q-1} \sup_{0 \leq u \leq t} \sum_{j=1}^n \left| \int_0^u \int_{\mathbb{R}_0} H^{:j}(s, z) \tilde{N}(ds, dz) \right|^q \\ &\leq n^{q-1} \sum_{j=1}^n \int_0^t \|H^{:j}(s)\|_{L^q(\Omega, L^q(\nu_j, \mathbb{R}^d))} + \|H^{:j}(s)\|_{L^q(\Omega, L^2(\nu_j, \mathbb{R}^d))} ds \\ &= n^{q-1} \int_0^t E \left[\sum_{j=1}^n \|H^{:j}(s)\|_{L^q(\nu_j, \mathbb{R}^d)}^q + \sum_{j=1}^n \|H^{:j}(s)\|_{L^2(\nu_j, \mathbb{R}^d)}^q \right] ds \\ &\leq n^{q-1} \int_0^t E \left[\|H(s)\|_{L^q(\nu, \mathbb{R}^{d \times n})}^q + C_0 \|H(s)\|_{L^2(\nu, \mathbb{R}^{d \times n})}^q \right] ds \\ &\leq n^{q-1} C_0 \int_0^t \|H(s)\|_{L^q(\Omega, L^q(\nu, \mathbb{R}^{d \times n}))}^q + \|H(s)\|_{L^q(\Omega, L^2(\nu, \mathbb{R}^{d \times n}))}^q ds. \end{aligned}$$

□

2.2 Itô's formula

In this section we aim at deriving Itô's formula for the SFDDE's studied in Section 2.1.1, which we recall, have the form (2.7),

$$\begin{aligned} dX(t) &= f(t, X_t)dt + g(t, X_t)dW(t) + \int_{\mathbb{R}_0} h(t, X_t, z) \tilde{N}(dt, dz) \\ X_0 &= \eta, \end{aligned}$$

where X_t is the segment of the process X in $[t-r, t]$, with $r > 0$ a finite delay, taking values in a suitable path space, and $X(t) \in \mathbb{R}^d$ the present value of the process X . For the whole section we work in the M^p -framework. See Section 2.1.3. Moreover, we assume that f , g and h are *deterministic* functionals, see Remark 2.1.12.

The main problem, when dealing with the SFDDE (2.7) is that the infinite dimensional process $(X_t)_{t \in [0, T]}$ fails, in general, to be a semimartingale and standard Itô calculus does not apply. In order to overcome this problem several approaches have been used in the literature.

The first attempts go back to [Moh98, YM05] where an Itô-type formula for continuous SFDDE was proved via Malliavin calculus. More recently, exploiting the concepts of *horizontal derivative* and *vertical derivative*, a path-wise Itô calculus for non-anticipative stochastic differential equation was derived in [CF10, CF13].

In [DGR14], an Itô formula for Banach-valued continuous processes was proved exploiting the *calculus via regularisation*, where an application to window processes (see [DGR14, Definition 1.4]) is also provided. Several works have followed studying Itô-type formulae for delay equations exploiting the *calculus via regularisation* and showing that the Banach-valued setting and the path-dependent setting can be in fact connected, see, e.g. [CDGR, CRa]. Eventually, the connection between the Banach space stochastic calculus and the path-wise calculus was made definitely clear in [CRb, FZ].

We remark that the literature on Itô formulae by the calculus via regularisation deals with equations driven by continuous noises. In this paper, we focus on the SFDDE's with jumps, thus extending the existing literature in this respect. We have chosen to consider the approach of the *calculus via regularisation*, first introduced in [RV95, RV96a], which was proved to be well-suited when dealing with infinite-dimensional processes or in the non-semimartingale case, see e.g. [CDGR, CRb, CDGR11, DGR12, DGR14]. In particular, we prove an Itô formula for the SFDDE (2.7) with values in M^p and we show that our result is coherent with those of [Moh98, YM05]. In the Appendix we provide a connection with the path-dependent calculus developed in [CF10, CF13].

Recall that, for a finite delay $r > 0$, $L^p := L^p([-r, 0]; \mathbb{R}^d)$ endowed with the standard norm $\|\cdot\|_{L^p}$, $p \in [2, \infty)$. In what follows we implicitly identify the topological dual of L^p , i.e. $(L^p)^*$, with L^q being $\frac{1}{p} + \frac{1}{q} = 1$, via the natural isomorphism given by

$$\begin{aligned} J : L^q &\rightarrow (L^p)^* \\ g &\mapsto J(g) = {}_q\langle g, \cdot \rangle_p, \end{aligned}$$

where $J(g)$ acts on L^p as follows

$$J(g)(f) = {}_q\langle g, f \rangle_p = \int_{-r}^0 g(s) \cdot f(s) ds, \quad g \in L^q, \quad f \in L^p,$$

being \cdot in the integral the usual scalar product in \mathbb{R}^d . It is well-known that J is a continuous linear isomorphism and hence, with a slight abuse of notation, we just write $h \in (L^p)^*$ when we actually mean $J^{-1}(h) \in L^q$, i.e. $(L^p)^* \cong L^q$.

Moreover, we denote by $C^1(L^p)$ the space of once Fréchet differentiable, not necessarily linear, functionals $F : L^p \rightarrow \mathbb{R}$ with continuous derivative, that is $DF : L^p \rightarrow L(L^p, \mathbb{R})$ where $L(L^p, \mathbb{R})$ denotes the space of continuous linear operators from L^p to \mathbb{R} . Now, since F is \mathbb{R} -valued, we actually have that $L(L^p, \mathbb{R}) = (L^p)^*$. Hence we can regard $DF(f)$, $f \in L^p$ as an element in L^q via J^{-1} . In a summary, we identify $DF(f)$ with $J^{-1}(DF(f))$ and simply write $DF : L^p \rightarrow L^q$ so that

$$DF(f)(g) = {}_q\langle DF(f), g \rangle_p = \int_{-r}^0 DF(f)(s) \cdot g(s) ds, \quad f \in L^p, \quad g \in L^q,$$

where the first equality is, by an abuse of notation, meant as an identification.

Also, recall that $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^d . Finally, recall that $M^p = L^p \times \mathbb{R}^d$ endowed with the standard product norm.

Let the SFDDE

$$\begin{cases} dX(t) = f(t, X_t, X(t))dt + g(t, X_t, X(t))dW(t) + \int_{\mathbb{R}_0} h(t, X_t, X(t))(z)\tilde{N}(dt, dz), \\ (X_0, X(0)) = (\eta, x) \in M^p, \end{cases} \quad (2.47)$$

for $t \in [0, T]$, $T < \infty$. We assume that $f : [0, T] \times M^p \rightarrow \mathbb{R}^d$, $g : [0, T] \times M^p \rightarrow \mathbb{R}^{d \times m}$ and $h : [0, T] \times M^p \times \mathbb{R}_0 \rightarrow \mathbb{R}^{d \times n}$ satisfy Assumptions **(L₁)** and **(L₂)** so that Theorem 2.1.17 holds and equation (2.47) admits a unique strong solution.

In the sequel every process is indexed by the time $t \in [0, T]$ and following [DGR14], if necessary, we extend the process $X = (X(t))_{t \in [0, T]}$ to the positive real line as follows

$$X(t) := \begin{cases} X(t) & \text{for } t \in [0, T], \\ X(T) & \text{for } t > T, \end{cases}.$$

Next, we consider the definition of forward integral.

Definition 2.2.1. Let $X = \{X(s), s \in [0, T]\}$ and $Y = \{Y(s), s \in [0, T]\}$ two \mathbb{R}^d -valued process. For every $t \in [0, T]$ we define the forward integral of Y w.r.t. X by $\int_0^t Y(s) \cdot dX(s)$ as the following limit,

$$\int_0^t Y(s) \cdot dX(s) := \lim_{\epsilon \downarrow 0} \int_0^t Y(s) \cdot \frac{X(s+\epsilon) - X(s)}{\epsilon} ds, \quad (2.48)$$

where the convergence is uniformly on compacts in probability (ucp).

Similarly, let $X = \{X_s, s \in [0, T]\}$, and $Y = \{Y_s, s \in [0, T]\}$, be L^p -valued and L^q -valued processes, respectively. For every $t \in [0, T]$ we define the L^p -forward integral of Y w.r.t. X as the following limit,

$$\int_0^t \int_{-r}^0 Y_s(\theta) \cdot dX_s(\theta) := \lim_{\epsilon \searrow 0} \int_0^t \int_{-r}^0 Y_s(\theta) \cdot \frac{X_{s+\epsilon}(\theta) - X_s(\theta)}{\epsilon} d\theta ds, \quad (2.49)$$

where the convergence is uniformly on compacts in probability. We introduce the short-hand notation:

$$\int_0^t \langle Y_s, dX_s \rangle_p := \int_0^t \int_{-r}^0 Y_s(\theta) \cdot dX_s(\theta).$$

Recall that a sequence of real-valued processes $\{X^n\}_{n \geq 1}$ converges to a process X *uniformly on compacts in probability* (ucp), if for each $t \in [0, T]$ we have that

$$\sup_{0 \leq s \leq t} |X_s^n - X_s| \rightarrow 0,$$

in probability. See e.g. [Pro05, p.57]. In this section, if not otherwise stated, any limit will be taken in the ucp sense.

Remark 2.2.2. Following [DGR14], in Definition 2.2.1 we have considered the ucp limit. In fact the space of càdlàg functions is a metrizable space with the metric induced by the ucp topology, see, e.g. [Pro05, p.57]. This implies that, being the approximating sequence in the right-hand-side of equation (2.48) càdlàg, the ucp convergence ensures that the limiting process, that is the forward integral, is also càdlàg.

Let us now introduce the notation we will use in the present work.

Definition 2.2.3. Let $F : [0, T] \times L^p \times \mathbb{R}^d \rightarrow \mathbb{R}$ a given function, we say that

$$F \in C^{1,1,2}([0, T] \times L^p \times \mathbb{R}^d),$$

if F is continuously differentiable w.r.t. the first variable, Fréchet differentiable with continuous derivative w.r.t. the second variable, and twice continuously differentiable w.r.t. the third variable.

We thus denote by ∂_t the derivative w.r.t. to time, $D_i F$ the Fréchet derivative w.r.t. the i -th component of the segment X_t and ∂_i the derivative w.r.t. the i -th component of the present state $X(t)$ and finally, $\partial_{i,j}$ the second order derivative w.r.t. the i, j -th component of $X(t)$.

We will then define the Fréchet gradient w.r.t. the segment as

$$D := (D_1, \dots, D_d) ,$$

the gradient w.r.t. the present state

$$\nabla_x := (\partial_1, \dots, \partial_d) ,$$

and the Hessian matrix w.r.t. the present state

$$\nabla_x^2 := (\partial_{i,j})_{i,j=1,\dots,d} .$$

Definition 2.2.4. Let $\eta \in W^{1,p}([-r, 0]; \mathbb{R}^d) =: W^{1,p}$, then we define by $\partial_{\theta,i}\eta$ and $\partial_{\theta,i}^+\eta$ the weak derivative and the right weak derivative, respectively, of the i -th component of η . Accordingly we define the gradient as

$$\nabla_{\theta} := (\partial_{\theta,1}, \dots, \partial_{\theta,d}) , \quad \text{resp. } \nabla_{\theta}^+ := (\partial_{\theta,1}^+, \dots, \partial_{\theta,d}^+) .$$

Eventually, in proving Itô's formula, we will need the notion of *modulus of continuity* of operators between infinite-dimensional normed spaces.

Definition 2.2.5 (Modulus of continuity). Let $(Y_1, \|\cdot\|_{Y_1})$ and $(Y_2, \|\cdot\|_{Y_2})$ be two normed spaces and $F : Y_1 \rightarrow Y_2$ a uniformly continuous function. We define the *modulus of continuity* of F as

$$\varpi(\epsilon) := \sup_{\|y-y'\|_{Y_1} \leq \epsilon} \|F(y) - F(y')\|_{Y_2}, \quad \epsilon > 0.$$

We thus have the following Itô's formula for SFDDE (2.47).

Theorem 2.2.6 (Itô's formula). *Let X be the solution to equation (2.47) and let $F : [0, T] \times L^p \times \mathbb{R}^d \rightarrow \mathbb{R}$ such that $F \in C^{1,1,2}([0, T] \times L^p \times \mathbb{R}^d)$ and such that $DF(t, \eta, x) \in W^{1,q}$, (q such that $\frac{1}{p} + \frac{1}{q} = 1$) for any $t \in [0, T]$, $\eta \in L^p$ and $x \in \mathbb{R}^d$ and $\nabla_{\theta} DF(t, \cdot, x) : L^p \rightarrow L^q$ is uniformly continuous. Then the following limit exists in the ucp sense,*

$$\lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \int_0^t {}_q \langle DF(s, X_s, X(s)), X_{s+\epsilon} - X_s \rangle_p ds =: \int_0^t {}_q \langle DF(s, X_s, X(s)), dX_s \rangle_p. \quad (2.50)$$

Moreover, for $t \in [0, T]$, we have that

$$\begin{aligned}
F(t, X_t, X(t)) &= F(0, X_0, X(0)) + \int_0^t \partial_t F(s, X_s, X(s)) ds + \int_0^t q \langle DF(s, X_s, X(s)), dX_s \rangle_p \\
&+ \int_0^t \nabla_x F(s, X_s, X(s)) \cdot dX(s) + \frac{1}{2} \int_0^t \text{Tr} [g^*(s, X_s, X(s)) \nabla_x^2 F(s, X_s, X(s)) g(s, X_s, X(s))] ds \\
&+ \int_0^t \int_{\mathbb{R}_0} (F(s, X_s, X(s) + h(s, X_s, X(s))(z)) - F(s, X_s, X(s))) N(ds, dz) \\
&- \int_0^t \int_{\mathbb{R}_0} \nabla_x F(s, X_s, X(s)) h(s, X_s, X(s))(z) N(ds, dz),
\end{aligned} \tag{2.51}$$

holds, P -a.s., where we have denoted by Tr the trace and by g^* the adjoint of g and the fourth term in equation (2.51) has to be intended component-wise, that is

$$\begin{aligned}
&\int_0^t \int_{\mathbb{R}_0} (F(s, X_s, X(s) + h(s, X_s, X(s))(z)) - F(s, X_s, X(s))) N(ds, dz) \\
&:= \sum_{i=1}^n \int_0^t \int_{\mathbb{R}_0} (F(s, X_s, X(s) + h^{:,i}(s, X_s, X(s))(z)) - F(s, X_s, X(s))) N^i(ds, dz).
\end{aligned}$$

Proof. Let $t \in [0, T]$. First, observe that for $\varepsilon > 0$ small enough, we have

$$\begin{aligned}
&\frac{1}{\varepsilon} \int_0^t F(s + \varepsilon, X_{s+\varepsilon}, X(s + \varepsilon)) - F(s, X_s, X(s)) ds = \\
&\frac{1}{\varepsilon} \int_\varepsilon^{t+\varepsilon} F(s, X_s, X(s)) ds - \frac{1}{\varepsilon} \int_0^t F(s, X_s, X(s)) ds = \\
&= \frac{1}{\varepsilon} \int_t^{t+\varepsilon} F(s, X_s, X(s)) ds - \frac{1}{\varepsilon} \int_0^\varepsilon F(s, X_s, X(s)) ds,
\end{aligned} \tag{2.52}$$

which, by the continuity of F and X_s and the right-continuity of $X(s)$, $s \in [0, T]$, recalling remark 2.2.2 and arguing as in [BR, eq. (4.6)], converges ucp to

$$F(t, X_t, X(t)) - F(0, X_0, X(0)).$$

The first part of (4.70) can be rewritten as

$$\begin{aligned}
&\frac{1}{\varepsilon} \int_0^t F(s + \varepsilon, X_{s+\varepsilon}, X(s + \varepsilon)) - F(s, X_s, X(s)) ds \\
&= \frac{1}{\varepsilon} \underbrace{\int_0^t F(s + \varepsilon, X_{s+\varepsilon}, X(s + \varepsilon)) - F(s, X_{s+\varepsilon}, X(s + \varepsilon)) ds}_{J_\varepsilon^1} \\
&\quad + \frac{1}{\varepsilon} \underbrace{\int_0^t F(s, X_{s+\varepsilon}, X(s + \varepsilon)) - F(s, X_{s+\varepsilon}, X(s)) ds}_{J_\varepsilon^2} \\
&\quad + \frac{1}{\varepsilon} \underbrace{\int_0^t F(s, X_{s+\varepsilon}, X(s)) - F(s, X_s, X(s)) ds}_{J_\varepsilon^3}.
\end{aligned}$$

Following the same arguments as in the proof of [DGR14, Theorem 5.2] we can show that

$$\lim_{\epsilon \searrow 0} J_\epsilon^1 = \int_0^t \partial_t F(s, X_s, X(s)) ds, \quad \text{ucp.}$$

Let us now consider J_ϵ^2 . A straightforward application of [BR, Corollary 4.4] implies that

$$\begin{aligned} J_\epsilon^2 &\rightarrow \int_0^t \nabla_x F(s, X_s, X(s)) \cdot dX(s) \\ &\quad + \frac{1}{2} \int_0^t \text{Tr} [g^*(s, X_s, X(s)) \nabla_x^2 F(s, X_s, X(s)) g(s, X_s, X(s))] ds \\ &\quad + \int_0^t \int_{\mathbb{R}_0} (F(s, X_s, X(s) + h(s, X_s, X(s), z)) - F(s, X_s, X(s))) N(ds, dz) \\ &\quad - \int_0^t \int_{\mathbb{R}_0} \nabla_x F(s, X_s, X(s)) h(s, X_s, X(s), z) N(ds, dz), \quad \text{as } \epsilon \searrow 0. \end{aligned}$$

Let us now show that

$$\lim_{\epsilon \searrow 0} J_\epsilon^3 = \int_0^t q \langle DF(s, X_s, X(s)), dX_s \rangle_p ds.$$

By an application of the infinite-dimensional version of Taylor's theorem of order one (see e.g. [Zei95, Ch. 4, Theorem 4.C]), we obtain

$$\begin{aligned} &\frac{1}{\epsilon} \int_0^t F(s, X_{s+\epsilon}, X(s)) - F(s, X_s, X(s)) ds \\ &= \frac{1}{\epsilon} \int_0^t \int_0^1 q \langle DF(s, X_s + \tau(X_{s+\epsilon} - X_s), X(s)), X_{s+\epsilon} - X_s \rangle_p d\tau ds \\ &= \frac{1}{\epsilon} \int_0^t \int_0^1 \int_{-r}^0 DF(s, X_s + \tau(X_{s+\epsilon} - X_s), X(s))(\alpha) \cdot (X(s + \epsilon + \alpha) - X(s + \alpha)) d\alpha d\tau ds \\ &= - \underbrace{\int_0^t \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 (DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha + \epsilon) - DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha)) \cdot X(s + \epsilon + \alpha) d\alpha d\tau ds}_{J_\epsilon^{3,1}} \\ &\quad + \underbrace{\int_0^t \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha + \epsilon) \cdot X(s + \alpha + \epsilon) d\alpha d\tau ds}_{J_\epsilon^{3,2}} \\ &\quad - \underbrace{\int_0^t \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha) \cdot X(s + \alpha) d\alpha d\tau ds}_{J_\epsilon^{3,2}}, \end{aligned} \tag{2.53}$$

where we have denoted by $X_{s,s+\epsilon}^\tau := X_s + \tau(X_{s+\epsilon} - X_s)$. We apply the change of variables

$g(\alpha) = \alpha + \epsilon$ to the first term of $J_\epsilon^{3,2}$ in Equation (2.53) in order to obtain

$$\begin{aligned}
J_\epsilon^{3,2} &= \frac{1}{\epsilon} \int_0^1 \int_{-r+\epsilon}^\epsilon DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha) \cdot X(s+\alpha) d\alpha d\tau \\
&\quad - \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha) \cdot X(s+\alpha) d\alpha d\tau \\
&= \frac{1}{\epsilon} \underbrace{\int_0^1 \int_0^\epsilon DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha) \cdot X(s+\alpha) d\alpha d\tau}_{J_\epsilon^{3,2,1}} \\
&\quad - \frac{1}{\epsilon} \underbrace{\int_0^1 \int_{-r}^{-r+\epsilon} DF(s, X_{s,s+\epsilon}^\tau, X(s))(\alpha) \cdot X(s+\alpha) d\alpha d\tau}_{J_\epsilon^{3,2,2}}.
\end{aligned}$$

We thus have, from the continuity of DF and X_s , that

$$\lim_{\epsilon \searrow 0} J_\epsilon^{3,2,1} = DF(s, X_s, X(s))(0) \cdot X(s), \quad \lim_{\epsilon \searrow 0} J_\epsilon^{3,2,2} = DF(s, X_s, X(s))(-r) \cdot X(s-r).$$

Let $\nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))$ denote a version of the weak derivative of $DF(s, X_{s,s+\epsilon}^\tau, X(s)) \in W^{1,q}$. Consider $J_\epsilon^{3,1}$. Using the mean value-theorem and interchanging the order of integration by Fubini's theorem we have

$$\begin{aligned}
J_\epsilon^{3,1} &= \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 \int_\alpha^{\alpha+\epsilon} \nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))(\beta) d\beta \cdot X(s+\epsilon+\alpha) d\alpha d\tau \\
&= \frac{1}{\epsilon} \int_0^1 \int_{-r}^{-r+\epsilon} \int_{-r}^\beta \nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))(\beta) \cdot X(s+\epsilon+\alpha) d\alpha d\beta d\tau \\
&\quad + \frac{1}{\epsilon} \int_0^1 \int_{-r+\epsilon}^0 \int_{\beta-\epsilon}^\beta \nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))(\beta) \cdot X(s+\epsilon+\alpha) d\alpha d\beta d\tau \\
&\quad + \frac{1}{\epsilon} \int_0^1 \int_0^\epsilon \int_{\beta-\epsilon}^0 \nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))(\beta) \cdot X(s+\epsilon+\alpha) d\alpha d\beta d\tau.
\end{aligned} \tag{2.54}$$

Now, we add and subtract integral terms so that the second integral on the right-hand side of (2.54) goes from $-r$ to 0, that is

$$\begin{aligned}
J_\epsilon^{3,1} &= \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 \int_{\beta-\epsilon}^\beta \nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))(\beta) \cdot X(s+\epsilon+\alpha) d\alpha d\beta d\tau \\
&\quad + \frac{1}{\epsilon} \int_0^1 \int_0^\epsilon \int_{\beta-\epsilon}^0 \nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))(\beta) \cdot X(s+\epsilon+\alpha) d\alpha d\beta d\tau \\
&\quad - \frac{1}{\epsilon} \int_0^1 \int_{-r}^{-r+\epsilon} \int_{\beta-\epsilon}^{-r} \nabla_\theta DF(s, X_{s,s+\epsilon}^\tau, X(s))(\beta) \cdot X(s+\epsilon+\alpha) d\alpha d\beta d\tau.
\end{aligned} \tag{2.55}$$

Then, Lebesgue's differentiation theorem implies that the second and third integral on the right-hand side of (2.55) converge to 0 as $\epsilon \rightarrow 0$. Concerning the first integral on the

right-hand side of (2.55), we add and subtract the corresponding terms in order to have

$$\begin{aligned}
& \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 \int_{\beta-\epsilon}^{\beta} \nabla_{\theta} DF(s, X_{s,s+\epsilon}^{\tau}, X(s))(\beta) \cdot X(s+\epsilon+\alpha) d\alpha d\beta d\tau \\
&= \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 \underbrace{\left(\nabla_{\theta} DF(s, X_{s,s+\epsilon}^{\tau}, X(s))(\beta) - \nabla_{\theta} DF(s, X_s, X(s))(\beta) \right)}_{J_{\epsilon}^{3,1,1}} \cdot \left(\int_{\beta-\epsilon}^{\beta} X(s+\epsilon+\alpha) d\alpha \right) d\beta d\tau \\
&+ \frac{1}{\epsilon} \int_0^1 \int_{-r}^0 \underbrace{\nabla_{\theta} DF(s, X_s, X(s))(\beta)}_{J_{\epsilon}^{3,1,2}} \cdot \left(\int_{\beta-\epsilon}^{\beta} X(s+\epsilon+\alpha) d\alpha \right) d\beta d\tau.
\end{aligned} \tag{2.56}$$

Using Hölder's inequality with exponents, $p, q \geq 2$, $\frac{1}{p} + \frac{1}{q} = 1$ on $J_{\epsilon}^{3,1,1}$ we have

$$\begin{aligned}
& |J_{\epsilon}^{3,1,1}| \\
&\leq \int_0^1 \|\nabla_{\theta} DF(s, X_{s,s+\epsilon}^{\tau}, X(s)) - \nabla_{\theta} DF(s, X_s, X(s))\|_{L^q} d\tau \left(\int_{-r}^0 \left| \frac{1}{\epsilon} \int_{\beta}^{\beta+\epsilon} X_s(\alpha) d\alpha \right|^p d\beta \right)^{1/p} \\
&\leq \int_0^1 \|\nabla_{\theta} DF(s, X_{s,s+\epsilon}^{\tau}, X(s)) - \nabla_{\theta} DF(s, X_s, X(s))\|_{L^q} d\tau \|M[X_s]\|_{L^p},
\end{aligned}$$

where

$$M[X_s](\beta) := \sup_{\epsilon > 0} \frac{1}{\epsilon} \int_{\beta}^{\beta+\epsilon} |X_s(\alpha)| d\alpha,$$

is the Hardy-Littlewood maximal operator, which is a (non-linear) bounded operator from L^p to L^p , $p > 1$. Hence, we can apply Lebesgue's dominated convergence theorem and the fact that by Lebesgue's differentiation we have

$$\lim_{\epsilon \searrow 0} \left(\int_{-r}^0 \frac{1}{\epsilon} \int_{\beta-\epsilon}^{\beta} |X(s+\epsilon+\alpha)|^p d\alpha d\beta \right)^{1/p} = \|X_s\|_{L^p}.$$

The above arguments in connection with the uniform continuity of $\nabla_{\theta} DF(s, X_s, X(s))$ implies the following estimate for $J_{\epsilon}^{3,1,1}$: there is a constant $C > 0$ independent of ϵ such that

$$|J_{\epsilon}^{3,1,1}| \leq C\varpi(\epsilon)\|X_s\|_{L^p} \rightarrow 0, \quad \text{as } \epsilon \rightarrow 0,$$

where $\varpi(\epsilon)$ denotes the modulus of continuity of $\nabla_{\theta} DF(s, \cdot, X(s))$ from Definition 2.2.5. Further, we can formally apply integration by parts to $J_{\epsilon}^{3,1,2}$ in order to obtain

$$\begin{aligned}
J_{\epsilon}^{3,1,2} &= \frac{1}{\epsilon} DF(s, X_s, X(s))(\beta) \cdot \int_{\beta-\epsilon}^{\beta} X_{s+\epsilon}(\alpha) d\alpha \Big|_{-r}^0 - \int_{-r}^0 DF(s, X_s, X(s))(\beta) \cdot \frac{X_{s+\epsilon}(\beta) - X_s(\beta)}{\epsilon} d\beta.
\end{aligned} \tag{2.57}$$

Then it follows that

$$J_\epsilon^{3,1,2} \rightarrow DF(s, X_s, X(s))(-r) \cdot X(s-r) - DF(s, X_s, X(s))(0) \cdot X(s) - {}_q \langle DF(s, X_s, X(s)), dX_s \rangle_p,$$

as $\epsilon \searrow 0$. Altogether, we have finally shown that

$$\lim_{\epsilon \searrow 0} J_\epsilon^3 = \int_0^t {}_q \langle DF(s, X_s, X(s)), dX_s \rangle_p.$$

This corresponds to (2.50). Hence, we conclude the proof. \square

In Section 2.5 it is shown, exploiting the results obtained in [FZ], that the Itô formula (2.51) is coherent with the Itô formula for path-dependent processes with jumps proved in [CF10], as well as the results obtained in [FZ].

Let us consider the *forward integral*

$$\int_0^t {}_q \langle DF(s, X_s, X(s)), dX_s \rangle_p,$$

introduced in Theorem 2.2.6, we want now to relate the *forward integral* to the operator introduced in [YM05]. In fact, since the right-derivative operator introduced in Definition 2.2.4 is the infinitesimal generator of the left-shift semigroup introduced in [YM05], it can be shown that the *forward integral* does coincide, under some suitable regularity conditions, with the operator $\mathcal{S}F(s, X_s, X(s))$ introduced in [YM05].

Proposition 2.2.7. *Let X be the solution to equation (2.47) and let $F : [0, T] \times L^p \times \mathbb{R}^d \rightarrow \mathbb{R}$ be such that $F \in C^{1,1,2}([0, T] \times L^p \times \mathbb{R}^d)$ and such that the forward integral defined in Equation (2.50) is well-defined as a limit in probability uniformly on compacts. Additionally, let us assume that $X_s \in W^{1,p}$ for every $s \in [0, T]$. Then*

$$\int_0^t {}_q \langle DF(s, X_s, X(s)), dX_s \rangle_p = \int_0^t {}_q \langle DF(s, X_s, X(s)), \nabla_\theta^+ X_s \rangle_p ds = \int_0^t \mathcal{S}F(s, X_s, X(s)) ds \quad (2.58)$$

holds P -a.s., where $\mathcal{S}F(s, X_s, X(s))$ is the operator introduced in [YM05, Section. 9].

Proof. Let $X_s \in W^{1,p}$. From the fundamental theorem of calculus for absolutely continuous functions we have, \mathbb{P} -a.s.,

$$\begin{aligned} & \lim_{\epsilon \searrow 0} {}_q \langle DF(s, X_s, X(s)), \frac{X_{s+\epsilon} - X_s}{\epsilon} - \nabla_\theta^+ X_s \rangle_p \\ &= \lim_{\epsilon \searrow 0} \int_{-r}^0 DF(s, X_s, X(s))(\beta) \cdot \left(\frac{(X(s+\epsilon+\beta) - X(s+\beta))}{\epsilon} - \nabla_\theta^+ X(s+\beta) \right) d\beta \\ &= \lim_{\epsilon \searrow 0} \int_{-r}^0 DF(s, X_s, X(s))(\beta) \cdot \left(\frac{1}{\epsilon} \int_{s+\beta}^{s+\beta+\epsilon} \nabla_\theta X(r) dr - \nabla_\theta^+ X(s+\beta) \right) d\beta \end{aligned}$$

Now, by Lebesgue's dominated convergence theorem, which is justified by analogous arguments as for the convergence of (2.56), we finally get

$$\lim_{\epsilon \searrow 0} \int_{-r}^0 DF(s, X_s, X(s))(\beta) \cdot \left(\frac{1}{\epsilon} \int_{s+\beta}^{s+\beta+\epsilon} \nabla_\theta X(r) dr - \nabla_\theta^+ X(s+\beta) \right) d\beta = 0.$$

\square

Exploiting the standard definition of the Poisson random measure, we can now give another formulation of the Itô formula (2.2.6).

Theorem 2.2.8 (Itô's formula). *Let the hypothesis of Theorem 2.2.6 hold, then*

$$\begin{aligned}
F(t, X_t, X(t)) &= F(0, X_0, X(0)) + \int_0^t \langle DF(s, X_s, X(s)), dX_s \rangle_p ds + \\
&+ \int_0^t \partial_t F(s, X_s, X(s)) ds + \int_0^t \nabla_x F(s, X_s, X(s)) \cdot dX(s) ds + \\
&+ \frac{1}{2} \int_0^t \int_0^t \text{Tr} [g^*(s, X_s, X(s)) \nabla_x^2 F(s, X_s, X(s)) g(s, X_s, X(s))] ds + \\
&+ \sum_{s \leq t} F(s, X_s, X(s)) - F(s, X_s, X(s-)) - \Delta X(s) \cdot \nabla_x F(s, X_s, X(s)),
\end{aligned} \tag{2.59}$$

holds P -a.s., where the notation is as in Theorem 2.2.6 and $\Delta X(s)$ is the jump of the process X at time s , namely

$$\Delta X(s) := X(s) - X(s-).$$

Proof. It immediately follows from Theorem 2.2.6 and [App09, Theorem 4.4.10]. \square

Hereinafter, we state a crucial probabilistic property of the solution of SFDDE (2.7) which is needed for the derivation of Feynman-Kac's formula also stated below. As it is perceptible, the finite-dimensional process $^{(\eta, x)}X(t)$, $t \in [0, T]$, $(\eta, x) \in M^p$ is not Markov, since the value of $^{(\eta, x)}X(t)$ depends on the *near* past. Nevertheless, if we enlarge the state space and regard the process X as a process of the segment, i.e. with infinite-dimensional state space, in the present case M^p , then such process is indeed Markovian.

The proof follows almost immediately given the fact that the Markov property of the solution is fully known for the case without jumps, i.e. $h = 0$, see [Moh84, Theorem (III. 1.1)], which actually follows from measure theoretical properties of the driving noise and not path or distributional properties of it. More concretely, one would expect the Markov property of the solution to hold if, for instance, the driving noises have independent increments which is the case in our setting.

In order to state the Markov property one is compelled to look at solutions starting at time $t_1 \geq 0$, that is we hereby consider $^{(t_1, \eta, x)}X_t$, $t \geq t_1$, $(\eta, x) \in M^p$, the segment of the solution starting in (η, x) at times $t_1 \geq 0$, i.e.

$$\begin{aligned}
^{(t_1, \eta, x)}X(t) &= \eta(0) + \int_{t_1}^t f(s, ^{(t_1, \eta, x)}X_s, ^{(t_1, \eta, x)}X(s)) ds + \int_{t_1}^t g(s, ^{(t_1, \eta, x)}X_s, ^{(t_1, \eta, x)}X(s)) dW(s) \\
&+ \int_{t_1}^t \int_{\mathbb{R}_0} h(s, ^{(t_1, \eta, x)}X_s, ^{(t_1, \eta, x)}X(s), z) \tilde{N}(ds, dz),
\end{aligned}$$

for every $t \in [t_1, T]$ and $^{(t_1, \eta, x)}X(t) = \eta(t - t_1)$ for every $t \in [t_1 - t, t_1]$. Define further the family of operators

$$\begin{aligned}
T_t^{t_1} : L^2(\Omega, \mathcal{F}_{t_1}; M^p) &\longrightarrow L^2(\Omega, \mathcal{F}_t; M^p) \\
(\eta, x) &\longmapsto \left(^{(t_1, \eta, x)}X_t, ^{(t_1, \eta, x)}X(t) \right).
\end{aligned}$$

We denote $T_t = T_t^0$. It turns out that, under hypotheses (\mathbf{L}_1) and (\mathbf{L}_2) , $T_t^{t_1}$ is Lipschitz continuous and the family of operators $\{T_t^{t_1}\}_{0 \leq t_1 \leq t \leq T}$ defines a semigroup on $L^2(\Omega, M^p)$, i.e.

$$T_{t_2}(\eta, x) = T_{t_2}^{t_1} \circ T_{t_1}(\eta, x),$$

for every $0 \leq t_1 \leq t_2 \leq T$ and $(\eta, x) \in L^2(\Omega, \mathcal{F}_0; M^p)$. The latter property can easily be obtained by showing that both sides of the identity solve the same SFDDE and the fact that solutions are unique, see [Moh84, Theorem (II. 2.2)] for the case $h = 0$.

Theorem 2.2.9 (The Markov property). *Assume hypotheses (\mathbf{L}_1) and (\mathbf{L}_2) hold and $(\eta, x)X(t)$, $t \in [0, T]$, $(\eta, x) \in M^p$ denotes the unique strong solution of the SFDDE (2.7). Then the random field*

$$\left\{ \left((\eta, x)X_t, (\eta, x)X(t) \right) : t \in [0, T], (\eta, x) \in M^p \right\}$$

describes a Markov process on M^p with transition probabilities given by

$$p(t_1, (\eta, x), t_2, B) = \mathbb{P}(\omega \in \Omega, T_{t_2}^{t_1}(\eta, x)(\omega) \in B),$$

for $0 \leq t_1 \leq t_2 \leq T$, $(\eta, x) \in M^p$ and Borel set $B \in \mathcal{B}(M^p)$. In other words, for any $(\eta, x) \in L^2(\Omega, \mathcal{F}_0; M^p)$ and Borel set $B \in \mathcal{B}(M^p)$, the Markov property

$$\mathbb{P}(T_{t_2}(\eta, x) \in B | \mathcal{F}_{t_1}) = p(t_1, T_{t_1}(\eta, x), t_2, B) = \mathbb{P}(T_{t_2}(\eta, x) \in B | T_{t_1}(\eta, x))$$

holds a.s. on Ω .

Proof. we can see that for every $0 \leq t_1 \leq t_2 \leq T$ and every $(\eta, x) \in M^p$, the mapping $B \mapsto p(t_1, (\eta, x), t_2, B) = \mathbb{P} \circ (T_{t_2}^{t_1}(\eta, x))^{-1}(B)$, $B \in \mathcal{B}(M^p)$ defines a probability measure on M^p , since the random variable $T_{t_2}^{t_1}(\eta, x) : \Omega \rightarrow M^p$ is $(\mathcal{F}, \mathcal{B}(M^p))$ -measurable. We would then like to show that, if $0 \leq t_1 \leq t_2 \leq T$, then

$$\mathbb{P}\left(\omega \in \Omega : T_{t_2}(\eta, x)(\omega) \in B | \mathcal{F}_{t_1}\right)(\omega') = p(t_1, T_{t_1}(\eta, x)(\omega'), t_2, B) \quad (2.60)$$

for almost all $\omega' \in \Omega$, every Borel set $B \in \mathcal{B}(M^p)$ and any $(\eta, x) \in L^2(\Omega, \mathcal{F}_0; M^p)$. The right-hand side of (2.60) equals

$$\int_{\Omega} 1_B \left((T_{t_2}^{t_1}(T_{t_1}(\eta)(\omega'))) (\omega) \right) \mathbb{P}(d\omega)$$

for almost all $\omega' \in \Omega$. Hence, by the definition of conditional expectation, identity (2.60) is synonymous with

$$\int_A 1_B (T_{t_2}(\eta, x)(\omega)) \mathbb{P}(d\omega) = \int_A \int_{\Omega} 1_B ((T_{t_2}^{t_1}(T_{t_1}(\eta, x)(\omega'))) (\omega)) \mathbb{P}(d\omega) \mathbb{P}(d\omega') \quad (2.61)$$

for all $A \in \mathcal{F}_{t_1}$ and all $B \in \mathcal{B}(M^p)$. In a summary, the main challenge is to verify relation (2.61) which is stated in quite general terms.

Let \mathcal{G}_t , $t \in [0, T]$ be the σ -algebra generated by $\{N((s, u], B), t < s \leq u \leq T, B \in \mathcal{B}(\mathbb{R}_0)\}$. The key steps in proving (2.61) according to [Moh84, Theorem (III. 1.1)] are the following. First, the family of operators $\{T_t\}_{t \in [0, T]}$ defines a semigroup on $L^2(\Omega, M^p)$. Secondly, the

σ -algebras \mathcal{F}_t and \mathcal{G}_t are independent for every $t \in [0, T]$, and $\{T_t^{t_1}(\eta, x)\}_{t \geq t_1}$ is adapted to $\{\mathcal{F}_t \cap \mathcal{G}_{t_1}\}_{t \geq t_1}$, being each $\mathcal{F}_t \cap \mathcal{G}_{t_1}$ independent of \mathcal{F}_{t_1} . Finally, a key point to prove (2.61) is that one can first prove

$$\int_A f(T_{t_2}(\eta, x)(\omega)) \mathbb{P}(d\omega) = \int_A \int_{\Omega} f((T_{t_2}^{t_1}(T_{t_1}(\eta, x)(\omega')))(\omega)) \mathbb{P}(d\omega) \mathbb{P}(d\omega') \quad (2.62)$$

for any bounded continuous function $f : M^p \rightarrow \mathbb{R}$. Then one can use a limit argument to show the relation (2.61) for the case $f = 1_B$ being B just an open set of M^p and eventually for any general indicator function on Borel sets. The argument which transfers (2.62) into the case $f = 1_B$ for any open set B in M^p requires that the state space in consideration is separable so that 1_B , being B an open set of M^p , can be approximated by uniformly continuous partitions of unity $\{f_m\}_{m \in \mathbb{N}}$, $f_m : M^p \rightarrow \mathbb{R}$ such that $\lim_{m \rightarrow \infty} f_m = 1_B$. All these properties above mentioned are indeed satisfied in our framework. \square

Exploiting Itô's formula from Theorem 2.2.6 together with the Markov property from Theorem 2.2.9, we can now prove a Feynman-Kac theorem for M^p -valued SFDDE's with jumps.

Theorem 2.2.10 (Feynman-Kac theorem). *Let the hypothesis of Theorem 2.2.6 hold, then the following path-dependent partial integro-differential equation (PPIDE) holds*

$$\begin{cases} \partial_t F(t, \eta, x) &= {}_q \langle DF(s, \eta, x), d\eta \rangle_p + \nabla_x F(t, \eta, x) \cdot f(t, \eta) \\ &+ \frac{1}{2} Tr [g(t, \eta, x) g^*(t, \eta, x) \nabla_x^2 F(t, \eta, x)] \\ &+ \int_{\mathbb{R}_0} (F(t, \eta, x + h(t, \eta, x)(z)) - F(t, \eta, x) - \nabla_x F(t, \eta, x) h(t, \eta, x)(z)) \nu(dz), \\ F(T, \eta, x) &= \Phi(\eta, x), \end{cases} \quad (2.63)$$

with $\Phi \in C^{1,2}(L^p \times \mathbb{R}^d)$, then we have

$$F(t, \eta, x) := \mathbb{E}[\Phi(X_T, X(T)) | X_t = \eta, X(t) = x], \quad t \in [0, T], \quad (2.64)$$

where $(X_t, X(t))$ solves the SFDDE (2.47). If further $\Phi \in C^{1,2}(L^p \times \mathbb{R}^d)$, then the converse holds true, as well.

Proof. We have to show that if a function $F : [0, T] \times L^p \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfies the PIDE (3.40), then we have that equation (2.64) holds. Let us assume X is the unique solution to equation (2.47), as in Section 2.1.1 we will use the notation $(\tau, \eta, x)X$ to denote the process with initial time $\tau \in [0, T]$ and initial value $(\eta, x) \in M^p$. If F satisfies equation (3.40) by Itô's formula (2.51) we have

$$\begin{aligned} & F(T, (\tau, \eta, x)X_T, (\tau, \eta, x)X(T)) - F(\tau, \eta, x) \\ &= \int_{\tau}^T \nabla_x F(s, (\tau, \eta, x)X_s, (\tau, \eta, x)X(s)) \cdot g(s, (\tau, \eta, x)X_s, (\tau, \eta, x)X(s)) dW(s) \\ &+ \int_{\tau}^T \int_{\mathbb{R}_0} \left(F(s, (\tau, \eta, x)X_s, (\tau, \eta, x)X(s)) + h(s, (\tau, \eta, x)X_s, (\tau, \eta, x)X(s))(z) \tilde{N}(ds, dz) \right. \\ &\left. - \int_{\tau}^T \int_{\mathbb{R}_0} F(s, (\tau, \eta, x)X_s, (\tau, \eta, x)X(s)) \tilde{N}(ds, dz) \right) \end{aligned} \quad (2.65)$$

Taking now the expectation, exploiting the fact that the right-hand side of equation (2.65) has null conditional expectation and using the terminal condition we have for any $0 \leq \tau < t \leq T$,

$$F(t, \eta, x) = \mathbb{E}[\Phi(X_T, X(T)) | X_t = \eta, X(t) = x],$$

and the first implication is proved.

Conversely, let us now suppose equation (2.64) holds true, then from the Markov property from Theorem 2.2.9 of the M^p -valued process and the tower rule for the conditional expectation, we have that for $0 \leq \tau < t \leq T$,

$$\begin{aligned} & \mathbb{E}[F(t, X_t, X(t)) - F(\tau, X_\tau, X(\tau)) | X_\tau = \eta, X(\tau) = x] = \\ & = \mathbb{E}[\mathbb{E}[\Phi(X_T, X(T)) | X_t, X(t)] - \mathbb{E}[\Phi(X_T, X(T)) | X_\tau, X(\tau)] | X_\tau, X(\tau)] = \\ & = \mathbb{E}[\Phi(X_T, X(T)) | X_\tau, X(\tau)] - \mathbb{E}[\Phi(X_T, X(T)) | X_\tau, X(\tau)] = 0. \end{aligned}$$

On the other side, we can apply Itô's formula (Theorem 2.2.6) to the function F , then we have that for $0 \leq \tau < t \leq T$,

$$\begin{aligned} F(t, X_t, X(t)) &= F(\tau, X_\tau, X(\tau)) + \int_\tau^t q \langle DF(s, X_s, X(s)), dX_s \rangle_p + \int_\tau^t \partial_t F(s, X_s, X(s)) ds \\ &+ \int_\tau^t \nabla_x F(s, X_s, X(s)) \cdot dX(s) + \frac{1}{2} \int_\tau^t \text{Tr} [g^*(s, X_s, X(s)) \nabla_x^2 F(s, X_s, X(s)) g(s, X_s, X(s))] ds \\ &+ \int_\tau^t \int_{\mathbb{R}_0} (F(s, X_s, X(s) + h(s, X_s, X(s))(z)) - F(s, X_s, X(s))) \nu(dz) ds \\ &- \int_\tau^t \int_{\mathbb{R}_0} \nabla_x F(s, X_s, X(s)) h(s, X_s, X(s))(z) \nu(dz) ds \\ &+ \int_\tau^t \int_{\mathbb{R}_0} (F(s, X_s, X(s) + h(s, X_s, X(s))(z)) - F(s, X_s, X(s))) \tilde{N}(ds, dz). \end{aligned}$$

Then taking conditional expectation the PPIDE (3.40) holds true. \square

2.3 A nonlinear Kolmogorov equation for stochastic delay differential equations with jumps

During recent years, an increasing attention has been paid to stochastic equations whose evolution depends not only on the present state, but also on the past history. In particular, it has been shown that memory effects cannot be neglected when dealing with many natural phenomena. As examples, let us mention the coupled *atmosphere-ocean models*, see, e.g., [BTR07], and their applications in describing climate changes in the environmental sciences setting, or the effect of time delay considering population dynamics, when suitable growth models are considered, see, e.g., [AF90]. Nevertheless, assumptions that will be made throughout the work are mainly taken into account having in mind concrete financial applications. For instance, in [KP07, Swi05] the authors pointed out how delay arises in commodity markets and energy markets, when it is necessary to take into account the impact of production and transportation. Similarly, delay naturally arises when dealing with financial instruments as, e.g., *Asian options* or *lookback options*, as studied in, e.g. [AHMP07, CY99, CDP15c, KSW07b, KSW04] and references therein.

For the mathematical foundations of the theory of stochastic functional delay differential equations (SFDDEs) we refer to [AHMP07], as well as to [Moh98] to many motivating examples concerning the treatment of equations with delay. In particular the monograph [AHMP07] represents an early and deep treatment of SFDDE's, where several results concerning existence and uniqueness of solutions to SFDDE's as well as regularity results are derived. The theory of delay equations has seen a renewed attention recently, in particular in [CF10, CF13] an ad hoc stochastic calculus, known as *functional Itô's calculus*, has been derived, based on a suitable Itô's formula for delay equations. Also, in past few years several different works have appeared deriving fundamental results on delay equations based on semigroup theory and infinite dimensional analysis, see, e.g. [FZ], or based on the calculus via regularization, see, e.g. [CR15, FMT10]. Eventually, in [FZ, CR15], it has been shown that SFDDE's, *path-dependent calculus* and delay equations via semigroup theory, are in fact closely connected.

Having in mind possible financial applications, the aim of the present work is to extend some results concerning the *non-linear Feynman-Kac formula* for a forward-backward system with delay, where the driving noise is a non Gaussian Lévy process, using the theory of SFDDE's first introduced in [AHMP07]. It is worth to mention that, particularly during last decades, asset price dynamics and, more generally, financial instruments processes, have been widely characterized by trajectories showing sudden changes and ample jumps. It follows that the classical Black and Scholes picture has to be refined by allowing to consider random components constituted by both diffusive and jump components.

We thus consider the following \mathbb{R} -valued SFDDE with jumps

$$dX(t) = \mu(t, X(t + \cdot), X(t))dt + \sigma(t, X(t + \cdot), X(t))dW(t) + \int_{\mathbb{R}_0} \gamma(t, X(t + \cdot), X(t), z)\tilde{N}(dt, dz), \quad (2.66)$$

where $W(t)$ is a standard Brownian motion, $\tilde{N}(dt, dz)$ is a compensated Poisson random measure with associated Lévy measure ν . Also the notation $X(t + \cdot)$ means that the *coefficients* μ, σ and γ depend not only on the present state, at time t , of the process X but also on its past values. We consider the finite dimensional \mathbb{R} -valued process solution to (2.66) as an infinite dimensional stochastic process with values in a suitable path-space, by exploiting the concept of *segment* of a process X , see, e.g., [AHMP07, Moh98]. More precisely, in what follows we will denote by $r > 0$ the maximum delay taken into account and $T < \infty$ a fixed finite time horizon. Thus, for an \mathbb{R} -valued stochastic process X , we indicate with $X(t)$ the value in \mathbb{R} at time $t \in [0, T]$ and with X_t the corresponding segment, i.e. the trajectory in the time interval $[t - r, t]$, that is $X_t(\cdot) : [-r, 0] \rightarrow \mathbb{R}$ is such that $X_t(\theta) := X(t + \theta)$ for all $\theta \in [-r, 0]$.

Then equation (2.66) can be rewritten as

$$\begin{cases} dX(t) = & \mu(t, X_t, X(t))dt + \sigma(t, X_t, X(t))dW(t) + \int_{\mathbb{R}_0} \gamma(t, X_t, X(t), z)\tilde{N}(dt, dz) \\ (X_0, X(0)) = & (\eta(\theta), x) \end{cases}, \quad (2.67)$$

for all $t \in [0, T]$, $\theta \in [-r, 0]$, $x \in \mathbb{R}$ and η a suitable \mathbb{R} -valued function on $[-r, 0]$.

Remark 2.3.1. In what follows we will only consider the 1-dimensional case, the case of a \mathbb{R}^d -valued stochastic process, perturbed by a general \mathbb{R}^m -dimensional Wiener process and

a \mathbb{R}^n -dimensional Poisson random measure, with $d > 1$, $m > 1$ and $n > 1$, can be easily obtained from the present one.

In order to take into account the delay component, we study the equation (2.67) in the *Delfour-Mitter space* defined as follows $\mathcal{D} := L^2([-r, 0]; \mathbb{R}) \times \mathbb{R}$, endowed with the scalar product

$$\langle (X_t, X(t)), (Y_t, Y(t)) \rangle_{M_2} = \langle X_t, Y_t \rangle_{L^2} + X(t) \cdot Y(t),$$

and norm

$$\|(X_t, X(t))\|_{M_2}^2 = \|X_t\|_{L^2}^2 + |X(t)|^2, \quad (X_t, X(t)) \in M_2, \quad (2.68)$$

where \cdot , resp. $|\cdot|$, stands for the scalar product in \mathbb{R} , resp. the absolute value, and $\langle \cdot, \cdot \rangle_{L^2}$, resp. $\|\cdot\|_{L^2}$, is the scalar product, resp. norm, in $L^2([-r, 0]; \mathbb{R}) =: L^2$. Note that the space M^2 is a separable Hilbert space, see, e.g., [BCDNR16]. The *Delfour-Mitter space* can be generalized to be a separable Banach space if we consider $p \in (1, \infty)$, equipped with the appropriate norm. In this work we will consider the case $p = 2$.

Alternatively, we could consider the space $\mathbf{D} := D([-r, 0]; \mathbb{R})$ of *càdlàg random variables*, which is a non separable Banach space if endowed with the sup norm $\|\cdot\|_{\mathbf{D}} = \sup_{t \in [-r, 0]} |\cdot|$. In fact we have that $\mathbf{D} \subset M^2$, with continuous injection, hence, by choosing M^p as state space, we decide to work on a more general space. Nevertheless, choosing M^2 as state space we cannot deal with the case of discrete delays, see, e.g., [BCDNR16, AHMP07]. The choice of taking $p = 2$, and then to consider the Hilbert space M^2 instead of the general Banach space M^p , or even the *Skorkhod space* \mathbf{D} , is mainly due to the extensive use we will do of the *Malliavin derivative*. In fact Malliavin calculus provides a powerful tool to study general regularity properties of a process or, as in the present case, to obtain representation theorem under mild regularity assumptions for the process. Nevertheless its generalization to the infinite dimensional setting, in particular when the driving noise is a general Lévy process, is rather technical and the theory, to our best knowledge, is still not completely developed. For these reasons we will only focus in the present work on the particular case of the state space being a Hilbert space, and we leave the case of \mathbf{D} -valued random variable for future investigations.

Despite the fact that the process (2.67) exhibits memory effects, in [AHMP07] the author shows that, in the pure diffusive case, lifting the problem to consider a \mathcal{D} -value solution leads to obtain a solution which is a *Markov process*. Taking in mind latter results and in order to derive the *Kolmogorov equation* associated to equation (2.67), we will consider, following [FMT10, FT02, FT05], a classical \mathbb{R} -valued backward stochastic differential equation (BSDE), coupled with the forward equation (2.67), which evolves according to

$$\begin{cases} dY(t) &= \psi \left(t, X_t, X(t), Y(t), Z(t), \int_{\mathbb{R}_0} U(t, z) \delta(z) \nu(dz) \right) dt \\ &+ Z(t) dW(t) + \int_{\mathbb{R}_0} U(t, z) \tilde{N}(dt, dz) \\ Y(T) &= \phi(X_T, X(T)) \end{cases}, \quad (2.69)$$

where ψ and ϕ are given suitable functions to be specified later on. We recall that a solution to equation (2.69) is a triplet (Y, Z, U) , where Y is the *state process*, while Z and U are the *control processes*.

It is well known that, when the delay is not involved, there exists a Feynman-Kac representation theorem that connects the solution of the coupled forward-backward system

(2.67) and (2.69), to a deterministic semi-linear partial integro-differential equation, see, e.g., [Del13] for further details. When the delay is taken into consideration, previous result holds in the Brownian case, see, e.g., [FMT10, FT05]. In the present paper we extend latter results to the case previously introduced, namely taking into consideration a non Gaussian Lévy noise. In particular, exploiting notations already introduced, we will consider the following coupled forward-backward stochastic differential equation (FBSDE) with delay, for $t \in [\tau, T] \subset [0, T]$,

$$\begin{cases} dX^{\tau, \eta, x}(t) &= \mu(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t))dt + \sigma(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t))dW(s) \\ &\quad + \int_{\mathbb{R}_0} \gamma(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t), z)\tilde{N}(dt, dz) \\ (X_\tau^{\tau, \eta, x}, X^{\tau, \eta, x}(\tau)) &= (\eta, x) \in \mathcal{D} \\ dY^{\tau, \eta, x}(t) &= \psi\left(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t), Y^{\tau, \eta, x}(t), Z^{\tau, \eta, x}(t), \tilde{U}^{\tau, \eta, x}(t)\right)dt \\ &\quad + Z^{\tau, \eta, x}(t)dW(t) + \int_{\mathbb{R}_0} U^{\tau, \eta, x}(t, z)\tilde{N}(dt, dz) \\ (Y^{\tau, \eta, x}(T)) &= \phi(X_T^{\tau, \eta, x}, X^{\tau, \eta, x}(T)) \end{cases}, \quad (2.70)$$

where we have denoted for short by

$$\tilde{U}^{\tau, \eta, x}(t) := \int_{\mathbb{R}_0} U^{\tau, \eta, x}(t, z)\delta(z)\nu(dz).$$

Moreover we have denoted by $X^{\tau, \eta, x}$ the value of the process with starting time $\tau \in [0, T]$ and initial value $(\eta, x) \in \mathcal{D}$. In what follows we omit the dependence on the initial value point (η, x) and we assume that the process starts at time $\tau = 0$, i.e. $X_t^{0, \eta, x} =: X_t$, if not stated otherwise. In order to simplify notation, most of the results will be proved taking $\tau = 0$, the extension to the general case of $\tau \neq 0$ being straightforward.

We are going to connect the solution to the FBSDE (2.70) to the following partial integro-differential equation

$$\begin{cases} \frac{\partial}{\partial t}u(t, \eta, x) + \mathcal{L}_t u(t, \eta, x) = \psi(t, \eta, x, u(t, \eta, x), \partial_x u(t, \eta, x)\sigma(t, \eta, x), \mathcal{J}u(t, \eta, x)) \\ u(T, \eta, x) = \phi(\eta, x), \quad t \in [0, T], \quad (\eta, x) \in \mathcal{D}. \end{cases} \quad (2.71)$$

where \mathcal{L}_t is the *infinitesimal generator* of the forward \mathcal{D} -valued process in equation (2.156), ∂_x is the derivative with respect to the present state $X(t)$ and \mathcal{J} is the operator

$$\mathcal{J}u(t, \eta, x) := \int_{\mathbb{R}_0} [u(t, \eta, x + \gamma(t, \eta, x, z)) - u(t, \eta, x)]\delta(z)\nu(dz).$$

In particular, we will consider a mild notion of solution to equation (2.71), so that we say that the function $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ is a *mild solution* to equation (2.71) if there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0, T]$ and any $(\eta_1, x_1), (\eta_2, x_2) \in \mathcal{D}$, u satisfies

$$\begin{aligned} |u(t, \eta_1, x_1) - u(t, \eta_2, x_2)| &\leq C|(\eta_1, x_1) - (\eta_2, x_2)|_2(1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2)^m, \\ |u(t, 0, 0)| &\leq C, \end{aligned} \quad (2.72)$$

and the following equality holds true

$$u(t, \eta, x) = P_{t, T}\phi(\eta, x) + \int_t^T P_{t, s}[\psi(\cdot, u(s, \cdot), \partial_x u(s, \cdot)\sigma(s, \cdot), \mathcal{J}u(s, \cdot))](\eta, x)ds, \quad (2.73)$$

for all $t \in [0, T]$, and $(\eta, x) \in \mathcal{D}$, $P_{t,s}$ being the *Markov semigroup* related to the equation (2.67). In particular, we considered a locally Lipschitz function $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ with respect to the second variable with at most polynomial growth so that the derivative in equation (2.73) is to be defined as in [FT05, Section 4], see also section 2.3.3 of the present work.

If we define

$$\begin{cases} Y^{\tau, \eta, x}(t) & := u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)) \\ Z^{\tau, \eta, x}(t) & := \partial_x u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)) \sigma(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)) \\ U^{\tau, \eta, x}(t, z) & := u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t) + \gamma(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t), z)) \\ & \quad - u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)) \end{cases}$$

then the triplet $(Y^{\tau, \eta, x}, Z^{\tau, \eta, x}, U^{\tau, \eta, x})$ is the unique solution to the backward equation (2.69), where ∂_x is the derivative with respect to the \mathbb{R} -valued present state $X(s)$ of $(X_s, X(s))$, u being the *mild solution* to the following *Kolmogorov equation*

$$\begin{cases} \frac{\partial}{\partial t} u(t, \eta, x) + \mathcal{L}_t u(t, \eta, x) = \psi(t, \eta, x, u(t, \eta, x), \partial_x u(t, \eta, x) \sigma(t, \eta, x), \mathcal{J}u(t, \eta, x)) \\ u(T, \eta, x) = \phi(\eta, x), \quad t \in [0, T], \quad (\eta, x) \in \mathcal{D} \end{cases}$$

The paper is organized as follows: in Section 2.3.1 we recall necessary notations also deriving results on the Malliavin derivative for Lévy processes that will be later used; Section 2.3.2 is devoted to the characterization of a FSFDE with delay, as well as to the characterization of the infinitesimal generator of the forward process; in Section 2.3.3 we describe our main results related to the study of the joint quadratic variation of the forward equation in (2.70) and a suitable function to be better specified later; in Section 2.3.4 we give the *non-linear Feynman-Kac theorem* that is later used to derive a deterministic representation to the FBSDE (2.70).

2.3.1 Notations and preliminaries

In this Section we introduce the notation used throughout the paper, also presenting basic definitions and main results related to the mathematical techniques involved in our approach.

Let us consider a probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \in [0, T]}, \mathbb{P})$, where $\{\mathcal{F}_t\}_{t \in [0, T]}$ is the *natural* filtration jointly generated by the random variables $W(s)$ and $N(ds, dz)$, for all $z \in \mathbb{R} \setminus \{0\} =: \mathbb{R}_0$ and for all $s \in [0, T]$, augmented by all \mathbb{P} -null sets, $W := \{W(t), t \in [0, T]\}$ being a 1-dimensional Brownian motion, while N is a 1-dimensional Poisson random measure, independent from W , with Lévy measure $\nu(dz)$ and associated compensated measure $\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt$. We will assume that the Lévy measure ν satisfies

$$\int_{\mathbb{R}_0} \min\{1, z^2\} \nu(dz) < \infty \quad , \quad \int_{\mathbb{R}_0} |z|^2 \nu(dz) < \infty ,$$

we underline that the former condition is a standard assumption in the definition of a Lévy measure ν , whereas the latter assumption implies that the process has a finite second moment, a natural assumption if one has in mind financial applications, see, e.g., [DNØP⁺09].

In the following, we fix a *delay* $r > 0$, we will use the notation $X(t)$ to denote the state at time t of the real valued process X , whereas we use X_t to denote the *segment* of the path

described by X during the time interval $[t-r, t]$ with values in a suitable infinite dimensional path space. In particular, we refer to the couple

$$\left((X(t+\theta))_{\theta \in [-r, 0]}, X(t) \right) =: (X_t, X(t)) .$$

From now on, we define $\mathcal{D} := L^2 \times \mathbb{R} := L^2([-r, 0]; \mathbb{R}) \times \mathbb{R}$ as the space of square integrable random variables, endowed with the scalar product

$$\langle (X_t, X(t)), (Y_t, Y(t)) \rangle_{M_2} = \langle X_t, Y_t \rangle_{L^2} + X(t) \cdot Y(t) ,$$

and norm

$$\| (X_t, X(t)) \|_{\mathcal{D}}^2 = \| X_t \|_{L^2}^2 + |X(t)|^2 , \quad (2.74)$$

namely the *Delfour-Mitter space*, which is a separable Hilbert space, see, e.g., [AHMP07] and reference therein for details.

Furthermore, for any $p \in [2, \infty)$, we denote by $\mathcal{S}^p(t) := \mathcal{S}^p([0, t]; \mathcal{D})$ and we say that a \mathcal{D} -valued stochastic process $(X_s, X(s))_{s \in [0, t]}$ belongs to $\mathcal{S}^p(t)$ if

$$\| X \|_{\mathcal{S}^p(t)}^p := \mathbb{E} \left[\sup_{s \in [0, t]} \| (X_s, X(s)) \|_{\mathcal{D}}^p \right] < \infty .$$

We denote for short $\mathcal{S}^p := \mathcal{S}^p(T)$. For the sake of simplicity, the following notation is used throughout the paper: $\| \cdot \|_2$ denotes the norm in \mathcal{D} and $|\cdot|$ the absolute value in \mathbb{R} .

Remark 2.3.2. Let us stress that for the sake of brevity we will consider here a \mathbb{R} -valued stochastic process X , nevertheless any result that follows can be easily generalized to the case of an \mathbb{R}^d -valued stochastic process. In particular we would have considered the *Delfour-Mitter space* $M^2([-r, 0]; \mathbb{R}^d) := L^2([-r, 0]; \mathbb{R}^d) \times \mathbb{R}^d$, see, e.g. [BCDNR16].

2.3.2 Forward stochastic functional differential equation with delay

This Section is devoted to the characterization of the delay equation (2.67). Some results are already established in literature, such as existence and uniqueness of solutions, whereas others are proved here for the first time.

As briefly said in Sec. 2.3.1, the main goal of this work is to focus on a stochastic functional delay differential equation (SFDDE) of the form

$$\begin{cases} dX(t) = \mu(t, X_t, X(t))dt + \sigma(t, X_t, X(t))dW(t) + \int_{\mathbb{R}_0} \gamma(t, X_t, X(t), z)\tilde{N}(dt, dz) \\ (X_0, X(0)) = (\eta, x) \in \mathcal{D} \end{cases} , \quad (2.75)$$

for all $t \in [0, T]$. We will assume the functionals μ, σ and γ fulfils the following assumptions.

Hypothesis 2.3.3. (A1) the coefficients

$$\mu : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}, \quad \sigma : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}, \quad \gamma : [0, T] \times \mathcal{D} \times \mathbb{R}_0 \rightarrow \mathbb{R}$$

are continuous.

(A2) There exists $K > 0$ such that for all $t \in [0, T]$ and for all $(\eta_1, x_1), (\eta_2, x_2) \in \mathcal{D}$,

$$\begin{aligned} & |\mu(t, \eta_1, x_1) - \mu(t, \eta_2, x_2)|^2 + |\sigma(t, \eta_1, x_1) - \sigma(t, \eta_2, x_2)|^2 \\ & + \int_{\mathbb{R}_0} |\gamma(t, \eta_1, x_1, z) - \gamma(t, \eta_2, x_2, z)|^2 \nu(dz) \\ & \leq K |(\eta_1, x_1) - (\eta_2, x_2)|_2^2 (1 + |(\eta_1, x_1)|_2^2 + |(\eta_2, x_2)|_2^2). \end{aligned}$$

Throughout the paper, we will look for *strong solution* to equation (2.75) in the following sense.

Definition 2.3.4. We say that $X := (X_t, X(t))_{t \in [0, T]}$ is a *strong solution* to equation (2.75) if for any $t \in [0, T]$ X is indistinguishably unique and $\{\mathcal{F}_t\}_{t \in [0, T]}$ -adapted and it holds \mathbb{P} -a.s.

$$\begin{aligned} X(t) &= x + \int_0^t \mu(s, X_s, X(s)) ds + \int_0^t \sigma(s, X_s, X(s)) dW(s) \\ &+ \int_0^t \int_{\mathbb{R}_0} \gamma(s, X_s, X(s), z) \tilde{N}(ds, dz), \\ X_0 &= \eta. \end{aligned}$$

In what follows we will denote by $(X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t))$ the \mathcal{D} -value of the process at time $t \in [\tau, T]$, with initial value $(\eta, x) \in \mathcal{D}$ at initial time $\tau \in [0, T]$. However, for the sake of brevity, in most of the results, we will avoid to state the dependence on the initial value (τ, η, x) writing for short (X_t, X^t) instead of $(X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t))$.

Now we provide an existence and uniqueness result for equation (2.67).

Theorem 2.3.5. *Suppose that μ, σ and γ satisfy conditions (A1) – (A2) in Assumptions 2.3.3. Then, for all $t \in [0, T]$ and $(\eta, x) \in \mathcal{D}$, there exists a unique strong solution to the SFDDE (2.67) in \mathcal{S}^p and there exists $C_1 := C_1(K, L, T, p)$ such that*

$$\|X^{\eta, x}\|_{\mathcal{S}^p}^p \leq C_1 (1 + |(\eta, x)|_2^p). \quad (2.76)$$

Moreover, the map $(\eta, x) \mapsto X^{\eta, x}$ is Lipschitz continuous from \mathcal{D} to \mathcal{S}^p and it exists $C_2 := C_2(K, L, T)$ such that

$$\|X^{\eta_1, x_1} - X^{\eta_2, x_2}\|_{\mathcal{S}^p}^p \leq C_2 |(\eta_1, x_1) - (\eta_2, x_2)|_2^p. \quad (2.77)$$

Proof. Existence and uniqueness of the solution to equation (2.75) are proved in [BCDNR16], as well as the estimate in equation (2.76).

As regards equation (2.77), exploiting the Burkholder-Davis-Gundy inequality, see, e.g. [App09, Section 4.4.], we have that, for any $t \in [0, T]$, denoting for short by C several

positive constants,

$$\begin{aligned}
& |X^{\eta_1, x_1} - X^{\eta_2, x_2}|_{\mathcal{S}^p}^p = \\
& = \mathbb{E} \sup_{t \in [0, T]} |(X_t^{\eta_1, x_1}, X^{\eta_1, x_1}(t)) - (X_t^{\eta_2, x_2}, X^{\eta_2, x_2}(t))|_2^p \leq \\
& \leq C |(\eta_1, x_1) - (\eta_2, x_2)|_2^p \\
& + C \left[\int_0^t |\mu(s, X_s^{\eta_1, x_1}, X^{\eta_1, x_1}(s)) - \mu(s, X_s^{\eta_2, x_2}, X^{\eta_2, x_2}(s))|^p ds \right. \\
& + \left. \left(\int_0^t |\sigma(s, X_s^{\eta_1, x_1}, X^{\eta_1, x_1}(s)) - \sigma(s, X_s^{\eta_2, x_2}, X^{\eta_2, x_2}(s))|^2 ds \right)^{\frac{p}{2}} \right. \\
& \left. + \int_0^t \int_{\mathbb{R}_0} |\gamma(s, X_s^{\eta_1, x_1}, X^{\eta_1, x_1}(s), z) - \gamma(s, X_s^{\eta_2, x_2}, X^{\eta_2, x_2}(s), z)|^p \nu(dz) ds \right],
\end{aligned}$$

so that from the Lipschitz continuity in assumption 2.3.3 A2, it follows

$$\begin{aligned}
& \mathbb{E} \sup_{t \in [0, T]} |(X_t^{\eta_1, x_1}, X^{\eta_1, x_1}(t)) - (X_t^{\eta_2, x_2}, X^{\eta_2, x_2}(t))|_2^p \leq \\
& \leq C |(\eta_1, x_1) - (\eta_2, x_2)|_2^p + \int_0^T \sup_{s \in [0, q]} |(X_s^{\eta_1, x_1}, X^{\eta_1, x_1}(s)) - (X_s^{\eta_2, x_2}, X^{\eta_2, x_2}(s))|_2^p,
\end{aligned}$$

and the claim follows from Gronwall's inequality. \square

Remark 2.3.6. We want to stress that a result analogous to Thm. 2.3.5 can be obtained by replacing the Delfour-Mitter space \mathcal{D} with the space \mathbf{D} of càdlàg functions, with the corresponding sup norm $\|\cdot\|_{\mathbf{D}} = \sup_{t \in [-r, 0]} |\cdot|$, see e.g. [BCDNR16, RvG06].

When the SFDDE is driven by a Gaussian noise, a result analogous to Thm. 2.3.5 has been derived, in different and yet related setting, in [AHMP07, Moh98].

One of the major results, when one is to lift the delay equation into an infinite dimensional setting exploiting the notion of *segment*, is that one is able to recover the Markov property of the driving equation, see, e.g [Moh98, Theorem II.1], similarly equation (2.75) results to be an \mathcal{D} -valued Markov process.

Proposition 2.3.7. *Let $X = \{(X_t, X(t))\}_{t \in [0, T]}$ be the strong solution to equation (2.75), then the process X is a Markov process in the sense that*

$$\mathbb{P}((X_t, X(t)) \in B | \mathcal{F}_s) = \mathbb{P}((X_t, X(t)) \in B | (X_s, X(s)) = (\eta, x)), \quad \mathbb{P} - a.s.,$$

for all $0 \leq s \leq t \leq T$ and for all B a Borel set $B \in \mathcal{B}(\mathcal{D})$.

Proof. See, e.g. [BCDNR16, Th. 3.10], or also, see, e.g., [RvG06, Prop. 3.3] or [PZ07, Sec.9.6]. \square

Having shown in Proposition 2.3.7 that X is a \mathcal{D} -valued Markov process, we can therefore introduce the transition semigroup $P_{\tau, t}$, acting on the space of Borel bounded function on \mathcal{D} , denoted by $B_b(\mathcal{D})$, namely, we define

$$P_{\tau, t} : B_b(\mathcal{D}) \rightarrow B_b(\mathcal{D}), \quad P_{\tau, t}[\varphi](x) := \mathbb{E}[\varphi(X_t^{\tau, \eta, x})], \quad \varphi \in B_b(\mathcal{D}). \quad (2.78)$$

Concerning the infinitesimal generator we have the following proposition, we refer to [App09] to a deep introduction on Markov semigroups.

Proposition 2.3.8. *Let us consider equation (2.75) and a function $\varphi : [0, T] \times L^2 \times \mathbb{R} \rightarrow \mathbb{R}$ which is differentiable with continuous derivative w.r.t. the first variable t , Fréchet differentiable with continuous Fréchet derivative w.r.t. the second variable η and twice differentiable w.r.t. the third variable x . Also assume that $\eta \in W^{1,2}$.*

We denote by ∂_t , resp. D , resp. ∂_x , the derivative w.r.t. the first variable, resp. the second variable, resp. the third variable. Also we will denote by ∂_θ the (weak) derivative of η .

Then the infinitesimal generator \mathcal{L}_t of equation (2.75) is

$$\begin{aligned} \mathcal{L}_t \varphi(t, \eta, x) &= \langle D\varphi(t, \eta, x), \partial_\theta^+ \eta \rangle_{L^2} \\ &+ \partial_x \varphi(t, \eta, x) \mu(t, \eta, x) + \frac{1}{2} \partial_{xx} \varphi(t, \eta, x) \sigma^2(t, \eta, x) \\ &+ \int_{\mathbb{R}_0} (\varphi(t, \eta, x + \gamma(t, \eta, x, z)) - \varphi(t, \eta, x) - \gamma(t, \eta, x, z) \partial_x \varphi(t, \eta, x)) \nu(dz). \end{aligned} \quad (2.79)$$

where $\partial_\theta^+ : W^{1,2} \subset L^2 \rightarrow L^2$ is the right derivative in L^2 .

Proof. The proof follows from a direct application of Itô's formula, see, e.g. [BCDNR16]. \square

Similarly, proposition 2.3.8, can be proved alongside [YM05, Lemma 9.3]. Nevertheless, following [FMT10, FT02, FT05] we are not interested in a full characterization of the domain of the infinitesimal generator since this goes beyond the aim of the present work.

Malliavin calculus for jump processes with delay

In this subsection we recall some definitions and main results concerning Malliavin operator and Skorokhod integral for jump processes. We will give fundamental definition in order to fix the notation and to recall the most effective result, we refer to [DNØP⁺09, Pet06, SUV07] for further references and proofs of some results. We will also derive some ad hoc results concerning the differentiability of SFDDE's, which will turn to be necessary in the rest of the paper. Topics covered in this Subsection largely follow [DNØP⁺09, Pet06] and, to a lesser extent, [DI10a, LSUV02, SUV07].

We would like to stress that, for the sake of brevity, we will state the results just for the jump component and we refer to [FT05, Moh98] for the Malliavin derivative for stochastic processes without jumps, namely in what follows we will set $D_{s,0} =: D_s$.

We denote by $I_n(f)$ the n -fold iterated stochastic integral w.r.t. the random measure \tilde{N} , as

$$I_n(f_n) := \int_{([0,T] \times \mathbb{R}_0)^n} f((t_1, z_1), \dots, (t_n, z_n)) \tilde{N}(dt_1, dz_1) \dots \tilde{N}(dt_n, dz_n) \in L^2(\Omega), \quad (2.80)$$

where $f \in L^2((([0, T] \times \mathbb{R}_0)^n) = L^2((([0, T] \times \mathbb{R}_0)^n), \otimes \nu(dz)dt)$ is a deterministic function.

Thus, every random variable $F \in L^2(\Omega)$ can be represented as an infinite sum of iterated integrals of the form (2.80). This representation is known as *chaos expansion*, see, e.g. [DNØP⁺09, Def. 12.1] or [Pet06, Th. 1].

Theorem 2.3.9. *The stochastic Sobolev space $\mathcal{D}^{1,2}$ consists of \mathcal{F} -measurable random variable $F \in L^2(\Omega)$ such that, for $(f_n)_{n \geq 0}$, with $f_n \in L^2([0, T] \times \mathbb{R}_0^n)$, it holds*

$$F = \sum_{n=0}^{\infty} I_n(f_n), \quad (2.81)$$

with the following norm

$$\|F\|^2 = \sum_{n=0}^{\infty} n! \|I_n(f_n)\|_{L^2([0, T] \times \mathbb{R}_0^n)}^2.$$

Given the *chaos expansion* in equation (2.81), we can introduce the *Malliavin derivative* $D_{t,z}$ and its domain $\mathbb{D}^{1,2}$, see, e.g. [DNØP⁺09, Def. 12.2].

Definition 2.3.10. Let us consider a random variable $F \in \mathcal{D}^{1,2}$, the *Malliavin derivative* is the operator $D : \mathcal{D}^{1,2} \subset L^2(\Omega) \rightarrow L^2(\Omega \times [0, T] \times \mathbb{R}_0)$ defined as

$$D_{t,z}F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t, z)), \quad F \in \mathcal{D}^{1,2}, \quad z \neq 0. \quad (2.82)$$

Since the operator D is closable, see, e.g., [DNØP⁺09, Thm. 3.3 and Thm 12.6], we denote by $\mathcal{D}^{1,2}$ the domain of its closure.

The following result, see [DNØP⁺09, Thm. 12.8] for the proof, represents a *chain rule* for Malliavin derivative.

Theorem 2.3.11. *Let $F \in \mathcal{D}^{1,2}$ and let ϕ be a real continuous function on \mathbb{R} . Suppose $\phi(F) \in L^2(\Omega)$ and $\phi(F + D_{t,z}F) \in L^2(\Omega \times [0, T] \times \mathbb{R}_0)$. Then, $\phi \in \mathcal{D}^{1,2}$ and*

$$D_{t,z}\phi(F) = \phi(F + D_{t,z}F) - \phi(F). \quad (2.83)$$

Once the Malliavin derivative has been defined, we are able to introduce its adjoint operator, the *Skorokhod integral*, in particular next definition is taken from [DNØP⁺09, Def. 11.1], see, also [Pet06, Sec. 3] for details.

Definition 2.3.12. Let $\delta : L^2(\Omega \times [0, T] \times \mathbb{R}_0) \rightarrow L^2(\Omega)$ be the adjoint operator of the derivative D . The set of processes $h \in L^2(\Omega \times [0, T] \times \mathbb{R}_0)$ such that

$$\left| \mathbb{E} \int_0^T \int_{\mathbb{R}_0} D_{s,z}F h_t(z) \nu(dz) ds \right| \leq C \|F\|,$$

for all $F \in \mathbb{D}^{1,2}$, form the domain of δ , denoted by $\text{dom } \delta$.

For every $h \in \text{dom } \delta$ we define the *Skorokhod integral* as

$$\delta(h) := \int_0^T \int_{\mathbb{R}_0} h_t(z) \tilde{N}(\hat{d}t, dz),$$

for any $F \in \mathbb{D}^{1,2}$.

Definition 2.3.13. We denote by $\mathbb{L}^{1,2}$ the space of \mathcal{F} -adapted processes $h : \Omega \times [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}$ such that $h_t \in \mathbb{D}^{1,2}$ and

$$\begin{aligned} \mathbb{E} \int_0^T \int_{\mathbb{R}_0} |h_t(z)| \nu(dz) dt &< \infty \\ \mathbb{E} \int_{([0, T] \times \mathbb{R}_0)^2} |D_{t,z} h_s(\zeta)| \nu(d\zeta) ds \nu(dz) dt &< \infty. \end{aligned}$$

From Definition above, We have $\mathbb{L}^{1,2} \subset \text{dom } \delta$. If $h \in \mathbb{L}^{1,2}$ and $D_{t,z} h \in \text{dom } \delta$, then $\delta(h) \in \mathbb{D}^{1,2}$ and

$$D_{t,z} \delta(h) = h(z) + \delta(D_{t,z} h), \quad (2.84)$$

see, e.g. [LSUV02]. Notice also that $\mathbb{L}^{1,2} \simeq L^2([0, T]; \mathbb{D}^{1,2})$.

Next result will be needed in order to prove subsequent results, see, e.g. [Pet06, Prop. 6].

Proposition 2.3.14. *Let h_t be a predictable square integrable process. Then, if $h \in \mathbb{D}^{1,2}$, we have*

$$\begin{aligned} D_{s,z} \int_0^t h_\tau d\tau &= \int_s^t D_{\tau,z} h_\tau d\tau, \\ D_{s,z} \int_0^t h_\tau dW(\tau) &= \int_s^t D_{\tau,z} h_\tau dW(\tau), \\ D_{s,z} \int_0^t \int_{\mathbb{R}_0} h_\tau \tilde{N}(d\tau, dz) &= h_s + \int_s^t \int_{\mathbb{R}_0} D_{\tau,\zeta} h_\tau \tilde{N}(d\tau, d\zeta). \end{aligned}$$

The next is the main result of the current subsection concerning the differentiability of the SFDDE (2.75).

Theorem 2.3.15. *Let us suppose that Assumptions 2.3.3 (A1)-(A2) hold and $X = \{X(t)\}_{t \in [-r, T]}$ is the solution to equation (2.75). Then, $X \in L^2([-r, T]; \mathbb{D}^{1,2})$ and, for every $s \in [0, T]$ and $z \in \mathbb{R}_0$, the stochastic process $\{D_{s,z} X(t) : t \in [s, T]\}$ satisfies*

$$\mathbb{E} \int_0^T \int_{\mathbb{R}_0} \sup_{t \in [s, T]} |D_{s,z} X(t)|^2 \nu(dz) ds < \infty. \quad (2.85)$$

In particular, for any $t \in [0, T]$, $X(t) \in \mathbb{D}^{1,2}$ and it holds

$$\begin{cases} D_{s,z} X(t) &= \gamma(s, X_s, X(s), z) + \int_s^t D_{s,z} \mu(u, X_u, X(u)) du \\ &+ \int_s^t D_{s,z} \sigma(u, X_u, X(u)) dW(u) + \int_s^t \int_{\mathbb{R}_0} D_{s,z} \gamma(u, X_u, X(u), \zeta) \tilde{N}(du, d\zeta), \\ D_s X(t) &= 0, \quad t \in [-r, s) \end{cases}$$

Moreover, for any $z \in \mathbb{R}_0$, there exists a measurable version of the two-parameter process

$$D_{s,z} X_t = \{D_{s,z} X_t(\theta) : s \in [0, T], \theta \in [-r, 0]\}.$$

Proof. We will use a standard Picard's approximation scheme, see, e.g. [DNØP⁺09, Th. 17.2]. Let $X^0(t) = x$ and $X_t^0 = \eta$, then set, for $n > 0$,

$$\begin{aligned} X^{n+1}(t) &= x + \int_0^t \mu(s, X_s^n, X^n(s))ds + \int_0^t \sigma(s, X_s^n, X^n(s))dW(s) \\ &\quad + \int_0^t \int_{\mathbb{R}_0} \gamma(s, X_s^n, X^n(s), z)\tilde{N}(ds, dz), \\ X_0^{n+1} &= \eta, \end{aligned}$$

where we use the notation $X_s^n := \{X^n(s + \theta)\}_{\theta \in [-r, 0]}$.

We are going to prove by induction over n that $X^n(t) \in \mathbb{D}^{1,2}$ for any $t \in [0, T]$, $D_{s,z}X(t)$ is a predictable process and that

$$\xi_{n+1}(t) \leq C_1 + C_2 \int_{-r}^t \xi_n(s)ds,$$

where C_1, C_2 are some suitable constants and

$$\xi_n(s) := \sup_{0 \leq s \leq t} \mathbb{E} \int_{\mathbb{R}_0} \sup_{s \leq \tau \leq t} |D_{s,z}X^n(\tau)|^2 \nu(dz) < \infty.$$

For $n = 0$ the above claim is trivially satisfied. Let us thus assume that the previous assumptions hold for n , we have to show that they hold also for $n + 1$. Indeed we have that $\int_0^t \mu(s, X_s^n, X^n(s))ds$, $\int_0^t \sigma(s, X_s^n, X^n(s))dW(s)$ and $\int_0^t \gamma(s, X_s^n, X^n(s), z)\tilde{N}(ds, dz) \in \mathbb{D}^{1,2}$, and proposition 2.3.14 guarantees that

$$\begin{aligned} D_{s,z} \int_0^t \mu(s, X_s^n, X^n(s))ds &= \int_s^t D_{\tau,z} \mu(\tau, X_\tau^n, X^n(\tau))d\tau \\ D_{s,z} \int_0^t \sigma(s, X_s^n, X^n(s))dW(s) &= \int_s^t D_{\tau,z} \sigma(\tau, X_\tau^n, X^n(\tau))dW(\tau) \end{aligned}$$

and

$$\begin{aligned} D_{s,z} \int_0^t \gamma(s, X_s^n, X^n(s), z)\tilde{N}(ds, dz) &= \gamma(s, X_s^n, X^n(s), z) \\ &\quad + \int_s^t \int_{\mathbb{R}_0} D_{\tau,\zeta} \gamma(\tau, X_\tau^n, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta) \end{aligned}$$

for $s \leq t$. Consequently, for any $t \in [0, T]$, $X^{n+1}(t) \in \mathbb{D}^{1,2}$ and

$$\begin{aligned} D_{s,z}X^{n+1}(t) &= \gamma(s, X_s^n, X^n(s), z) + \int_s^t D_{\tau,z} \mu(\tau, X_\tau^n, X^n(\tau))d\tau \\ &\quad + \int_s^t D_{\tau,z} \sigma(\tau, X_\tau^n, X^n(\tau))dW(\tau) \\ &\quad + \int_s^t \int_{\mathbb{R}_0} D_{\tau,\zeta} \gamma(\tau, X_\tau^n, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta). \end{aligned} \tag{2.86}$$

By squaring both sides of equation (2.86), we have

$$\begin{aligned} |D_{s,z}X^{n+1}(t)|^2 &\leq 4|\gamma(s, X_s^n, X^n(s), z)|^2 + \left| \int_s^t D_{\tau,z}\mu(\tau, X_\tau^n, X^n(\tau))d\tau \right|^2 \\ &\quad + 4 \left| \int_s^t D_{\tau,z}\sigma(\tau, X_\tau^n, X^n(\tau))dW(\tau) \right|^2 \\ &\quad + 4 \left| \int_s^t \int_{\mathbb{R}_0} D_{\tau,\zeta}\gamma(\tau, X_s^\tau, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta) \right|^2. \end{aligned} \quad (2.87)$$

By exploiting Doob maximal inequality, stochastic Fubini's theorem and Itô isometry, we get

$$\begin{aligned} \mathbb{E} \int_{\mathbb{R}_0} \sup_{s \leq \tau \leq t} |D_{s,z}X^{n+1}(t)|^2 \nu(dz) &\leq C \left[\mathbb{E} \int_{\mathbb{R}_0} |\gamma(s, X_s^n, X^n(s), z)|^2 \nu(dz) \right. \\ &\quad + \mathbb{E} \left| \int_s^t D_{\tau,z}\mu(\tau, X_\tau^n, X^n(\tau))d\tau \right|^2 + \mathbb{E} \left| \int_s^t D_{\tau,z}\sigma(\tau, X_\tau^n, X^n(\tau))dW(\tau) \right|^2 \\ &\quad \left. + \mathbb{E} \left| \int_s^t \int_{\mathbb{R}_0} D_{\tau,\zeta}\gamma(\tau, X_s^\tau, X^n(\tau), \zeta)\tilde{N}(d\tau, d\zeta) \right|^2 \right] \\ &\leq C \left[\mathbb{E} \int_{\mathbb{R}_0} |\gamma(s, X_s^n, X^n(s), z)|^2 \nu(dz) \right. \\ &\quad + \mathbb{E} \int_s^t |D_{\tau,z}\mu(\tau, X_\tau^n, X^n(\tau))|^2 d\tau + \mathbb{E} \int_s^t |D_{\tau,z}\sigma(\tau, X_\tau^n, X^n(\tau))|^2 d\tau \\ &\quad \left. + \mathbb{E} \int_s^t \int_{\mathbb{R}_0} |D_{\tau,\zeta}\gamma(\tau, X_s^\tau, X^n(\tau), \zeta)|^2 \nu(dz)d\tau \right], \end{aligned} \quad (2.88)$$

where we denote for short by $C > 0$ a suitable constant.

Exploiting Assumptions 2.3.3 together with Theorem 2.3.11, we get

$$\begin{aligned} \mathbb{E} \int_{\mathbb{R}_0} \sup_{s \leq \tau \leq t} |D_{s,z}X^{n+1}(\tau)|^2 \nu(dz) &\leq \\ &\leq C_1 \int_s^t \mathbb{E} \int_{\mathbb{R}_0} |(D_{s,z}X_\tau^n, D_{s,z}X^n(\tau))|_2^2 \nu(dz)d\tau + C_2 (1 + \mathbb{E}|(X_\tau^n, X^n(\tau))|_2^2) \\ &\leq C_1 \left(\mathbb{E} \int_s^t \int_{-r}^0 \int_{\mathbb{R}_0} |D_{s,z}X^n(\tau + \theta)|^2 \nu(dz)d\theta d\tau + \mathbb{E} \int_s^t \int_{\mathbb{R}_0} |D_{s,z}X^n(\tau)|^2 \nu(dz)d\tau \right) + C_3(1 + \lambda) \\ &\leq C_1 \left(\mathbb{E} \int_{-r}^0 \int_s^{t+\theta} \int_{\mathbb{R}_0} |D_{s,z}X^n(p)|^2 \nu(dz)dpd\theta + \mathbb{E} \int_s^t \int_{\mathbb{R}_0} |D_{s,z}X^n(\tau)|^2 \nu(dz)d\tau \right) + C_3(1 + \lambda) \\ &\leq C_4 \mathbb{E} \int_s^t \int_{\mathbb{R}_0} |D_{s,z}X^n(\tau)|^2 \nu(dz)d\tau + C_3(1 + \lambda), \end{aligned} \quad (2.89)$$

where C_1, C_2, C_3 and C_4 denote some suitable constants and λ is such that

$$\lambda = \sup_n \mathbb{E} \sup_{-r \leq s \leq T} |X^n(s)|_2^2 < \infty.$$

What is more, we obtain

$$X^{n+1} = \{X^{n+1}(t)\}_{t \in [-r, T]} \in L^2(\Omega \times [-r, T]),$$

and for any t , $X^{n+1}(t) \in \mathbb{D}^{1,2}$, so that $X^{n+1} \in L^2(\Omega \times [-r, T]; \mathbb{D}^{1,2})$ and, for $p \leq s$, $D_{s,z}X^{n+1}(p) = 0$.

It follows that, for any $z \in \mathbb{R}_0$, it exists a measurable version of the two-parameter process

$$D_{s,z}X_t^{n+1} = \{D_{s,z}X_t^{n+1}(\theta) : s \in [0, T], \theta \in [-r, 0]\},$$

such that $D_{s,z}X_t^{n+1} \in L^2(\Omega \times [0, T] \times [-r, 0])$, see, e.g. [Moh98, Sec. 4].

Therefore, the inductive hypothesis is fulfilled by X^{n+1} and

$$\mathbb{E} \sup_{s \leq T} |X^n(s) - X(s)|^2 \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Finally, thanks to a discrete version of Gronwall's lemma, see, e.g. [DNØP⁺09, Th. 17.2] and applying equation (2.89), we have

$$\sup_{n \geq 0} \mathbb{E} \int_{-r}^T |D_{s,z}X^n(\tau)|^2 d\tau < \infty,$$

so that $X(t) \in \mathbb{D}^{1,2}$, see, e.g. [DNØP⁺09, Th. 17.2].

By repeating the same reasoning as before, we have

$$X = \{X(t)\}_{t \in [-r, T]} \in L^2(\Omega \times [-r, T]), \quad X(t) \in \mathbb{D}^{1,2},$$

for any t , so that $X \in L^2(\Omega \times [-r, T]; \mathbb{D}^{1,2})$. The proof is complete by observing that, for any $z \in \mathbb{R}_0$, there exists a measurable version of the two-parameter process

$$D_{s,z}X_t = \{D_{s,z}X_t(\theta) : s \in [0, T], \theta \in [-r, 0]\},$$

such that $D_{s,z}X_t \in L^2(\Omega \times [0, T] \times [-r, 0])$. □

2.3.3 Joint quadratic variation

In order to prove the main result of this work, which consists in giving an explicit *Feynman-Kac* representation formula for a coupled forward-backward system with delay, we need first to prove a joint quadratic variation result. The main advantage of such an approach is to overcome difficulties that may arise in dealing with the Itô's formula in infinite dimension, since, in general, the process X_t fails to be a semi-martingale, so we cannot rely in standard Itô calculus. Furthermore, we are able to relax hypothesis concerning the differentiability of the coefficients.

In particular, following [FMT10, FT05], we introduce a *generalized covariation* process. The definition of joint generalized quadratic variation we consider in the present paper has been first introduced in [RV95] and [RV93, RV96b] with the difference that they consider the limit to hold uniformly on compacts in probability. We have chosen here, following [FMT10, FT05], to consider the limit in probability because that the limiting procedure is easier with a stronger notion of convergence, such as the convergence in probability.

Definition 2.3.16. Given a couple of \mathbb{R} -valued stochastic processes $(X(t), Y(t))$, $t \geq 0$, their *joint quadratic variation* on $[0, T]$, is given by

$$\langle X(t), Y(t) \rangle_{[0, T']} := p - \lim_{\epsilon \downarrow 0} C_{[0, T']}^\epsilon(X(t), Y(t)) ,$$

where $p - \lim$ denotes the limit to be taken in probability and

$$C_{[0, T']}^\epsilon(X(t), Y(t)) := \frac{1}{\epsilon} \int_0^{T'} (X(t + \epsilon) - X(t))(Y(t + \epsilon) - Y(t)) dt, \quad \epsilon > 0, \quad (2.90)$$

with $0 \leq T' + \epsilon < T$.

It is shown in [RV95, Prop. 1.1] the definition of joint quadratic variation coincides with the standard notion of quadratic variation, see, e.g. [App09].

Before stating our main result we are to better introduce a mild notion of derivative we will use throughout the paper. In what follows we will consider a function $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$, there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0, T]$ and any $(\eta_1, x_1), (\eta_2, x_2) \in \mathcal{D}$, u satisfies

$$\begin{aligned} |u(t, \eta_1, x_1) - u(t, \eta_2, x_2)| &\leq C |(\eta_1, x_1) - (\eta_2, x_2)|_2 (1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2)^m, \\ |u(t, 0, 0)| &\leq C. \end{aligned} \quad (2.91)$$

that is we require the function u to be Lipschitz continuous without requiring any further regularity concerning differentiability. Nevertheless, in what follows, we will use the notation of ∂_x^σ . In particular following [FT05] we will introduce a mild notion of derivative, that is the *generalized directional gradient* ∂_x^σ . When u is sufficiently regular, it can be shown that the *generalized directional gradient*, in the direction $\sigma(t, \eta, x)$, of a function u , coincides with $\partial_x u(t, \eta, x) \sigma(t, \eta, x)$.

The definition, as well as the characterization of several properties, for the *generalized directional gradient* has been provided in [FT05]. We will only state the definition of *generalized directional gradient*, whereas we refer to [FT05] to a complete treatment of the topic. The definition of *generalized directional gradient* relies on the definition of *generalized quadratic variation* of the function u with the Brownian motion W , introduced in [RV93, RV95], and denoted by $\langle u(\cdot, X_\cdot, X(\cdot), W(\cdot)) \rangle_{\tau, t}$, $0 \leq \tau < t \leq T$.

In particular it has been shown in [FT05] that the following holds

$$\langle u(\cdot, X_\cdot, X(\cdot), W(\cdot)) \rangle_{\tau, t} = \int_\tau^t \zeta(s, X_s, X(s)) ds, \quad (2.92)$$

for $\zeta : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ a suitable measurable map, see also [FMT10, FT02, FT05] for details. Under suitable hypothesis of regularity, in [FT05] the authors show that

$$\langle u(\cdot, X_\cdot, X(\cdot), W(\cdot)) \rangle_{\tau, t} = \int_\tau^t \partial_x u(t, X_s, X(s)) \sigma(t, X_s, X(s)) ds, \quad \mathbb{P} - a.s., \quad (2.93)$$

where we denote by ∂_x the derivative w.r.t. the present state. Hence, equation (2.92) can be considered as the definition of the *generalized directional gradient* of the function u along the direction σ . We say that the map $\zeta : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ belongs to the directional gradient

of u , or equivalently that $\zeta \in \partial_x^\sigma$, if equation (2.92) holds. Therefore, we use for short the notation $\partial_x^\sigma u$ to represent an element of the *generalized directional gradient*. Since this topic lies outside our goals, having been deeply studied in a more general setting in [FT05], we skip every technicality and invite the interested reader to refer to [FT05].

The following result represents the core of this paper.

Theorem 2.3.17. *Let us assume that $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ is locally Lipschitz w.r.t. the second variable and with at most polynomial growth, namely, there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0, T]$ and any $(\eta_1, x_1), (\eta_2, x_2) \in \mathcal{D}$, u satisfies*

$$\begin{aligned} |u(t, \eta_1, x_1) - u(t, \eta_2, x_2)| &\leq C |(\eta_1, x_1) - (\eta_2, x_2)|_2 (1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2)^m, \\ |u(t, 0, 0)| &\leq C. \end{aligned} \quad (2.94)$$

Then, for every $(\eta, x) \in \mathcal{D}$ and $0 \leq \tau \leq T' \leq T$, the process

$$\{u(t, X_t^{(\tau, \eta, x)}), X^{\tau, \eta, x}(t), t \in [\tau, T']\}$$

admits a joint quadratic variation on the interval $[\tau, T']$ with

$$J(t) := \int_0^t \int_{\mathbb{R}_0} z \tilde{N}(ds, dz),$$

given by

$$\begin{aligned} \left\langle u(\cdot, X^{(\tau, \eta, x)}, X^{\tau, \eta, x}(\cdot)), J(\cdot) \right\rangle_{[\tau, T']} &= \int_\tau^{T'} \int_{\mathbb{R}_0} z \left[u(s, X_s^{(\tau, \eta, x)}, X^{\tau, \eta, x}(s)) \right. \\ &\quad \left. + \gamma(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s), z) - u(s, X_s^{(\tau, \eta, x)}, X^{\tau, \eta, x}(s)) \right] N(ds, dz). \end{aligned} \quad (2.95)$$

Remark 2.3.18. An analogous of [FT05, Prop. 4.4] is valid in the present case, that is the following representation holds

$$\left\langle u(\cdot, X^{\tau, \eta, x}, X^{(\tau, \eta, x)}(\cdot)), W(\cdot) \right\rangle_{[\tau, T']} = \int_\tau^{T'} \partial_x^\sigma u(s, X_s^{(\tau, \eta, x)}, X^{\tau, \eta, x}(s)) ds,$$

where ∂_x^σ is the *generalized directional gradient*. The claim follows from [FT05] by observing that the Poisson random measure does not affect the result and the proof follows exactly the same steps as in [FT05].

Proof. Without loss of generality, we prove the result with $\tau = 0$, as the case of a general initial time $\tau \neq 0$ can be proved by using the same techniques. Fix $(\eta, x) \in \mathcal{D}$ and a time horizon $T' \in [0, T]$ and denote for brevity $X^{(0, \eta, x)}$ by X . In what follows we will denote with $\tilde{N}(\hat{dt}, dz)$ the Skorokhod integral.

In order to shorten the notation set

$$v_t := (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) \mathbb{1}_{[0, T']}(t),$$

and

$$A^\epsilon := \{(t, s) \in [0, T'] \times [0, T'] : 0 \leq t \leq T', t \leq s \leq t + \epsilon\}.$$

Thanks to equation (2.94) and theorem 2.3.15, we have $v_t \in \mathbb{L}^{1,2}$, so that, for any t , $v_t \in \mathcal{D}^{1,2}$ and then $v_t \mathbb{1}_{A^\epsilon}(t, \cdot) \in L^2(\Omega \times [0, T])$. Furthermore, equation (2.84) implies that v_t is Skorokhod integrable and from [DNØP⁺09, Th. 12.11] we have

$$\begin{aligned} \int_0^{T'} \int_{\mathbb{R}_0} z v_t \mathbb{1}_{A^\epsilon}(t, s) \tilde{N}(\hat{d}t, dz) &= v_t \int_0^{T'} \int_{\mathbb{R}_0} z \mathbb{1}_{A^\epsilon}(t, s) \tilde{N}(\hat{d}t, dz) \\ &\quad - \int_0^{T'} \int_{\mathbb{R}_0} z D_{s,z} v_t \mathbb{1}_{A^\epsilon}(t, s) N(ds, dz) =: z_t, \end{aligned} \quad (2.96)$$

which holds since $z \in L^2(\Omega \times [0, T])$.

equation (2.96) implies, for a.a. $t \in [0, T']$,

$$\begin{aligned} &u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))(J_{t+\epsilon} - J_t) \\ &= u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t)) \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z \tilde{N}(ds, dz) \\ &= \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z D_{s,z} (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) N(ds, dz) \\ &\quad + \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) \tilde{N}(ds, dz). \end{aligned} \quad (2.97)$$

Let us integrate the right-hand side of equation (2.97) in $[0, T']$ w.r.t. t . By noticing that the left-hand side equals to ϵC^ϵ , we write the right-hand side as follows

$$\begin{aligned} &\int_0^{T'} \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z D_{s,z} (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) N(ds, dz) dt \\ &+ \int_0^{T'} \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) \tilde{N}(\hat{d}s, dz) dt \\ &= \int_0^{T'} \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z D_{s,z} (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) N(ds, dz) dt \\ &+ \int_0^{T'+\epsilon} \int_{\mathbb{R}_0} \int_{(s-\epsilon)^+}^{s \wedge T'} z (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) dt \tilde{N}(\hat{d}s, dz). \end{aligned} \quad (2.98)$$

It remains to verify that $\int_0^{T'} z v_t \mathbb{1}_{A^\epsilon}(t, \cdot) dt$ appearing in equation (2.98) is Skorokhod integrable. From the definition of Skorokhod integral, by using equation (2.96) for $G \in \mathcal{D}^{1,2}$ and the *duality formula*, see e.g. [DNØP⁺09, equation (12.14)], we have

$$\begin{aligned} &\mathbb{E} \int_0^T \int_{\mathbb{R}_0} \int_0^T z v_t \mathbb{1}_{A^\epsilon}(t, s) dt D_{s,z} G \nu(dz) ds \\ &= \int_0^T \mathbb{E} \int_0^T \int_{\mathbb{R}_0} z v_t \mathbb{1}_{A^\epsilon}(t, s) D_{s,z} G \nu(dz) ds dt \\ &= \int_0^T \mathbb{E} G \int_0^T \int_{\mathbb{R}_0} z v_t \mathbb{1}_{A^\epsilon}(t, s) D_{s,z} \tilde{N}(\hat{d}s, dz) dt = \mathbb{E} G \int_0^T z_t dt, \end{aligned}$$

so that $\int_0^{T'} v_t \mathbb{1}_{A^\epsilon}(t, \cdot) dt$ is Skorokhod integrable. Hence,

$$\begin{aligned} & \int_0^T \int_0^T \int_{\mathbb{R}_0} z v_t \mathbb{1}_{A^\epsilon}(t, s) dt \tilde{N}(\hat{ds}, dz) \\ &= \int_0^T z_t dt = \int_0^T \int_0^T \int_{\mathbb{R}_0} z v_t \mathbb{1}_{A^\epsilon}(t, s) \tilde{N}(\hat{ds}, dz) dt. \end{aligned}$$

Exploiting again equation (2.96) we have

$$\begin{aligned} & \int_0^T \int_0^{T'} \int_{\mathbb{R}_0} z v_t \mathbb{1}_{A^\epsilon}(t, s) dt \tilde{N}(\hat{ds}, dz) = \int_0^T z v_t (J_{t+\epsilon} - J_t) \mathbb{1}_{[t, T]}(t) dt \\ & \quad - \int_0^{T'} \int_0^T \int_{\mathbb{R}_0} z D_{s,z} v_t \mathbb{1}_{A^\epsilon}(t, s) N(ds, dz) dt, \end{aligned}$$

and then equation (2.98) is proved.

On the other hand, thanks to the chain rule 2.3.11, see also [DNØP⁺09, Th. 12.8], and from theorem 2.3.15 together with the adeptness property of the Malliavin derivative, i.e. $D_{s,z}X(t) = 0$ if $s > t$, we have that, for a.a. $s \in [t, t + \epsilon]$,

$$\begin{aligned} D_{s,z}v_t &= D_{s,z}[u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))] \\ &= D_{s,z}[u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon))] \\ &= u(t + \epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t + \epsilon) + D_{s,z}X(t + \epsilon)) \\ & \quad - u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)). \end{aligned}$$

Now, we apply equation (2.98) and we get

$$\begin{aligned} C^\epsilon &= \frac{1}{\epsilon} \int_0^{T'} \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z \left[u(t + \epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t + \epsilon) + D_{s,z}X(t + \epsilon)) N(\hat{ds}, dz) dt \right. \\ & \quad - \frac{1}{\epsilon} \int_0^{T'} \int_t^{t+\epsilon} u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) N(\hat{ds}, dz) dt \\ & \quad \left. + \frac{1}{\epsilon} \int_0^{T'+\epsilon} \int_{\mathbb{R}_0} \int_{(s-\epsilon)^+}^{s \wedge T'} z (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) dt \tilde{N}(\hat{ds}, dz) \right]. \end{aligned}$$

Let us consider separately the two terms

$$\begin{aligned} I_1^\epsilon &:= \frac{1}{\epsilon} \int_0^{T'} \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z \left[u(t + \epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t + \epsilon) + D_{s,z}X(t + \epsilon)) N(\hat{ds}, dz) dt \right. \\ & \quad \left. - \frac{1}{\epsilon} \int_0^{T'} \int_t^{t+\epsilon} u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) N(\hat{ds}, dz) dt \right], \\ I_2^\epsilon &:= \frac{1}{\epsilon} \int_0^{T'+\epsilon} \int_{\mathbb{R}_0} \int_{(s-\epsilon)^+}^{s \wedge T'} z (u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) - u(t, X_t, X(t))) dt \tilde{N}(\hat{ds}, dz). \end{aligned}$$

As regards I_2^ϵ , the proof proceed exactly as in [FT05, Prop. 4.4.] or in [FMT10, Th. 3.1], for the sake of completeness we state the proof. We have to show that

$$\frac{1}{\epsilon} \int_0^{T'} v_t \mathbb{1}_{[A^\epsilon]}(t, s) dt \rightarrow 0,$$

in $\mathbb{L}^{1,2}$, since this implies $I_2^\epsilon \rightarrow 0$ in $L^2(\Omega)$, together with the boundedness of the Skorokhod integral. Thus, for a general $y \in \mathbb{L}^{1,2}$, we have

$$T^\epsilon(y)_s = \frac{1}{\epsilon} \int_0^{T'} (y_{t+\epsilon} - y_t) \mathbb{1}_{[A^\epsilon]}(t, s) dt = \frac{1}{\epsilon} \int_{(s-\epsilon) \vee t}^{s \wedge T} (y_{t+\epsilon} - y_t) dt,$$

so that we have to show that $T^\epsilon(y) \rightarrow 0$ in $\mathbb{L}^{1,2}$.

Let us recall the isomorphism

$$L^2([0, T]; \mathbb{D}^{1,2}(\mathbb{R})) \simeq \mathbb{L}^{1,2}.$$

Following [FT05], we have to show that $\|T^\epsilon\|_{\mathbb{L}^{1,2}(\mathbb{R})}$ is bounded uniformly w.r.t. ϵ . In fact,

$$\begin{aligned} \|T^\epsilon(y)_s\|_{\mathbb{D}^{1,2}(\mathbb{R})}^2 &\leq \frac{1}{\epsilon^2} \int_0^{T'} \mathbb{1}_{[A^\epsilon]}(t, s) dt \int_0^{T'} |y_{t+\epsilon} - y_t|_{\mathbb{D}^{1,2}(\mathbb{R})}^2 \mathbb{1}_{[A^\epsilon]}(t, s) dt \\ &\leq \int_0^{T'} |y_{t+\epsilon} - y_t|_{\mathbb{D}^{1,2}(\mathbb{R})}^2 \mathbb{1}_{[A^\epsilon]}(t, s) dt, \\ \|T^\epsilon(y)_s\|_{\mathbb{L}^{1,2}(\mathbb{R})}^2 &= \int_0^{T'} \|T^\epsilon(y)_s\|_{\mathbb{D}^{1,2}(\mathbb{R})}^2 ds \\ &\leq \int_0^{T'} |y_{t+\epsilon} - y_t|_{\mathbb{D}^{1,2}(\mathbb{R})}^2 \int_0^{T'} \mathbb{1}_{[A^\epsilon]}(t, s) ds dt \\ &\leq \int_0^{T'} |y_{t+\epsilon} - y_t|_{\mathbb{D}^{1,2}(\mathbb{R})}^2 dt \leq 2\|y\|_{\mathbb{L}^{1,2}(\mathbb{R})}^2, \end{aligned}$$

so the claim follows, see, e.g. [FT05, Prop. 4.4.] or [FMT10, Th. 3.1].

As regards I_1^ϵ , we have

$$\begin{aligned} I_1^\epsilon &= \frac{1}{\epsilon} \int_0^T \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z u(t + \epsilon, X_{t+\epsilon} + D_{s,z} X_{t+\epsilon}, X(t + \epsilon) + D_{s,z} X(t + \epsilon)) N(\hat{d}s, dz) dt \\ &\quad - \frac{1}{\epsilon} \int_0^T \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) N(\hat{d}s, dz) dt := K_1^\epsilon - K_2^\epsilon. \end{aligned}$$

Let us first prove that

$$K_2^\epsilon \rightarrow \int_0^{T'} \int_{\mathbb{R}_0} z u(t, X_t, X(t)) N(\hat{d}t, dz), \quad \mathbb{P} - a.s. \quad (2.99)$$

as $\epsilon \rightarrow 0$.

From assumption (2.94) on the function u and the right-continuity of X , it follows from the Lebesgue differentiation theorem and the dominated convergence theorem that

$$\begin{aligned} &\frac{1}{\epsilon} \int_0^{T'} \int_t^{t+\epsilon} \int_{\mathbb{R}_0} z u(t + \epsilon, X_{t+\epsilon}, X(t + \epsilon)) N(\hat{d}s, dz) dt \\ &= \int_0^{T'+\epsilon} \int_{\mathbb{R}_0} z \frac{1}{\epsilon} \int_{s \vee \epsilon}^{(s+\epsilon) \wedge T'} u(t, X_t, X(t)) dt N(\hat{d}s, dz) \\ &\rightarrow \int_0^{T'} \int_{\mathbb{R}_0} z u(s, X_s, X(s)) N(\hat{d}s, dz), \end{aligned} \quad (2.100)$$

\mathbb{P} -a.s., as $\epsilon \rightarrow 0$.

Let us now prove that

$$K_1^\epsilon \rightarrow \int_0^{T'} \int_{\mathbb{R}_0} zu(t, X_t, X(t) + \gamma(t, X_t, z))N(\hat{dt}, dz). \quad (2.101)$$

Theorem 2.3.15 assures

$$\begin{aligned} D_{s,z}X(t + \epsilon) &= \gamma(s, X_s, X(s), z) + \int_s^{t+\epsilon} D_{s,z}[\mu(q, X_q, X(q))]dq \\ &\quad + \int_s^{t+\epsilon} D_{s,z}[\sigma(q, X_q, X(q))]dW(q) \\ &\quad + \int_s^{t+\epsilon} \int_{\mathbb{R}_0} D_{s,z}[\gamma(q, X_q, X(q)\zeta)]\tilde{N}(dq, d\zeta). \end{aligned} \quad (2.102)$$

Proceeding as above, we get

$$\begin{aligned} &\frac{1}{\epsilon} \int_0^T \int_t^{t+\epsilon} \int_{\mathbb{R}_0} zu(t + \epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon})N(\hat{ds}, dz)dt \\ &= \int_0^{T'+\epsilon} \int_{\mathbb{R}_0} z \frac{1}{\epsilon} \int_{(s-\epsilon)^+}^{s \wedge T'} u(t + \epsilon, X_{t+\epsilon} + D_{s,z}X_{t+\epsilon}, X(t + \epsilon) \\ &\quad + D_{s,z}X(t + \epsilon)) dt N(\hat{ds}, dz). \end{aligned} \quad (2.103)$$

The continuity of u , together with the right-continuity of X , and the Lebesgue differentiation theorem provide that

$$\begin{aligned} &\int_\epsilon^{T'+\epsilon} \int_{\mathbb{R}_0} z \frac{1}{\epsilon} \int_{s \vee \epsilon}^{(s+\epsilon) \wedge T'} u(t, X_t + D_{s,z}X_t, X(t) + D_{s,z}X(t)) dt N(\hat{ds}, dz) \\ &\rightarrow \int_0^{T'} \int_{\mathbb{R}_0} zu(t, X_t + D_{t,z}X_t, X(t) + D_{t,z}X(t))N(\hat{dt}, dz). \end{aligned} \quad (2.104)$$

Moreover from theorem 2.3.15, together with the adaptedness of the Malliavin derivative, it follows that

$$\begin{aligned} D_{s,z}X(t + \theta) &= \gamma(s, X_s, z) + \int_s^{t+\theta} D_{s,z}[\mu(q, X_q)]dq \\ &\quad + \int_s^{t+\theta} D_{s,z}[\sigma(q, X_q)]dW(q) \\ &\quad + \int_s^{t+\theta} \int_{\mathbb{R}_0} D_{s,z}[\gamma(q, X_q, \zeta)]\tilde{N}(dq, d\zeta), \quad \theta \in [-r, 0], \\ D_{s,z}X(t + \theta) &= 0, \quad s > t + \theta. \end{aligned}$$

Exploiting thus the adaptedness of the Malliavin derivative, that is

$$\begin{aligned} D_{t,z}X_t(\theta) &= D_{t,z}X(t + \theta) = 0, \quad \text{for } \theta \in [-r, 0], \\ D_{t,z}X(t) &= \gamma(t, X_t, X(t), z), \end{aligned} \quad (2.105)$$

so that substituting equation (2.105) into equation (2.104) we then get the claim and then equation (2.101) is proved. equation (2.95) thus follows and the proof is then complete. \square

2.3.4 Existence of mild solutions of Kolmogorov equation

The main goal of this section is to provide an existence and uniqueness result of a *mild* solution, whose meaning will be specified later, of a non-linear path-dependent partial integro-differential equation. Such a solution is connected to a forward-backward system with delay of the form

$$\left\{ \begin{array}{l} dX^{\tau,\eta,x}(t) \\ (X^{\tau,\eta,x}, X^{\tau,\eta,x}(\tau)) \\ dY^{\tau,\eta,x}(t) \\ Y^{\tau,\eta,x}(T) \end{array} \right. = \begin{array}{l} \mu(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))dt + \sigma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))dW(s) \\ \quad + \int_{\mathbb{R}_0} \gamma(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), z)\tilde{N}(dt, dz) \\ (\eta, x) \in \mathcal{D} \\ \psi \left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), Y^{\tau,\eta,x}(t), Z^{\tau,\eta,x}(t), \tilde{U}^{\tau,\eta,x}(t) \right) dt \\ \quad + Z^{\tau,\eta,x}(t)dW(t) + \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t, z)\tilde{N}(dt, dz) \\ \phi(X_T^{\tau,\eta,x}, X^{\tau,\eta,x}(T)) \end{array}, \quad (2.106)$$

where we have set for short

$$\tilde{U}^{\tau,\eta,x}(t) := \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t, z)\delta(z)\nu(dz).$$

In particular the solution to the SFDDE (2.156) is the quadruple (X, Y, Z, U) taking values in $\mathcal{D} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. We refer to [Del13] for a detailed introduction to forward-backward system with jumps.

We now assume the following assumptions to hold.

Hypothesis 2.3.19.

(B1) The map $\psi : [0, T] \times \mathcal{D} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ is continuous and there exists $K > 0$ and $m \geq 0$ such that

$$\begin{aligned} & |\psi(t, \eta_1, x_1, y_1, z_1, u_1) - \psi(t, \eta_2, x_2, y_2, z_2, u_2)| \leq K|(\eta_1, x_1) - (\eta_2, x_2)|_2 \\ & \quad + K(|y_1 - y_2| + |z_1 - z_2| + |u_1 - u_2|); \\ & |\psi(t, \eta_1, x_1, y, z, u) - \psi(t, \eta_2, x_2, y, z, u)| \\ & \quad \leq K(1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2 + |y|)^m \\ & \quad \cdot (1 + |z| + |u|)(|(\eta_1, x_1) - (\eta_2, x_2)|_2); \\ & |\psi(t, 0, 0, 0, 0, 0)| \leq K, \end{aligned}$$

for all $(t, \eta_1, x_1, y_1, z_1, u_1), (t, \eta_2, x_2, y_2, z_2, u_2) \in [0, T] \times \mathcal{D} \times \mathbb{R}^3$;

(B2) the map $\phi : \mathcal{D} \rightarrow \mathbb{R}$ is measurable and there exist $K > 0$ and $m \geq 0$ such that

$$|\phi(\eta_1, x_1) - \phi(\eta_2, x_2)| \leq K(1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2)^m |(\eta_1, x_1) - (\eta_2, x_2)|_2,$$

for all $(\eta_1, x_1), (\eta_2, x_2) \in \mathcal{D}$;

(B3) there exists $K > 0$ such that the function $\delta : \mathbb{R}_0 \rightarrow \mathbb{R}$ satisfies

$$|\delta(z)| \leq K|(1 \wedge |z|), \quad \delta(z) \geq 0, \quad z \in \mathbb{R}_0.$$

Remark 2.3.20. Following [Del13], we have chosen this particular form for the generator ψ of the backward component in equation (2.156), due to the fact that it results to be convenient in many applications.

Remark 2.3.21. We want to stress that assumptions 2.3.19 imply that there exists a suitable constant $C > 0$ such that

$$\begin{aligned} |\psi(t, \eta, x, y, z, u)| &\leq C(1 + |(\eta, x)|_2^{m+1} + |y| + |z| + |u|), \\ |\phi(\eta, x)| &\leq C(1 + |(\eta, x)|_2^{m+1}). \end{aligned}$$

In what follows we will denote by $\mathbb{K}([0, T])$ the space of all triplet (Y, Z, U) of predictable stochastic processes taking value in $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$ and such that

$$\begin{aligned} \|(Y, Z, U)\|_{\mathbb{K}}^2 &:= \mathbb{E} \left[\sup_{t \in [0, T]} |Y(t)|^2 \right] + \mathbb{E} \left[\int_0^T |Z(t)|^2 d\tau \right] \\ &+ \mathbb{E} \left[\int_0^T \int_{\mathbb{R}_0} |U(t, z)|^2 \nu(dz) dt \right] < \infty. \end{aligned} \quad (2.107)$$

The following Proposition ensures the existence and the uniqueness of the solution to the system (2.156), under suitable properties of the coefficients.

Proposition 2.3.22. *Let us consider the coupled forward-backward system (2.156) which satisfies Assumptions 2.3.3 and Assumptions 2.3.19.*

Then, the coupled forward-backward system admits a unique solution

$$(X^{\tau, \eta, x}, Y^{\tau, \eta, x}, Z^{\tau, \eta, x}, U^{\tau, \eta, x}) \in \mathcal{S}^p \times \mathbb{K}([0, T]).$$

Eventually we have that the map

$$(\tau, \eta, x) \mapsto (X^{\tau, \eta, x}, Y^{\tau, \eta, x}, Z^{\tau, \eta, x}, U^{\tau, \eta, x}),$$

is continuous.

Proof. The existence and uniqueness of the solution to the forward component follows from theorem 2.3.5, since Assumptions 2.3.3 hold true by hypothesis, whereas for the existence and uniqueness of the backward component under Assumptions 2.3.19 follows [BBP97, Cor. 2.3] or [Del13, Thm. 4.1.3].

The continuity of the map $(\tau, \eta, x) \mapsto X^{\tau, \eta, x}$ is guaranteed by theorem 2.3.5, whereas the continuity of $(\tau, \eta, x) \mapsto (Y^{\tau, \eta, x}, Z^{\tau, \eta, x}, U^{\tau, \eta, x})$ follows from [BBP97, Prop. 1.1]. \square

Theorem 2.3.23. *Let us consider the coupled forward-backward system (2.156) which satisfies Assumptions 2.3.3 and 2.3.19. Let us define the function $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$,*

$$u(t, \eta, x) := Y_t^{t, \eta, x},$$

with $t \in [0, T]$ and $(\eta, x) \in \mathcal{D}$.

Then, there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0, T]$ and any $(\eta_1, x_1), (\eta_2, x_2) \in \mathcal{D}$, u satisfies

$$\begin{aligned} |u(t, \eta_1, x_1) - u(t, \eta_2, x_2)| &\leq C(|(\eta_1, x_1) - (\eta_2, x_2)|_2(1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2))^m, \\ |u(t, 0, 0)| &\leq C. \end{aligned} \quad (2.108)$$

Moreover, for every $t \in [0, T]$ and $(\eta, x) \in \mathcal{D}$ we have \mathbb{P} -a.s. and for a.e. $t \in [\tau, T]$

$$\begin{aligned} Y^{\tau, \eta, x}(t) &= u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)), \\ Z^{\tau, \eta, x}(t) &= \partial_x^\sigma u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)), \\ U^{\tau, \eta, x}(t, z) &= u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t) + \gamma(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t), z)) \\ &\quad - u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)), \end{aligned} \quad (2.109)$$

where ∂_x^σ is the generalized directional gradient in the sense of equation (2.92).

Remark 2.3.24. Let us recall that, if u is sufficiently regular, then

$$Z^{(\tau, \eta, x)}(t) = \partial_x u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)) \sigma(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)).$$

Proof. The fact that $u(t, \eta, x) := Y_t^{t, \eta, x}$ satisfies (2.108) immediately follows from the continuity of the map

$$(\tau, \eta, x) \mapsto (X^{\tau, \eta, x}, Y^{\tau, \tau, \eta, x}, Z^{\tau, \tau, \eta, x}, U^{\tau, \tau, \eta, x}),$$

proved in proposition 2.3.22 together with assumptions 2.3.3.

The representation of Y and Z follow from [FMT10, Cor. 4.3]. As regards the process U , we use the standard notion of joint variation

$$\langle Y^{\tau, \eta, x}(\cdot), J(\cdot) \rangle_{[\tau, T]} = \int_\tau^T \int_{\mathbb{R}_0} z U^{\tau, \eta, x}(s, z) N(ds, dz). \quad (2.110)$$

On the other hand, theorem 2.3.17 implies

$$\begin{aligned} &\langle u(\cdot, X^{\tau, \eta, x}, X^{\tau, \eta, x}(\cdot)), J(\cdot) \rangle_{[\tau, T]} \\ &= \int_\tau^T \int_{\mathbb{R}_0} z (u(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s) + \gamma(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s), z)) N(ds, dz) \\ &\quad - \int_\tau^T \int_{\mathbb{R}_0} z (u(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s))) N(ds, dz). \end{aligned} \quad (2.111)$$

By comparing equation (2.110) and equation (2.111), the thesis follows. \square

The non-linear Kolmogorov equation

The present section is devoted to prove that the solution to the forward-backward system (2.156) can be connected to a solution, in a suitable sense, to a of path-dependent partial integro-differential equations on the space \mathcal{D} , driven by a suitable generator.

More precisely, let us consider the Markov process $(X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t))$ defined as the solution of equation (2.75), and the corresponding Markov generator \mathbb{L}_t defined in (2.79).

The path-dependent partial-integro differential equation we want to investigate has the following form

$$\begin{cases} \frac{\partial}{\partial t} u(t, \eta, x) + \mathbb{L}_t u(t, \eta, x) = \psi(t, \eta, x, u(t, \eta, x), \partial_x^\sigma u(t, \eta, x), \mathcal{J}u(t, \eta, x)) , \\ u(T, \eta, x) = \phi(\eta, x), \end{cases} \quad (2.112)$$

for all $t \in [0, T]$, and $(\eta, x) \in \mathcal{D}$, where $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ is an unknown function, ψ and ϕ are two given functions such that $\psi : [0, T] \times \mathcal{D} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ and $\psi : \mathcal{D} \rightarrow \mathbb{R}$, $\partial_x^\sigma u$ is the *generalized directional gradient* and \mathcal{J} is a functional acting as

$$\mathcal{J}u(t, \eta, x) = \int_{\mathbb{R}_0} (u(t, \eta, x + \gamma(t, \eta, x, z)) - u(t, \eta, x)) \delta(z) \nu(dz).$$

In particular, we want to look for a *mild solution* of equation (2.151), according to the following definition.

Definition 2.3.25. Let us consider the partial integro-differential equation (2.151). A *mild solution* to equation (2.151) is a function $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ if there exist $C > 0$ and $m \geq 0$, such that, for any $t \in [0, T]$ and any $(\eta_1, x_1), (\eta_2, x_2) \in \mathcal{D}$, u satisfies

$$\begin{aligned} |u(t, \eta_1, x_1) - u(t, \eta_2, x_2)| &\leq C |(\eta_1, x_1) - (\eta_2, x_2)|_2 (1 + |(\eta_1, x_1)|_2 + |(\eta_2, x_2)|_2)^m \\ |u(t, 0, 0)| &\leq C \end{aligned} \quad (2.113)$$

and the following identity hold true

$$u(t, \eta, x) = P_{t, T} \phi(\eta, x) + \int_t^T P_{t, s} [\psi(\cdot, u(s, \cdot), \partial_x^\sigma u(s, \cdot), \mathcal{J}u(s, \cdot))](\eta, x) ds, \quad (2.114)$$

for all $t \in [0, T]$, and $(\eta, x) \in \mathcal{D}$ and where $P_{t, s}$ is the Markov semigroup for equation (2.67) introduced in equation (2.78).

Theorem 2.3.26. Assume that Assumptions 2.3.3 and Assumptions 2.3.19 hold true. Then, the path-dependent partial integro-differential equation (2.151) admits a unique mild solution u , in the sense of definition 2.3.25. In particular, the mild solution u coincide with the function u introduced in theorem 2.3.23.

Proof. In what follows, as above, we will denote for short

$$\tilde{U}^{\tau, \eta, x}(s) := \int_{\mathbb{R}_0} U^{\tau, \eta, x}(s, z) \delta(z) \nu(dz),$$

Let us consider the backward stochastic differential equation in equation (2.156), namely,

$$\begin{aligned} Y^{t, \eta, x}(t) &= \phi(X_T^{t, \eta, x}, X^{t, \eta, x}(T)) + \int_t^T \psi \left(X_s^{t, \eta, x}, X^{t, \eta, x}(s), Y^{t, \eta, x}(s), Z^{t, \eta, x}(s), \tilde{U}^{t, \eta, x}(s) \right) ds \\ &\quad + \int_t^T Z^{t, \eta, x}(s) dW(s) + \int_t^T \int_{\mathbb{R}_0} U^{t, \eta, x}(s, z) \tilde{N}(ds, dz). \end{aligned}$$

Taking the expectation and exploiting equation (2.152), then Y satisfies equation (2.153).

In order to show the uniqueness let $u(t, \eta, x)$, $0 \leq \tau \leq t \leq T$, be a mild solution of equation (2.151), so that

$$\begin{aligned} u(t, \eta, x) &= \mathbb{E} [\phi(X_T^{t, \eta, x}, X^{t, \eta, x}(T))] \\ &+ \mathbb{E} \left[\int_t^T \psi \left(X_s^{t, \eta, x}, X^{t, \eta, x}(s), Y^{t, \eta, x}(s), Z^{t, \eta, x}(s), \tilde{U}^{t, \eta, x}(s) \right) ds \right]. \end{aligned}$$

By recalling that $X_t^{(\tau, \eta, x)}$ is a Markov process, and denoting by \mathbb{E}^t the conditional expectation w.r.t. the filtration \mathbb{F}_t , we can write

$$\begin{aligned} u(t, X_t^{t, \eta, x}, X^{t, \eta, x}(t)) &= \mathbb{E}^t [\phi(X_T^{t, \eta, x}, X^{t, \eta, x}(T))] \\ &+ \mathbb{E}^t \left[\int_\tau^T \psi \left(X_s^{t, \eta, x}, X^{t, \eta, x}(s), Y^{t, \eta, x}(s), Z^{t, \eta, x}(s), \tilde{U}^{t, \eta, x}(s) \right) ds \right] \\ &- \mathbb{E}^t \left[\int_\tau^t \psi \left(X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s), Y^{\tau, \eta, x}(s), Z^{\tau, \eta, x}(s), \tilde{U}^{\tau, \eta, x}(s) \right) ds \right]. \end{aligned}$$

We set, for short,

$$\begin{aligned} \xi &:= \phi(X_T^{t, \eta, x}, X^{t, \eta, x}(T)) \\ &+ \int_\tau^T \psi \left(X_s^{t, \eta, x}, X^{t, \eta, x}(s), Y^{t, \eta, x}(s), Z^{t, \eta, x}(s), \tilde{U}^{t, \eta, x}(s) \right) ds. \end{aligned}$$

Thanks to the representation theorem of martingales, see [App09, Thm. 5.3.5], there exist two predictable processes $\bar{Z} \in L^2(\Omega \times [0, T])$ and $\bar{U} \in L^2(\Omega \times [0, T] \times \mathbb{R}_0)$ such that

$$\begin{aligned} u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)) &= u(\tau, \eta, x) \\ &+ \int_\tau^t \bar{Z}^{\tau, \eta, x}(s) dW(s) + \int_\tau^t \int_{\mathbb{R}_0} \bar{U}^{\tau, \eta, x}(s, z) \tilde{N}(ds, dz) \\ &- \int_\tau^t \psi \left(X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s), Y^{\tau, \eta, x}(s), Z^{\tau, \eta, x}(s), \tilde{U}^{\tau, \eta, x}(s) \right) ds. \end{aligned}$$

By applying theorem 2.3.17, we have

$$\begin{aligned} u(t, X_t^{\tau, \eta, x}, X^{\tau, \eta, x}(t)) &= \phi(X_T^{\tau, \eta, x}, X^{\tau, \eta, x}(T)) + \\ &- \int_t^T \sigma(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s)) \partial_x u(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s)) dW(s) \\ &- \int_t^T \int_{\mathbb{R}_0} [(u(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s)) + \gamma(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s), z), z)) \\ &- u(s, X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s))] \tilde{N}(ds, dz) \\ &+ \int_t^T \psi \left(X_s^{\tau, \eta, x}, X^{\tau, \eta, x}(s), Y^{\tau, \eta, x}(s), Z^{\tau, \eta, x}(s), \tilde{U}^{\tau, \eta, x}(s) \right) ds. \end{aligned}$$

By comparing last equation with the backward component of equation (2.156), we note that $(Y^{\tau,\eta,x}(t), Z^{\tau,\eta,x}(t), U^{\tau,\eta,x}(t, z))$ and the following three functions

$$\begin{aligned} & u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)), \\ & \partial_x^\sigma(u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))), \\ & u(t, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s) + \gamma(s, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s), z)) - u(t, X_s^{\tau,\eta,x}, X^{\tau,\eta,x}(s)) \end{aligned}$$

solve the same equation. Therefore, due to the uniqueness of the solution, we have that

$$Y^{\tau,\eta,x}(t) = u(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)).$$

Setting $\tau = t$, we obtain $Y^{(t,\eta,x)}(t) = u(t, \eta, x)$ and the proof is complete. \square

2.4 Financial applications

During recent years an increasing attention has been devoted to the study of stochastic differential equations with delay (SDDEs), which are equations defining the dynamic of a stochastic process whose time evolution depends on the past history of the process itself. Delay equations arise naturally in many fields of applied mathematics, see, e.g., [IKS03] and references therein, spanning from engineering to biology, from computer science to finance, etc. Concerning the latter type of applications, SDDEs can be used, e.g., to price a wide class of so-called *path-dependent options*, or to characterize financial market with memory, see, e.g., [AHMP07, CDPMZ16, KSW05a, KSW07a, KP07] and references therein for further details.

Even if we do not aim to give a complete survey of problems, solutions and method related to the study of SDDE's, we would like to recall at least some of the most relevant approaches recently proposed to analyse such kind of stochastic differential equations. In particular in [FZ] an infinite dimensional approach based on semigroup theory has been used, providing a deterministic Kolmogorov equation associated to a SDDE driven by a Brownian motion. In [CF10, CF13] a new type of *path-dependent calculus* has been derived *ad hoc*, introducing new types of derivatives, namely the *horizontal derivative* and the *vertical derivative*, which are then used in dealing with *path-dependent stochastic equations*. In [FPS00] latter results have been used in the computation of the *greeks* for various types of *path-dependent options*.

In what follows we will exploit a different approach, first introduced in [Moh84] and then developed in [Moh98, YM05], which is based on the theory of *stochastic functional delay differential equations* (SFDDEs), in particular in the light of [BCDNR16] where SFDDEs with jumps have been introduced. We would also like to mention that in [FZ] a connection between the three previously cited approaches, has been provided.

We recall that in [BCDNR16, CDPO16] existence and uniqueness results for SFDDEs have been shown together with a suitable related Itô-Döblin formula. Moreover, in [CDPO16], a non-linear *Feynman-Kac theorem* for SFDDEs arising from a forward-backward system with delay under mild assumptions of regularity is provided. Latter result is particularly useful in financial applications, thanks to the mild assumptions needed on differentiability. In the present work we start from the latter property, and we exploit the mild notion of gradient first introduced in [FT05] in order to connect a forward-backward system with delay to an infinite dimensional partial integro-differential equation (PIDE) whose solution is only required to be Lipschitz. Such mild regularity request is rather advantageous

in financial applications, where typical payoff functions are Lipschitz, without being also differentiable.

In particular we will exploit previously cited result in order to derive a set of infinite dimensional PIDE to price a rather general class of stochastic volatility models with jumps arising in finance. In particular we will focus our attention on continuously monitored *Asian options*, even if analogous results hold for discretely monitored *Asian options* as well as for other type of *path-dependent options*, often referred as *exotic options*, such as the *realized volatility options*, the *variance swap options*, etc., see, e.g., [R⁺91, WDH93] for a general introduction to exotic options. It is worth to mention that a similar approach can be extended to market models with memory effect, see, e.g., [KSW05a, KSW07a, KP07, Swi07], or to time changed stochastic volatility models, see, e.g., [CGMY03, HS11], the latter subject being the object of our future works.

The present paper is organized as follows: in section 2.3 we introduce the main theoretical results derived in [BCDNR16, CDPO16] also providing the main *non-linear Feynman-Kac theorem*; in Sec. 2.5.1 we introduce the financial setting also showing some related example of possible payoffs; in Sec. 2.4.1 we derive the infinite dimensional pricing PIDE for the Bates models, see [Bat96], whereas in Sec. 2.4.1 we derive the same result for the Barndorff-Nielsen-Shephard (BNS) model, see later for its definition, with respect to two different specifications for the stochastic law of the volatility process, namely the *Gaussian Inverse law* and the *Variance Gamma law*.

2.4.1 Path-dependent derivative pricing in stochastic volatility models with jumps

Let us consider an asset S evolving under the *risk-neutral probability* (RNP) \mathbb{Q} as

$$S(t) = e^{(\rho-q)t+X(t)}, \quad (2.115)$$

being $\rho > 0$ the fixed deterministic interest rate of a given riskless asset, e.g., a certain *Bank account*, and $q > 0$ the dividend yield. We will denote by X the discounted log return of the asset S whose evolution will be stated later on. We will assume that the volatility of the process S is described by a second, stochastic, process V , so that we will be interested in studying the couple (X, V) , which defines what is usually called a *stochastic volatility model* (SVM). For the sake of simplicity, in what follows we will state our results in the *risk-neutral world* characterized by the RNP \mathbb{Q} , so that we will not take into account the problem of choosing a particular risk neutral measure. Let us underline that, since we will consider SVM with jumps, the related markets fail to be complete, hence we lost the uniqueness of the RNM, see, e.g., [BMB05, H§, HS09] for detailed discussions concerning the topic of identify the *right* RNM, in the present setting.

We are interested in pricing *path-dependent options* and, in particular, we will focus on considering *Asian options*, even if many other types of *path-dependent options* such as the *realized volatility options*, or the *variance swap options* can be treated by our approach as well.

Being the payoff of such type of options path-dependent, in general we are not allowed to exploit standard results based on the Markov property of the underlying process, instead we will use the theory developed in Sec. 2.3, considering the process of interest as an infinite dimensional process taking values in a suitable *path-space*. The latter idea will allow to recover the Markov property of the couple (X, V) , as an infinite dimensional process. In

particular, according to the notation introduced in Se. 2.3, we will set the maximum delay taken into account to be $r = T$, $T < \infty$ being the maturity time and we will consider the solution to take values in the infinite dimensional space $M^2 := L^2([-T, 0]; \mathbb{R}^2) \times \mathbb{R}^2$. Moreover, introducing an appropriate backward stochastic differential equations (BSDE), and also by means of the recovered Markov property, we can establish a *Feynman-Kac representation theorem* for the path-dependent pricing PIDE.

In what follows we will not specify a particular form for the terminal payoff function. In fact the pricing PIDE we are interested in, still remains valid for different type of *path-dependent options*. Therefore we treat a representative case of the latter class of options, namely we study the Asian type options, setting $K > 0$ to be the fixed strike price, S to be the asset price, and $X = \log S$ to be the log return of the asset price, see, e.g. [BBP97], while possible payoffs are

(i) *arithmetic average floating strike call*

$$\Phi(S(T), S_T) = \max \left\{ S(T) - \frac{1}{T} \int_0^T S(s) ds, 0 \right\};$$

(ii) *arithmetic average fixed strike call*

$$\Phi(S(T), S_T) = \max \left\{ \frac{1}{T} \int_0^T S(s) ds - K, 0 \right\};$$

(iii) *geometric average floating strike call*

$$\Phi(X(T), X_T) = \max \left\{ e^{S(T)} - e^{\frac{1}{T} \int_0^T X(s) ds}, 0 \right\};$$

(iv) *arithmetic average fixed strike call*

$$\Phi(X(T), X_T) = \max \left\{ e^{\frac{1}{T} \int_0^T X(s) ds} - K, 0 \right\}.$$

Let us notice that the above payoff function are Lipschitz in the sense of Assumption 2.3.19, but they fail to be differentiable, so that such type of options perfectly fits the setting developed in Sec. 2.3. In what follows, we will derive the pricing PIDE for general *path-dependent options* written in some relevant *stochastic volatility models* used in finance, nevertheless our approach can be extended to other SVM, and related results can be generalized in order to consider multidimensional underlyings as well as multiscale market models.

The Bates model

The *Bates model* has been introduced in [Bat96] and it is perhaps the most simple SVM with jumps, moreover, provided suitable choices for its parameters, we can use it to recover the *Heston model* as well as the *Merton jump model*. In the sequel we will use the notation introduced throughout the previous sections.

Let us consider an asset $S(t)$ whose dynamics is described by eq. (2.115), and let us assume that the discounted log price X evolves according to the following equation

$$dX(t) = - \left(\kappa(1) + \frac{1}{2}V(t) \right) dt + \sqrt{V(t)}dW^1(t) + dZ(t), \quad (2.116)$$

while the volatility V is defined by

$$dV(t) = \lambda(\theta - V(t))dt + \zeta\sqrt{V(t)}dW^2(t), \quad (2.117)$$

where λ , θ and ζ are fixed, positive parameters, W^1 and W^2 are two Brownian motion with correlation $\rho \in [-1, 1]$, and

$$Z(t) = \sum_{i=1}^{N_t} J_i,$$

is assumed to be independent of both W^1 and W^2 , N being a Poisson process with intensity μ , while the random variables J_i are as follows $J_i \sim \mathcal{N}(\gamma, \delta)$. In the present setting we have that the Lévy measure $\nu(dz)$ of Z is

$$\nu(dz) = \frac{\mu}{\delta\sqrt{2\pi}} e^{-\frac{(z-\gamma)^2}{2\delta^2}} dz,$$

and its cumulant generating function reads as follows

$$\kappa(\xi) = \log \mathbb{E}e^{\xi Z(1)} = \mu \left(e^{\gamma\xi + \frac{\delta^2\xi^2}{2}} - 1 \right).$$

Remark 2.4.1. Let us note that from equations (2.116) and (2.117), we can recover the *Heston model* setting $\mu = 0$, as well as the *Merton jump model*, taking $\zeta = 0$.

Proposition 2.4.2. *Let us consider an asset $S(t) = e^{(\rho-q)t+X(t)}$ where the log return X evolves according to eq. (2.116), and the volatility V evolves according to eq. (2.117). Let us also consider a path-dependent options with a terminal payoff at maturity T given by a (possibly path-dependent) function Φ satisfying assumptions 2.3.19. The the fair price $u(t, x, v, \eta)$ is the mild solution to the pricing PIDE*

$$\begin{cases} \partial_t u(t, x, v, \eta) + \mathcal{L}_t u(t, x, v, \eta) = (\rho - q)u(t, x, v, \eta), \\ u(T, x, v, \eta) = \Phi(x, v, \eta), \end{cases} \quad (2.118)$$

where the infinitesimal generator of the triple (X, V, \mathcal{A}) , is

$$\begin{aligned} \mathcal{L}_t u(t, x, v, \eta) &= \int_0^T \partial_s \eta D_\eta(s) u(t, x, v, \eta(s)) ds + \frac{1}{2} v \partial_{x,x} u(t, x, v, \eta) + \\ &+ \frac{1}{2} \zeta v \partial_{v,v} u(t, x, v, \eta) + \rho \zeta v \partial_{x,v} u(t, x, v, \eta) + \\ &- \left(\kappa(1) + \frac{1}{2} v \right) \partial_x u(t, x, v, \eta) + \lambda(\theta - v) \partial_v u(t, x, v, \eta) + \\ &+ \int_0^\infty (u(t, x+z, v, \eta, \omega) - u(t, x, v, \eta, \omega)) \nu(dz) = \\ &= \int_0^T \partial_s \eta(s) D_\eta u(t, x, v, \eta(s)) ds + \frac{1}{2} v \partial_{x,x} u(t, x, v, \eta) + \\ &+ \frac{1}{2} \zeta v \partial_{v,v} u(t, x, v, \eta) + \rho \zeta v \partial_{x,v} u(t, x, v, \eta) + \\ &- \left(\mu \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \frac{1}{2} v \right) \partial_x u(t, x, v, \eta) + \lambda(\theta - v) \partial_v u(t, x, v, \eta) + \\ &+ \int_0^\infty (u(t, x+z, v, \eta) - u(t, x, v, \eta)) \frac{\mu}{\delta\sqrt{2\pi}} e^{-\frac{(z-\gamma)^2}{2\delta^2}} dz. \end{aligned}$$

Proof. Let us consider a general derivative whose payoff, possibly *path-dependent*, is given by

$$Y(t, x, v, \eta) = \mathbb{E} \left[e^{-(\rho-q)(T-t)} \Phi(X(T), X_T) \middle| X(t) = x, V(t) = v, X_t = \eta \right]. \quad (2.119)$$

If we introduce the discounted option price $e^{-(\rho-q)t}Y(t)$, then applying the Itô-Döblin formula we have

$$dY(t) = e^{(\rho-q)t} d(e^{-(\rho-q)t}Y(t)) + (\rho - q)Y(t)dt,$$

so that the process Y is a solution to the following BSDE

$$\begin{aligned} Y(t) = & \Phi(X(T), X_T) + \int_t^T (\rho - q)Y(s)ds + \\ & - \int_t^T Z(s)dW(s) - \int_t^T \int_{\mathbb{R}_0} U(s, z)\tilde{N}(ds, dz), \end{aligned} \quad (2.120)$$

which satisfies assumptions 2.3.19, hence there exists a unique solution $(Y, Z, U) \in \mathbb{K}([0, T])$ to the eq. (2.120).

Let us then consider the forward-backward system

$$\begin{cases} dX(t) & = -(\kappa(1) + \frac{1}{2}V(t))dt + \sqrt{V(t)}dW^1(t) + dZ(t), \\ dV(t) & = \lambda(\theta - V(t))dt + \zeta\sqrt{V(t)}dW^2(t), \\ (X(0), V(0), X_0) & = (x, v, \eta) \in M^2, \\ dY(t) & = (\rho - q)Y(t)dt - Z(t)dW(t) - \int_{\mathbb{R}_0} U(t, z)\tilde{N}(dt, dz), \\ Y(T) & = \Phi(X(T), X_T), \end{cases} \quad (2.121)$$

then, applying Th. 2.5.5, we have that the unique solution $(X, V, \mathcal{A}, Y, Z, U)$ to the forward backward system (2.121), is given by the *mild solution* to the deterministic path-dependent PIDE (2.127). \square

Remark 2.4.3. Let us notice that even if the forward equation for the log price does not satisfies Lipschitz assumption 2.3.3, due to the square root appearing in the volatility, nevertheless its existence and uniqueness is proven in [Bat96].

We have not specified any particular payoff function for the option to be priced, however we easily have that the payoff functions introduced in 2.4.1 fits into assumptions 2.3.19 so that the above theorem holds for this particular type of options. For instance, in the case of a *geometric average floating strike* option we have that the *path-component* $\mathcal{A} := \int_0^T X(s)ds$. As briefly said above, in the present setting we can derive the pricing PIDE also for other notable type of *path-dependent options* just taking into account the different path-component. For instance, if we are to treat a *realized volatility options* or *variance swap options* we would have to consider $\mathcal{A} := \int_0^T V(s)ds$, so that eq. (2.119) would read

$$Y(t, x, v, \eta) = \mathbb{E} \left[e^{-r(T-t)} \Phi(X(T), X_T) \middle| X(T) = x, V(T) = v, V_T = \eta \right].$$

The Barndorff-Nielsen-Shephard model

In what follows we consider a second relevant SVM which is widely used in finance, namely the Barndorff-Nielsen-Shephard (BNS) model, see [BNNS⁺02, BN97, BMB05, HS11, HS09, NV03] for a complete treatment of it.

Let us consider the discounted log price X evolving as

$$dX(t) = - \left(\kappa(\rho) + \frac{1}{2}V(t) \right) dt + \sqrt{V(t)}dW^1(t) + dL_\lambda(t), \quad (2.122)$$

and let us assume that the volatility V is a mean-reverting Ornstein-Uhlenbeck process driven by a pure-jump subordinator L with Lévy measure

$$dV(t) = -\lambda V(t)dt + \rho dL_\lambda(t), \quad (2.123)$$

where $\lambda > 0$ is a fixed positive parameter, and $\rho < 0$ is the correlation parameter. As in Sec. 2.4.1 we have denoted by κ the cumulant generating function, that is

$$\kappa(\xi) = \log \mathbb{E}e^{\xi L(1)}.$$

We will now focus on two particular laws under which the volatility process is assumed to evolve.

The Inverse Gaussian law

We first consider the BNS model introduced above with the further assumption that the stationary distribution of the volatility process V follows an *inverse Gaussian law* of parameter γ and δ , that is $V \sim IG(\gamma, \delta)$, see, e.g., [BN97], then the associated Lévy measure of the subordinator L , becomes

$$\nu(dz) = \frac{\gamma}{2\sqrt{2\pi}} e^{-\frac{3}{2}} (1 + \delta z) e^{-\frac{1}{2}\delta z} dz, \quad (2.124)$$

see, e.g., [BN97, BMB05].

Proposition 2.4.4. *Let us consider an asset $S(t) = e^{(\rho-q)t+X(t)}$ where the log return X evolves according to the eq. (2.122), and the volatility V evolves according to the eq. (2.123). We further assume that $V \sim IG(\gamma, \delta)$, so that the Lévy measure is as in eq. (2.124).*

Let us then consider a path-dependent options with a terminal payoff at maturity T given by a (possibly path-dependent) function Φ satisfying assumptions 2.3.19. then the fair price $u(t, x, v, \eta)$ is the mild solution to the pricing PIDE

$$\begin{cases} \partial_t u(t, x, v, \eta) + \mathcal{L}_t u(t, x, v, \eta) = (\rho - q)u(t, x, v, \eta), \\ u(T, x, v, \eta) = \Phi(x, v, \eta), \end{cases} \quad (2.125)$$

where the infinitesimal generator of the triple (X, V, \mathcal{A}) , is

$$\begin{aligned}
\mathcal{L}_t u(t, x, v, \eta) &= \int_0^T \partial_s \eta(s) D_\eta u(t, x, v, \eta(s)) ds + \frac{1}{2} v \partial_{x,x} u(t, x, v, \eta) + \\
&\quad - \left(\kappa(\rho) + \frac{1}{2} v \right) \partial_x u(t, x, v, \eta) + \lambda(\theta - v) \partial_v u(t, x, v, \eta) + \\
&\quad + \int_0^\infty (u(t, x + z, v + \rho z, \eta) - u(t, x, v, \eta)) \nu(dz) = \\
&= \int_0^T \partial_s \eta(s) D_\eta u(t, x, v, \eta(s)) ds + \frac{1}{2} v \partial_{x,x} u(t, x, v, \eta) + \\
&\quad - \left(\kappa(\rho) + \frac{1}{2} v \right) \partial_x u(t, x, v, \eta) + \lambda(\theta - v) \partial_v u(t, x, v, \eta) + \\
&\quad + \int_0^\infty (u(t, x + z, v + \rho z) - u(t, x, v, \eta)) \frac{\gamma}{2\sqrt{2\pi}} e^{-\frac{3}{2} - \frac{1}{2}\delta z} (1 + \delta z) z.
\end{aligned}$$

Proof. The proof is analogous to the one given for Prop. 2.4.5. □

The Variance Gamma law

We consider in the present subsection the BNS model introduced above with the further assumption that the stationary distribution of the volatility process V follows a *variance gamma law*, with parameters γ and δ , that is $V \sim \Gamma(\gamma, \delta)$. In the present setting the Lévy measure of L becomes

$$\nu(dz) = \delta \gamma e^{\delta z} dz, \quad (2.126)$$

see, e.g., [BMB05, MS90].

Proposition 2.4.5. *Let us consider an asset $S(t) = e^{(\rho-q)t+X(t)}$ where the log return X evolves according to the eq. (2.122) and the volatility V evolves according to the eq. (2.123). We further assume that $V \sim \Gamma(\gamma, \delta)$, so that the Lévy measure is as in eq. (2.126).*

Let us then consider a path-dependent options with a terminal payoff at maturity T given by a (possibly path-dependent) function Φ satisfying assumptions 2.3.19. Then the fair price $u(t, x, v, \eta)$ is the mild solution to the pricing PIDE

$$\begin{cases} \partial_t u(t, x, v, \eta) + \mathcal{L}_t u(t, x, v, \eta) = (\rho - q)u(t, x, v, \eta), \\ u(T, x, v, \eta) = \Phi(x, v, \eta), \end{cases} \quad (2.127)$$

where the infinitesimal generator of the triple (X, V, \mathcal{A}) , is

$$\begin{aligned}
\mathcal{L}_t u(t, x, v, \eta) &= \int_0^T \partial_s \eta(s) D_\eta u(t, x, v, \eta(s)) ds + \frac{1}{2} v \partial_{x,x} u(t, x, v, \eta) + \\
&\quad - \left(\kappa(\rho) + \frac{1}{2} v \right) \partial_x u(t, x, v, \eta) + \lambda(\theta - v) \partial_v u(t, x, v, \eta) + \\
&\quad + \int_0^\infty (u(t, x + z, v + \rho z, \eta) - u(t, x, v, \eta)) \nu(dz) = \\
&= \int_0^T \partial_s \eta(s) D_\eta u(t, x, v, \eta(s)) ds + \frac{1}{2} v \partial_{x,x} u(t, x, v, \eta) + \\
&\quad - \left(\kappa(\rho) + \frac{1}{2} v \right) \partial_x u(t, x, v, \eta) + \lambda(\theta - v) \partial_v u(t, x, v, \eta) + \\
&\quad + \int_0^\infty (u(t, x + z, v + \rho z, \eta) - u(t, x, v, \eta)) \delta \gamma e^{\delta z} dz .
\end{aligned}$$

Proof. The proof follows as in Prop. 2.4.5. □

2.4.2 Application to option pricing in market with memory

Recent years have seen an increasing attention towards the study of delay differential equations (DDE), mainly because of the large number of their applications ranging from mathematical biology to mathematical finance, see, e.g., [Kua93, Moh98]. Concerning financial applications, examples are given, e.g., in [KP07] where it has been shown how, due to some inherent factors such as time to transport, delay is a key factor to be taken into account in commodity markets, while in [AHMP07, CY07, KSW07b] applications to markets with delay and to option pricing in such markets have been provided, and in [KSW05b, KSW05a, SX11] stochastic volatility models with delays are treated.

We would like to stress that the aforementioned applications can be treated in our framework. In particular, to the best of our knowledge, very few applications concerning DDE with jumps has been provided. The aim of the current section is to give possible applications of DDE paying particular attention to the problem of pricing a contingent claim in a delayed market with jumps.

In what follows we will denote by $\mathcal{D} := \mathcal{D}([-r, 0], \mathbb{R})$ the space of the \mathbb{R} -valued càdlàg-functions on $[-r, 0]$ endowed with the norm $\|\cdot\|_{\mathcal{D}} := \sup_{\theta \in [-r, 0]} |\cdot|$, and by $M_p := L^p([-r, 0], \mathbb{R}) \times \mathbb{R}$, for $p \geq 2$, endowed with the norm $\|(\cdot, \cdot)\|_{M_p([-r, 0], \mathbb{R})}^p := \|\cdot\|_{L^p([-r, 0], \mathbb{R})}^p + |\cdot|^p$.

Let us consider a positive maturity time $T < \infty$, a fixed delay $r \geq 0$, and a market composed by one riskless bond B and one risky asset S , the case of d risky assets being easily derived from the current example.

We assume that the bank account B satisfies the following delayed deterministic linear equation

$$\begin{cases} dB(t) = R(t, B_t) dt, & t \in [0, T], \\ B_0 = b, \end{cases} \quad , \quad (2.128)$$

for a suitable functional $R : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$, while for the risky asset $S := \{S(t) : t \geq 0\}$,

we consider its price at time $t \geq 0$ determined by

$$\begin{cases} dS(t) &= S(t_-) \left[\mu(t, S_t)dt + \sigma(t, S_t)dW(t) + \int_{\mathbb{R}_0} \gamma(t, S_t)\tilde{N}(t) \right], \quad t \in [0, T], \\ S_0 &= s \in S^p(\mathcal{F}_0), \end{cases} \quad (2.129)$$

with $\mu, \sigma : \mathbf{S}_p \rightarrow L^p(\Omega; \mathbb{R})$ and $\gamma : \mathbf{S}_p \rightarrow L^p(\Omega; L^2(\mathbb{R}))$ some suitable functional which will be specified later. Furthermore we have denoted by B_t , resp. S_t , the segment of the process B , resp. S , i.e. $B_t = \{B(t + \theta)\}_{\theta \in [-r, 0]}$, resp. $S_t = \{S(t + \theta)\}_{\theta \in [-r, 0]}$. The space \mathcal{D} can be replaced with M_p , just taking into account that existence and uniqueness of a solution to eq. (2.129) does not hold for discrete delay. In particular we would like to stress that in the present setting the functionals μ, σ and γ are allowed to be stochastic processes by themselves, namely they can have the general form

$$f(t, S_t) = \bar{\mu}(t, S_t)dt + \bar{\sigma}(t, S_t)dW^1(t) + \int_{\mathbb{R}_0} \bar{\gamma}(t, S_t)\tilde{N}^1(dt, dz), \quad f = \mu, \sigma, \gamma,$$

for some suitable functional $\bar{\mu}, \bar{\sigma}$ and $\bar{\gamma}$.

The present section is structured as follows: in Sec. 2.1.1 we give some examples of delays that can be treated in our frameworks and related concrete models, then, in Sec. 4.2.4, we give a concrete example of a market model with jumps and provide a bound for the price of a European call option when the underlying exhibits a delay.

Let us consider a market model with delay, our main goal being to provide an upper and a lower bound for the price of a European call option written on a delayed underlying.

Let us thus assume that the bank account B satisfies the following delayed deterministic linear equation

$$\begin{cases} dB(t) = R(B_t)dt, \quad t \geq 0, \\ B_0 = b \in \mathcal{C}([-r, 0]; \mathbb{R}), \end{cases}, \quad (2.130)$$

with $b \in \mathcal{C}([-r, 0]; \mathbb{R})$, while $R : [0, T] \times \mathcal{C}([-r, 0]; \mathbb{R}) \rightarrow \mathbb{R}$ is a deterministic linear and continuous functional that can be represented as

$$R(B_t) = \int_{-r}^0 B(t + \theta)\xi(d\theta),$$

for ξ a non-decreasing finite Borel measure on $[-r, 0]$ such that $\xi(0) - \xi(-r) > 0$.

We will further assume, following [CY07], that it exists a unique positive real constant ρ such that the solution to eq. (5.85) takes the particular form

$$B(t) = b(0)e^{\rho t}, \quad t \geq 0, \quad (2.131)$$

and ρ satisfies

$$\rho = \int_{-r}^0 e^{\rho\theta}\xi(d\theta),$$

we refer to [CY07, Prop. 2.5] for a proof of the existence and uniqueness of such ρ .

Let us further consider an asset S whose price evolves according to eq. (2.132). In particular let us consider some deterministic time homogeneous functionals μ, σ and $\gamma_0 : \mathcal{D} \rightarrow \mathbb{R}$ satisfying Lipschitz continuous and linear growth assumptions.

Furthermore let $\lambda : \mathbb{R}_0 \rightarrow \mathbb{R}$ be such that $\lambda \in \mathcal{L}^2(\nu)$. We will assume that the asset price S evolves according with the discrete delay introduced in (ii), namely, for $t \in [0, T]$, we have

$$\begin{cases} dS(t) &= S(t_-) \left[\mu(S(t-a))dt + \sigma(S(t-b))dW(t) + \int_{\mathbb{R}_0} \gamma(S(t-c))(z)\tilde{N}(t) \right], \\ S_0 &= s \in \mathcal{D}, \end{cases} \quad (2.132)$$

with a, b and $c \in [-r, 0)$ and $\gamma(S)(z) := \gamma_0(S)\lambda(z)$. In what follows we will omit to specify the dependence of $\gamma(S)$ by z , writing for short $\gamma(S)$ instead of $\gamma(S)(z)$.

Proposition 2.4.6. *For any initial process $s \in \mathcal{D}$ there exists a unique solution to eq. (2.132). If we further assume that $s(0) > 0$ and $\gamma(S) \geq -1$, then $S(t) > 0$ a.s., for any $t \in [0, T]$.*

Proof. The existence and uniqueness of the solution immediately follows from the assumptions on the functional μ, σ, γ_0 and λ and from Section 2.1.

An explicit form for the solution can be retrieved, following [AHMP07], via iterated steps. Let us in particular denote by

$$dX(t) = \mu(S(t-a))dt + \sigma(S(t-b))dW(t) + \int_{\mathbb{R}_0} \gamma(S(t-c))\tilde{N}(t),$$

then eq. (2.132) can be written in a more compact form as

$$\begin{cases} dS(t) &= S(t_-)dX(t), t \in [0, T], \\ S_0 &= s \in \mathcal{D}, \end{cases},$$

Let us now set $l := \min\{a, b, c\}$, then, for any $t \in [0, l]$, we have

$$X(t) = \int_0^t \mu(s(q-a))dq + \int_0^t \sigma(s(q-b))dW(q) + \int_0^t \int_{\mathbb{R}_0} \gamma(s(q-c))\tilde{N}(q);$$

a solution to eq. (2.132) can be explicitly computed via Itô's lemma as

$$\begin{aligned} S(t) &= s(0)e^{\int_0^t \mu(s(q-a))dq} e^{\int_0^t \sigma(s(q-b))dW(q) - \frac{1}{2} \int_0^t \sigma(s(q-b))^2 dq} \times \\ &\times e^{\int_0^t \int_{\mathbb{R}_0} \ln(1+\gamma(s(q-c)))\tilde{N}(q) - \int_0^t \int_{\mathbb{R}_0} (\gamma(s(q-c)) - \ln(1+\gamma(s(q-c))))\nu(dz) dq}. \end{aligned} \quad (2.133)$$

If we then assume $s(0) > 0$ and from the fact that $\gamma \geq -1$ we have that, for any $t \in [0, l]$, the solution S is well defined and it holds $S(t) > 0$.

Eventually by iteration we can compute the solution for any $t \in [kl, (k+1)l \wedge T]$, $k = 1, 2, \dots$. \square

In what follows we will price a given contingent claim written on the underlying defined by (2.132). In particular it is well known that in a standard Black-Scholes model, where the prices evolve according to a pure diffusive process, and under suitable assumptions on the coefficients, completeness of the model holds. Thus we have a unique (*risk neutral*) measure \mathbb{Q} equivalent to the real world measure \mathbb{P} , and such that the discounted stock price process $e^{-\rho t}S$ is a \mathbb{Q} -martingale. Unfortunately the same does not hold in the present case where the driving process is a general, possibly discontinuous, martingale.

Such a drawback can be tackled exploiting the Girsanov theorem for general Lévy processes, see, e.g., [DNØP⁺09], namely

Theorem 2.4.7 (Girsanov). *Let $\theta(t, z) \leq 1$, $t \in [0, T]$, $z \in \mathbb{R}_0$ and $u(t)$, $t \in [0, T]$ be \mathbb{F} -predictable processes such that*

$$\int_0^T \int_{\mathbb{R}_0} (|\log(1 + \theta(t, z))| + \theta^2(t, z)) \nu(dz) dt < \infty, \quad P - a.s., \quad (2.134)$$

$$\int_0^T u^2(t) dt < \infty, \quad P - a.s. \quad (2.135)$$

Let

$$\begin{aligned} M(t) := \exp \left\{ \int_0^t u(s) dW(s) - \frac{1}{2} \int_0^t u^2(s) ds + \int_0^t \int_{\mathbb{R}_0} (\log(1 - \theta(s, z)) + \theta(s, z)) \nu(dz) ds \right. \\ \left. + \int_0^t \int_{\mathbb{R}_0} \log(1 - \theta(s, z)) \tilde{N}(ds, dz) \right\}, \quad t \in [0, T]. \end{aligned} \quad (2.136)$$

Define a measure Q on \mathcal{F} by

$$dQ(\omega) = M(T, \omega) dP(\omega).$$

Assume that Z satisfies the Novikov condition so that Q is a probability measure on (Ω, \mathcal{F}) . Define

$$\tilde{N}_Q(dt, dz) = \theta(t, z) \nu(dz) dt + \tilde{N}(dt, dz)$$

and

$$dW_Q(t) = u(t) dt + dW(t).$$

Then $\tilde{N}_Q(\cdot, \cdot)$ and $W_Q(\cdot)$ are compensated Poisson measure of $N(\cdot, \cdot)$ Brownian motion under Q , respectively.

By Girsanov theorem 2.4.7, we have that our setting is characterized by an infinite number of possible choices of equivalent measure \mathbb{Q} , depending on the particular form for the process θ . In what follows, in order to stress the dependence of the measure \mathbb{Q} on the process θ , we will denote by W^θ , resp. \tilde{N}^θ , the Brownian motion, resp. the Poisson compensated random measure, under the measure \mathbb{Q}^θ .

Let us now consider two suitable processes u and θ satisfying the assumptions (2.134)-(2.135) of Th. 2.4.7 such that the stock price S , under \mathbb{Q}^θ , evolves according to

$$dS(t) = S(t_-) \left[\rho dt + \sigma(S(t-b)) dW^\theta(t) + \int_{\mathbb{R}_0} \gamma(S(t-c)) \tilde{N}^\theta(dt, dz) \right],$$

whose solution can be explicitly computed as in eq. (5.17) via Itô's formula. Note that the discounted stock price $\hat{S}(t) := e^{-\rho t} S(t)$ is a \mathbb{Q}^θ -martingale.

Thus the price of a contingent claim written on the underlying S depends on the choice of the process θ that we use in order to define the new equivalent measure \mathbb{Q}^θ . In particular let us denote by Γ the set of measure \mathbb{Q}^θ equivalent to the real world measure \mathbb{P} , and such that u and θ satisfy assumptions (2.134)-(2.135) of Th. 2.4.7 and the discounted stock price S is a \mathbb{Q}^θ -martingale. Our aim is to give an upper and lower bound for the price $P(t)$ at

time $t \in [0, T]$ of a contingent claim written on the asset S evolving according to (2.132), namely

$$P^\theta(t) \in \left[\inf_{\mathbb{Q}^\theta \in \Gamma} P^\theta(t), \sup_{\mathbb{Q}^\theta \in \Gamma} P^\theta(t) \right].$$

In what follows we will use for short the notation \mathbb{E}_t^θ to denote the conditional expectation w.r.t. \mathcal{F}_t evaluated under the measure \mathbb{Q}^θ , namely

$$\mathbb{E}_t^\theta [\cdot] := \mathbb{E}^\theta [\cdot | \mathcal{F}_t].$$

Theorem 2.4.8. *Let us fix a measure $\mathbb{Q}^\theta \in \Gamma$ and let us define*

$$C^\theta(t) := \mathbb{E}_t^\theta \left[e^{-\rho(T-t)} \Phi(S(T)) \right],$$

the price at time $t \in [0, T]$ of a European contingent claim C with terminal payoff $\Phi(S) := (S(T) - K)^+$, written on the underlying S evolving according to eq. (2.132).

Then for any $t \in [0, T]$ and any $\mathbb{Q}^\theta \in \Gamma$ we have

$$C^\theta(t) \in [P_{DBS}(t), S(t)],$$

where we have denoted by P_{DBS} the delayed Black-Scholes price computed in [AHMP07].

Proof. Applying the Feynman-Kac theorem for jump processes, see, e.g., [BJ00], we have that the fair price of a European contingent claim can be retrieved solving the following deterministic problem

$$\begin{cases} \partial_t C^\theta(t, S) = -\frac{1}{2} \sigma(S(t-b))^2 S^2 \partial_{SS} C^\theta(t, S) - \rho S \partial_S C^\theta(t, S) + \rho C^\theta(t, S) + \\ \quad + \int_{\mathbb{R}_0} [C^\theta(t, S(1 + \gamma(S(t-c)))) - C^\theta(t, S) - S\gamma(S(t-c)) \partial_S C^\theta(t, S)] (1 - \theta) \nu(dz), \\ C^\theta(T, S) = (S(T) - K)^+, \end{cases} \quad (2.137)$$

where we have denoted for short by ∂_t , resp. ∂_S , resp. ∂_{SS} , the partial derivative w.r.t. t , resp. the partial derivative w.r.t. S , resp. the second order partial derivative w.r.t. S .

However, as stressed in [AHMP07], we cannot a priori apply the Feynman-Kac theorem for any $t \in [0, T]$. In fact if $t < m$, with $m := \min\{b, c\}$, we have that the solution to the deterministic problem (2.137) is anticipating w.r.t. the filtration \mathcal{F}_t . However previous deterministic representation is still valid for any $t \in [T - m, T]$.

Let us then denote by

$$\begin{aligned} \mathcal{L}_{DBS} C^\theta(t, S) &= -\frac{1}{2} \sigma(S(t-b))^2 S^2 \partial_{SS} C^\theta(t, S) - \rho S \partial_S C^\theta(t, S) + \rho C^\theta(t, S), \\ \mathcal{L}_J C^\theta(t, S) &= \int_{\mathbb{R}_0} [C^\theta(t, S(1 + \gamma(S(t-c)))) - C^\theta(t, S) - S\gamma(S(t-c)) \partial_S C^\theta(t, S)] (1 - \theta(t, z)) \nu(dz), \end{aligned}$$

the delayed Black-Scholes operator introduced in [AHMP07, eq. (11)], resp. the jump operator, see, e.g., [BJ00].

We have that $0 \leq \Phi(S) \leq S$, hence for any t it holds $0 \leq C(t, S) \leq S$. Exploiting now the fact that $\theta(t, z) \leq 1$ and $\gamma \geq -1$, and from the convexity of C , we have that, for any $t \in [T - m, T]$, $\mathcal{L}_J C(t, S) \geq 0$. Therefore the minimum is achieved when C solves

$$\partial_t C(t, S) = \mathcal{L}_{DBS} C(t, S), \quad (2.138)$$

whose solution $C(t, S)$, $t \in [T - m, T]$, is the delayed Black-Scholes price $P_{DBS}(t)$ computed explicitly in [AHMP07, Th. 4]. We have thus proved the lower bound, namely

$$C^\theta(t, S) \geq P_{DBS}(t), \quad t \in [T - m, T].$$

The claim for any $t \in [0, T]$ can be now achieved via iteration. In particular, for $t \in [T - 2m, T - m]$ we have that the price $C(t, S)$ solves the deterministic problem (2.138) with terminal condition $C(T - m, S) = P_{DBS}(T - m)$ explicitly computed in the previous step. The solution of such an equation is given in [AHMP07, Th. 4] in terms of the conditional expectation w.r.t. \mathcal{F}_t , $t \in [T - 2m, T - m]$. Iteratively proceeding, for any $t \in [T - (k + 1)m \vee 0, T - km]$, $k = 1, 2, \dots$, and exploiting the conditional expectation tower rule, we have that

$$C^\theta(t, S) \geq P_{DBS}(t), \quad t \in [0, T],$$

where $P_{DBS}(t)$ is the delayed Black-Scholes formula given in [AHMP07, Th. 4], and the upper bound immediately follows from the estimate $\Phi(S) \leq S$. \square

Remark 2.4.9. The interval for the price of a European put option can be computed similarly or equivalently exploiting the Put-Call parity.

Remark 2.4.10. Previous result can be extended to any contingent claim with convex terminal payoff Φ such that $0 \leq \Phi(S) \leq S$ and the function $g(S) := S - \Phi(S)$ is bounded, see, e.g., [BJ00].

Although we have provided an interval for the fair price of a European option, the problem of finding a price, and thus an hedging strategy, is only partially solved. In fact Th. 2.4.8 shows that the cost of superhedging a contingent claim in this type of market is too high, since the cost of superhedging would be the highest price that ensures absence of arbitrage, that is $S(0)$.

Thus alternative solution has to be found. One would be to chose only one among all the feasible martingale measures $\mathbb{Q}^\theta \in \Gamma$, but how to choose the most natural measure \mathbb{Q}^θ in Th. 2.4.7 in order to compute the fair price of a given contingent claim.

Remark 2.4.11. A first trivial choice would be $\theta = 0$, which can be easily seen to satisfies assumption of Th. 2.4.7. In particular this choice of the process θ has been used by Merton in [Mer76]. Therefore in this particular setting we have that the deterministic problem (2.137) becomes

$$\begin{cases} \partial_t C^0(t, S) = -\frac{1}{2}\sigma(S(t-b))^2 S^2 \partial_{SS} C^0(t, S) - \rho S \partial_S C^0(t, S) + \rho C^0(t, S) + \\ \quad + \int_{\mathbb{R}_0} [C^0(t, S(1 + \gamma(S(t-c)))) - C^0(t, S) - S\gamma(S(t-c))\partial_S C^0(t, S)] \nu(dz), \\ C^0(T, S) = (S - K)^+, t \in [T - m, T], \end{cases} \quad (2.139)$$

It is easy to see that C^0 belongs to the interval $[P_{DBS}(t), S(t)]$.

2.5 Connection with path-dependent calculus

In what follows we provide a connection between Itô's formula (2.2.6) and the path-dependent Itô's formula given in [CF10, CF13] which relies on the concepts of vertical and horizontal derivative, there introduced. Let us first set the notation we use in the current section.

Let (Ω, \mathcal{F}, P) be the probability space with $\Omega = \mathcal{D}([0, T], \mathbb{R}^d)$ endowed with the P -augmented (right-continuous) filtration $\{\mathcal{F}_t\}_{t \in [0, T]}$ generated by the canonical process $Y : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, $Y(t, \omega) = \omega(t)$ and here $\mathcal{F} := \mathcal{F}_T$. In this setting we define, for every $\omega \in \Omega$ and $t \in [0, T]$, $\omega_t := \{\omega(s), 0 \leq s \leq t\} \in \mathcal{D}([0, t])$, the trajectory up to time t . A stochastic process is a function $\varphi : [0, T] \times \Omega \rightarrow \mathbb{R}^d$, $(t, \omega) \mapsto \varphi(t, \omega)$. In addition, we say φ is non-anticipative if it is defined on $\mathcal{D}([0, t]; \mathbb{R}^d)$, i.e. $\varphi(t, \omega) = \varphi(t, \omega_t) := \varphi_t(\omega_t)$.

Let $\varphi = \{\varphi_t, t \in [0, T]\}$ be a non-anticipative stochastic process and $\{e_i\}_{i=1}^d \subset \mathbb{R}^d$ the canonical basis, we define the so-called *vertical derivative* as the following (path-wise) limit

$$\mathcal{D}^{V,i} \varphi_t(\omega_t) = \lim_{h \rightarrow 0} \frac{\varphi_t(\omega_t^{he_i}) - \varphi_t(\omega_t)}{h},$$

where $\omega_t^{he_i}(s) := \omega_t(s) + he_i 1_{\{t\}}(s)$, for every $s \in [0, t]$. Here, $\omega_t^{he_i}$ means adding a jump of size h at time t on the direction of e_i and hence the name. We then define the *vertical gradient* of φ_t as

$$\mathcal{D}^V \varphi_t = (\mathcal{D}^{V,1} \varphi_t, \dots, \mathcal{D}^{V,d} \varphi_t).$$

Furthermore, we define the *horizontal derivative* as the following (path-wise) limit

$$\mathcal{D}^H \varphi_t(\omega_t) = \lim_{h \searrow 0} \frac{\varphi_{t+h}(\omega_{t,h}) - \varphi_t(\omega_t)}{h},$$

where $\omega_{t,h}(s) := \omega_t(s) 1_{[0,t]}(s) + \omega_t(t) 1_{(t,t+h]}(s)$, for every $s \in [0, t+h]$. Here, $\omega_{t,h}$ is the extension of the trajectory ω_t on $[0, t]$ to $[0, t+h]$ by an horizontal line of length h at $\omega_t(t)$ and hence the name.

We consider a functional $F : [0, T] \times \mathcal{D}([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ which will act on processes φ_t . We say F is non-anticipative if

$$F(t, \psi) = F(t, \psi_t) =: F_t(\psi_t),$$

for every non-anticipative stochastic process ψ_t . Next, we state an Itô formula for $F_t(\psi_t)$ where F_t is a non-anticipative functional which is once horizontally and twice vertically differentiable. This result is taken from [CF10, Proposition 6].

Theorem 2.5.1 (Functional Itô's formula). *Consider an \mathbb{R}^d -valued non-anticipative stochastic process φ_t which admits the following càdlàg semimartingale representation*

$$\varphi_t = \varphi_0 + \int_0^t \mu(s) ds + \int_0^t \sigma(s) dW(s) + \int_0^t \int_{\mathbb{R}_0} \gamma(s-, z) \tilde{N}(ds, dz)$$

for processes $\mu : [0, T] \rightarrow \mathbb{R}^d$, $\sigma : [0, T] \rightarrow \mathbb{R}^{d \times m}$ and $\gamma : [0, T] \times \mathbb{R}_0 \rightarrow \mathbb{R}^{d \times n}$ such that $\int_0^T E \left[|\mu(s)| + \|\sigma(s)\|^2 + \int_{\mathbb{R}_0} \|\gamma(s, z)\|^2 \nu(dz) \right] ds < \infty$ being $\|\cdot\|$ a matrix norm.

Let F be a once horizontally and twice vertically differentiable non-anticipative functional satisfying some technical continuity conditions on F (see [CF10, Proposition 6]), $\mathcal{D}^V F_t$,

$\mathcal{D}^V \mathcal{D}^V F_t$ and $\mathcal{D}^H F_t$. Then for any t the following functional Itô formula holds P -a.s.

$$\begin{aligned} F_t(\varphi_t) &= F_0(\varphi_0) + \int_{(0,t]} \mathcal{D}^H F_s(\varphi_{s-}) ds + \int_{(0,t]} \mathcal{D}^V F_s(\varphi_{s-}) dX(s) \\ &+ \int_{(0,t]} \frac{1}{2} Tr [\sigma^*(s) \mathcal{D}^V \mathcal{D}^V F_s(\varphi_{s-}) \sigma(s)] ds \\ &+ \int_{(0,t]} \int_{\mathbb{R}^0} \mathcal{D}^V F_s(\varphi_{s-}) (F_s(\varphi_{s-} + \gamma(s-, z) 1_{\{s\}}) - F_s(\varphi_{s-}) - \gamma(s-, z)) N(ds, dz). \end{aligned} \quad (2.140)$$

To be able to show that the path-dependent Itô's formula (2.149) and the Itô's formula from Theorem 2.2.6 do coincide we need first to connect the two settings. Such a link can be established following [FZ], where the following operators are considered:

- the *restriction operator*, for every $t \in [0, r]$

$$\begin{aligned} M_t &: \mathcal{D}([-r, 0], \mathbb{R}^d) \rightarrow \mathcal{D}([0, t], \mathbb{R}^d), \\ M_t(f)(s) &= f(s-t), \quad s \in [0, t], \end{aligned}$$

- the *backward extension operator*, for every $t \in (0, r)$

$$\begin{aligned} L_t &: \mathcal{D}([0, t], \mathbb{R}^d) \rightarrow \mathcal{D}([-r, 0], \mathbb{R}^d), \\ L_t(f)(s) &= f(0) \mathbb{1}_{[-r, -t)}(s) + f(t+s) \mathbb{1}_{[-t, 0)}(s), \quad s \in [-r, 0), \end{aligned}$$

Let us consider a non-anticipative functional $b : [0, T] \times \mathcal{D}([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$, $b(t, \psi) = b(t, \psi_t) =: b_t(\psi_t)$ for any non-anticipative stochastic process ψ , then one can define a different functional \widehat{b} on $[0, T] \times \mathcal{D}([-r, 0]; \mathbb{R}^d) \times \mathbb{R}^d$ as

$$\widehat{b}(t, X_t, X(t)) := b_t(\widetilde{M}_t X_t), \quad (X_t, X(t)) \in \mathcal{D}([-r, 0]; \mathbb{R}^d) \times \mathbb{R}^d,$$

with

$$\widetilde{M}_t X_t(s) := \begin{cases} M_t(X_t)(s) & \text{if } s \in [0, t) \\ X_t(s) & \text{if } s = t \end{cases}.$$

The converse holds true as well, in fact let us consider a given functional \widehat{b} on $[0, T] \times \mathcal{D}([-r, 0]; \mathbb{R}^d)$, then we can obtain a corresponding functional b_t on $\mathcal{D}([0, t]; \mathbb{R}^d)$ as

$$b_t(\varphi_t) := \widehat{b}(t, L_t \varphi_t, \varphi_t(t)), \quad \varphi_t \in \mathcal{D}([0, t]; \mathbb{R}^d), \quad (2.141)$$

see [FZ] for details.

We can now show how the vertical and horizontal derivatives can be written in terms of the Fréchet derivative D and the derivative with respect to the present state. Part of the next theorem was already established in [FZ, Theorem 6.1].

Proposition 2.5.2. *Consider a function $F : [0, T] \times \mathcal{D}([-r, 0]; \mathbb{R}^d) \rightarrow \mathbb{R}$ and let us define $u_t : \mathcal{D}([0, t]; \mathbb{R}^d) \rightarrow \mathbb{R}$ as above in (2.160) $u_t(X_t) := F(t, L_t X_t, X(t))$. Then the i -th vertical*

derivative $\mathcal{D}^{V,i}$ of u_t coincides with the derivative with respect to the present state $X^i(t)$ of F , namely

$$\mathcal{D}^{V,i}u_t(X_t) = \partial_{x_i}F(t, L_tX_t, X(t)). \quad (2.142)$$

Furthermore, we have

$$u_t(X_t^{h^i}) - u_t(X_t) = F(t, L_tX_t, X(t) + h^i(t, L_tX_t, X(t))) - F(t, L_tX_t, X(t)). \quad (2.143)$$

If we assume that $X_t \in W^{1,p}$, then

$$\mathcal{D}^H u_t(X_t) = \partial_t F(t, L_tX_t, X(t)) + \langle DF(t, L_tX_t, X(t)), \nabla_\theta^+ L_tX_t \rangle_{\mathcal{D}}, \quad (2.144)$$

holds, where the notation is given in Section 2.2.

Proof. Concerning (2.161) we have

$$\begin{aligned} \mathcal{D}^{V,i}u_t(X_t) &= \lim_{h \rightarrow 0} \frac{1}{h} (u_t(X_t^h) - u_t(X_t)) = \lim_{h \rightarrow 0} \frac{1}{h} (F(t, L_tX_t^h, X^h(t)) - F(t, L_tX_t, X(t))) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (F(t, X(t) + h, L_tX_t^h) - F(t, X(t), L_tX_t)) = \partial_i F(t, X(t), L_tX_t). \end{aligned} \quad (2.145)$$

For what concerns (2.143), proceeding as in (2.145), we immediately have

$$\begin{aligned} u_t(X_t^{h^i}) - u_t(X_t) &= F(t, L_tX_t^{h^i}, X^{h^i}(t)) - F(t, L_tX_t, X(t)) = \\ &= F(t, X(t) + h^i, L_tX_t^{h^i}) - F(t, X(t), L_tX_t). \end{aligned}$$

We refer to [FZ, Theorem 6.1] for a proof of equation (2.162) \square

In the framework of this section, exploiting the previous Propositionosition we have that, for suitable regular coefficients, Itô's formula from Theorem 2.2.6 and the path-dependent Itô's formula in Theorem 2.5.4 coincide. In particular let us consider a process X evolving according to

$$\begin{cases} dX_t = f(t, X_t)dt + g(t, X_t)dW(t) + \int_{\mathbb{R}_0} h(t, X_t, z)\tilde{N}(dt, dz), \\ X_0 = \eta, \end{cases} \quad (2.146)$$

for some suitably regular enough coefficients f , g and h . Then proceeding as above we have that Equation (2.146) can be written as a path dependent process

$$\begin{cases} dX_t = \hat{f}_t(X_t)dt + \hat{g}_t(X_t)dW(t) + \int_{\mathbb{R}_0} \hat{h}_t(X_t)\tilde{N}(dt, dz), \\ X_0 = \eta, \end{cases}$$

with \hat{f}_t , \hat{g}_t and \hat{h}_t defined as in (2.160). Then we have the following result.

Theorem 2.5.3. *Let $F : [0, T] \times M^p \rightarrow \mathbb{R}$, $F \in C^{1,1,2}([0, T] \times \mathcal{D} \times \mathbb{R}^d)$ and let us define $u_t : \mathcal{D}([0, t]; \mathbb{R}^d) \rightarrow \mathbb{R}$ as in (2.160) $u_t(X_t) := F(t, L_tX_t, X(t))$. Then Itô's formula from Theorem 2.2.6 and the path dependent Itô's formula from Theorem 2.5.4 coincide.*

Proof. It is straightforward from Proposition 2.5.6 exploiting the backward extension operator L_t and eventually using Itô's formula from Theorem 2.2.6 and the *path-dependent* Itô formula from Theorem 2.5.4. \square

During last years an increasing interest has been devoted to the study of stochastic delay differential equations (SDDE's), particularly in the framework of mathematical finance and concerning, e.g., the analysis of markets with memory, path-dependent options, trading in presence of asymmetric information, etc.

The standard approach to SDDE's, first introduced in [Moh84], consists in considering the so called *segment* of the \mathbb{R} -valued stochastic process of interest in order to introduce a suitable infinite dimensional path-space E as, e.g., $E := \mathbb{R} \times L^p([-r, 0]; \mathbb{R})$ or $E : C([-r, 0]; \mathbb{R})$, $r > 0$ being the maximum delay allowed. Latter approach turns the original setting into an infinite dimensional one allowing to successfully treat a large class of problems both theoretical and of applied nature, see, e.g., [FZ, FMT10, Moh84]. Recently in [BCDNR16, CDPO16] the above framework has been extended to the case of jump processes, providing also a suitable infinite dimensional Itô's formula for SDDE exploiting the *calculus via regularization* that has proved to be extremely powerful when one is to prove an infinite dimensional Itô's formula, see, e.g. [CRb] and reference therein.

A different approach is given by the *path-wise calculus*, see, e.g., [CF10, CF13, Dup09], where DDSEs are studied exploiting the notions of *vertical derivative* and *horizontal derivative*. One of the main result obtained within the latter framework concerns the well-posedness of a *path-wise* counterpart of the standard Itô's formula which allows to obtain an ample class of fundamental results in a rather straightforward manner. In what follows we will use the just mentioned technique to avoid the standard approach in order to provide an explicit expression for the infinitesimal generator of the process of interest.

2.5.1 Pathwise derivatives and functional Itô's formula

In what follows we recall basic notions about pathwise Itô calculus following [CF10, CF13, Dup09]. Without loss of generality, we consider the one dimensional case from which the \mathbb{R}^d , $d > 1$, case can be easily derived.

Let us fix a finite time $T > 0$ and a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t \in [0, T]}, \mathbb{P})$. For any $t \in [0, T]$ we will denote by $\mathcal{D}_t := \mathcal{D}([0, t]; \mathbb{R})$ the set of right continuous and with left-hand limits, or càdlàg, \mathbb{R} -valued processes that are \mathcal{F}_t -adapted. In what follows for each process $\varphi_t \in \mathcal{D}_t$ we will employ the notation $\varphi(t) := \varphi_t(s) \mathbb{1}_{[s=t]}$ to denote the present value of the path $\varphi_t \in \mathcal{D}_t$, moreover we set $\mathcal{D} := \bigcup_{t \in [0, T]} \mathcal{D}_t$.

Let $\varphi_t \in \mathcal{D}$ and $u : \mathcal{D} \rightarrow \mathbb{R}$, then we say that u is vertically differentiable at $\varphi_t \in \mathcal{D}$, and we indicate such derivative by $\nabla^V u_t(\varphi_t)$, if the following limit exist

$$\lim_{h \rightarrow 0} \frac{u_t(\varphi_t^h) - u_t(\varphi_t)}{h}, \quad \varphi_t^h(s) := \varphi_t(s) + h \mathbb{1}_{[s=t]}(s), \quad (2.147)$$

any higher order vertical derivative is analogously defined. We say that u is horizontally differentiable at $\varphi_t \in \mathcal{D}$ and we indicate such derivative by $\nabla^H u(\varphi_t)$, if the following limit exist

$$\lim_{h \rightarrow 0^+} \frac{u_t(\varphi_{t,h}) - u_t(\varphi_t)}{h}, \quad \varphi_{t,h}(s) := \varphi(s) \mathbb{1}_{[0,t]}(s) + \varphi(t) \mathbb{1}_{[t,t+h]}(s) \in \mathcal{D}_{t+h}.$$

We refer to [CF10, CF13, Dup09] for further details about previously introduced derivatives.

Equivalently we denote by $\mathcal{C}^{1,2}(\mathcal{D})$ the space of functions $u : \mathcal{D} \rightarrow \mathbb{R}$ that are continuous on \mathcal{D} and admit continuous and bounded first and second order vertical derivative and first order horizontal derivative, see, e.g., [CF10, CF13] for details.

Let us recall the functional version of the Itô's formula, see, e.g., [CF10, Prop. 5], which plays a crucial role for stating our main result

Theorem 2.5.4 (Functional Itô's formula). [CF10, Prop. 5] *Let $(\Omega, \mathcal{F}_t, \mathcal{F}_{t \in [0, T]}, \mathbb{P})$ be a filtered probability space and let $\varphi_t \in \mathcal{D}$ be a \mathbb{R} -valued semimartingale with representation*

$$\varphi_t = \varphi_0 + \int_0^t \mu(s) ds + \int_0^t \sigma(s) dW(s) + \int_0^t \int_{\mathbb{R}_0} \gamma(s, z) \tilde{N}(ds, dz), \quad (2.148)$$

where μ , σ and γ are suitable stochastic processes, and let F_t with $F_t : \mathcal{D} \rightarrow \mathbb{R}$, be a non-anticipative functional such that $F \in \mathcal{C}^{1,2}(\mathcal{D})$. Then for any $t \in [0, T]$ the following change of variable formula holds true:

$$\begin{aligned} F_t(\varphi_t) &= F_0(\varphi_0) + \int_0^t \nabla^H F_s(\varphi_{s-}) ds + \int_0^t \nabla^V F_s(\varphi_{s-}) dX(s) + \int_0^t \frac{1}{2} \sigma^2(s) \nabla^V F_s(\varphi_{s-}) ds + \\ &+ \int_0^t \int_{\mathbb{R}_0} \left(F_s(\varphi_{s-}^\gamma) - F_s(\varphi_{s-}) - \gamma(s, z) \nabla^V F_s(\varphi_{s-}) \right) N(ds, dz), \end{aligned} \quad (2.149)$$

W being a \mathbb{R} -valued Wiener process, while N is a \mathbb{R} -valued Poisson measure with \tilde{N} its compensated measure and ν its compensator.

2.5.2 The nonlinear Feynman-Kac formula and the connection with PPIDE

In the present section we will consider the following path-dependent forward-backward system

$$\begin{cases} dX^{\varphi_\tau}(t) &= \mu_t(X_t^{\varphi_\tau}) dt + \sigma_t(X_t^{\varphi_\tau}) dW(s) + \int_{\mathbb{R}_0} \gamma_t(X_t^{\varphi_\tau}, z) \tilde{N}(dt, dz), \quad t \in [\tau, T] \subset [0, T], \\ X_\tau^{\varphi_\tau} &= \varphi_\tau \in \mathcal{D}, \\ dY^{\varphi_\tau}(t) &= f\left(X_t^{\varphi_\tau}, Y^{\varphi_\tau}(t), Z^{\varphi_\tau}(t), \int_{\mathbb{R}_0} U^{\varphi_\tau}(t, z) \delta(X^{\varphi_\tau}(t), z) \nu(dz)\right) dt + \\ &+ Z^{\varphi_\tau}(t) dW(t) + \int_{\mathbb{R}_0} U^{\varphi_\tau}(t, z) \tilde{N}(dt, dz), \\ Y^{\varphi_\tau}(T) &= \phi(X_T^{\varphi_\tau}) \end{cases}, \quad (2.150)$$

with μ , σ , γ , f , δ and ϕ some regular enough given functions. We refer assume that they satisfy suitable regularity conditions, such as Lipschitz continuity and at most linear growth at infinity. We will not prove the existence and uniqueness of a solution to the system (2.150) since it goes beyond the aim of the present work, we refer instead to [Kel14]. Also above we have denoted by X^{φ_τ} the value of the process X with initial value φ_τ at initial time τ .

The main goal of the present section is to show that, as standard in a Markovian setting, the solution to (2.150) can be retrieved as a solution of the following *path-dependent partial*

integro-differential equation (PPIDE)

$$\begin{cases} \nabla^H u_t(\varphi_t) + \mathbb{L}_t u_t(\varphi_t) = f(\phi, u_t(\varphi_t), \nabla^V u_t(\varphi_t) \sigma(\varphi_t), \mathcal{J} u_t(\varphi_t)) \\ u_t(\varphi_T) = \phi(\varphi_T), \quad \varphi_T \in \mathcal{D} \end{cases}, \quad (2.151)$$

for all $\varphi_t \in \mathcal{D}$, where $u_t : \mathcal{D} \rightarrow \mathbb{R}$ is an unknown function, \mathcal{J} is a functional defined by

$$\mathcal{J} u_t(\varphi_t) := \int_{\mathbb{R}_0} (u_t(\varphi_t^\gamma) - u_t(\varphi_t)) \delta(\varphi_t, z) \nu(dz),$$

and $u_t(\varphi_t^\gamma)$ is defined as in equation (2.147), \mathcal{L} is the infinitesimal generator of the forward component in (2.150), i.e.

$$\mathcal{L}_t u_t(\varphi_t) := \mu_t(\varphi_t) \nabla^V v(\varphi_t) + \sigma_t^2(\varphi_t) (\varphi_t) \nabla^{VV} u_t(\varphi_t) + \int_{\mathbb{R}_0} (u_t(\varphi_t^\gamma) - u_t(\varphi_t) - \gamma_t(\varphi_t, z) \nabla^V u_t(\varphi_t)) \nu(dz),$$

while ∇^H and ∇^V are the horizontal and vertical derivatives introduced in Sec. (2.5.1). We refer to [Kel14] for a complete treatment of forward-backward system of the form (2.151).

Theorem 2.5.5. *Let us consider $u \in C^{1,2}(\mathcal{D})$ solution to (2.151) satisfying*

$$|u_t(\varphi_t)| + |\nabla^V u_t(\varphi_t)| \leq C(1 + |\varphi_t|),$$

then the triplet

$$\begin{aligned} Y^{\varphi\tau}(t) &:= u_t(X_t^{\varphi\tau}) \quad \text{for every } t \in [\tau, T] \\ Z^{\varphi\tau}(t) &:= \nabla^V u_t(X_t^{\varphi\tau}) \sigma(X_t^{\varphi\tau}) \quad \text{for a.e. } t \in [\tau, T], \\ U^{\varphi\tau}(t, z) &:= u_t(\gamma X_t^{\varphi\tau}) - u_t(X_t^{\varphi\tau}) \quad \text{for a.e. } t \in [\tau, T] \end{aligned} \quad (2.152)$$

with

$$\gamma X_t^{\varphi\tau} = (X^{\varphi\tau}(t) + \gamma(X_t^{\varphi\tau}, z), X_t^{\varphi\tau}),$$

is the unique solution to the backward component in eq. (2.150).

Moreover the following representation holds

$$u_t(\varphi_t) = \mathbb{E}[\phi(X_T^{\varphi\tau})] + \int_t^T \mathbb{E} \left[f \left(X_s^{\varphi\tau}, Y^{\varphi\tau}(s), Z^{\varphi\tau}(s), \int_{\mathbb{R}_0} U^{\varphi\tau}(s, z) \delta(X^{\varphi\tau}(t), z) \nu(dz) \right) \right] ds. \quad (2.153)$$

Proof. By the pathwise Itô's formula (2.149), we have

$$\begin{aligned} u_T(X_T^{\varphi\tau}) &= u_t(X_t^{\varphi\tau}) + \int_t^T \nabla^H u_s(X_{s-}^{\varphi\tau}) ds + \frac{1}{2} \int_t^T \nabla^{VV} u_s(X_{s-}^{\varphi\tau}) \sigma^2(X_{s-}^{\varphi\tau}) ds \\ &\quad + \int_t^T \nabla^V u_s(X_{s-}^{\varphi\tau}) \mu(X_{s-}^{\varphi\tau}) ds + \int_t^T \int_{\mathbb{R}_0} \nabla^V u_s(X_{s-}^{\varphi\tau}) \gamma(X_{s-}^{\varphi\tau}) \tilde{N}(ds, dz) \\ &\quad + \int_t^T \nabla^V u(X_{s-}^{\varphi\tau}) \sigma(X_{s-}^{\varphi\tau}) dW(s) \\ &\quad + \int_t^T \int_{\mathbb{R}_0} (u_s(\gamma X_{s-}^{\varphi\tau}) - u_s(X_{s-}^{\varphi\tau}) - \gamma(s, X_{s-}^{\varphi\tau}) \nabla^V u_s(X_{s-}^{\varphi\tau})) N(ds, dz), \end{aligned} \quad (2.154)$$

furthermore since u solves eq. (2.151), then

$$\begin{aligned} \phi(X_T^{\varphi_\tau}) &= u_t(X_t^{\varphi_\tau}) - \int_t^T f\left(X_{s-}^{\varphi_\tau}, u(X_{s-}^{\varphi_\tau}), \nabla^V u(X_{s-}^{\varphi_\tau})\sigma(X_{s-}^{\varphi_\tau}), \mathcal{J}u(X_{s-}^{\varphi_\tau})\right) ds \\ &\quad + \int_t^T \nabla^V u(X_{s-}^{\varphi_\tau})\sigma(X_{s-}^{\varphi_\tau})dW(s) + \int_t^T \int_{\mathbb{R}_0} u(\gamma X_{s-}^{\varphi_\tau}) - u(X_{s-}^{\varphi_\tau})\tilde{N}(ds, dz), \end{aligned} \quad (2.155)$$

therefore $(Y^{\varphi_\tau}(t), Z^{\varphi_\tau}(t), U^{\varphi_\tau}(t, z))$, see eq. (2.152), is the unique solution to the backward component in (2.150), while eq. (2.153) is retrieved taking the expectation in eq. (2.155). \square

2.5.3 Comparison with the representation given in [CDPO16]

In what follows we show how present setting can be connected to functional one used, e.g., in [BCDNR16, CDPO16]. In particular we will denote by $\mathfrak{D} = \mathcal{D}([-r, 0]; \mathbb{R})$ the space of real-valued càdlàg functions on $[-r, 0]$. When endowed with the norm $\|\cdot\|_{\mathfrak{D}} = \sup_{s \in [-r, 0]} |\cdot|$. We have that $(\mathfrak{D}, \|\cdot\|_{\mathfrak{D}})$ is a non separable Banach space. Also we will use the notation $X(t)$ in order to indicate the present state of the process X at time t whereas X_t denote the path in \mathfrak{D} . Let us further fix an initial time $\tau \in [0, T]$, in what follows we will use the index τ in order to indicate the starting time $\tau \in [0, T]$ and $x \in \mathfrak{D}$ to indicate the initial value.

We will thus consider the following delayed forward-backward system

$$\left\{ \begin{aligned} dX^{\tau,x}(t) &= \hat{\mu}(t, X_t^{\tau,x}, X^{\tau,x}(t))dt + \hat{\sigma}(t, X_t^{\tau,x}, X^{\tau,x}(t))dW(s) + \\ &\quad + \int_{\mathbb{R}_0} \hat{\gamma}(t, X_t^{\tau,x}, X^{\tau,x}(t), z)\tilde{N}(dt, dz), \quad t \in [\tau, T] \subset [0, T], \\ X_{\tau}^{\tau,x} &= x \in \mathfrak{D}, \\ dY^{\tau,x}(t) &= \hat{f}\left(t, X_t^{\tau,x}, X^{\tau,x}(t), Y^{\tau,x}(t), Z^{\tau,x}(t), \int_{\mathbb{R}_0} U^{\tau,x}(t, z)\hat{\delta}(X_t^{\tau,x}, X^{\tau,x}(t), z)\nu(dz)\right)dt + \\ &\quad + Z^{\tau,x}(t)dW(t) + \int_{\mathbb{R}_0} U^{\tau,x}(t, z)\tilde{N}(dt, dz), \\ Y^{\tau,x}(T) &= \hat{\phi}(X_T^{\tau,x}) \end{aligned} \right. \quad (2.156)$$

where $\hat{\mu}$, $\hat{\sigma}$, $\hat{\gamma}$, \hat{f} , $\hat{\delta}$ and $\hat{\phi}$ are some given regular enough functions satisfying a suitable Lipschitz condition and linear growth at infinity, so that existence and uniqueness of a solution to (2.156) follows from [BCDNR16, CDPO16], we refer to [BCDNR16, CDPO16] for a complete characterization of the aforementioned coefficients in order to guarantee the existence and uniqueness of (2.156).

In particular, following the approach provided in [FZ], it can be shown that eq. (2.156) can be connected to eq. (2.150) exploiting the following operators

the *restriction* operator

$$M_t : D([-r, 0]; \mathbb{R}) \rightarrow D([0, t]; \mathbb{R}) \quad , \quad M_t(f)(s) = f(s - t), \quad s \in [0, t], \quad (2.157)$$

the *backward extension* operator

$$\begin{aligned} L_t &: D([0, t]; \mathbb{R}) \rightarrow D([-r, 0]; \mathbb{R}), \\ L_t(f)(s) &= f(0)\mathbb{1}_{[-r, -t]}(s) + f(t + s)\mathbb{1}_{[-t, 0]}(s), \quad s \in [-r, 0], \quad s \in [0, t], \end{aligned} \quad (2.158)$$

Let us consider a given functional $b \in \mathcal{D}$, one can define a different functional \hat{b} on $[0, T] \times \mathfrak{D}$ as

$$\hat{b}(t, X_t) := b_t \left(\tilde{M}_t X_t \right), \quad X_t \in \mathfrak{D}, \quad (2.159)$$

with

$$\tilde{M}_t X_t(s) := \begin{cases} M_t X_t(s) & \text{if } s \in [0, t) \\ X_t(s) & \text{if } s = t \end{cases}.$$

The converse holds true as well, in fact let us consider a given functional \hat{b} on \mathfrak{D} , then we can obtain a corresponding functional b on \mathcal{D} as

$$b_t(\varphi_t) := \hat{b}(t, L_t \varphi_t, \varphi(t)), \quad (\varphi_t, \varphi(t)) \in \mathcal{D}, \quad (2.160)$$

see [FZ] for details.

Next result shows that the representation given in theorem 2.5.5 coincide with the results obtained in [BCDNR16, CDPO16]. In what follows, we will use the notation introduced in [BCDNR16]. In particular given a function $v : [0, T] \times \mathfrak{D} \times \mathbb{R} \rightarrow \mathbb{R}$, we will denote by $\partial_t v(t, \eta, x)$ the classical \mathbb{R} -dimensional derivative w.r.t. to the first variable, that is the time-variable, $Dv(t, \eta, x)$ will be the Fréchet derivative w.r.t. the second variable, that is $\eta \in D([-r, 0]; \mathbb{R})$, and by $\partial_x v(t, \eta, x)$ the classical \mathbb{R} -dimensional derivative w.r.t. the third variable, that is the present value $x \in \mathbb{R}$. Also we will denote by $\partial_\theta \eta$, $\eta \in C^1 := C^1([-r, 0]; \mathbb{R})$ the weak derivative in distributional sense. Eventually we will denote by $\langle \cdot, \cdot \rangle_{\mathfrak{D}}$ the pairing in \mathfrak{D} . Any other notation is as already introduced.

Proposition 2.5.6. *Let us consider two given function $u : \mathcal{D} \rightarrow \mathbb{R}$ and $v : [0, T] \times \mathfrak{D} \times \mathbb{R} \rightarrow \mathbb{R}$ such that $u_t(\varphi) = v(t, \varphi(t), L_t \varphi)$, where L_t is the backward extension operator defined in (2.158). Then we have that the vertical derivative ∇^V of u coincide with the derivative w.r.t. the present state $\varphi(t)$ of v , namely*

$$\nabla^V u_t(\varphi_t) = \partial_x v(t, L_t \varphi_t, \varphi(t)). \quad (2.161)$$

and we have that it holds

$$u(\varphi_t^\gamma) = v(t, L_t \varphi_t, \varphi(t) + \gamma) - v(t, L_t \varphi_t, \varphi(t)). \quad (2.162)$$

Also we have that, if $L_t \varphi_t \in C^1$, it holds

$$\nabla^H u_t(\varphi_t) = \partial_t v(t, L_t \varphi_t, \varphi(t)) + \langle Dv(t, L_t \varphi_t, \varphi(t)), \partial_\theta L_t \varphi_t \rangle_{\mathfrak{D}}, \quad (2.163)$$

Proof. Following [FZ, Theorem 6.1], we have

$$\begin{aligned} \nabla_x u_t(\varphi) &= \lim_{h \rightarrow 0} \frac{1}{h} (u(t, \varphi^h) - u(t, \varphi)) = \lim_{h \rightarrow 0} \frac{1}{h} (v(t, L_t \varphi_t^h, \varphi^h(t)) - v(t, L_t \varphi_t, \varphi(t))) \\ &= \lim_{h \rightarrow 0} \frac{1}{h} (v(t, L_t \varphi_t^h, \varphi(t) + h) - v(t, L_t \varphi_t, \varphi(t))) = \frac{\partial}{\partial \varphi(t)} v(t, L_t \varphi_t, \varphi(t)). \end{aligned}$$

Equation follows also from [FZ, Theorem 6.1] whereas equation (2.162) it immediately follows from eq. (2.147) taking into account eq. (2.158). \square

Theorem 2.5.7. *Let us consider two given function $u : \mathcal{D} \rightarrow \mathbb{R}$ and $v : [0, T] \times \mathfrak{D} \times \mathbb{R} \rightarrow \mathbb{R}$ such that $u_t(\varphi_t) = v(t, L_t \varphi_t, \varphi(t))$, where L_t is the backward extension operator defined in (2.158). Then the representation given in theorem 2.5.5 and the representation given in [CDPO16, Th. 5.5] coincide.*

Proof. It immediately follows from Prop. 2.5.6. \square

Part II

Infinite dimensional analysis and network models

3 | Infinite dimensional analysis

Section 3.1 is taken from [CDP15a], Section 3.2 is taken from [BCP15] and Section 3.3 is taken from [CDP16c].

Abstract

We first study a particular class of forward rate problems, related to the Vasicek model, where the driving equation is a linear Gaussian stochastic partial differential equation. We thus give an existence and uniqueness results of the related mild solution in infinite dimensional setting, then we study the related *Ornstein-Uhlenbeck semigroup* with respect to the determination of a unique invariant measure for the associated Heath-Jarrow-Morton-Musiela model.

Thus we are concerned with existence and uniqueness of solution for the the optimal control problem governed by the stochastic FitzHugh-Nagumo equation, with and without recovery variable, driven by a Gaussian noise. First order conditions of optimality are also obtained.

3.1 Invariant measure for the Vasicek model in the Heath–Jarrow–Morton–Musiela framework

A typical problem in financial mathematics consists in determining the *right price* a certain financial product has to have from both the buyer and the seller point of view. Between the huge plethora of such contracts there are those related to forward rates determining the price a buyer has to pay now to purchase some type of financial asset at a later date. In the latter scenario buyers and sellers underwrite a contract to finalize a sales transaction on a determined future (*maturity*) date, for a specific price, hence they reach a legal agreement on what is called a *future* (contract), namely they sign a *Forward rate agreement* (FRA). Suppose we have to deal with a future with maturity time $T > 0$, hence the price of the related FRA depends on the actual time $t \in [0, T)$ we choose to enter in the investment. A standard way to formalize the latter is to define the forward rate $F(t, T)$ as the annual interest rate on a FRA starting at t and ending at T , namely the pre-agreed (fixed) interest rate on a FRA. It follows that we can consider the function $t \mapsto F(t, T)$ which determines the so called *forward curve* describing the time evolution of forward rates that share the same maturity date. Such type of curves can be used to manage portfolio risk, to determine

the present value of future returns linked to specific financial instrument, to analyse price fluctuations of raw materials against seasonal effects, to determine the time value of money against future inflation/deflation periods, to reveals commodity prices based on demand expectations and related managing expenses, to fix the interest rate at which banks borrow large amounts of money from each other as in the LIBOR (London Interbank Offered Rate) case, etc. There exists a huge amount of techniques used to study forward curve, e.g. bootstrapping techniques, see, e.g. [HW06], splines methods, see, e.g. [NS87, Wag97], etc. , each of which has has own pros and cons. Moreover, starting form the seminal work of Musiela, see [Mus93], a growing attention has been given to approaches based on infinite dimensional (stochastic) analysis. In particular during the last decade, see, e.g., [Con05, Fil01, Var99], a well structured theory for forward rates has been developed in the framework, see, e.g., [DPZ14], of the theory of stochastic partial differential equations (SPDEs) taking values in separable Hilbert spaces. In such approaches a fundamental issue consists in determining the right Hilbert space to work with, so that the related developed mathematical model could have a concrete financial meaning. In the present work we will focus our attention on the asymptotic behaviour of the *instantaneous* forward rate, see, e.g., [SN88], in the setting developed in [Fil01], where the main issue is the consistency problem related to the driving SPDE. In particular we provide a better and more general financial interpretation of results similar to those given in [Var99], where the author underlines that the chosen Hilbert spaces do not seem to be the appropriate ones for a concrete financial characterization of the developed theoretical results.

Namely we study a particular class of forward rate, where the driving equation is a linear Gaussian stochastic partial differential equation. In particular let $\beta(t)$, $t \geq 0$, be a standard Brownian motion on the real line on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$, \mathcal{F}_t being the σ -algebra generated by $\beta(s)$ with $s \leq t$, and let r_t represents the interest rate which behaves according to a Vasicek model, i.e.

$$r_t = r_0 + \int_0^t \kappa(\theta - r_s) ds + \int_0^t \sigma_{vas} d\beta(s),$$

for some real constants σ_{vas} , θ , κ and r_0 . The associated Heath-Jarrow-Morton-Musiela (HJMM) model, developed in [Mus93] as a generalization of the standard Heath-Jarrow-Morton (HJM) model, satisfies the following stochastic equation

$$\begin{cases} \frac{\partial f}{\partial t}(t, \tau) = \frac{\partial f}{\partial \tau}(t, \tau) + g(\tau) + \sigma(\tau) \frac{\partial \beta}{\partial t}(t), & \tau \geq 0 \\ f(0, \tau) = f_0(\tau), & \tau \geq 0 \end{cases}, \quad (3.1)$$

where $g, \sigma : \mathbb{R}_+ \rightarrow \mathbb{R}$ are given deterministic functions and $f_0 : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the actual market forward rate at initial ime $t = 0$, see, e.g. [CT07, Fil01, Sch04], for details about its rigorous derivation. Following the approach developed in [DPZ14], see also [DP06], eq. (3.1) can be rewritten as an Ornstein-Uhlenbeck (OU) equation in a suitable Hilbert space H_ω which will be later specified, see sec.3.1.1, namely

$$\begin{cases} df_t = (Af_t + g) dt + \sigma dW_t \\ f_0 = x \in H_\omega \end{cases}, \quad (3.2)$$

where W_t is a Wiener process on the real line, $A : D(A) \subset H_\omega \rightarrow H_\omega$ with

$$D(A) = \{\varphi \in H_\omega : \varphi' \in H_\omega\}, \quad A\varphi = \varphi',$$

is an unbounded operator whose domain will be specified later in sec.3.1.1, while g and σ are deterministic functions, independent of t , belonging to the space H_ω , namely

$$\sigma(\tau) := \sigma_{vas} e^{-\kappa\tau}, \quad g(\tau) := \frac{\sigma_{vas}^2}{\kappa} e^{-\kappa\tau} (1 - e^{-\kappa\tau}). \quad (3.3)$$

Exploiting the theory developed in [DP12, DPZ96, DPZ14], in [Var99] it is shown that there exists an invariant measure for the above described Hilbert space valued forward rate problem, but such an invariant measure is not uniquely determined.

In order to obtain a uniqueness result we first take into account the rescaled semigroup $e^{tA^n} := e^{tA - \frac{t}{n}}$, also perturbing the drift term g by a suitable coefficient. Then we will show that the operator A^n , for every $n \geq 1$, generates a strongly continuous semigroup of negative type, so that, see, e.g., [DPZ96], it admits a unique invariant measure on the *suitably chosen* Hilbert space H_ω . Such an approach is linked to the one developed in [Teh05], where it is shown that every HJMM-model admits a family of invariant measures parametrized by the distribution of the long rate. In particular we provide a uniquely determined extension of the Ornstein-Uhlenbeck (OU) semigroup of interest to a strongly continuous semigroup of contraction on the space $L^2(H_\omega, \mu)$, for a suitable measure μ on the state space H_ω , also giving an explicit expression for its infinitesimal generator, hence providing the well posedness of the associated Kolmogorov equation.

The work is structured as follows, we will first give existence and uniqueness results of a mild solution to eq. (3.1), then we will introduce the related transition semigroup R_t , the *Ornstein-Uhlenbeck semigroup* or OU semigroup for short, with an associated infinite family of invariant measures, since we do not have a uniqueness result. Then, considering the rescaled semigroup generated by $A^n = A - \frac{1}{n}$, we will show that, denoting by R_t^n the OU semigroup of the rescaled equation, a unique invariant measure μ^n does exist for R_t^n . We shall then prove that the OU semigroup R_t^n can be uniquely extended to a strongly continuous semigroup of contraction on the space $L^2(H_\omega, \mu^n)$ of square integrable functions $\varphi : H_\omega \rightarrow \mathbb{R}$ with respect to the unique invariant measure μ^n , and we give an explicit expression for the infinitesimal generator L_2^n of R_t^n on $L^2(H_\omega, \mu^n)$, also providing a convergence result for R_t^n to R_t .

3.1.1 Preliminary results

Infinite dimensional framework

A deep treatment of forward rate taking values in infinite dimensional Hilbert space has been developed in [Fil01] and [Var95], where particular attention has been given to the choice of the *right* Hilbert space that has to be chosen in order to effectively model the forward rates structures of interest. In [Var95], the author noticed that the choice of a weighted L^2 space as well as the one of a weighted Sobolev space $W^{1,2}$, fails to give an effective financial characterization of forward rates. In order to outdo such a drawback, we chose to develop our approach in the state space introduced in [Fil01]. In what follows we introduce the main definitions and results needed to establish the mathematical framework of our study, see [Fil01, §3] for details. In particular we consider the Heath-Jarrow-Morton-Musiela framework (HJMM), see, e.g., [CT07, Mus93, Sch04], for a detailed introduction to HJMM models and Musiela parametrization. Let us denote by $f_t(\tau)$ the forward curve at time $t \in \mathbb{R}_+ := [0, +\infty]$ with time to maturity $\tau \in \mathbb{R}_+$.

In [Fil01] it is underlined that, since the forward curve is obtained by smoothing data, it is reasonable to assume both that $\int_{\mathbb{R}_+} |f'_t(\tau)|^2 d\tau < \infty$, and that the interest rate today for a loan starting in 10 years does not differ much from a loan starting in 10 years and one day, hence we assume that the forward curve flattens for big time to maturity and that it can be modelled by penalizing its irregularities of for large τ , namely

$$\int_{\mathbb{R}_+} |f_t(\tau)|^2 \omega(\tau) d\tau < \infty,$$

$\omega(\tau)$ being a decreasing penalizing function to be properly chosen. In [Fil01] the following specification of H_ω has been proposed

Definition 3.1.1 (Def. 3.1 [Fil01]). Let be $\omega : \mathbb{R}_+ \rightarrow [0, \infty)$ a non decreasing C^1 -function s.t.

$$\omega^{-\frac{1}{3}} \in L^1(\mathbb{R}_+);$$

we write

$$\|\varphi\|_\omega := |\varphi(\infty)|^2 + \int_{\mathbb{R}_+} |\varphi'(\tau)|^2 \omega(\tau) d\tau, \quad (3.4)$$

and define

$$H_\omega := \{ \varphi \in L^1_{loc}(\mathbb{R}_+) : \exists \varphi' \in L^1_{loc}(\mathbb{R}_+) \text{ and } \|\varphi\|_\omega < \infty \}.$$

Remark 3.1.2. It is worth to mention that in [Fil01], the author suggests to replace the norm (3.4) by the equivalent norm

$$\|\varphi\|_\omega := |\varphi(0)|^2 + \int_{\mathbb{R}_+} |\varphi'(\tau)|^2 \omega(\tau) d\tau,$$

anyhow, as noted in [Teh05], since we will focus our attention on the limit at infinity of the forward curve, the norm defined by eq. (3.4) turns to be a more convenient choice.

With respect to previous definition 3.1.1, we have

Theorem 3.1.3 (Th. 3.2 [Fil01]). *The space $(H_\omega, \|\cdot\|_\omega)$ in Def. 3.1.1 is a infinite dimensional, separable Hilbert space.*

Moreover we also have

Corollary 3.1.4 (Cor. 3.6 [Fil01]). *We have*

$$D(A) = \{ \varphi \in H_\omega : \varphi' \in H_\omega \}, \quad A\varphi = \varphi'.$$

In what follows we will consider a weight function ω of the form $\omega(\tau) := e^{\kappa\tau}$, other choices could be made in the class of functions ω_1 satisfying the assumptions in Def. 3.1.1 and such that

$$\exists T \in \mathbb{R}_+ \quad \text{s.t.} \quad \forall \tau \geq T \quad \omega(\tau) \geq \omega_1(\tau).$$

Markov semigroups

Let us indicate by H_ω the infinite dimensional separable Hilbert space introduced in Def. 3.1.1 equipped with the scalar product $\langle \cdot, \cdot \rangle$ with associated norm denoted by $|\cdot|$. In what follows We will state some definition and well known results concerning *Markov semigroups* which we will use later on.

Let $C_b(H_\omega)$, resp. $B_b(H_\omega)$, the space of uniformly continuous and bounded, resp. bounded, functions $\varphi : H_\omega \rightarrow \mathbb{R}$, endowed with the sup norm $\|\cdot\|_\infty$

$$\|\varphi\|_\infty := \sup_{x \in H_\omega} \varphi(x), \quad \varphi \in C_b(H_\omega), \quad (\text{resp. } B_b(H_\omega)),$$

and let us denote by $C_b^*(H_\omega)$, resp. $B_b^*(H_\omega)$, its topological dual, while we indicate by $\mathcal{L}(C_b(H_\omega))$, resp. $\mathcal{L}(B_b(H_\omega))$, the space of linear and bounded operators from $C_b(H_\omega)$, resp. $B_b(H_\omega)$, into itself equipped with the standard operator norm

$$\|T\| := \sup_{\varphi \neq 0} \frac{\|T\varphi\|_\infty}{\|\varphi\|_\infty}, \quad \varphi \in C_b(H_\omega), \quad (\text{resp. } B_b(H_\omega)),$$

and $M(H_\omega)$ is the space of all probability measure on H_ω .

Definition 3.1.5 (Probability kernel). Let H_ω be a Hilbert space, a *probability kernel* $p(\cdot, \cdot)$ on H_ω is a mapping

$$[0, +\infty] \times H_\omega \rightarrow M(H_\omega), \quad (t, x) \mapsto p_{t,x},$$

such that

- (i) $p_{t+s,x}(A) = \int_{H_\omega} p_{s,y}(A) p_{t,x}(dy)$, for all $t, s \geq 0$, $x \in H$, $A \in \mathcal{B}(H_\omega)$;
- (ii) $p_x(A) := p_{0,x}(A) = \mathbb{1}_{[A]}(x)$, for all $x \in H_\omega$, $A \in \mathcal{B}(H_\omega)$.

Given a *probability kernel* we can define a semigroup of linear operators P_t on the space $B_b(H_\omega)$ of Borel bounded functions on H_ω , as follows

Definition 3.1.6 (Markov semigroup). A Markov semigroup P_t on $B_b(H_\omega)$ is a mapping

$$[0, \infty) \rightarrow \mathcal{L}(B_b(H_\omega)), \quad t \mapsto P_t,$$

such that

- (i) $P_0 = 1$, $P_{t+s} = P_t P_s$;
- (ii) for any $t \geq 0$ and $x \in H_\omega$ exists a probability kernel $p_{t,x}(\cdot) \in M(H_\omega)$ s.t.

$$P_t \varphi(x) = \int_{H_\omega} \varphi(y) p_{t,x}(dy), \quad \forall \varphi \in B_b(H_\omega);$$

- (iii) for any $\varphi \in C_b(H_\omega)$, resp. $B_b(H_\omega)$, and $\forall x \in H_\omega$ the mapping $t \mapsto P_t \varphi(x)$ is continuous, resp. of Borel type.

A *Markov semigroup* P_t is said to be

Feller : $\varphi \in C_b(H_\omega) \Rightarrow P_t \varphi \in C_b(H_\omega)$, for all $t \geq 0$;

strong Feller : $\varphi \in B_b(H_\omega) \Rightarrow P_t\varphi \in C_b(H_\omega)$, for all $t \geq 0$;

Definition 3.1.7 (Invariant measure). A probability measure μ in H_ω is said to be invariant for the *Markov semigroup* P_t if

$$\int_{H_\omega} P_t\varphi(x)\mu(dx) = \int_{H_\omega} \varphi(x)\mu(dx), \quad \forall t > 0, \varphi \in C_b(H_\omega). \quad (3.5)$$

Denoting by $M(H_\omega)$ the space of all probability measures on $(H, \mathcal{B}(H_\omega))$, there is natural embedding of $M(H_\omega)$ into $C_b^*(H_\omega)$, see, e.g., [DP06], indeed for any $\mu \in M(H_\omega)$ we can defined $F_\mu(\varphi) \in C_b^*(H_\omega)$ by

$$F_\mu(\varphi) := \int_{H_\omega} \varphi(x)\mu(dx),$$

so that μ turns to be identifiable with F_μ , moreover, if P_t is of Feller type, we can read eq. (3.5) as

$$P_t^*F_\mu = F_\mu, \quad t \geq 0,$$

where

$$\langle \varphi, P_t^*F_\mu \rangle = \langle P_t\varphi, F_\mu \rangle, \quad \varphi \in C_b(H_\omega), F_\mu \in C_b^*(H_\omega).$$

If an invariant measure μ exists, we are interested in the asymptotic behaviour, as $T \rightarrow \infty$, of the following quantity

$$M(T)\varphi(x) := \frac{1}{T} \int_0^T P_t\varphi(x)dt, \quad \varphi \in L^2(H_\omega, \mu), x \in H, T > 0. \quad (3.6)$$

Let us denote by Δ the linear subspace of $L^2(H, \mu)$ of all stationary points of P_t , namely

$$\Delta := \{ \varphi \in L^2(H_\omega, \mu) : P_t\varphi = \varphi, \forall t \geq 0 \}.$$

The following, due to Von Neumann, is a fundamental result that ensures the existence of the limit in eq. (3.6)

Theorem 3.1.8 (Th. 3.11 [DP12]). *Let P_t be a Markov semigroup and let μ be an invariant measure for P_t , and consider the following quantity*

$$M(T)\varphi(x) = \frac{1}{T} \int_0^T P_t\varphi(x)dt, \quad \varphi \in L^2(H_\omega, \mu), x \in H, T > 0,$$

then the following limit exists

$$\lim_{T \rightarrow \infty} M(T)\varphi =: M_\infty\varphi, \quad \text{in } L^2(H_\omega, \mu),$$

moreover $M_\infty(L^2(H_\omega, \mu)) = \Delta$ and $\int_H M_\infty\varphi(x)\mu(dx) = \int_H \mu(dx)$.

Definition 3.1.9 (Ergodic measure). Let P_t be a *Markov semigroup*, then an invariant measure μ for P_t is said to be *ergodic* if

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T P_t\varphi dt = M_\infty\varphi = \bar{\varphi}, \quad \varphi \in L^2(H_\omega, \mu), \quad (3.7)$$

where $\bar{\varphi} := \int_{H_\omega} \varphi(x)\mu(dx)$.

In the case of μ being ergodic we will say that the temporal average of $P_t\varphi$ coincides with the spatial average of φ .

Definition 3.1.10 (Strongly mixing). Let P_t be a *Markov semigroup*, then an invariant measure μ for P_t is said to be *strongly mixing* if

$$\lim_{t \rightarrow \infty} P_t\varphi = \bar{\varphi}, \quad \varphi \in L^2(H_\omega, \mu), \mu \text{ a.e.} \quad (3.8)$$

3.1.2 The Heat-Jarrow-Morton-Musiela SPDE

As pointed out in the introduction, we are interested in the study of the H_ω -valued OU eq. of the form

$$\begin{cases} df_t = (Af_t + g) dt + \sigma dW_t \\ f_0 = x \in H_\omega \end{cases}, \quad (3.9)$$

where W_t is a Wiener process on the real line, $A : D(A) \subset H_\omega \rightarrow H_\omega$ with

$$D(A) = \{\varphi \in H_\omega : \varphi' \in H_\omega, \}, \quad A\varphi = \varphi',$$

is an unbounded operator, and g and σ are t -independent, deterministic functions belonging to H_ω , and defined by

$$\sigma(\tau) = \sigma_{vas} e^{-\kappa\tau}, \quad g(\tau) = \frac{\sigma_{vas}^2}{\kappa} e^{-\kappa\tau} (1 - e^{-\kappa\tau}). \quad (3.10)$$

Remark 3.1.11. Let us notice that, following the infinite dimensional approach developed in [DPZ14], eq. (3.9) can be rewritten as follows

$$\begin{cases} dX_t = (AX_t + g(t)) dt + \sqrt{Q} dW_t \\ X_0 = x \in H \end{cases},$$

with $\sqrt{Q} \in \mathcal{L}(U, H)$, while W_t is a Wiener process on U being a generic Hilbert space which has to be *appropriately* chosen according with the particular model we are dealing with.

In what follows we shall chose $U := \mathbb{R}$ since $\sigma \in H_\omega$ is a deterministic function and the infinite dimensional problem comes from a one-dimensional noise. In particular $\sqrt{Q} \in \mathcal{L}(\mathbb{R}, H_\omega)$ acts as

$$(\sqrt{Q}\alpha)(\tau) = \alpha\sigma(\tau), \quad \sigma \in H_\omega, \tau \in \mathbb{R}_+, \alpha \in \mathbb{R}. \quad (3.11)$$

Following [DP12] we will denote the space of *mean square continuous processes* $F : [0, T] \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; H_\omega)$ adapted to W_t for any t and taking values in H_ω , as follows

$$C_W([0, T]; L^2(\Omega, \mathcal{F}, \mathbb{P}; H_\omega)) =: C_W([0, T]; H_\omega).$$

The space $C_W([0, T]; H_\omega)$ equipped with the norm

$$\|F\|_{C_W([0, T]; H_\omega)} := \left(\sup_{t \in [0, T]} \mathbb{E}|F(t)|^2 \right)^{\frac{1}{2}},$$

is a Banach space.

Definition 3.1.12. Given a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ and an \mathcal{F}_t -adapted cylindrical Wiener process W_t , a *mild solution* to eq. (3.9) is a mean square continuous H_ω -valued process adapted to W_t , such that for any $t \geq 0$ it holds

$$f_{x,t} = e^{tA}x + \int_0^t e^{(t-s)A}g ds + \int_0^t e^{(t-s)A}\sigma dW_s, \quad t \geq 0, \quad (3.12)$$

where the last term in eq. (4.59) is called *stochastic convolution*, often denoted by

$$W_t^A := \int_0^t e^{(t-s)A}\sigma dW_s.$$

Theorem 3.1.13. Eq. (3.9) admits a unique mild solution $f \in C_W([0, T], H_\omega)$, $T > 0$, of the form

$$f_{x,t} = e^{tA}x + \int_0^t e^{(t-s)A}g ds + \int_0^t e^{(t-s)A}\sigma dW_s, \quad t \geq 0, \quad (3.13)$$

moreover for any given $t > 0$, the mild solution (3.13) of eq. (3.9) is Gaussian with law $\mathcal{N}(e^{tA}x + G_t, Q_t)$, where the covariance operator Q_t reads as follow

$$Q_t := \int_0^t e^{sA}Qe^{sA^*} ds, \quad (3.14)$$

with $Q := \sigma\sigma^*$, while G_t is given by

$$G_t := \int_0^t e^{(t-s)A}g ds. \quad (3.15)$$

Furthermore for any given $T \in \mathbb{R}_+$, we have that $f \in C_W([0, T]; H_\omega)$.

Let us notice that we have previously denoted by σ^* the adjoint operator associated to \sqrt{Q} defined in eq. (3.11), exploiting the standard definition of the adjoint operator, namely, given $\sqrt{Q} \in \mathcal{L}(\mathbb{R}, H_\omega)$, we define the adjoint operator $\sqrt{Q}^* \in \mathcal{L}(H_\omega, \mathbb{R})$ as the unique operator such that the following holds

$$\left\langle \sqrt{Q}\alpha, \varphi \right\rangle_{H_\omega} = \left\langle \alpha, \sqrt{Q}^*\varphi \right\rangle_{\mathbb{R}}, \quad \forall \alpha \in \mathbb{R}, \forall \varphi \in H_\omega,$$

analogously, we have defined by A^* the adjoint operator of A . In order to prove Th. 3.1.13 we first need some auxiliary results concerning the covariance operator Q_t , the function g and the operator A .

Proposition 3.1.14. The operator $A : D(A) \subset H_\omega \rightarrow H_\omega$ generates a strongly continuous semigroup of contractions on the Hilbert space H_ω .

Proof. For the proof that A generates a strongly continuous semigroup on H_ω we refer to [Fil01, Th.3.2]. Let us then prove that the semigroup e^{tA} is in fact a contraction on H_ω , namely $\|e^{tA}\| \leq 1$. Since for any $f \in H_\omega$ it holds

$$|e^{tA}f| = c^2 + \int_{\mathbb{R}_+} |f'(\tau + t)|^2 e^{\omega\tau} d\tau \leq c^2 + \int_{\mathbb{R}_+} |f'(\tau)|^2 e^{\omega\tau} d\tau = |f|,$$

it follows that

$$\|e^{tA}f\| = \sup_{f \neq 0} \frac{|e^{tA}f|}{|f|} \leq \sup_{f \neq 0} \frac{|f|}{|f|} = 1,$$

and thus the claim. \square

Proposition 3.1.15. *For any $t > 0$ the operator*

$$Q_t := \int_0^t e^{sA} Q e^{sA^*} ds,$$

is a positive defined operator of trace class from the Hilbert space H_ω into itself, moreover the following holds

$$\sup_{t \geq 0} \text{Tr} Q_t < \infty.$$

Proof. For a proof that $f_{x,t} \sim \mathcal{N}(e^{tA}x + G_t, Q_t)$ we refer to [DPZ14], hence we are left with the proof that the operator $Q_t \in \mathcal{L}_1^+(H_\omega)$, where $\mathcal{L}_1^+(H_\omega)$ is the space of trace class positive defined operators, see, e.g., [DP12], for details. Let us denote by $\{e_k\}_k$ a basis for the Hilbert space H_ω , then applying [Var95, Lemma 8.5], we have

$$\begin{aligned} \text{Tr} Q_t &= \sum_{k=0}^{\infty} \langle Q_t e_k, e_k \rangle = \int_0^t \sum_{k=0}^{\infty} \langle e^{sA} Q e^{sA^*} e_k, e_k \rangle ds \\ &= \int_0^t \sum_{k=0}^{\infty} \langle (e^{sA} \sigma) (e^{sA} \sigma)^* e_k, e_k \rangle ds \\ &= \int_0^t \sum_{k=0}^{\infty} \langle \langle e^{sA} \sigma, e_k \rangle e^{sA} \sigma, e_k \rangle ds = \int_0^t \sum_{k=0}^{\infty} \langle e^{sA} \sigma, e_k \rangle \langle e^{sA} \sigma, e_k \rangle ds \\ &= \int_0^t \langle e^{sA} \sigma, e^{sA} \sigma \rangle ds, \end{aligned}$$

which implies that $\text{Tr} Q_t = \int_0^t |e^{sA} \sigma|^2 ds$, and since

$$|e^{sA} \sigma|^2 = \int_{\mathbb{R}_+} \sigma_{vas} \kappa e^{-2\kappa(\tau+s)} e^{\kappa\tau} d\tau = \sigma_{vas} e^{-2\kappa s} \int_{\mathbb{R}_+} \kappa e^{-\kappa\tau} d\tau = \sigma_{vas} e^{-2\kappa s},$$

we get

$$\text{Tr} Q_t = \int_0^t |e^{sA} \sigma|^2 ds = \int_0^t \sigma_{vas} e^{-2\kappa s} ds < \infty,$$

so that the operator Q_t is of trace class and actually the following stronger result holds

$$\sup_{t \geq 0} \text{Tr} Q_t = \lim_{t \rightarrow \infty} \int_0^t |e^{sA} \sigma|^2 ds = \int_{\mathbb{R}_+} \sigma_{vas} e^{-2\kappa s} ds < \infty. \quad (3.16)$$

□

Proposition 3.1.16. *The functions $\sigma := \sigma_{vas} e^{-\kappa\tau}$, $g := \frac{\sigma_{vas}^2}{\kappa} e^{-\kappa\tau} (1 - e^{-\kappa\tau})$, and $G_t := \int_0^t e^{(t-s)A} g ds$, all belong to H_ω .*

Proof. It clearly holds that

$$\lim_{\tau \rightarrow \infty} \sigma(\tau) = \lim_{\tau \rightarrow \infty} g(\tau) = \lim_{\tau \rightarrow \infty} G(\tau) = 0,$$

furthermore $\sigma, \sigma', g, g', G, G'$ are all elements of $L^1_{loc}(\mathbb{R}_+)$, hence we are left to prove that $|\iota| < \infty$, $\iota = \sigma, g, G$. We have

$$\begin{aligned} |\sigma| &= \int_{\mathbb{R}_+} \kappa^2 \sigma_{vas}^2 e^{-2\kappa\tau} e^{\kappa\tau} d\tau = \int_{\mathbb{R}_+} \kappa^2 \sigma_{vas}^2 e^{-\kappa\tau} d\tau < \infty, \\ |g| &= \int_{\mathbb{R}_+} \sigma_{vas}^4 (4e^{-4\kappa\tau} + e^{-2\kappa\tau} - 4e^{-3\kappa\tau}) e^{\kappa\tau} d\tau < \int_{\mathbb{R}_+} \sigma_{vas}^4 e^{-\kappa\tau} d\tau < \infty, \end{aligned}$$

moreover, for any $t > 0$, the following holds

$$\begin{aligned} G_t(\tau) &= \int_0^t e^{(t-s)A} g(\tau) ds = \int_\tau^{\tau+t} g(x) dx \\ &= -\frac{\sigma_{vas}^2}{2\kappa^2} \left(2e^{-\kappa(t+\tau)} - 2e^{-\kappa\tau} - e^{-2\kappa(t+\tau)} - e^{-2\kappa\tau} \right) \leq \frac{3\sigma_{vas}^2}{2\kappa^2} = G_\infty(\tau), \end{aligned}$$

so that

$$|G_t| < \frac{3\sigma_{vas}^2}{2\kappa^2} < \infty.$$

□

Proposition 3.1.17. *For any given $T > 0$, we have that $W_t^A \in C_W([0, T]; H_\omega)$, $\forall t \in [0, T]$. Furthermore, for any $m \in \mathbb{N}$, it holds*

$$\mathbb{E} \left[\sup_{t \geq 0} |W_t^A|^{2m} \right] \leq C_{m,T}, \quad (3.17)$$

where $C_{m,T}$ is a positive constant depending on m and T .

Proof. Let us first prove that $\forall t \geq 0$ $W_t^A \in C_W([0, T]; H_\omega)$. With no loss of generality we can take $t > s > 0$, and we have

$$W_t^A - W_s^A = \underbrace{\int_0^s (e^{(t-u)A} - e^{(s-u)A}) \sigma dW_s}_{I_1} + \underbrace{\int_s^t e^{(t-u)A} \sigma dW_s}_{I_2},$$

where I_1 and I_2 are independent random variables, hence

$$\mathbb{E}|I_1 + I_2|^2 = \mathbb{E}I_1^2 + \mathbb{E}I_2^2 - 2\mathbb{E}I_1\mathbb{E}I_2 = \mathbb{E}I_1^2 + \mathbb{E}I_2^2,$$

the last equality being implied by the fact that the stochastic integral has zero mean. Therefore, by the Itô isometry, we get

$$\begin{aligned} \mathbb{E}|W_t^A - W_s^A|^2 &= \mathbb{E}I_1^2 + \mathbb{E}I_2^2 \\ &= \int_0^s (e^{(t-u)A} - e^{(s-u)A})^2 \sigma^2 du + \int_s^t e^{2(t-u)A} \sigma^2 du, \end{aligned}$$

and taking the limit as $t \rightarrow s$, we have

$$\lim_{t \rightarrow s} \mathbb{E}|W_t^A - W_s^A|^2 = 0,$$

hence $W_t^A \in C_W([0, T]; H_\omega)$ and the stochastic convolution W_t^A is continuous \mathbb{P} -a.s. Concerning the estimate stated by eq. (3.17), we have that for any $u \in (0, 1)$, the following holds

$$\int_0^\infty s^{-u} Tr \left[e^{sA} Q e^{sA^*} \right] ds = \int_0^\infty s^{-u} |e^{sA} \sigma|^2 ds,$$

where

$$|e^{sA} \sigma|^2 = \int_0^\infty \sigma_{vas} \kappa e^{-2\kappa(\tau+s)} e^{\kappa\tau} d\tau = \sigma_{vas} \xi e^{-\kappa s},$$

then we get

$$\int_0^\infty s^{-u} |e^{sA} \sigma|^2 ds = \int_0^\infty \sigma_{vas} e^{-\kappa s} s^{-u} < \infty,$$

and the claim follows from [DPZ96, Th. 5.2.6]. \square

Proof of Th. 3.1.13 . Exploiting Prop.3.1.17, we have that $f_{x,t} \in C_W([0, T]; H_\omega)$, where $f_{x,t}$ is the solution to the OU eq. (3.9). Moreover using propositions 3.1.14, 3.1.15 and 3.1.16, we obtain that, for any given $t > 0$, the Gaussian measure $N_{e^{tA}x + G_t, Q_t}$ is well defined on the Hilbert space H_ω , see, e.g., [DPZ14], hence it follows, by contraction mapping principle, see [DPZ14, Th. 5.4], that there exists a unique mild solution to eq. (3.9) and it is given by eq. (3.13). \square

The Ornstein-Uhlenbeck semigroup

Let us consider the *Ornstein-Uhlenbeck (OU) process* (3.9). We have previously shown that the solution $f_{x,t}$ of the problem represented by eq. (3.9) is normally distributed, i.e. $f_{x,t} \sim \mathcal{N}(e^{tA}x + G_t, Q_t)$. Let us then define the *Ornstein-Uhlenbeck (OU) semigroup*, in the sense of definition 3.1.6, as follows

$$R_t \varphi(x) = \mathbb{E} \varphi(f_{x,t}) = \int_{H_\omega} \varphi(y) N_{e^{tA}x + G_t, Q_t}(dy), \quad \varphi \in B_b(H_\omega), t \geq 0, x \in H_\omega, \quad (3.18)$$

where we have denoted by $N_{a,Q}$ the infinite dimensional Gaussian measure with mean a and covariance operator Q , see, e.g., [DPZ14] for a detailed treatment of the topic. By an appropriate change of variable, expression in eq. (3.18) can be rewritten as follows

$$R_t \varphi(x) = \int_{H_\omega} \varphi(e^{tA}x + y) N_{G_t, Q_t}(dy), \quad \varphi \in B_b(H_\omega), t \geq 0, x \in H_\omega. \quad (3.19)$$

Definition 3.1.18. Let us denote by $\mathcal{E}(H_\omega)$ the *space of exponential functions*, namely

$$\mathcal{E}(H_\omega) := \text{span} \left\{ \Re \varphi_h(x), \Im \varphi_h(x), \text{ where } \varphi_h(x) = e^{i\langle h, x \rangle} : h, x \in H_\omega \right\},$$

then the following holds

Proposition 3.1.19. *The space $\mathcal{E}(H_\omega)$ is stable under the action of the semigroup R_t , namely $R_t \mathcal{E}(H_\omega) \subset \mathcal{E}(H_\omega)$.*

Proof. For any given $h, x \in H_\omega$, by eq. (3.19), we have

$$\begin{aligned} R_t \varphi_h(x) &= \int_{H_\omega} e^{i\langle e^{tA}x+y, h \rangle} N_{G_t, Q_t}(dy) = e^{i\langle e^{tA}x, h \rangle} \int_{H_\omega} e^{i\langle y, h \rangle} N_{G_t, Q_t}(dy) \\ &= e^{i\langle e^{tA}x, h \rangle} e^{i\langle G_t, h \rangle - \frac{1}{2}\langle Q_t h, h \rangle} = e^{i\langle G_t, h \rangle - \frac{1}{2}\langle Q_t h, h \rangle} e^{i\langle x, e^{tA^*} h \rangle} \\ &= e^{i\langle G_t, h \rangle - \frac{1}{2}\langle Q_t h, h \rangle} \varphi_{e^{tA^*} h}(x) \in \mathcal{E}(H_\omega). \end{aligned} \quad (3.20)$$

□

Moreover the OU semigroup R_t is Feller, in the sense of Def. 3.1.6, indeed we have

Proposition 3.1.20. *For all $t \geq 0$, the OU semigroup R_t defined by eq. (3.19) is Feller and the following estimate holds*

$$\|R_t \varphi\|_0 \leq \|\varphi\|_0, \quad t \geq 0, \varphi \in C_b(H_\omega). \quad (3.21)$$

Proof. For any given $t > 0$, let us consider $\varphi \in C_b^1(H_\omega)$, then we have

$$\forall \epsilon > 0, \exists \delta(\epsilon) : \forall x, z \in H_\omega \quad |x - z| < \delta(\epsilon) \Rightarrow |\varphi(x) - \varphi(z)| < \epsilon,$$

hence, for an appropriate constant C , if we set $\delta := \frac{\epsilon}{C}$, by the mean value theorem, we obtain

$$\begin{aligned} |R_t \varphi(x) - R_t \varphi(z)| &\leq \int_{H_\omega} |\varphi(e^{tA}x + y) - \varphi(e^{tA}z + y)| N_{G_t, Q_t}(dy) \\ &\leq \int_{H_\omega} |e^{tA}x + y - e^{tA}z - y| |\varphi'(\xi)| N_{G_t, Q_t}(dy) \\ &\leq e^{tA} |x - z| |\varphi'(\xi)| \leq C \delta(\epsilon) = \epsilon, \end{aligned}$$

therefore since $C_b^1(H_\omega) \stackrel{d}{\subset} C_b(H_\omega)$, we get that $R_t \varphi \in C_b(H_\omega)$, and estimate (3.21) easily follows. □

Exploiting Prop. 3.1.15, the following theorem is a byproduct of [DPZ96, Th.9.3.1], see also [Var99, Th.5],

Theorem 3.1.21. *Let us assume that*

(i) $\sup_{t \geq 0} \int_0^t e^{sA} Q e^{sA^*} ds < \infty;$

(ii) $\exists \bar{g} \in D(A)$ s.t. $A\bar{g} + g = 0$ and there exists an invariant measure ν for the equation

$$dZ_t = AZ_t dt, \quad (3.22)$$

the problem stated in eq. (3.9) admits invariant measures which are all of the form $\nu * \mathcal{N}(\bar{g}, Q_\infty)$, for $Q_\infty := \int_0^\infty e^{tA} Q e^{tA^*} dt$.

Furthermore from [Teh05], we also have that

Proposition 3.1.22. *Let us consider the functions*

$$\bar{g}(\tau) = - \int_0^\tau g(s) ds, \quad g_0 = \int_{\mathbb{R}_+} g(s); ds,$$

then for any $b \in \mathbb{R}$, we have that the measure $\lambda_b := \delta_{\bar{b}} * N_{\bar{g}+g_0, Q_\infty} = N_{\bar{g}+g_0+\bar{b}, Q_\infty}$, where δ_a is Dirac mass centred at a and $\bar{b} \in H_\omega$ is the constant function with value $b \in \mathbb{R}$ $\bar{b}(\tau) := b \mathbb{1}_{[\tau \geq 0]}$, is an invariant measure for the semigroup R_t on $C_b(H_\omega)$.

Previous results show that the problem stated by eq. (3.9) does not have a unique invariant measure with respect to the space H_ω . In particular for any $\bar{b} \in H_\omega$ as defined in Prop. 3.1.22, the measure $\{\lambda_b\}_b := \delta_{\bar{b}} * N_{\bar{g}+g_0, Q_\infty}$ is still an invariant measure. In Sec. 3.1.3 we will show that the measure $\mu := N_{\bar{g}+g_0+c, Q_\infty}$ can be chosen as the *natural* invariant measure associated to eq. (3.9) with respect to the space H_ω .

3.1.3 An approximation problem

Recalling the results given in Sec. 3.1.2, we know that the problem stated by eq. (3.9) admits an infinite number of invariant measures in the Hilbert space H_ω . We will now consider an *approximation problem* related to the one established in eq. (3.9) and such that it generates a sequence of solutions, uniquely determined at each step, which converges to the solution of eq. (3.9), for any $t > 0$, allowing us to select the most *natural exponent* within the set of invariant measures given in Prop. 3.1.22. Let us then consider the OU eq. of the form

$$\begin{cases} df_t^n = (A^n f_t^n + g^n) dt + \sigma dW_t \\ f_0^n = x \in H_\omega \end{cases}, \quad (3.23)$$

where $A^n := A - \frac{1}{n}$, $n \in \mathbb{N}_0 := \mathbb{N} \setminus \{0\}$ is the rescaled semigroup, while $g^n := g + \frac{1}{n}\bar{c}$, and where we have defined $\lim_{\tau \rightarrow \infty} x(\tau) = c < \infty$ and $c \mathbb{1}_{[\tau \geq 0]} =: \bar{c}(\tau) \in H_\omega$, and σ is as in eq. (3.10), then we have

Proposition 3.1.23. *For any $n \in \mathbb{N}$, the operator $A^n := A - \frac{1}{n}$ generates a strongly continuous semigroup of strict contraction of negative type over the Hilbert space H_ω , in particular, we have that $e^{tA^n} := e^{tA} e^{-\frac{1}{n}t}$, and $D(A) = D(A^n)$, furthermore $e^{tA^n} f \rightarrow e^{tA} f$ as $n \rightarrow \infty$, for all $f \in H_\omega$, uniformly for $t \in [0, T]$.*

Proof. Given $n \in \mathbb{N}$, the operator A^n generates the rescaled semigroup $e^{tA^n} := e^{-t\frac{1}{n}} e^{tA}$, which is strongly continuous on H_ω since e^{tA} is strongly continuous on H_ω , see, e.g., [EN00b, §1]. Furthermore, exploiting Prop. 3.1.14, we have that e^{tA} is a contraction on H_ω , see, e.g., [Fil01], therefore it has growth bound $\gamma_0 = 0$, the growth bound being defined as follows

$$\gamma_0 := \inf \left\{ \gamma \in \mathbb{R} : \exists M > 1 : \|e^{tA}\| \leq M e^{\gamma t}, \quad \forall t \geq 0 \right\},$$

see, e.g., [EN00b], therefore e^{tA^n} has strictly negative growth bound given by $\gamma_0 = -\frac{1}{n}$, $n \in \mathbb{N}_0$, so that e^{tA^n} is a C_0 -semigroup of contractions of negative type on the Hilbert space H_ω , see also [EN00b] for details about rescaled semigroup. In particular, since $A^n := A - \frac{1}{n}$, we have that, for any $n \in \mathbb{N}_0$, $D(A) = D(A^n)$ and $A^n f \rightarrow A f$, as $n \rightarrow \infty$, for all $f \in D(A)$, indeed

$$|A^n f - A f| = \left| f' - \frac{1}{n} f - f' \right| = -\frac{1}{n} |f| \rightarrow 0 \quad \text{as } n \rightarrow \infty, \quad \forall f \in D(A),$$

moreover, by Trotter-Kato theorem, see, e.g., [EN00b, Th. 4.8], it follows that $e^{tA^n} f \rightarrow e^{tA} f$ as $n \rightarrow \infty$, for all $f \in H_\omega$, uniformly for $t \in [0, T]$, $T > 0$. \square

Lemma 3.1.24. *The adjoint operators A^* and $(A^n)^*$ generate strongly continuous semigroups, moreover we have that $\lim_{n \rightarrow \infty} (A^n)^* \rightarrow A^*$.*

Proof. Recalling that the *weak* and *weak** topologies coincide on a separable Hilbert space, and since both A and A^n generate a strongly continuous semigroup on H_ω , then also A^* and $(A^n)^*$ generate a strongly continuous semigroup, moreover we have

$$\lim_{n \rightarrow \infty} \langle f, (A^n)^* g \rangle = \lim_{n \rightarrow \infty} \langle A^n f, g \rangle = \langle Af, g \rangle = \langle f, A^* g \rangle .$$

\square

Analogously to what is stated in 3.1.2, we have the following existence and uniqueness result

Theorem 3.1.25. *The problem stated by (3.23) admits a unique mild solution $f \in C_W([0, T]; H_\omega)$, which is given by*

$$f_{x,t}^n = e^{tA^n} x + \int_0^t e^{(t-s)A^n} g ds + c \left(1 - e^{-\frac{t}{n}}\right) + \int_0^t e^{(t-s)A^n} \sigma dW_s, \quad t \geq 0. \quad (3.24)$$

Proof. Analogously to what has been shown in the proof of Th. 3.1.13, we exploit results given by propositions 3.1.15, 3.1.16 and 3.1.17. Since Prop. 3.1.23 implies that A^n generates a strongly continuous semigroup, then, for any $n \in \mathbb{N}$, we have that $\frac{1}{n}c \in H_\omega$ and exploiting the fact that the constant function c is invariant under the translation semigroup e^{tA} , we have

$$\int_0^t e^{(t-s)A^n} \frac{\bar{c}}{n} ds = \int_0^t e^{-\frac{s}{n}} e^{sA} \frac{\bar{c}}{n} ds = \bar{c} \int_0^t \frac{1}{n} e^{-\frac{s}{n}} ds = \bar{c} \left(1 - e^{-\frac{t}{n}}\right),$$

so that eq. (3.24) follows. \square

In particular for any $t > 0$, the solution of eq. (3.24) is distributed as $\mathcal{N}(e^{tA^n} x + G_t^n, Q_t)$ with $G_t^n := G_t + \bar{c} \left(1 - e^{-\frac{t}{n}}\right)$, with G_t as in eq. (4.58), and the operator Q_t given as in eq. (3.14). Furthermore $f \in C_W([0, T]; H_\omega)$, for any $T \in \mathbb{R}_+$.

Proposition 3.1.26. *Let be $x \in H_\omega$ and $n \in \mathbb{N}$. Then we have that, for any $T \geq 0$,*

$$\lim_{n \rightarrow \infty} f_{x,t}^n = f_{x,t} \quad \text{in } C_W([0, T]; H_\omega),$$

namely

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, T]} \mathbb{E} |f_{x,t}^n - f_{x,t}| = 0,$$

where $f_{x,t}^n$ is the solution to eq. (3.23) and $f_{x,t}$ is the solution to eq. (3.9).

Proof. For any $n \in \mathbb{N}$ we have

$$\begin{aligned} f_{x,t} - f_{x,t}^n &= \underbrace{e^{tA}x - e^{tA^n}x}_{I_1} + \underbrace{\int_0^t (e^{(t-s)A} - e^{(t-s)A^n})g ds}_{I_2} \\ &\quad + \underbrace{\bar{c} \left(1 - e^{-\frac{t}{n}}\right)}_{I_3} + \underbrace{\int_0^t (e^{(t-s)A} - e^{(t-s)A^n})\sigma dW_s}_{I_4}, \quad t \geq 0, \end{aligned}$$

but I_1, I_2 and I_3 are deterministic quantities and the stochastic integral has zero mean, so that we have

$$\mathbb{E}|I_1 + I_2 + I_3 + I_4|^2 = (I_1 + I_2 + I_3)^2 + \mathbb{E}I_4^2,$$

therefore, by Prop. 3.1.23, since $e^{tA^n}f \rightarrow e^{tA}f$, for any $f, \in H_\omega$ as $n \rightarrow \infty$, uniformly in $t \in [0, T]$, and also exploiting the dominated convergence theorem, for any $T < \infty$, we have

$$\begin{aligned} |I_1| &\leq \sup_{t \in [0, T]} |e^{tA}x - e^{tA^n}x| := \mathcal{I}_1 \rightarrow 0 \quad \text{as } n \rightarrow \infty, \\ |I_2| &\leq \int_0^T |(e^{(T-s)A} - e^{(T-s)A^n})g| ds := \mathcal{I}_2 \rightarrow 0 \quad \text{as } n \rightarrow \infty, \\ |I_3| &\leq \bar{c} \left(1 - e^{-\frac{T}{n}}\right) := \mathcal{I}_3 \rightarrow 0 \quad \text{as } n \rightarrow \infty, \\ \mathbb{E}|I_4|^2 &\leq \int_0^T (e^{(t-s)A} - e^{(t-s)A^n})^2 \sigma^2 ds := \mathcal{I}_4 \rightarrow 0 \quad \text{as } n \rightarrow \infty. \end{aligned}$$

The latter implies that

$$\sup_{t \in [0, T]} \mathbb{E}|f_{x,t} - f_{x,t}^n|^2 \leq (\mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3)^2 + \mathcal{I}_4 \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

for any $T < \infty$, and taking the limit $T \rightarrow \infty$, we obtain

$$\begin{aligned} |I_1| &\leq \sup_{t \in \mathbb{R}_+} |e^{tA}x - e^{tA^n}x| \leq x(\infty) = c, \\ |I_2| &\leq \int_{\mathbb{R}_+} |(e^{(T-s)A} - e^{(T-s)A^n})g| ds \leq 0, \\ |I_3| &\leq \lim_{T \rightarrow \infty} \bar{c} \left(1 - e^{-\frac{T}{n}}\right) = \bar{c}, \\ \mathbb{E}|I_4|^2 &\leq \int_{\mathbb{R}_+} (e^{(t-s)A} - e^{(t-s)A^n})^2 \sigma^2 ds \leq 0, \end{aligned}$$

hence

$$\lim_{t \rightarrow \infty} \mathbb{E}|f_{x,t} - f_{x,t}^n|^2 \leq \bar{c} - \bar{c} = 0,$$

where the last inequality follows from the dominated convergence theorem. \square

Let us now consider the approximated OU-process (3.23), then the rescaled (OU) semi-group, in the sense of definition 3.1.6, is defined as follows

$$R_t^n \varphi(x) := \mathbb{E}\varphi(f_{x,t}^n) = \int_{H_\omega} \varphi(y) N_{e^{tA^n}x + G_t^n, Q_t^n}(dy), \quad \varphi \in B_b(H_\omega), t \geq 0, x \in H_\omega,$$

or, changing variable, we can also write

$$R_t^n \varphi(x) = \int_{H_\omega} \varphi(e^{tA^n} x + y) N_{G_t^n, Q_t}^n(dy), \quad \varphi \in B_b(H_\omega), t \geq 0, x \in H_\omega. \quad (3.25)$$

Proposition 3.1.27. *The linear operator*

$$Q_\infty^n x := \int_0^\infty e^{tA^n} Q e^{t(A^n)^*} x dt, \quad x \in H_\omega,$$

is well defined, trace class operator from H_ω into itself.

Proof. We have that

$$\begin{aligned} Q_\infty^n x &= \sum_{k=1}^\infty \int_0^1 e^{(s+k-1)A^n} Q e^{(s+k-1)(A^n)^*} x ds \\ &= \sum_{k=1}^\infty e^{(k-1)A^n} Q_1 e^{(k-1)(A^n)^*} x, \quad x \in H_\omega, \end{aligned} \quad (3.26)$$

therefore

$$\text{Tr} Q_\infty \leq \sum_{k=1}^\infty e^{-\frac{2}{n}(k-1)} \text{Tr} Q_1 < \infty,$$

where the boundedness of $\text{Tr} Q_1$ follows from Prop. 3.1.15, taking $t = 1$. \square

Proposition 3.1.28. *For every $n \in \mathbb{N}$, let \bar{g}^n, g_0^n be defined as follows*

$$\bar{g}^n(\tau) = - \int_0^\tau e^{-\frac{1}{n}(s-\tau)} g(s) ds, \quad g_0^n = \int_{\mathbb{R}_+} e^{-\frac{1}{n}(s-\tau)} g(s) ds, \quad (3.27)$$

then the vector $\bar{g}^n + g_0^n + \bar{c} \in H_\omega$.

Proof. By a direct computation, we obtain that

$$\bar{g}^n + g_0^n + \bar{c} = \frac{\sigma_{vas}^2}{\kappa} \frac{n}{n\kappa + 1} e^{-\kappa\tau} - \frac{\sigma_{vas}^2}{\kappa} \frac{2n}{2n\kappa + 1} e^{-2\kappa\tau} + \bar{c},$$

so that $\lim_{\tau \rightarrow \infty} \bar{g}^n + g_0^n + \bar{c} = c < \infty$. Eventually it holds

$$|\bar{g}^n + g_0^n + \bar{c}| = c + \int_{\mathbb{R}_+} \left(-\sigma_{vas}^2 \frac{n}{n\kappa + 1} e^{-\kappa\tau} + 2\sigma_{vas}^2 \frac{2n}{2n\kappa + 1} e^{-2\kappa\tau} \right)^2 e^{\kappa\tau} d\tau < \infty,$$

and the claim follows. \square

Proposition 3.1.29. *There exists a unique invariant measure μ^n for the OU transition semigroup R_t^n on the Hilbert space H_ω . In particular we have $\mu^n := N_{\bar{g}^n + g_0^n + \bar{c}, Q_\infty^n}$, with*

$$\bar{g}^n(\tau) = - \int_0^\tau e^{-\frac{1}{n}(s-\tau)} g(s) ds, \quad g_0^n = \int_{\mathbb{R}_+} e^{-\frac{1}{n}(s-\tau)} g(s) ds,$$

and $c = \lim_{\tau \rightarrow \infty} x(\tau)$. Furthermore, for any $\varphi \in C_b(H_\omega)$ and any $x \in H_\omega$, we have that R_t^n is strongly mixing, namely

$$\lim_{t \rightarrow \infty} R_t^n \varphi(x) = \int_H \varphi(y) \mu^n(dy). \quad (3.28)$$

Proof. Propositions 3.1.27 , Prop. 3.1.28 imply that $\bar{g}^n + g_0^n + \bar{c} \in H_\omega$ and $Q_\infty^n \in \mathcal{L}_1^+(H_\omega)$, hence the measure μ^n is well defined, moreover such a measure is the unique invariant measure associated to eq. (3.9) on the space H_ω and it is explicitly given by $\mu^n = N_{\bar{g}^n + g_0^n + \bar{c}, Q_\infty^n}$. Indeed, assuming that $\mu^n = N_{\bar{g}^n + g_0^n + \bar{c}, Q_\infty^n}$, then if $\forall t \geq 0, \varphi \in C_b(H_\omega)$, it holds

$$\int_{H_\omega} \varphi(x) \mu^n(dx) = \int_{H_\omega} R_t^n \varphi(x) \mu^n(dx), \quad (3.29)$$

then we have that μ^n is in fact an invariant measure for eq. (3.9). By density of $\mathcal{E}(H_\omega)$ in $C_b(H_\omega)$, it is sufficient to prove eq. (3.29) for exponential functions, then, by eq. (3.20), we have that eq. (3.29) reads as

$$\int_{H_\omega} \varphi_h(x) \mu^n(dx) = \int_{H_\omega} e^{i\langle G_t^n, h \rangle - \frac{1}{2} \langle Q_t^n h, h \rangle} \varphi_{e^{t(A^n)^*} h}(x) \mu^n(dx), \quad (3.30)$$

and since the characteristic function of a Gaussian measure of mean a and covariance Q is given by

$$N_{a, Q}^{\wedge}(h) = e^{i\langle a, h \rangle - \frac{1}{2} \langle Qh, h \rangle},$$

we have that eq. (3.30) holds if and only if

$$\hat{\mu}^n(h) = e^{i\langle G_t^n, h \rangle - \frac{1}{2} \langle Q_t^n h, h \rangle} \hat{\mu}^n(e^{t(A^n)^*} h), \quad (3.31)$$

that explicitly reads as

$$e^{i\langle \bar{g}^n + g_0^n + \bar{c}, h \rangle - \frac{1}{2} \langle Q_\infty^n h, h \rangle} = e^{i\langle G_t^n, h \rangle - \frac{1}{2} \langle Q_t^n h, h \rangle} e^{i\langle \bar{g}^n + g_0^n + \bar{c}, e^{t(A^n)^*} h \rangle - \frac{1}{2} \langle Q_\infty^n e^{t(A^n)^*} h, e^{t(A^n)^*} h \rangle}, \quad (3.32)$$

and, by the monotonicity of the exponential, $\forall h \in H_\omega$, eq. (3.32) reads as follow

$$\begin{aligned} & i\langle \bar{g}^n + g_0^n + \bar{c}, h \rangle - \frac{1}{2} \langle Q_\infty^n h, h \rangle = \\ & = i\langle G_t^n, h \rangle - \frac{1}{2} \langle Q_t^n h, h \rangle + i\langle \bar{g}^n + g_0^n + \bar{c}, e^{t(A^n)^*} h \rangle - \frac{1}{2} \langle Q_\infty^n e^{t(A^n)^*} h, e^{t(A^n)^*} h \rangle, \end{aligned} \quad (3.33)$$

which implies

$$\begin{aligned} & i\langle \bar{g}^n + g_0^n + \bar{c}, h \rangle - \frac{1}{2} \langle Q_\infty^n h, h \rangle = \\ & = i\langle G_t^n, h \rangle - \frac{1}{2} \langle Q_t^n h, h \rangle + i\langle e^{tA^n} (\bar{g}^n + g_0^n + \bar{c}), h \rangle - \frac{1}{2} \langle e^{tA^n} Q_\infty^n e^{t(A^n)^*} h, h \rangle. \end{aligned} \quad (3.34)$$

We have that, for any $t \geq 0$, it holds

$$\begin{aligned} Q_\infty^n &= \int_0^\infty e^{sA^n} Q e^{s(A^n)^*} ds = \int_0^t e^{sA^n} Q e^{s(A^n)^*} ds + \int_t^\infty e^{sA^n} Q e^{s(A^n)^*} ds = \\ &= Q_t^n + \int_0^\infty e^{(x+t)A^n} Q e^{(x+t)(A^n)^*} ds = Q_t^n + e^{tA^n} Q_\infty^n e^{t(A^n)^*}, \end{aligned} \quad (3.35)$$

furthermore we have

$$\begin{aligned} G_t^n &= \int_0^t e^{(t-s)A^n} g(\tau) ds + \bar{c} \left(1 - e^{-\frac{t}{n}}\right) \\ &= \int_0^{t+\tau} e^{-\frac{1}{n}(s-\tau)} g(s) ds - \int_0^\tau e^{-\frac{1}{n}(s-\tau)} g(s) ds + \bar{c} \left(1 - e^{-\frac{t}{n}}\right), \end{aligned} \quad (3.36)$$

so that it follows

$$\begin{aligned}
G_t^n + e^{tA^n}(\bar{g}^n + g_0^n + \bar{c}) &= \int_0^{t+\tau} e^{-\frac{1}{n}(s-\tau)} g(s) ds - \int_0^\tau e^{-\frac{1}{n}(s-\tau)} g(s) ds + \\
&- \int_0^{t+\tau} e^{-\frac{1}{n}(s-\tau)} g(s) ds + \int_{\mathbb{R}_+} e^{-\frac{1}{n}(s-\tau)} g(s) ds + \bar{c} + \bar{c}e^{-\frac{1}{n}t} - \bar{c}e^{-\frac{1}{n}t} \\
&= \bar{g}^n + g_0^n + \bar{c}.
\end{aligned} \tag{3.37}$$

Exploiting identities (3.37) and (3.35), we have that eq. (3.34) holds true and eq. (3.29) follows. The uniqueness is implied by eq. (3.31), taking the limit $t \rightarrow \infty$ and by the uniqueness of the Fourier transform, so that we get

$$\hat{\mu}^n(h) = e^{i\langle G_\infty^n, h \rangle - \frac{1}{2}\langle Q_\infty^n h, h \rangle} = e^{i\langle \bar{g}^n + g_0^n + \bar{c}, h \rangle - \frac{1}{2}\langle Q_\infty^n h, h \rangle}.$$

Let us then prove that μ^n is strongly mixing. Let $\varphi_h \in \mathcal{E}(H_\omega)$, then we have

$$\begin{aligned}
\lim_{t \rightarrow \infty} R_t^n \varphi_h(x) &= \lim_{t \rightarrow \infty} e^{i\langle e^{tA^n} h, x \rangle + i\langle G_t^n, h \rangle - \frac{1}{2}\langle Q_t^n h, h \rangle} \\
&= e^{i\langle G_\infty^n, h \rangle - \frac{1}{2}\langle Q_\infty^n h, h \rangle} = \int_{H_\omega} \varphi_h(x) \mu^n(dx),
\end{aligned}$$

and the result follows since $\mathcal{E}(H_\omega)$ is dense in $C_b(H_\omega)$. \square

Moreover we have the following convergence result

Proposition 3.1.30. *Let μ^n be the unique invariant measure for eq. (3.23) on H_ω , then the sequence $\{\mu^n\}_{n \in \mathbb{N}}$ converges, as $n \rightarrow \infty$, to the invariant measure μ for eq. (3.9) on H_ω , μ being defined as $N_{\bar{g}+g_0+\bar{c}, Q_\infty}$ with*

$$\bar{g}(\tau) = - \int_0^\tau g(s) ds, \quad g_0 = \int_{\mathbb{R}_+} g(s) ds,$$

moreover $\mu \in \{\lambda_b\}_b$.

Proof. Let us consider the Fourier transform $\hat{\mu}^n$ and $\hat{\mu}$. Then we have

$$\lim_{n \rightarrow \infty} \hat{\mu}^n = e^{i\langle \bar{g}^n + g_0^n + \bar{c}, h \rangle - \frac{1}{2}\langle Q_\infty^n h, h \rangle} = e^{i\langle \bar{g} + g_0 + \bar{c}, h \rangle - \frac{1}{2}\langle Q_\infty h, h \rangle} = \hat{\mu},$$

and propositions 3.1.28 , 3.1.27 , imply that

$$\begin{aligned}
\bar{g}^n + g_0^n + \bar{c} &= \frac{\sigma_{vas}^2}{\kappa} \frac{n}{n\kappa + 1} e^{-\kappa\tau} - \frac{\sigma_{vas}^2}{\kappa} \frac{2n}{2n\kappa + 1} e^{-2\kappa\tau} \\
&+ \bar{c} \rightarrow \frac{\sigma_{vas}^2}{\kappa} e^{-\kappa\tau} - \frac{\sigma_{vas}^2}{\kappa} e^{-2\kappa\tau} + \bar{c} = \bar{g} + g_0 + \bar{c}, \\
Q_\infty^n &= \int_0^\infty e^{tA^n} Q e^{t(A^n)^*} x dt \rightarrow \int_0^\infty e^{tA} Q e^{tA^*} x dt = Q_\infty,
\end{aligned}$$

where the limits follow by the convergence $e^{tA^n} \rightarrow e^{tA}$, for $n \rightarrow \infty$ as stated in Prop. 3.1.23. In particular by density of $\mathcal{E}(H_\omega)$ into $C_b(H_\omega)$, we have that the result holds for any bounded and continuous functions, so that $\mu^n \rightarrow \mu$ weakly, as $n \rightarrow \infty$. \square

The transition semigroup in $L^2(H_\omega, \mu)$

Proposition 3.1.31. *Let be μ^n the invariant measure in Prop. (3.1.29), then the semigroup R_t^n admits a unique extension to a strongly continuous semigroup of contractions in $L^2(H_\omega, \mu^n)$.*

Proof. By Holder inequality and from the invariance of μ^n , we have that

$$\int_{H_\omega} |R_t^n \varphi_h(x)|^2 \mu^n(dx) \leq \int_{H_\omega} R_t^n |\varphi_h|^2(x) \mu^n(dx) = \int_{H_\omega} |\varphi_h(x)|^2 \mu^n(dx),$$

and since $C_b(H_\omega) \stackrel{d}{\subset} L^2(H_\omega, \mu^n)$, see, e.g., [DP06], R_t^n is uniquely extendible to $L^2(H_\omega, \mu^n)$, therefore

$$\|R_t^n \varphi_h\|_{L^2(H_\omega, \mu^n)} \leq \|\varphi_h\|_{L^2(H_\omega, \mu^n)}, \quad t \geq 0, \varphi_h(x) \in L^2(H_\omega, \mu^n),$$

and strong continuity follows from dominated convergence theorem. \square

With a slight abuse of notation, we still denote the previously determined extension by R_t^n , and its generator by L_2^n with associated domain given by $D(L_2^n)$. Even if we do not know the domain $D(L_2^n)$ we are able to give an explicit form for the core of the infinitesimal generator L_2^n , see, e.g., [EN00b, Def. 1.6] for details on the core for an operator. Since for any $h \in H_\omega$ we have that

$$\varphi_h(x) \in D(L_2^n) \Leftrightarrow h \in D((A^n)^*),$$

then we introduce the following space

$$\mathcal{E}_{A^n}(H_\omega) := \text{span} \left\{ \Re e \varphi_h(x), \Im m \varphi_h(x), \varphi_h(x) = e^{i\langle h, x \rangle} : h \in D((A^n)^*) \right\} \subset \mathcal{E}(H_\omega).$$

Lemma 3.1.32. *For any $n \in \mathbb{N}$, we have that $\mathcal{E}_A(H_\omega) = \mathcal{E}_{A^n}(H_\omega)$. Furthermore for any $n \in \mathbb{N}$ we have that $\mathcal{E}_A(H_\omega) \stackrel{d}{\subset} L^2(H_\omega, \mu^n)$.*

Proof. Prop. 3.1.23 implies that that $D(A) = D(A^n)$, hence $\mathcal{E}_A(H_\omega) = \mathcal{E}_{A^n}(H_\omega)$, while the density result is assured by, e.g., [DP12, DPZ14]. \square

Let us notice that we also have that, for any $n \in \mathbb{N}_0$, $\mathcal{E}_A(H)$ is stable for R_t^n and furthermore that it is dense in $L^p(H_\omega, \mu^n)$ by [EN00b, Prop. 1.7], moreover the following holds

Theorem 3.1.33. *Let be R_t^n be the strongly continuous semigroup of contraction on $L^2(H_\omega, \mu^n)$ introduced in Prop. 3.1.31 whereas μ^n is the unique invariant measure introduced in Prop. 3.1.29, then we have that $\mathcal{E}_A(H_\omega)$ is a core for L_2^n . Moreover, for every $x \in H_\omega, \varphi \in \mathcal{E}_{A^n}(H_\omega)$, the infinitesimal generator L_2^n is of the form*

$$L_2^n \varphi(x) = \frac{1}{2} \text{Tr} [QD^2 \varphi(x)] + \langle x, (A^n)^* D\varphi(x) \rangle + \langle g, D\varphi(x) \rangle,$$

where we have denoted by D and D^2 , respectively the Fréchet derivative and the second order Fréchet derivative.

Proof. The stability of $\mathcal{E}_A(H_\omega)$ under R_t^n comes from eq. (3.20), namely

$$R_t^n \mathcal{E}_A(H_\omega) \subset \mathcal{E}_A(H_\omega).$$

The density of $\mathcal{E}_A(H_\omega)$ into $L^2(H, \mu^n)$ follows since $C_b(H_\omega) \stackrel{d}{\subset} L^p(H_\omega, \mu^n)$, $\mathcal{E}(H_\omega)$ is dense in $C_b(H_\omega)$, and by the dominated convergence theorem, while $\mathcal{E}_A(H_\omega)$ is a core for L_2^n by [EN00b, Prop. 1.7]. Explicitly computing the infinitesimal generator L_2^n we obtain that if $h \in D(A^*)$, then, $\forall x \in H_\omega$, $\varphi_h \in \mathcal{E}_{A^n}(H_\omega)$, we have

$$\begin{aligned} L_2^n &= \lim_{t \downarrow 0} \frac{1}{t} (R_t^n \varphi_h - \varphi_h) = \lim_{t \downarrow 0} \frac{1}{t} \left(e^{i\langle G_t^n, h \rangle - \frac{1}{2}\langle Q_t^n h, h \rangle} \varphi_{e^{t(A^n)^*} h}(x) - \varphi_h \right) \\ &= \lim_{t \downarrow 0} \frac{1}{t} \left(i\langle G_t^n, h \rangle - \frac{1}{2}\langle Q_t^n h, h \rangle + i\langle e^{t(A^n)^*} h - h, x \rangle \right) \varphi_h(x) \\ &= \frac{1}{2} Tr [QD_x^2 \varphi_h(x)] + \langle x, (A^n)^* D_x \varphi_h(x) \rangle + \langle g, D_x \varphi_h(x) \rangle. \end{aligned}$$

□

We have proven in Prop. 3.1.22 that the semigroup R_t on $C_b(H_\omega)$ admits an infinite family of invariant measures of the form $\delta_b * \mu$. In particular it means that R_t can be extended, even if not uniquely, following Prop. 3.1.31, to a strongly continuous semigroup on $L^2(H_\omega, \lambda_b)$ for any λ_b of the form $\lambda_b = \delta_b * \mu$, with the same techniques employed in 3.1.3. In particular we have that Thm. 3.1.33 holds also for R_t on the space $L^2(H_\omega, \lambda_b)$, with λ_b an invariant measure of the form given in Prop. 3.1.22, indeed we have that $\mathcal{E}_A(H_\omega)$ is still a core for L_2 , where L_2 is the infinitesimal generator for the semigroup R_t over any space $L^2(H_\omega, \lambda_b)$ and, analogously to what has been made in Thm. 3.1.33, we have that L_2 is

$$L_2 \varphi(x) = \frac{1}{2} Tr [QD^2 \varphi(x)] + \langle x, A^* D \varphi(x) \rangle + \langle g, D \varphi(x) \rangle, \quad x \in H_\omega, \varphi \in \mathcal{E}_A(H_\omega).$$

The following proposition aims at showing that the most natural invariant measure for the semigroup R_t to consider is the measure μ , since it is the unique invariant measure for the rescaled semigroup R_t^n , which converges to R_t for any $\varphi_h(x) \in \mathcal{E}_A(H_\omega)$.

Proposition 3.1.34. *We have that $R_t^n \varphi_h(x) \rightarrow R_t \varphi_h(x)$ as $n \rightarrow \infty$, for any $\varphi_h \in \mathcal{E}_A(H_\omega)$, uniformly for $t \geq 0$.*

Proof. By Thm. 3.1.33 we have that, for any $k \in \mathbb{N}_0$, $\mathcal{E}_A(H_\omega)$ is a core for the semigroup R_t^k , hence, for any $\varphi_h(x) \in \mathcal{E}_A(H_\omega)$, we have that

$$R_t^k \varphi_{2h}(x) = e^{i\langle G_t^k, 2h \rangle - \frac{1}{2}\langle Q_t^k h, 2h \rangle} \varphi_{e^{t(A^k)^*} 2h}(x) \in \mathcal{E}_A(H_\omega),$$

and the same holds for R_t , so that we have

$$R_t \varphi_{2h}(x) = e^{i\langle G_t, 2h \rangle - \frac{1}{2}\langle Q_t h, 2h \rangle} \varphi_{e^{tA^*} 2h}(x) \in \mathcal{E}_A(H_\omega).$$

By Prop. 3.1.23 and Lemma 3.1.24, we have that for, any $h \in H_\omega$, $e^{tA^k} h \rightarrow e^{tA} h$ and $e^{t(A^k)^*} h \rightarrow e^{tA^*} h$, therefore by the dominated convergence theorem, it holds that for any

fixed $m \in \mathbb{N}_0$, we have

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_{H_\omega} e^{i\langle G_t^k, 2h \rangle - \frac{1}{2} \langle Q_t^k h, 2h \rangle} \varphi_{e^{t(A^k)^*} 2h}(x) \mu^m(dx) &= \\ &= \int_{H_\omega} e^{i\langle G_t, 2h \rangle - \frac{1}{2} \langle Q_t 2h, 2h \rangle} \mu^m(dx), \end{aligned} \quad (3.38)$$

conversely, for any fixed $k \in \mathbb{N}_0$, Prop. 3.1.30 implies that

$$\begin{aligned} \lim_{m \rightarrow \infty} \int_{H_\omega} e^{i\langle G_t^k, 2h \rangle - \frac{1}{2} \langle Q_t^k h, 2h \rangle} \varphi_{e^{t(A^k)^*} 2h}(x) \mu^m(dx) &= \\ &= \int_{H_\omega} e^{i\langle G_t^k, 2h \rangle - \frac{1}{2} \langle Q_t^k h, 2h \rangle} \varphi_{e^{t(A^k)^*} 2h}(x) \mu(dx), \end{aligned} \quad (3.39)$$

therefore, exploiting both eq. (3.38) and eq. (3.39), we have that, for any $\varphi_h(x) \in \mathcal{E}_A(H_\omega)$,

$$\begin{aligned} \lim_{k \rightarrow \infty} \lim_{m \rightarrow \infty} \|R_t^k \varphi_h(x)\|_{L^2(H_\omega, \mu^m)} &= \lim_{m \rightarrow \infty} \lim_{k \rightarrow \infty} \|R_t^k \varphi_h(x)\|_{L^2(H_\omega, \mu^m)} \\ &= \|R_t \varphi_h(x)\|_{L^2(H_\omega, \mu)}, \end{aligned}$$

and the uniform convergence for $t \geq 0$ follows from the uniform convergence in t of $e^{tA^n} \rightarrow e^{tA}$ proved in Prop. 3.1.23. \square

Exploiting the result given by Prop. 3.1.34 we are in position to select a specific invariant measure among the infinite $\{\lambda_b\}_b$, uniquely extendible to the OU semigroup R_t to $L^2(H_\omega, \mu)$. In particular, defining

$$u(t, x) := \mathbb{E} \varphi(f_{x,T}) | \mathcal{F}_t, \quad \varphi \in L^2(H_\omega, \mu),$$

we have that u solves the following backward Kolmogorov equation

$$\begin{cases} \frac{\partial u}{\partial t}(x, t) + \frac{1}{2} \text{Tr} [QD^2 \varphi(x)] + \langle Ax + g, D\varphi(x) \rangle \\ u(T, \cdot) = \varphi(\cdot) \end{cases}.$$

In what follows we consider a specific case of latter *extension result*, particularly useful when dealing with financial applications. Let us define

$$u(t, x) := \mathbb{E} e^{\int_t^T -V(f_{x,s}) ds} \varphi(f_{x,T}) | \mathcal{F}_t, \quad \varphi \in L^2(H_\omega, \mu),$$

then u solves

$$\begin{cases} \frac{\partial u}{\partial t}(x, t) + \frac{1}{2} \text{Tr} [QD^2 \varphi(x)] + \langle Ax + g, D\varphi(x) \rangle - V(x)u(x, t) \\ u(T, \cdot) = \varphi(\cdot) \end{cases}, \quad (3.40)$$

with $V \in C_b(H_\omega)$ being the so called *discounting process*, while φ represents the final payoff of a financial contract. In particular the solution $u(0, x)$ of eq. (3.40) is the so called *fair price* of a contingent claim with terminal value φ .

3.1.4 Conclusions

Since there exists an infinite number of invariant measures associated to the problem stated by eq. (3.9), see, e.g., [Teh05, Var99], the semigroup R_t can not be uniquely extended to a strongly continuous semigroup on a suitable weighted L^2 space. Nevertheless we have proven that, when properly rescaled, the original Ornstein-Uhlenbeck eq. (3.9) can be studied via a series of approximating problems represented by eq. (3.23) whose solutions uniformly converge to the solution of eq. (3.9) and such that there exists an associated, unique invariant measure for each approximating step, see Prop. 3.1.29. Therefore, following the approach given in [DP12, DPZ96], we can uniquely extend the approximated OU semigroup R_t^n to a strongly continuous semigroup over the space of square integrable functions with respect to the above mentioned unique invariant measure, see Prop. 3.1.31, besides an explicit representation for both the associated semigroup core and the respective infinitesimal generator, it is shown. Exploiting previous results, we have also proven that, among the infinite number of invariant measure associated to eq. (3.9), it is possible to select one candidate which turns to be the *natural choice*. Latter result is of particular relevance when we are dealing with problems coming from the mathematical finance arena and, as an example, this is the case represented by the so called Heath-Jarrow-Morton model used to describe the time behaviour of both interest rate curve and instantaneous forward rate curve.

3.2 Optimal control of stochastic FitzHugh-Nagumo equation

Consider here the reaction-diffusion equation

$$\begin{cases} dX(t, \xi) - \Delta_\xi X(t, \xi) dt + f(X(t, \xi)) dt = \sqrt{Q} dW(t) + F(t, \xi) dt & (t, \xi) \in [0, T] \times \mathcal{O}, \\ X(t, \xi)|_{\partial\mathcal{O}} = 0, & t \in [0, T], \\ X(0, \xi) = x(\xi), & \xi \in \mathcal{O}, x \in L^2(\mathcal{O}) \end{cases}, \quad (3.41)$$

in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $f(u) = u(u - a)(u - b)$, $\forall u \in \mathbb{R}$, $\mathcal{O} \subset \mathbb{R}^d$, $d = 1, 2, 3$ is a bounded and open set with smooth boundary $\partial\mathcal{O}$, $W(t)$ is a cylindrical Wiener process and $Q \in \mathcal{L}(L^2(\mathcal{O}), L^2(\mathcal{O}))$ (the space of linear and continuous operator from $L^2(\mathcal{O})$ into itself equipped with the operator norm) is a self-adjoint positive operator with $\text{Tr}Q < \infty$. Here $a, b \in L^\infty([0, T] \times \mathcal{O})$ and $x \in L^2(\mathcal{O})$ are given. Also $F \in L^2([0, T] \times \mathcal{O})$. We shall denote by $(\mathcal{F}_t)_{t \geq 0}$ the natural filtration induced by $W(t)$. Equation (3.102) can be rewritten as follows

$$\begin{cases} dX(t) + AX(t) dt + f(X(t)) dt = \sqrt{Q} dW(t) + F(t) dt, & t \in [0, T], \\ X(0) = x, & x \in L^2(\mathcal{O}) \end{cases}, \quad (3.42)$$

A being the Laplace operator $-\Delta_\xi$ with domain $D(A) := H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$. In the special case $a, b \in \mathbb{R}$, (3.102) is the dimensionless form of the celebrated *FitzHugh-Nagumo* equation, see, e.g., [ADP11] and reference therein, perturbed by a coloured Gaussian noise $\sqrt{Q}W$. Its deterministic counterpart has been introduced by FitzHugh (1922–2007) and Nagumo, see [Fit61, NAY62] in order to model the conduction of electrical impulses in a nerve axon. In particular X is the nerve membrane potential and $F := -V + I$ where V is the ion concentration and I is the applied current. The Gaussian perturbation is the effect of

random input currents in neurons and their source is the random opening or closing of ion channels, see, e.g. [Tuc92]. In 2-D and 3-D equation (3.102) is relevant in statistical mechanics where it is called *Ginzburg-Landau* equation and also in phase transition models of *Ginzburg-Landau* type, see, e.g. [Die97]. We would like to underline that nonlinear potential of the form $f(u) = u(u - a)(u - b)$ arising here are specific for diffusion processes in excitable media or for phase transition.

In what follows we will study the optimal control problem for (3.102) providing an existence and uniqueness result as well as the first order necessary conditions for optimality, namely the maximum principle. In Sec. 3.2.1 we shall prove the well-posedness of problem (3.102), see [BMZ⁺08] for other results of this type.

The existence of a solution to optimal control problem (P) will be proved under suitable conditions on time interval $[0, T]$ and the cost functional in Sec. 3.2.1. It should be mentioned that there exists a large literature concerning the optimal control problems governed by the deterministic FitzHugh-Nagumo equation, see, e.g., [CRT13, KW13], while to the best of our knowledge, the stochastic case that we are interested in, lacks of such results. The motivation is that existence of an optimal control for the stochastic problem we consider here is quite a delicate problem which cannot be solved with standard optimization arguments which require the weak lower semicontinuity of cost functional in the control basic space and a more subtle argument based on Eckelands's variational principle was used. The existence result we obtain here is the main novelty of this work. To prove the existence of an optimal control an essential property of nonlinear function f is that it is ultimately monotonically increasing, that is outside a bounded interval.

We shall use the basic notions and standard notation $L^p(\mathcal{O})$, $1 \leq p \leq \infty$ and $H^k(\mathcal{O})$, $k = 1, 2$, $H_0^1(\mathcal{O})$ for spaces of Lebesgue p -integrable functions on \mathcal{O} and respectively, Sobolev spaces on \mathcal{O} . The norm in $L^p(\mathcal{O})$ will be denoted by $\|\cdot\|_p = \|\cdot\|_{L^p(\mathcal{O})}$ and the scalar product in $L^2(\mathcal{O})$ by $\langle \cdot, \cdot \rangle_2$. Given a Banach space Y we shall denote by $\|\cdot\|_Y$ its norm. By $C([0, T]; Y)$ we denote the space of Y -valued continuous functions on $[0, T]$ and by $L^p([0, T]; Y)$ the space of p -integrable Y -valued functions on $[0, T]$. By $W^{1,p}([0, T]; Y)$, $1 \leq p \leq \infty$ we shall denote the space of absolutely continuous functions $u : [0, T] \rightarrow Y$ such that $\frac{du}{dt} \in L^p([0, T]; Y)$.

We shall use the standard notations, see, e.g. [DP12], for spaces of processes defined in probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0}, W)$. $C_W([0, T]; L^2(\mathcal{O}))$ is the space of all $L^2(\mathcal{O})$ -valued $(\mathcal{F}_t)_{t \geq 0}$ -adapted process such that $u \in C([0, T]; L^2(\Omega, L^2(\mathcal{O})))$. Similarly, $L_W^2([0, T]; H_0^1(\mathcal{O}))$ is the space of all $(\mathcal{F}_t)_{t \geq 0}$ -adapted processes $u \in L^2([0, T]; L^2(\Omega, H_0^1(\mathcal{O})))$.

We denote by W_A the stochastic convolution defined by

$$W_A(t) := \int_0^t e^{-(t-s)A} \sqrt{Q} dW(s), \quad \forall t \geq 0.$$

In the following we shall assume that

$$\mathbb{E} \sup_{(t, \xi) \in [0, T] \times \mathcal{O}} |W_A(t, \xi)|^{2m} < \infty, \quad m \in [1, 2]. \quad (3.43)$$

Sufficient conditions for (3.43) to hold are given in [DP12, Th.2.13]. We refer to [BP12] for standard results on convex analysis which will be used in the following.

3.2.1 Existence for equation (3.102)

Definition 3.2.1. We say that the function $X \in C_W([0, T]; L^2(\mathcal{O}))$ is called a *mild solution* to (3.102) if $X(t) : [0, T] \rightarrow L^2(\mathcal{O})$ is continuous \mathbb{P} -a.s., $\forall t \in [0, T]$ and it satisfies the stochastic integral equation

$$X(t) = e^{-At}x - \int_0^t e^{-(t-s)A} (f(X(s)) - F(s)) ds + W_A(t), \quad \forall t \in [0, T].$$

Theorem 3.2.2. Assume that assumption (3.43) holds and that $x \in H_0^1(\mathcal{O})$. Then there exists a unique solution X to (3.102) which satisfies

$$X \in C_W([0, T]; H_0^1(\mathcal{O})) \cap L_W^2([0, T]; H^2(\mathcal{O})) \cap L^2(\Omega; C([0, T]; H_0^1(\mathcal{O}))).$$

We note in particular that assumption (3.43) holds if $Q = A^{-\frac{\gamma}{2}}$, with $\gamma > \frac{d}{2} - 1$, see, e.g., [DP12, Prop.4.3].

If we define the stochastic process $y := X - W_A$, then equation (3.102) reduces to the random parabolic equation

$$\begin{cases} y_t(t, \xi) - \Delta_\xi y(t, \xi) + y^3(t, \xi) + f_1(t, \xi)y^2(t, \xi) + f_2(t, \xi)y(t, \xi) = f_3(t, \xi) & \text{in } [0, T] \times \mathcal{O}, \\ y(t, \xi) = 0 & \text{in } [0, T] \times \partial\mathcal{O}, \\ y(0, \xi) = y_0(\xi), \xi \in \mathcal{O} \end{cases}, \quad (3.44)$$

where $f_1, f_2 \in L^\infty([0, T] \times \mathcal{O})$, $f_3 \in L^2([0, T] \times \mathcal{O})$ are $(\mathcal{F}_t)_{t \geq 0}$ -adapted $L^2(\mathcal{O})$ -valued processes on $[0, T]$. More precisely $f_1 = a + 3W_A$, $f_2 = b + 3W_A^2 + 2W_A$, $f_3 = -W_A^3 - 9W_A^2 - bW_A + F$.

The following proposition states an existence and uniqueness result for equation (3.44)

Proposition 3.2.3. Assume $x \in H_0^1(\mathcal{O})$. Then there is a unique solutions to equation (3.44) satisfying \mathbb{P} -a.s.

$$y \in C([0, T]; H_0^1(\mathcal{O})) \cap L^2([0, T]; H^2(\mathcal{O})). \quad (3.45)$$

Moreover the process $t \mapsto y(t)$ is $(\mathcal{F}_t)_{t \geq 0}$ -adapted.

Proof. Let us consider, for fixed $\omega \in \Omega$, the set

$$K = \left\{ y \in C([0, T^*]; L^2(\mathcal{O})) : \|y\|_{L^\infty([0, T^*]; H_0^1(\mathcal{O}))} \leq R, \quad 0 \leq T^* \leq T \right\}, \quad (3.46)$$

where R is a positive real constant and T^* has to be chosen later on.

The set K is closed in $C([0, T^*]; L^2(\mathcal{O}))$ and therefore it is a complete metric space when equipped with the metric

$$\rho(y, v) = \sup_{t \in [0, T^*]} |y(t) - v(t)|_2. \quad (3.47)$$

Let $z \in K$ and let us consider the operation $F : K \rightarrow K$ defined by $Fz = y$, where y is solution to

$$\begin{cases} y_t(t, \xi) - \Delta_\xi y(t, \xi) + y^3(t, \xi) = -f_1(t, \xi)z^2(t, \xi) - f_2(t, \xi)z(t, \xi) + f_3(t, \xi) & \text{in } [0, T] \times \mathcal{O}, \\ y(0, \xi) = x(\xi) & \text{in } \mathcal{O}, \\ y(t, \xi) = 0, \quad (t, \xi) \in [0, T] \times \partial\mathcal{O} \end{cases}. \quad (3.48)$$

By standard existence and uniqueness results, see, e.g., [Bar10], problem (3.48) has a unique solution

$$\begin{aligned} y &\in C([0, T^*]; H_0^1(\mathcal{O})) \cap L^2([0, T^*]; H^2(\mathcal{O})), \quad \mathbb{P} - a.s., \\ y_t &\in L^2([0, T]; L^2(\mathcal{O})), \quad \mathbb{P} - a.s., \end{aligned}$$

and by the Sobolev embedding theorem the following estimate holds

$$\begin{aligned} &\|y(t)\|_{H_0^1(\mathcal{O})}^2 + \int_0^t |\Delta_\xi y(s)|_2^2 ds + \int_0^t |y(s)|_6^6 ds \leq \\ &\leq C_1 \left(\int_0^t \int_{\mathcal{O}} (f_1^2 z^4 + f_2^2 z^2 + f_3^2) d\xi ds + \|x\|_{H_0^1(\mathcal{O})}^2 \right). \end{aligned} \quad (3.49)$$

By multiplying (3.48) by y , respectively Δy , and integrating on $(0, t) \times \mathcal{O}$ it also follows that

$$\begin{aligned} &\|y\|_{C([0, T^*]; H_0^1(\mathcal{O}))}^2 + \int_0^t (|y(s)|_6^6 + |y(s)|_{H^2(\mathcal{O})}^2) ds \leq \\ &\leq C_2 \int_0^t \int_{\mathcal{O}} (|z|^4 + |z|^2 + 1) d\xi ds \leq C_3 T^* (R^4 + R^2 + 1), \end{aligned} \quad (3.50)$$

because $\|y\|_{H^2(\mathcal{O})} \leq C \|\Delta y\|_2$ and by the Sobolev embedding theorems $H_0^1(\mathcal{O}) \subset L^6(\mathcal{O})$. This yields

$$\|y\|_{C([0, T^*], H_0^1(\mathcal{O}))} \leq C_3 \sqrt{T^* (R^4 + R^2 + 1)},$$

and so for T^* small enough we have that $y = Fz \in K$. Hence F maps K into itself. Moreover F is a contraction on K under the metric (3.47). Indeed we have by (3.48)

$$\begin{aligned} &\frac{1}{2} \frac{d}{dt} |y(t) - \bar{y}(t)|_2^2 + \|y(t) - \bar{y}(t)\|_{H_0^1(\mathcal{O})}^2 \leq C \int_{\mathcal{O}} (|z - \bar{z}| |y - \bar{y}| (|z| + |\bar{z}| + 1)) d\xi \leq \\ &\leq C (|z(t) - \bar{z}(t)|_2 |y(t) - \bar{y}(t)|_3 (|z(t)|_6 + |\bar{z}(t)|_6 + 1)), \quad a.e. \quad t \in [0, T], \end{aligned}$$

the last being implied by the Hölder inequality, namely, $|\int_{\mathcal{O}} uvz| \leq |u|_2 |v|_3 |z|_6$, therefore, we have

$$\begin{aligned} &|y(t) - \bar{y}(t)|_2^2 + \int_0^t \|y(s) - \bar{y}(s)\|_{H_0^1(\mathcal{O})}^2 ds \\ &\leq C(R+1) \left(\int_0^t |z(s) - \bar{z}(s)|_2^2 ds \right)^{\frac{1}{2}} \left(\int_0^t |y(s) - \bar{y}(s)|_3^2 ds \right)^{\frac{1}{2}} \leq \\ &\leq \frac{C^2}{4} (R+1)^2 \int_0^t |z(s) - \bar{z}(s)|_2^2 ds + \int_0^t |y(s) - \bar{y}(s)|_{H_0^1(\mathcal{O})}^2 ds, \end{aligned} \quad (3.51)$$

so that

$$\rho(y, \bar{y}) \leq \frac{C}{2} (R+1) T^* \rho(z, \bar{z}),$$

and taking $T^* < \frac{2}{C(R+1)}$, we have that F is a contraction on K . Then by the Banach fixed point theorem on $[0, T^*]$, there is exists a unique solution y to (3.48) providing that $T^* \in [0, T]$ is sufficiently small.

Let us now show by contradiction, that such a solution exists on a fixed interval $[0, T]$. Indeed if $[0, T^*]$ is the maximal interval on which y exists, by (3.48) we have, as mentioned above, that the following estimate holds

$$\begin{aligned} & \|y(t)\|_{H_0^1(\mathcal{O})}^2 + \int_0^t \left(\|y(s)\|_{H^2(\mathcal{O})}^2 + |y(s)|_6^6 \right) ds \leq \\ & \leq C \int_0^t \left(|y(s)|_4^4 + |y(s)|_2^2 + \|f_3\|_2^2 \right) ds, \quad \forall t \in [0, T^*]. \end{aligned} \quad (3.52)$$

Taking into account that

$$|y|_2^2 + |y|_4^4 \leq \epsilon |u|_6^6 + C_\epsilon, \quad \forall \epsilon > 0,$$

we get by (3.52) that

$$\|u(t)\|_{H_0^1(\mathcal{O})} + \int_0^t \|u(s)\|_{H^2(\mathcal{O})}^2 ds + \int_0^t |u(s)|_6^6 ds \leq C, \quad \forall t \in [0, T^*],$$

where C is independent of T^* . Therefore we also have that

$$\left| \frac{d}{dt} y(t) \right|_2 \leq C_1, \quad \forall t \in [0, T^*],$$

and the limit $\lim_{t \rightarrow T^*} y(t) = y(T^*)$ exists with $u(T^*) \in H_0^1(\mathcal{O})$. Then we can apply the above local existence result, extending y as a solution to (3.44) on $[T^*, T^* + \delta]$, which contradicts the assumption that $[0, T^*]$ is the maximal interval of existence, hence $T^* = T$.

Since the right hand side of (3.105) where $z = y$ is in $L^2(0, T; L^2(\mathcal{O}))$, we infer that

$$y \in C([0, T]; H_0^1(\mathcal{O})) \cap L^2([0, T]; H^2(\mathcal{O})), \quad \mathbb{P} - a.s., \quad (3.53)$$

moreover, since the contraction principle implies that the limit $y = \lim_{n \rightarrow \infty} y_n$ belongs to $C([0, T]; L^2(\mathcal{O}))$, where $y_n = F(y_{n-1})$ are $(\mathcal{F}_t)_{t \geq 0}$ -adapted, we can conclude that y is in fact an $(\mathcal{F}_t)_{t \geq 0}$ -adapted process and so y satisfies (3.45), as claimed. \square

Proof of Theorem 3.3.3(continued). We set $X := y + W_A$, where y is the solution to (3.44) given by Proposition 3.2.3, that is

$$\begin{cases} y_t - \Delta y + y^3 + ay^2 + by + 3W_A y^2 + 3W_A y^2 + 2W_A y = F - W_A^3 - aW_A^2 - bW_A & \text{in } [0, T] \times \mathcal{O}, \\ y = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \\ y(0, \xi) = x(\xi), \quad \xi \in \mathcal{O} \end{cases}, \quad (3.54)$$

By assumption (3.43) on W_A we see that

$$\begin{aligned} & \mathbb{E} \sup_{(t, \xi) \in [0, T] \times \mathcal{O}} (|f_1|^{2m} + |f_2|^{2m}) < \infty, \quad m = 1, 2, \\ & \mathbb{E} \|f_3\|_{L^2([0, T] \times \mathcal{O})} < \infty. \end{aligned} \quad (3.55)$$

Taking into account (3.54) and (3.44) we get that

$$y \in L^2(\Omega; C([0, T]; L^2(\mathcal{O})) \cap L_W^2([0, T]; H_0^1(\mathcal{O}) \cap H^2(\mathcal{O}))), \quad (3.56)$$

which implies (3.45) as claimed. \square

3.2.2 The optimal control of stochastic FitzHugh-Nagumo equation

Let U be a real Hilbert space with the norm $|\cdot|_U$ and $B \in L(U; L^2(\mathcal{O}))$. We shall denote by \mathcal{U} the space of all $(\mathcal{F}_t)_{t \geq 0}$ -adapted processes $u : [0, T] \rightarrow U$ s.t. $\mathbb{E} \int_0^T |u(t)|_U^2 dt < \infty$.

The space \mathcal{U} is a Hilbert space with the norm $|u|_{\mathcal{U}} = \left(\mathbb{E} \int_0^T |u(t)|_U^2 dt \right)^{\frac{1}{2}}$ and scalar product

$$\langle u, v \rangle_{\mathcal{U}} = \left(\mathbb{E} \int_0^T \langle u(t), v(t) \rangle_U dt \right)^{\frac{1}{2}}, \quad \forall u, v \in \mathcal{U},$$

where $\langle \cdot, \cdot \rangle_U$ is the scalar product of U .

Consider the functions $g, g_0 : \mathbb{R} \rightarrow \mathbb{R}$ and $h : U \rightarrow \bar{\mathbb{R}} :=]-\infty, \infty]$, which satisfy the following conditions

- (i) $g, g_0 \in C^1(L^2(\mathcal{O}))$ and $Dg, Dg_0 \in Lip(L^2(\mathcal{O}); L^2(\mathcal{O}))$ (where D stands for the Fréchet differential) and $Lip(L^2(\mathcal{O}); L^2(\mathcal{O}))$ is the space of Lipschitz continuous function from $L^2(\mathcal{O})$ to $L^2(\mathcal{O})$ with the norm defined denoted $\|\cdot\|_{Lip(L^2(\mathcal{O}))}$.
- (ii) h is convex, lower-semicontinuous and $(\partial h)^{-1} \in Lip(U)$ where $\partial h : U \rightarrow U$ is the subdifferential of h (see, e.g. [BP12, p. 82]). Moreover assume that $\exists \alpha_1 > 0$ and $\alpha_2 \in \mathbb{R}$ s.t. $h(u) \geq \alpha_1 |u|_U^2 + \alpha_2, \forall u \in U$. We set $L = \|(\partial h)^{-1}\|_{Lip(U)}$ (Here $Lip(U)$ is the space of Lipschitz operators on U).

We consider the following optimal control problem

$$\text{Minimize } \mathbb{E} \int_0^T (g(X(t)) + h(u(t))) dt + \mathbb{E} g_0(X(T)), \quad (\text{P})$$

subject to $u \in \mathcal{U}$ and

$$\begin{cases} dX(t) - \Delta_{\xi} X(t) dt + f(X(t)) dt = \sqrt{Q} dW(t) + Bu(t) dt + f_0 dt, & \text{in } [0, T] \times \mathcal{O}, \\ X = 0 & \text{on } [0, T] \times \partial \mathcal{O}, \\ X(0) = x & \text{in } \mathcal{O}, \end{cases}, \quad (3.57)$$

where $f_0 \in L^{\infty}([0, T] \times \mathcal{O})$.

In the following we shall assume both (3.55) and $Tr[QA] < \infty$, where A is as above the Laplace operator with domain $H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$.

Theorem 3.2.4. *Let $x \in H_0^1(\mathcal{O})$. Then there exists $C^* > 0$ independent of x such that for $LT + \|Dg_0\|_{Lip(L^2(\mathcal{O}))} < C^*$ there is a unique solution (u^*, X^*) to problem (P).*

Proof. The proof is based on Ekeland's variational principle already used in a similar deterministic context (See, e.g. [BI99]). Namely, we consider the function $\Psi : \mathcal{U} \rightarrow \bar{\mathbb{R}}$ defined by

$$\Psi(u) = \mathbb{E} \int_0^T (g(X^u(t)) + h(u(t))) dt + \mathbb{E} g_0(X^u(T)),$$

where X^u is the solution to (3.57). It is easily seen by (3.44) that Ψ is lower-semicontinuous and $\Psi(u) \rightarrow +\infty$ as $|u|_{\mathcal{U}} \rightarrow +\infty$.

If Ψ is weakly lower continuous on \mathcal{U} this is sufficient for the existence of a minimum of Ψ on \mathcal{U} . In the deterministic case, that is, if $Q = 0$ the weak lower continuity of Ψ is a direct consequence of compactness of the map $u \mapsto X^u$ from \mathcal{U} to $C([0, T]; L^2(\mathcal{O}))$ which is not the case here, that is, this map is not compact from \mathcal{U} to $L^2(\Omega; C([0, T]; L^2(\mathcal{O})))$. So the existence in problem (P) does not follow by standard minimization techniques. However, by the Ekeland variational principle, see, e.g., [Eke74], there is a sequence $\{u_\epsilon\} \subset \mathcal{U}$ such that

$$\begin{aligned}\Psi(u_\epsilon) &\leq \inf\{\Psi(u); u \in \mathcal{U}\} + \epsilon, \\ \Psi(u_\epsilon) &\leq \Psi(u) + \sqrt{\epsilon}|u_\epsilon - u|_{\mathcal{U}}, \quad \forall u \in \mathcal{U}.\end{aligned}\tag{3.58}$$

In other words,

$$u_\epsilon = \arg \min_{u \in \mathcal{U}} \{\Psi(u) + \sqrt{\epsilon}|u_\epsilon - u|_{\mathcal{U}}\}.$$

Hence $(X^{u_\epsilon}, u_\epsilon)$ is a solution to the optimal control problem

$$\begin{aligned}\min \left\{ \mathbb{E} \int_0^T (g(X^u(t)) + h(u(t))) dt + \mathbb{E} g_0(X^u(T)) + \right. \\ \left. + \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |u(t) - u_\epsilon(t)|_U^2 dt \right)^{\frac{1}{2}} ; u \in \mathcal{U} \right\}.\end{aligned}\tag{3.59}$$

The latter means that for all $v \in \mathcal{U}$ and $\lambda > 0$

$$\begin{aligned}\mathbb{E} \int_0^T (g(X^{u_\epsilon + \lambda v}(t)) + h((u_\epsilon + \lambda v)(t))) dt + \mathbb{E} g_0(X^{u_\epsilon + \lambda v}(T)) + \\ + \lambda \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |v(t)|_U^2 dt \right)^{\frac{1}{2}} \leq \\ \leq \mathbb{E} \int_0^T (g(X_\epsilon(t)) + h(u_\epsilon(t))) dt + \mathbb{E} g_0(X_\epsilon(T)).\end{aligned}$$

This yields

$$\begin{aligned}\mathbb{E} \int_0^T \langle Dg(X_\epsilon(t)), Z^v(t) \rangle_2 dt + \mathbb{E} \int_0^T h'(u_\epsilon(t), v(t)) dt + \\ + \mathbb{E} \langle Dg_0(X_\epsilon(T)), Z^v(T) \rangle_2 + \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |v(t)|_U^2 dt \right)^{\frac{1}{2}} \leq 0, \quad \forall v \in \mathcal{U},\end{aligned}\tag{3.60}$$

where Z^v solves the system in variations associated with (3.57), that is

$$\begin{cases} \frac{\partial}{\partial t} Z^v - \Delta Z^v + f'(X_\epsilon) Z^v = Bv & \text{in } [0, T] \times \mathcal{O}, \\ Z^v(0) = 0 & \text{in } \mathcal{O}, \\ Z^v = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \end{cases}\tag{3.61}$$

and $h' : U \times U \rightarrow \mathbb{R}$ is the directional derivatives of h , see, e.g., [BP12, p.81], namely

$$h'(u_\epsilon, v) = \lim_{\lambda \downarrow 0} \frac{h(u_\epsilon + \lambda v) - h(u_\epsilon)}{\lambda}, \quad \forall v \in U.$$

We associate with (3.57) the dual stochastic backward equation

$$\begin{cases} dp_\epsilon + \Delta p_\epsilon dt - f'(X_\epsilon)p_\epsilon dt = \kappa_\epsilon \sqrt{Q} dW(t) + Dg(X_\epsilon)dt & \text{in } [0, T] \times \mathcal{O}, \\ p_\epsilon(T) = -Dg_0(X_\epsilon(T)) & \text{in } \mathcal{O}, \\ p_\epsilon = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \end{cases} \quad (3.62)$$

It is well-known that equation (3.108) has a unique solution $(p_\epsilon, \kappa_\epsilon)$ satisfying

$$\begin{aligned} p_\epsilon &\in L_W^\infty([0, T]; L^2(\mathcal{O})) \cap L_W^2([0, T]; H_0^1(\mathcal{O})), \\ \kappa_\epsilon &\in L_W^2([0, T]; L^2(\mathcal{O})), \end{aligned}$$

(See, e.g., [FT02, Prop. 4.3] or [Tes96]). By Itô's formula we have

$$d\langle p_\epsilon, Z^v \rangle_2 = \langle dp_\epsilon, Z^v \rangle_2 + \langle p_\epsilon, dZ^v \rangle_2,$$

and this yields

$$\mathbb{E} \int_0^T \langle Dg(X_\epsilon(t)), Z^v(t) \rangle_2 dt + \mathbb{E} \langle Dg_0(X_\epsilon(T)), Z^v(T) \rangle_2 = 0,$$

and substituting in (5.25), we obtain, $\forall v \in \mathcal{U}$, the following inequality

$$\begin{aligned} \mathbb{E} \int_0^T h'(u_\epsilon(t), v(t)) dt + \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |v(t)|_U^2 dt \right)^{\frac{1}{2}} &\leq \\ &\leq \mathbb{E} \int_0^T \langle B^* p_\epsilon(t), v(t) \rangle_U dt. \end{aligned}$$

Let $G(u) := \mathbb{E} \int_0^T h(u(t)) dt$, then its subdifferential $\partial G : \mathcal{U} \rightarrow \mathcal{U}$, evaluated in u_ϵ is given by

$$\partial G(u_\epsilon) = \left\{ v^* \in \mathcal{U} : \langle v, v^* \rangle_{\mathcal{U}} \leq \mathbb{E} \int_0^T h'(u_\epsilon(t), v(t)) dt, \forall v \in \mathcal{U} \right\}.$$

(See, e.g., [BP12, p.81]). Then we infer that

$$u_\epsilon(t) = (\partial h)^{-1} \left(B^* p_\epsilon(t) + \sqrt{\epsilon} \tilde{\theta}_\epsilon \right), \quad t \in [0, T], \quad \mathbb{P} - a.s.,$$

where $\tilde{\theta}_\epsilon \in \mathcal{U}$ and $|\tilde{\theta}_\epsilon|_{\mathcal{U}} \leq 1, \forall \epsilon > 0$.

Therefore, we have shown that

$$\begin{aligned} u_\epsilon &= (\partial h)^{-1} (B^* p_\epsilon + \theta_\epsilon), \quad \|\theta_\epsilon\|_{L^2([0, T] \times \Omega; U)} \leq \sqrt{\epsilon}, \\ dp_\epsilon + \Delta p_\epsilon dt - f'(X_\epsilon)p_\epsilon dt &= Dg_\epsilon(X_\epsilon)dt + \kappa_\epsilon \sqrt{Q} dW(t) & \text{in } [0, T] \times \mathcal{O}, \\ p_\epsilon(T) &= -Dg_0(X_\epsilon(T)) & \text{in } \mathcal{O}, \\ p_\epsilon &= 0 & \text{in } [0, T] \times \partial\mathcal{O}, \end{aligned} \quad (3.63)$$

By (3.104) and by assumptions (ii) it follows also that $\{u_\epsilon\}_{\epsilon > 0}$ is bounded in \mathcal{U} . Moreover, by (3.57) we have that

$$\begin{cases} dX_\epsilon(t) - \Delta X_\epsilon(t)dt + f(X_\epsilon(t))dt = \sqrt{Q}dW(t) + f_0dt + Bu_\epsilon(t)dt, & \text{in } [0, T] \times \mathcal{O}, \\ X_\epsilon = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \\ X_\epsilon(0) = x & \text{in } \mathcal{O} \end{cases}, \quad (3.64)$$

which by (3.105), assumption (ii) and exploiting the Itô formula, implies that

$$\mathbb{E} \int_0^T \left(|X_\epsilon(t)|_2^2 + |\nabla X_\epsilon(t)|_2^2 + L^{-1} |u_\epsilon(t)|_U^2 \right) dt \leq C, \quad \forall \epsilon > 0. \quad (3.65)$$

Moreover by (3.64) and (3.104), again using the Itô formula applied to $|X|_2^2$, we have that $\forall \epsilon > 0$

$$\begin{aligned} & \mathbb{E} \sup_{t \in [0, T]} |X_\epsilon(t)|_2^2 + \mathbb{E} \int_0^T |X_\epsilon(t)|_{H_0^1(\mathcal{O})}^2 dt + \mathbb{E} \int_0^T |X_\epsilon(t)|_2^4 dt \leq \\ & \leq C(1 + |x|_2^2) \end{aligned} \quad (3.66)$$

If we now apply the Itô formula in (3.64) to the function $X \rightarrow \frac{1}{2}|X|_{H_0^1(\mathcal{O})}^2$, taking into account that $\text{Tr}[QA] < \infty$ and that

$$- \int_{\mathcal{O}} f(X_\epsilon) \Delta X_\epsilon d\xi \geq ab \int_{\mathcal{O}} |\nabla X_\epsilon|^2 d\xi - \int_{\mathcal{O}} |\Delta X_\epsilon| |X_\epsilon|^2 d\xi,$$

we obtain by (3.111) that

$$\mathbb{E} \sup_{t \in [0, T]} |X_\epsilon(t)|_{H_0^1(\mathcal{O})}^2 + \mathbb{E} \int_0^T |\Delta X_\epsilon(t)|_2^2 dt \leq C(1 + |x|_{H_0^1(\mathcal{O})}^2). \quad (3.67)$$

Similarly, by (3.109) we obtain that

$$\begin{aligned} & \frac{1}{2} d|p_\epsilon(t)|_2^2 - \int_{\mathcal{O}} |\nabla p_\epsilon(t)|^2 d\xi - \int_{\mathcal{O}} f'(X_\epsilon) p_\epsilon^2(t) d\xi = \\ & = \int_{\mathcal{O}} Dg(X_\epsilon(t)) p_\epsilon(t) d\xi + \frac{1}{2} \int_{\mathcal{O}} |\kappa_\epsilon|^2 d\xi + \int_{\mathcal{O}} p_\epsilon \kappa_\epsilon \sqrt{Q} dW(t). \end{aligned}$$

which yields

$$\begin{aligned} & \mathbb{E} \sup_{t \in [0, T]} |p_\epsilon(t)|_2^2 + \mathbb{E} \int_0^T |p_\epsilon(t)|_{H_0^1(\mathcal{O})}^2 dt + \mathbb{E} \int_0^T \int_{\mathcal{O}} |X_\epsilon|^2 |p_\epsilon|^2 d\xi dt \\ & + \mathbb{E} \int_0^T |\kappa_\epsilon(t)|_2^2 dt \leq C + \mathbb{E} |X_\epsilon(T)|_2^2 \leq C, \quad \forall \epsilon > 0. \end{aligned} \quad (3.68)$$

(Here and everywhere in the following we shall denote by C several positive constants independent of ϵ). In particular, it follows by (3.109) and (3.112) that $\{u_\epsilon\}_{\epsilon > 0}$ is bounded in $L^2(\Omega; L^\infty([0, T]; U))$.

Equation (3.64) implies that

$$\begin{aligned} & \frac{\partial}{\partial t} (X_\epsilon(t) - X_\lambda(t)) - \Delta (X_\epsilon(t) - X_\lambda(t)) + (f(X_\epsilon(t)) - f(X_\lambda(t))) = \\ & = BB^*(p_\epsilon(t) - p_\lambda(t)) + B(\theta_\epsilon(t) - \theta_\lambda(t)). \end{aligned} \quad (3.69)$$

In virtue of (3.112) this yields

$$\begin{aligned}
 & \frac{1}{2} |X_\epsilon(t) - X_\lambda(t)|_2^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds \leq \\
 & \leq - \int_0^t \int_{\mathcal{O}} (f(X_\epsilon(s)) - f(X_\lambda(s))) (X_\epsilon(s) - X_\lambda(s)) d\xi ds \\
 & \quad + L \int_0^t |p_\epsilon(s) - p_\lambda(s)|_2 |X_\epsilon(s) - X_\lambda(s)|_2 ds \\
 & \quad + C \int_0^t |\theta_\epsilon(s) - \theta_\lambda(s)|_U |X_\epsilon(s) - X_\lambda(s)|_2 ds, \quad \forall t \in [0, T],
 \end{aligned}$$

where $L = \|(\partial h)^{-1}\|_{Lip}$.

We further have

$$(f(X_\epsilon) - f(X_\lambda)) (X_\epsilon - X_\lambda) = f'(\alpha X_\epsilon + (1 - \alpha)X_\lambda)(X_\epsilon - X_\lambda)^2,$$

where $\alpha \in [0, 1]$ and assuming that $0 < a < b$,

$$\begin{aligned}
 & f'(u) \geq 0 \text{ for } u \notin [0, b], \\
 & |f'(u)| \leq C \text{ for } u \in [0, b],
 \end{aligned}$$

then

$$- \int_0^t \int_{\mathcal{O}} (f(X_\epsilon) - f(X_\lambda)) (X_\epsilon - X_\lambda) d\xi ds \leq C \int_0^t |X_\epsilon(s) - X_\lambda(s)|_2^2 ds, \quad \forall \epsilon, \lambda > 0,$$

which yields, for $t \in [0, T]$

$$\begin{aligned}
 & |X_\epsilon(t) - X_\lambda(t)|_2^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds \leq \\
 & \leq C \left(L \int_0^t |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \int_0^t \int_{\mathcal{O}} (X_\epsilon(s) - X_\lambda(s))^2 d\xi ds + \epsilon + \lambda \right).
 \end{aligned} \tag{3.70}$$

Applying Gronwall's lemma in (3.114), we have

$$\begin{aligned}
 & |X_\epsilon(t) - X_\lambda(t)|_2^2 + \frac{1}{2} \int_0^t |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds \leq \\
 & \leq C \left(L \int_0^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \epsilon + \lambda \right), \quad \forall \epsilon, \lambda > 0, t \in [0, T].
 \end{aligned} \tag{3.71}$$

Similarly we get by (3.52) and the Itô formula

$$\begin{aligned}
& |p_\epsilon(t) - p_\lambda(t)|_2^2 + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds + \frac{1}{2} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 ds = \\
& = |Dg_0(X_\epsilon(T)) - Dg_0(X_\lambda(T))|_2^2 + \\
& \quad + \int_t^T \int_{\mathcal{O}} (f'(X_\epsilon(s))p_\epsilon(s) - f'(X_\lambda(s))p_\lambda(s)) (p_\epsilon(s) - p_\lambda(s)) d\xi ds + \\
& \quad - \int_t^T \left\langle \kappa_\epsilon(s) - \kappa_\lambda(s) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \right\rangle_2 = \\
& = \int_t^T \int_{\mathcal{O}} f'(X_\epsilon(s))(p_\epsilon(s) - p_\lambda(s))^2 d\xi ds + \\
& \quad - \int_t^T \int_{\mathcal{O}} (f'(X_\epsilon(s)) - f'(X_\lambda(s))) (p_\epsilon(s) - p_\lambda(s)) d\xi ds + \\
& \quad - \int_t^T \left\langle \kappa_\epsilon(s) - \kappa_\lambda(s) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \right\rangle_2 + \\
& \quad + |Dg_0(X_\epsilon(T)) - Dg_0(X_\lambda(T))|_2^2 \leq \\
& \leq C \left(\int_t^T \int_{\mathcal{O}} (|X_\epsilon| + 1)(p_\epsilon(s) - p_\lambda(s))^2 d\xi ds \right) + \\
& \quad + \left(\int_t^T \int_{\mathcal{O}} (X_\epsilon(s) - X_\lambda(s)) (p_\epsilon(s) - p_\lambda(s)) (1 + |X_\epsilon| + |X_\lambda|) |p_\epsilon| d\xi ds \right) + \\
& \quad - \int_t^T \left\langle \kappa_\epsilon(s) - \kappa_\lambda(s) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \right\rangle_2 + \\
& \quad + \|Dg_0\|_{Lip(L^2(\mathcal{O}))} |X_\epsilon(T) - X_\lambda(T)|_2^2, \quad t \in [0, T], \mathbb{P} - a.s..
\end{aligned} \tag{3.72}$$

Proceeding as above, we also have

$$\begin{aligned}
& \int_{\mathcal{O}} |X_\epsilon(s)| |p_\epsilon(s) - p_\lambda(s)|^2 d\xi \leq |p_\epsilon(s) - p_\lambda(s)|_4 |p_\epsilon(s) - p_\lambda(s)|_2 |X_\epsilon(s)|_4 \leq \\
& \leq \frac{1}{2} |p_\epsilon(s) - p_\lambda(s)|_{H_0^1(\mathcal{O})}^2 + \frac{1}{2} |p_\epsilon(s) - p_\lambda(s)|_2^2 |X_\epsilon(s)|_4^2.
\end{aligned} \tag{3.73}$$

Moreover, exploiting both the Hölder and the interpolation inequality, we obtain

$$\begin{aligned}
& \int_{\mathcal{O}} |X_\epsilon - X_\lambda| |p_\epsilon - p_\lambda| (1 + |X_\epsilon| + |X_\lambda|) |p_\epsilon| d\xi \leq \\
& \leq |X_\epsilon - X_\lambda|_4 |p_\epsilon - p_\lambda|_4 \left(\int_{\mathcal{O}} (1 + |X_\epsilon| + |X_\lambda|)^2 |p_\epsilon|^2 d\xi \right)^{\frac{1}{2}} \leq \\
& \leq |X_\epsilon - X_\lambda|_2^{\frac{1}{2}} |X_\epsilon - X_\lambda|_6^{\frac{1}{2}} |p_\epsilon - p_\lambda|_2^{\frac{1}{2}} |p_\epsilon - p_\lambda|_6^{\frac{1}{2}} \times \\
& \quad \times \left(\int_{\mathcal{O}} (1 + |X_\epsilon| + |X_\lambda|)^2 |p_\epsilon|^2 d\xi \right)^{\frac{1}{2}} \leq \\
& \leq |X_\epsilon - X_\lambda|_2^{\frac{1}{2}} |X_\epsilon - X_\lambda|_{H_0^1(\mathcal{O})}^{\frac{1}{2}} |p_\epsilon - p_\lambda|_2^{\frac{1}{2}} |p_\epsilon - p_\lambda|_{H_0^1(\mathcal{O})}^{\frac{1}{2}} \times \\
& \quad \times \left(\int_{\mathcal{O}} (1 + |X_\epsilon| + |X_\lambda|)^2 |p_\epsilon|^2 d\xi \right)^{\frac{1}{2}} \leq \alpha \left(|X_\epsilon - X_\lambda|_{H_0^1(\mathcal{O})}^2 + |p_\epsilon - p_\lambda|_{H_0^1(\mathcal{O})}^2 \right) + \\
& \quad + \frac{C}{\alpha} \left(|X_\epsilon - X_\lambda|_2^2 + |p_\epsilon - p_\lambda|_2^2 \right) \left(\int_{\mathcal{O}} (1 + |X_\epsilon| + |X_\lambda|)^2 |p_\epsilon|^2 d\xi \right), \tag{3.74}
\end{aligned}$$

where α is arbitrary small. Substituting now (3.73), (3.117) into (3.114), (3.116), we obtain \mathbb{P} -a.s.

$$\begin{aligned}
& |X_\epsilon(t) - X_\lambda(t)|_2^2 + |p_\epsilon(t) - p_\lambda(t)|_2^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds + \\
& \quad + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds + \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 ds \leq \\
& \leq C \left(L \int_0^t |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \epsilon + \lambda \right) + C \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 |X_\epsilon(s)|_4^2 ds + \\
& \quad + \|Dg_0\|_{Lip} |X_\epsilon(T) - X_\lambda(T)|_2^2 + \\
& \quad + \frac{C}{\alpha} \left(\int_t^T |X_\epsilon(s) - X_\lambda(s)|_2^2 + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 T_{\epsilon,\lambda}(s) ds \right) + \\
& \quad - \int_t^T \left\langle \kappa_\epsilon(s) - \kappa_\lambda(s) \right\rangle \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \Big\rangle_2, \quad \forall t \in [0, T], \tag{3.75}
\end{aligned}$$

where

$$T_{\epsilon,\lambda} := \int_{\mathcal{O}} (1 + |X_\epsilon| + |X_\lambda|)^2 |p_\epsilon|^2 d\xi.$$

We note that the process $r \mapsto \int_t^r \langle (\kappa_\epsilon - \kappa_\lambda) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \rangle_2$ is a local martingale on $[t, T]$, hence by the Burkholder-Davis-Gundy inequality, see, e.g., [DPZ96, p.58], we have for all $r \in [t, T]$

$$\begin{aligned}
& \mathbb{E} \sup_{r \in [t, T]} \left| \int_t^r \left\langle (\kappa_\epsilon(s) - \kappa_\lambda(s)) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \right\rangle_2 \right| \leq \\
& \leq C \left(\mathbb{E} \int_0^r |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 |X_\epsilon(s) - X_\lambda(s)|_2^2 ds \right)^{\frac{1}{2}} \leq \\
& \leq C \mathbb{E} \sup_{s \in [t, r]} |X_\epsilon(s) - X_\lambda(s)|_2^2 + \frac{1}{2} \mathbb{E} \int_t^r |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 ds,
\end{aligned}$$

and by (3.117) we get

$$\begin{aligned}
& \mathbb{E} \sup_{s \in [t, T]} (|X_\epsilon(s) - X_\lambda(s)|_2^2 + |p_\epsilon(s) - p_\lambda(s)|_2^2) \\
& + \mathbb{E} \int_0^T |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds \\
& + \mathbb{E} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 ds \leq \\
& \leq \|Dg_0\| \mathbb{E} |X_\epsilon(T) - X_\lambda(T)|_2^2 + C \left(L \mathbb{E} \int_0^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \epsilon + \lambda \right) \\
& + C \mathbb{E} \sup_{s \in [t, T]} |X_\epsilon(s) - X_\lambda(s)|_2^2 \\
& + C \mathbb{E} \int_t^T (|p_\epsilon(s) - p_\lambda(s)|_2^2 + |X_\epsilon(s) - X_\lambda(s)|_2^2) (|X_\epsilon(s)|_4^2 + T_{\epsilon, \lambda}(s)) ds.
\end{aligned} \tag{3.76}$$

Taking into account estimates (3.111), (3.67) and (3.115), from (3.120) we have

$$\begin{aligned}
& \mathbb{E} \sup_{s \in [t, T]} (|X_\epsilon(s) - X_\lambda(s)|_2^2 + |p_\epsilon(s) - p_\lambda(s)|_2^2) \\
& + \mathbb{E} \int_0^T |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds \\
& + \mathbb{E} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 ds \leq \\
& \leq \tilde{C} \left(L \mathbb{E} \int_0^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds \right) \\
& + \tilde{C} \left(\mathbb{E} \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 (|X_\epsilon(s)|_4^2 + T_{\epsilon, \lambda}(s)) ds \right) \\
& + \tilde{C} \|Dg_0\|_{Lip} \mathbb{E} |X_\epsilon(T) - X_\lambda(T)|_2^2 + \tilde{C}(\epsilon + \lambda).
\end{aligned} \tag{3.77}$$

where \tilde{C} is a positive constant independent of ϵ and λ . It follows that if $\tilde{C}(LT + \|Dg_0\|_{Lip}) < 1$, then, for any $t \in [0, T]$,

$$\begin{aligned}
& \mathbb{E} \sup_{s \in [t, T]} (|X_\epsilon(s) - X_\lambda(s)|_2^2 + |p_\epsilon(s) - p_\lambda(s)|_2^2) \\
& + \mathbb{E} \int_0^T |X_\epsilon(s) - X_\lambda(s)|_{H_0^1(\mathcal{O})}^2 ds + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds \\
& + \mathbb{E} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_2^2 ds \leq \\
& \leq C \mathbb{E} \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 (|X_\epsilon(s)|_4^2 + T_{\epsilon, \lambda}(s)) ds + C(\epsilon + \lambda).
\end{aligned} \tag{3.78}$$

Let us define for $j \in \mathbb{N}$

$$\Omega_j := \left\{ \omega \in \Omega : \sup_{\epsilon} \sup_{t \in [0, T]} (|X_\epsilon(t)|_2^2 + |X_\epsilon(t)|_{H_0^1(\mathcal{O})}^2 + |X_\epsilon(t)|_4^2 + |p_\epsilon(t)|_2^2) dt \leq j \right\},$$

then estimates (3.111) and (3.67) implies that

$$\mathbb{P}(\Omega_j) \geq 1 - \frac{C}{j}, \quad \forall j \in \mathbb{N},$$

for some constant C independent of ϵ .

If we set $X_\epsilon^j := \mathbb{1}_{[\Omega_j]} X_\epsilon$, $p_\epsilon^j := \mathbb{1}_{[\Omega_j]} p_\epsilon$ and $\kappa_\epsilon^j := \mathbb{1}_{[\Omega_j]} \kappa_\epsilon$, then such quantities satisfy the system (3.109)-(3.64), with $\mathbb{1}_{[\Omega_j]} \sqrt{Q} dW$. The latter means that estimate (3.122) still holds in this context, so that we have

$$\begin{aligned} & \mathbb{E} \sup_{s \in [t, T]} |X_\epsilon^j(s) - X_\lambda^j(s)|_2^2 + \sup_{s \in [t, T]} |p_\epsilon^j(t) - p_\lambda^j(t)|_2^2 \\ & + \mathbb{E} \int_t^T |p_\epsilon^j(s) - p_\lambda^j(s)|_{H_0^1(\mathcal{O})}^2 ds + \mathbb{E} \int_t^T |(\kappa_\epsilon(s) - \kappa_\lambda(s)) \chi_j|_2^2 ds \leq \\ & \leq C_j \int_t^T \mathbb{E} |p_\epsilon^j(s) - p_\lambda^j(s)|_2^2 ds + C(\epsilon + \lambda), \quad j \in \mathbb{N}. \end{aligned} \quad (3.79)$$

By Gronwall's lemma we get, for any $t \in [0, T]$

$$\mathbb{E} \sup_{s \in [t, T]} |X_\epsilon^j(s) - X_\lambda^j(s)|_2^2 + \sup_{s \in [t, T]} |p_\epsilon^j(s) - p_\lambda^j(s)|_2^2 \leq C(\epsilon + \lambda) e^{C_j T}, \quad (3.80)$$

where $C_j = C(j^3 + 1)$, hence, for $\epsilon \rightarrow 0$ and all $j \in \mathbb{N}$, we obtain

$$\begin{aligned} X_\epsilon^j &\rightarrow X^j && \text{in } L^2(\Omega_j; L^2([0, T] \times \mathcal{O})), \\ p_\epsilon^j &\rightarrow p^j && \text{in } L^2(\Omega_j; L^2([0, T] \times \mathcal{O})), \end{aligned} \quad (3.81)$$

where \rightarrow means strong convergence. By estimates (3.111) and (3.112) it follows that taking related subsequences, still denoted by ϵ , we have

$$\begin{aligned} X_\epsilon &\rightharpoonup X^* && \text{in } L^2([0, T] \times \Omega; H_0^1(\mathcal{O})), \\ p_\epsilon &\rightharpoonup p^* && \text{in } L^\infty([0, T]; L^2(\Omega \times \mathcal{O})), \\ p_\epsilon &\rightharpoonup p^* && \text{in } L^2([0, T] \times \Omega \times \mathcal{O}), \\ p_\epsilon &\rightharpoonup p^* && \text{in } L^2([0, T] \times \Omega; H_0^1(\mathcal{O})), \\ u_\epsilon &\rightharpoonup u^* && \text{in } L^\infty([0, T]; L^2(\Omega; U)), \end{aligned} \quad (3.82)$$

where \rightharpoonup means weak (respectively, weak-star) convergence, so we have for $\epsilon \rightarrow 0$

$$X_\epsilon \rightarrow X^*, \quad p_\epsilon \rightarrow p^*, \quad a.e. \text{ in } [0, T] \times \Omega_j \times \mathcal{O}. \quad (3.83)$$

By (3.111) we see that

$$\mathbb{E} \int_0^T \int_{\mathcal{O}} |f(X_\epsilon(s, \xi))|^{\frac{4}{3}} d\xi ds \leq C, \quad \forall \epsilon > 0.$$

Since $\{f(X_\epsilon)\}$ is bounded in $L^{\frac{4}{3}}([0, T] \times \Omega \times \mathcal{O})$, then it is weakly compact in $L^1([0, T] \times \Omega \times \mathcal{O})$ and by (3.127) we have that for a subsequence $\{\epsilon\} \rightarrow 0$,

$$f(X_\epsilon) \rightarrow f(X^*), \quad a.e. \text{ in } [0, T] \times \Omega \times \mathcal{O},$$

which, in virtue of (3.127) and since

$$\mathbb{P}(\Omega_j) \geq 1 - \frac{C}{j}, \forall j \in \mathbb{N}_0,$$

we have

$$f(X_\epsilon) \rightarrow f(X^*) \quad \text{in } L^1([0, T] \times \Omega_j \times \mathcal{O}). \quad (3.84)$$

Then, letting $\epsilon \rightarrow 0$ in (3.64), we obtain

$$\begin{cases} dX^*(t) - \Delta X^*(t)dt + f(X^*(t))dt = \sqrt{Q}dW(t) + Bu^*(t)dt & \text{in } [0, T] \times \mathcal{O}, \\ X^* = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \\ X^*(0) = x & \text{in } \mathcal{O} \end{cases}.$$

Taking into account that Ψ is weakly lower semicontinuous in \mathcal{U} we infer by (3.104) that

$$\Psi(u^*) = \inf \{ \Psi(u); u \in \mathcal{U} \},$$

therefore (X^*, u^*) is optimal for the problem (P) and the proof of existence is therefore complete. \square

Concerning the uniqueness for the optimal pair (X^*, u^*) given by Th. 3.3.4, we have that it follows by the same argument via the maximum principle result for problem (P), namely one has the following result.

Theorem 3.2.5. *Let (X^*, u^*) be optimal in problem (P), then*

$$u^* = (\partial h)^{-1}(B^*p), \text{ a.e. } t \in [0, T], \quad (3.85)$$

where p is the solution to the backward stochastic equation

$$\begin{cases} dp + \Delta p dt + f'(X^*)p dt = g'(X)dt + \kappa\sqrt{Q}dW(t) & \text{in } [0, T] \times \mathcal{O}, \\ p(T) = -Dg_0(X^*(T)) & \text{in } \mathcal{O}, \\ p_\epsilon = 0 & \text{on } [0, T] \times \partial\mathcal{O} \end{cases}, \quad (3.86)$$

Proof. If (X^*, u^*) is optimal for the problem (P), then by the same argument used to prove Th. 3.3.4, see (5.25), we have

$$\begin{aligned} & \mathbb{E} \int_0^T \langle Dg(X^*(t)), Z^v(t) \rangle_2 dt + \mathbb{E} \int_0^T h'(u^*(t), v(t)) dt \\ & + \mathbb{E} \langle Dg_0(X^*(T)), Z^v(T) \rangle_2 \leq 0, \quad \forall v \in \mathcal{U}, \end{aligned} \quad (3.87)$$

where Z^v is solution to equation (3.107) with X_ϵ replaced by X^* . This implies as above that (3.129) holds. \square

The uniqueness in (P). If (X^*, u^*) is optimal in (P) then it satisfies systems (3.102), (3.129) and (3.130), so that arguing as in the proof of Th. 3.3.4, the same set of estimates implies that the previous system has at most one solution if $LT + \|Dg_0\|_{Lip} < C^*$, where C^* is sufficiently small. \square

Remark 3.2.6. Theorems 3.3.4 and 3.3.5 remain true if assumption (i) is relaxed to

(i)' $Dg_0 \in Lip(L^2(\mathcal{O}))$, $g = g(t, y) : [0, T] \times L^2(\mathcal{O}) \rightarrow \mathbb{R}$ is of class C^1 in y , $D_y g \in C([0, T] \times L^2(\mathcal{O}))$, and $\sup_{t \in [0, T]} \|D_y g(t, y)\|_{Lip(L^2(\mathcal{O}))} < \infty$.

Remark 3.2.7. As clear from the previous proof the constant C^* arising in conditions of Theorem 3.3.4 depends of f and g only and, as mentioned earlier, it is independent of initial data x .

3.2.3 An example

Roughly speaking the control objective in system (3.102) is to drive the potential X to track a given trajectory X^1 and an end potential X^0 . This can be reformulated as the optimal control problem

$$\text{Minimize } \mathbb{E} \int_0^T \alpha |u(t)|_2^2 + |X(t) - X^1(t)|_2^2 dt + \lambda \mathbb{E} |X(T) - X^0|_2^2, \quad (3.88)$$

subject to

$$u \in L^2_W([0, T]; L^2(\mathcal{O})), \quad m \leq u \leq M \text{ a.e. on } [0, T] \times \mathcal{O}. \quad (3.89)$$

$$\begin{cases} dX(t) - \Delta X(t)dt + f(X(t)) = \sqrt{Q}dW(t) + u(t)dt + f_0dt \text{ in } [0, T] \times \mathcal{O}, \\ X(0) = x \text{ in } \mathcal{O}, \\ X = 0 \text{ on } [0, T] \times \partial\mathcal{O} \end{cases}, \quad (3.90)$$

where $f(u) = u(u - a)(u - b)$, $\alpha, \lambda > 0$, $0 < m < M < \infty$ and the functions $X^1 \in L^2([0, T]; L^2(\mathcal{O}))$, $X^0 \in L^2(\mathcal{O})$ are given.

As mentioned above, the physical significance of the problem is the following: find an optimal current u applied to a nerve axon in such a way that the resulting potential X flows closely to a specified regime $X^0 = X^0(t, \xi)$ during the time interval $[0, T]$, and such that it is near to a given potential X_0 at the final time T (on these lines see also [PA10]).

Problem (3.88)-(3.90) is of the form (P) where

$$g(t, X) = |X - X^1(t)|_2^2, \quad g_0(X) = \lambda |X - X^0|_2^2, \quad (3.91)$$

and $h : L^2(\mathcal{O}) \rightarrow]-\infty, +\infty]$ is defined by

$$h(u) = \begin{cases} \alpha |u|_2^2 & \text{if } u \in U_0, \\ \infty & \text{if } u \notin U_0 \end{cases}, \quad (3.92)$$

$$U_0 := \{u \in L^2(\mathcal{O}) : m \leq u \leq M \text{ a.e. in } \mathcal{O}\}.$$

We have $\partial h(u) = 2\alpha u + N_{U_0}(u)$, where N_{U_0} (the normal cone to U_0) is given by

$$N_{U_0}(u) = \left\{ v \in L^2(\mathcal{O}) : \begin{array}{l} v(\xi) = 0 \text{ if } u(\xi) \in (m, M); \\ v(\xi) \geq 0 \text{ if } u(\xi) = M; \\ v(\xi) \leq 0 \text{ if } u(\xi) = m \end{array} \right\}.$$

Then $(\partial h)^{-1}(v) = 2\alpha P_{U_0}(v)$ where $P_{U_0} : L^2(\mathcal{O}) \rightarrow U_0$ is the projection operator

$$P_{U_0}(v)(\xi) := \begin{cases} M & \text{if } v(\xi) \geq M, \\ m & \text{if } v(\xi) \leq m, \\ v(\xi) & \text{if } m < v(\xi) < M, \end{cases} \quad \xi \in \mathcal{O}.$$

By Theorem 3.3.4, there exists a constant $C^* > 0$ such that if $\alpha T + \lambda < C^*$ the problem (3.88)-(3.90) has a unique solution (X^*, u^*) given by

$$\begin{cases} u^* = 2\alpha P_{U_0}(p^*) & \text{in } [0, T] \times \mathcal{O}, \quad \mathbb{P} - a.s. \\ dp^*(t) + \Delta p^*(t)dt - f'(X^*(t))p^*(t)dt = 2(X^*(t) - X^1)dt + \kappa\sqrt{Q}dW(t) & \text{in } [0, T] \times \mathcal{O}, \\ p^*(T) = -2\lambda(X^*(T) - X^0) & \text{in } \mathcal{O}, \\ p^* = 0 & \text{on } [0, T] \times \partial\mathcal{O} \end{cases} \quad (3.93)$$

3.2.4 The linear multiplicative noise perturbation

We briefly discuss here the case where the Gaussian perturbation is proportional with the nerve membrane potential. The neuron impulse dynamic is better described by the equation

$$\begin{cases} dX(t) - \Delta X(t)dt + f(X(t))dt = X(t)dW(t)dt + F(t)dt & \text{in } [0, T] \times \mathcal{O}, \\ X(0) = x & \text{in } \mathcal{O}, \\ X = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \end{cases} \quad (3.94)$$

where

$$W(t, \xi) = \sum_{j=1}^{\infty} \mu_j e_j(\xi) \beta_j(t), \quad t \geq 0, \quad \xi \in \mathcal{O},$$

$\mu_j \in \mathbb{R}$ and $\{e_j\}_{j=1}^{\infty}$ is an orthonormal basis in $L^2(\mathcal{O})$ of eigenfunctions for A corresponding to eigenvalues λ_j .

By the scaling transformation $X = e^W y$ equation (3.94) reduces to the random differential equation (see, e.g. [BR11])

$$\begin{cases} \frac{\partial y}{\partial t} - \Delta y + (\mu - \Delta W)y - 2\nabla W \cdot \nabla y = e^{-W} F & \text{in } [0, T] \times \mathcal{O}, \\ y(0, \xi) = x(\xi) & \xi \in \mathcal{O}, \\ y = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \end{cases} \quad (3.95)$$

where $\mu = \frac{1}{2} \sum_{j=1}^{\infty} \mu_j^2 e_j$.

We shall assume that

$$\sum_{j=1}^{\infty} \mu_j^2 |e_j|_{\infty}^2 < \infty. \quad (3.96)$$

Arguing as in Proposition 3.2.3 it follows by (3.96) that (3.95) has a unique solution y satisfying (3.56) and this implies that $X = e^W y$ is a strong solution to (3.94) which satisfies condition of Theorem 3.3.3. We omit the details.

As regards the corresponding optimal control problem \mathbf{P} governed by the equation

$$\begin{cases} dX(t) - \Delta_{\xi} X(t)dt + f(X(t))dt = X(t)dW(t) + Bu(t)dt + f_0 dt, & \text{in } [0, T] \times \mathcal{O}, \\ X(0) = x & \text{in } \mathcal{O}, X = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \end{cases} \quad (3.97)$$

the existence of an optimal control pair (X^*, u^*) follows as Theorem 3.3.4 by Eckeland variational principle (3.104) taking however in account that the corresponding dual backward equation (3.108) is in this case

$$\begin{cases} dp_{\epsilon} + \Delta p_{\epsilon} dt - f'(X_{\epsilon})p_{\epsilon} dt + \kappa_{\epsilon} dt = \kappa_{\epsilon} dW(t) + Dg(X_{\epsilon})dt & \text{in } [0, T] \times \mathcal{O}, \\ p_{\epsilon}(T) = -Dg_0(X_{\epsilon}(T)) & \text{in } \mathcal{O}, \\ p_{\epsilon} = 0 & \text{on } [0, T] \times \partial\mathcal{O}, \end{cases} \quad (3.98)$$

The details are left to the reader.

3.3 Optimal control for the stochastic FitzHugh-Nagumo equation with recovery variable

The mathematical formulation of the signal propagation in a neural cell has been firstly introduced in [HH52], in particular the authors proposed a system composed of four different equations. Clearly, in spite of being very realistic, this model has the disadvantage of being analytically not easy to deal with. This problem led to the works [Fit61, NAY62] to reduce the number of equations to two, where the first variable represents the voltage variable whilst the second one is the recovery variable. We will also generalize the above equations taking into account a stochastic perturbation.

In particular we will consider the equation

$$\begin{cases} \partial_t v(t, \xi) &= \Delta_\xi - I_{ion}(v(t, \xi)) - w(t, \xi) + f(\xi) + \partial_t \beta_1(t), \text{ in } [0, T] \times \mathcal{O}, \\ \partial_t w(t, \xi) &= \gamma v(t, \xi) - \delta w(t, \xi) + \partial_t \beta_2(t), \text{ in } [0, T] \times \mathcal{O}, \\ \partial_\nu v(t, \xi) &= 0, \quad \text{on } [0, T] \times \partial\mathcal{O}, \\ v(0, \xi) &= v_0(\xi), \quad w(0, \xi) = w_0(\xi), \text{ in } [0, T] \times \mathcal{O}. \end{cases}, \quad (3.99)$$

where as mentioned above the variable v represents the voltage variable and w is the recovery variable. All other notation is to better specified in a while. In particular the function I_{ion} is a polynomial of degree 3. This fact implies that standard existence and uniqueness results do not hold for equation (3.99), since the non-linear term I_{ion} fails to be Lipschitz continuous. This problem is often overcome taking into account some additional regularity properties of the infinitesimal generator, being in equation (3.99) the Laplacian Δ , such as the so-called m -dissipativity, we refer to [ADPM11] for details.

We will not concern in the present paper with the existence and uniqueness result, since it is an already established result in literature, but on the existence of an optimal control for the aforementioned equation. In particular in [BCP15], the existence and uniqueness of an optimal control has been proved for a similar equation, without the recovery variable w . To prove the existence of an optimal control in the stochastic case is quite a delicate point and necessitate of non trivial results, in particular the main results of the present work, is based, following [BCP15], on the Ekelands's variational principle.

The present work is so structured, in section 5.2 we introduce the main notation and assumptions used throughout the work, also the existence and uniqueness result for the main equation is stated. Then section 3.3.2 is devoted to the main result, that is to show the existence and uniqueness of an optimal control problem via the Ekelands's variational principle

3.3.1 The abstract setting

Let us consider the following controlled stochastic FitzHugh-Nagumo system of equations

$$\begin{cases} \partial_t v(t, \xi) &= \Delta_\xi - I_{ion}(v(t, \xi)) - w(t, \xi) + f(\xi) + B_v u(t, \xi) + \partial_t \beta_1(t), \text{ in } [0, T] \times \mathcal{O}, \\ \partial_t w(t, \xi) &= \gamma v(t, \xi) - \delta w(t, \xi) + \partial_t \beta_2(t), \text{ in } [0, T] \times \mathcal{O}, \\ \partial_\nu v(t, \xi) &= 0, \quad \text{on } [0, T] \times \partial\mathcal{O}, \\ v(0, \xi) &= v_0(\xi), \quad w(0, \xi) = w_0(\xi), \text{ in } [0, T] \times \mathcal{O}. \end{cases}, \quad (3.100)$$

where $v = v(t, \xi)$ represents the transmembrane electrical potential, $w = w(t, \xi)$ is a recovery variable, also known as gating variable and which can be used to describe the potassium conductance, $\mathcal{O} \subset \mathbb{R}^d$, $d = 2, 3$, is a bounded and open set with smooth boundary $\partial\mathcal{O}$. Furthermore Δ_ξ is the Laplacian operator with respect to the spatial variable ξ , while γ and δ are positive constants representing phenomenological coefficients, ν is the outer unit normal direction to the boundary $\partial\mathcal{O}$ and ∂_ν denotes the derivative in the direction ν , $f(\xi) \in L^\infty(\mathcal{O})$ is a given external forcing term, I_{ion} represents the *Ionic current* assumed to be as in the FitzHugh-Nagumo model, namely it is taken as a cubic nonlinearity of the following form $I_{ion}(v) = v(v-a)(v-b)$, $v_0, w_0 \in L^2(\mathcal{O})$. and β_1 and β_2 two independent Q_i -Brownian motions, $i = 1, 2$, Q_i being positive trace class commuting operators. Eventually we assume that the two operators Q_1 and Q_2 diagonalize on the same basis $\{e_k\}_{k \geq 1}$, namely we assume that there exists a sequence of positive real numbers $\{\lambda_k^i\}_{k \geq 1}$, $i = 1, 2$ such that

$$Q_i e_k = \lambda_k^i e_k, \quad i = 1, 2, \quad k \geq 1,$$

moreover we also assume that $Tr Q_i < \infty$, $i = 1, 2$. Eventually let U be a Hilbert space equipped with the scalar product $\langle \cdot, \cdot \rangle_U$, we have that $u : [0, T] \rightarrow U$ denotes the control and $B_v \in L(U, L^2(\mathcal{O}))$.

In order to rewrite (3.100) in a more compact form as an infinite dimensional stochastic evolution equation, let us define the Hilbert space $H := L^2(\mathcal{O}) \times L^2(\mathcal{O})$ endowed with the inner product

$$\langle (v_1, w_1), (v_2, w_2) \rangle_H = \gamma \langle v_1, v_2 \rangle_2 + \langle w_1, w_2 \rangle_2, \quad (3.101)$$

where $\langle \cdot, \cdot \rangle_2$ denotes the usual scalar product in $L^2(\mathcal{O})$, and the corresponding norm will be indicated by $|\cdot|_2$. Let us further introduce the space $V := H^1(\mathcal{O}) \times L^2(\mathcal{O})$ with the norm

$$|X|_V^2 = \gamma |v|_{H^1}^2 + |w|_2^2, \quad X = (v, w) \in H.$$

We then define the operator $A : D(A) \subset H \rightarrow H$ as follows

$$A = \begin{pmatrix} A_0 v & -w \\ \gamma v & -\delta w \end{pmatrix}, \quad A_0 = \Delta_\xi,$$

with domain given by

$$\begin{aligned} D(A) &:= D(A_0) \times L^2(\mathcal{O}), \\ D(A_0) &:= \{u \in H^2(\mathcal{O}) : \partial_\nu u(\xi) = 0 \text{ on } \partial\mathcal{O}\}, \end{aligned}$$

In particular we have that A generates a C_0 -semigroup satisfying

$$\|e^{tA}\| \leq e^{-\omega t}, \quad \omega > 0,$$

see, e.g. [BM08].

We further define the non-linear operator

$$F : D(F) := L^6(\mathcal{O}) \times L^2(\mathcal{O}) \rightarrow H,$$

as

$$F \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} I_{ion}(v) + f \\ 0 \end{pmatrix} = \begin{pmatrix} -v(v-a)(v-b) + f \\ 0 \end{pmatrix}.$$

In what follows we will assume that it exists a positive constant η such that

$$\langle F(x) - F(y) - \eta(x - y), x - y \rangle < 0, \quad x, y \in H,$$

and also that it holds $\omega - \eta > 0$. This implies that the term $A + F$ is m -dissipative in the sense of [DPZ14].

Let us thus consider the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, such that the two independent Wiener processes β_1 and β_2 are adapted to the filtration $\mathcal{F}_t, \forall t \geq 0$, and we define $W(t) = (\beta_1(t), \beta_2(t))$ a cylindrical Wiener process on H and by Q the operator

$$Q = \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix} \in \mathcal{L}(H; H).$$

Exploiting previously introduced notation, eq. (3.100) can be rewritten as follows

$$\begin{cases} dX(t) = [AX(t) + F(X(t))]dt + \sqrt{Q}dW(t), \\ X(0) = x_0 \in H, \quad t \in [0, T], \end{cases} \quad (3.102)$$

Definition 3.3.1. We say that the function $X \in C_W([0, T]; H)$ is called a *mild solution* to (3.102) if $X(t) : [0, T] \rightarrow H$ is continuous \mathbb{P} -a.s., $\forall t \in [0, T]$ and it satisfies the stochastic integral equation

$$X(t) = e^{-At}x + \int_0^t e^{-(t-s)A} (-F(s)) ds + \int_0^t e^{-(t-s)A} (\sqrt{Q}) dW(s), \quad \forall t \in [0, T].$$

Definition 3.3.2. sdfsadfsdfasd

The we have the following existence and uniqueness result concerning equation (3.102).

Theorem 3.3.3. For any $x \in D(F)$, there exists a unique mild solution X to (3.102) which satisfies

$$X \in L_W^2(\Omega; C([0, T]; H)) \cap L_W^2(\Omega; L^2([0, T]; V)).$$

Proof. Under above assumptions the proof follows from [ADPM11, Prop. 3.8] or [BM08, theorem 3.1]. \square

3.3.2 The optimal control problem

Let us now consider a controlled version of equation (3.102). Let then $B \in L(U; H)$ defined as

$$Bu = \begin{pmatrix} B_v u \\ 0 \end{pmatrix}, \quad B_v \in L(U; L^2(\mathcal{O})).$$

We shall denote by \mathcal{U} the space of all $(\mathcal{F}_t)_{t \geq 0}$ -adapted processes $u : [0, T] \rightarrow U$ s.t. $\mathbb{E} \int_0^T |u(t)|_U^2 dt < \infty$. The space \mathcal{U} is a Hilbert space with the norm $|u|_{\mathcal{U}} = \left(\mathbb{E} \int_0^T |u(t)|_U^2 dt \right)^{\frac{1}{2}}$ and scalar product

$$\langle u, v \rangle_{\mathcal{U}} = \left(\mathbb{E} \int_0^T \langle u(t), v(t) \rangle_U dt \right)^{\frac{1}{2}}, \quad \forall u, v \in \mathcal{U},$$

where $\langle \cdot, \cdot \rangle_U$ is the scalar product of U .

Consider the functions $g, g_0 : \mathbb{R} \rightarrow \mathbb{R}$ and $h : U \rightarrow \bar{\mathbb{R}} :=]-\infty, \infty]$, which satisfy the following conditions

- (i) $g, g_0 \in C^1(H)$ and $Dg, Dg_0 \in Lip(H; H)$, where D stands for the Fréchet differential
- (ii) h is convex, lower-semicontinuous and $(\partial h)^{-1} \in Lip(U)$ where $\partial h : U \rightarrow U$ is the subdifferential of h , see, e.g., [BP12, p. 82]. Moreover we assume that $\exists \alpha_1 > 0$ and $\alpha_2 \in \mathbb{R}$ s.t. $h(u) \geq \alpha_1 |u|_U^2 + \alpha_2, \forall u \in U$, and we set $L = \|(\partial h)^{-1}\|_{Lip(U)}$.

We consider the following optimal control problem

$$\text{Minimize } \mathbb{E} \int_0^T (g(X(t)) + h(u(t))) dt + \mathbb{E} g_0(X(T)), \quad (\text{P})$$

subject to $u \in \mathcal{U}$ and

$$\begin{cases} dX(t) = [AX(t) + F(X(t))]dt + Bu(t)dt + \sqrt{Q}dW(t), \\ X(0) = x_0 \in H, \quad t \in [0, T], \end{cases} \quad (3.103)$$

Theorem 3.3.4. *Let $x \in D(A)$. Then there exists $C^* > 0$ independent of x such that for $LT + \|Dg_0\|_{Lip} < C^*$ there is a unique solution (u^*, X^*) to problem (P).*

Proof. Let us consider the function $\Psi : \mathcal{U} \rightarrow \bar{\mathbb{R}}$ defined by

$$\Psi(u) = \mathbb{E} \int_0^T (g(X^u(t)) + h(u(t))) dt + \mathbb{E} g_0(X^u(T)),$$

where X^u is the solution to (3.103). Recall that Ψ is lower-semicontinuous.

We shall apply Ekeland's variational principle (See, e.g., [Eke74] or also [BCP15, BI99]), that is there is a sequence $\{u_\epsilon\} \subset \mathcal{U}$ such that

$$\begin{aligned} \Psi(u_\epsilon) &\leq \inf\{\Psi(u); u \in \mathcal{U}\} + \epsilon, \\ \Psi(u_\epsilon) &\leq \Psi(u) + \sqrt{\epsilon} |u_\epsilon - u|_{\mathcal{U}}, \quad \forall u \in \mathcal{U}. \end{aligned} \quad (3.104)$$

In other words,

$$u_\epsilon = \arg \min_{u \in \mathcal{U}} \{\Psi(u) + \sqrt{\epsilon} |u_\epsilon - u|_{\mathcal{U}}\}.$$

Hence $(X^{u_\epsilon}, u_\epsilon)$ is a solution to the optimal control problem

$$\begin{aligned} \min \left\{ \mathbb{E} \int_0^T (g(X^u(t)) + h(u(t))) dt + \mathbb{E} g_0(X^u(T)) + \right. \\ \left. + \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |u(t) - u_\epsilon(t)|_U^2 dt \right)^{\frac{1}{2}}; u \in \mathcal{U} \right\}. \end{aligned} \quad (3.105)$$

Equation (3.105) means that for all $v \in \mathcal{U}$ and $\lambda > 0$ it holds

$$\begin{aligned} &\mathbb{E} \int_0^T (g(X^{u_\epsilon + \lambda v}(t)) + h((u_\epsilon + \lambda v)(t))) dt + \mathbb{E} g_0(X^{u_\epsilon + \lambda v}(T)) + \\ &+ \lambda \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |v(t)|_U^2 dt \right)^{\frac{1}{2}} \leq \\ &\leq \mathbb{E} \int_0^T (g(X_\epsilon(t)) + h(u_\epsilon(t))) dt + \mathbb{E} g_0(X_\epsilon(T)), \end{aligned}$$

that is we get

$$\begin{aligned} & \mathbb{E} \int_0^T \langle Dg(X_\epsilon(t)), Z^v(t) \rangle_2 dt + \mathbb{E} \int_0^T h'(u_\epsilon(t), v(t)) dt + \\ & + \mathbb{E} \langle Dg_0(X_\epsilon(T)), Z^v(T) \rangle_2 + \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |v(t)|_U^2 dt \right)^{\frac{1}{2}} \leq 0, \quad \forall v \in \mathcal{U}, \end{aligned} \quad (3.106)$$

where Z^v solves the system in variations associated with (3.103),

$$\begin{cases} \frac{\partial}{\partial t} Z^v(t) = AZ^v(t) + DF(X_\epsilon(t))Z^v(t) + Bv(t), t \in [0, T], \\ Z^v(0) = 0, \end{cases} \quad (3.107)$$

and $h' : U \times U \rightarrow \mathbb{R}$ is the directional derivatives of h , see, e.g., [BP12, p.81], namely

$$h'(u_\epsilon, v) = \lim_{\lambda \downarrow 0} \frac{h(u_\epsilon + \lambda v) - h(u_\epsilon)}{\lambda}, \quad \forall v \in U.$$

We thus associate with (3.103) the dual stochastic backward equation

$$\begin{cases} dp_\epsilon(t) = -[Ap_\epsilon(t)dt + DF(X_\epsilon) p_\epsilon(t) - Dg(X_\epsilon(t))] dt + \kappa_\epsilon(t) \sqrt{Q} dW(t), t \in [0, T], \\ p_\epsilon(T) = -Dg_0(X_\epsilon(T)), \end{cases} \quad (3.108)$$

It is well-known that equation (3.108) has a unique solution $(p_\epsilon, \kappa_\epsilon)$ satisfying

$$\begin{aligned} p_\epsilon & \in L_W^\infty([0, T]; H) \cap L_W^2([0, T]; V), \\ \kappa_\epsilon & \in L_W^2([0, T]; H), \end{aligned}$$

(See, e.g., [FT02, Prop. 4.2] or [Tes96]). By Itô's formula we have from (3.107) and (3.108) that

$$d \langle p_\epsilon, Z^v \rangle_H = \langle dp_\epsilon, Z^v \rangle_H + \langle p_\epsilon, dZ^v \rangle_H,$$

and this immediately implies

$$\mathbb{E} \int_0^T \langle Dg(X_\epsilon(t)), Z^v(t) \rangle_H dt + \mathbb{E} \langle Dg_0(X_\epsilon(T)), Z^v(T) \rangle_H = \mathbb{E} \int_0^T \langle Bv(t), p_\epsilon(t) \rangle_H dt,$$

which substituted in (5.25) yields that $\forall v \in \mathcal{U}$, the following inequality holds

$$\begin{aligned} & \mathbb{E} \int_0^T h'(u_\epsilon(t), v(t)) dt + \sqrt{\epsilon} \left(\mathbb{E} \int_0^T |v(t)|_U^2 dt \right)^{\frac{1}{2}} \leq \\ & \leq \mathbb{E} \int_0^T \langle B^* p_\epsilon(t), v(t) \rangle_U dt. \end{aligned}$$

Let $G(u) := \mathbb{E} \int_0^T h(u(t)) dt$, then its subdifferential $\partial G : \mathcal{U} \rightarrow \mathcal{U}$, evaluated in u_ϵ is given by

$$\partial G(u_\epsilon) = \left\{ v^* \in \mathcal{U} : \langle v, v^* \rangle_{\mathcal{U}} \leq \mathbb{E} \int_0^T h'(u_\epsilon(t), v(t)) dt, \forall v \in \mathcal{U} \right\}.$$

(See, e.g., [BP12, p.81]). Then we infer that

$$u_\epsilon(t) = (\partial h)^{-1} \left(B^* p_\epsilon(t) + \sqrt{\epsilon} \tilde{\theta}_\epsilon \right), \quad t \in [0, T], \quad \mathbb{P} - a.s.,$$

where $\tilde{\theta}_\epsilon \in \mathcal{U}$ and $|\tilde{\theta}_\epsilon|_{\mathcal{U}} \leq 1, \forall \epsilon > 0$.

Therefore, we have shown that

$$\begin{aligned} u_\epsilon &= (\partial h)^{-1} (B^* p_\epsilon + \theta_\epsilon), \quad \|\theta_\epsilon\|_{L^2([0, T] \times \Omega; U)} \leq \sqrt{\epsilon}, \\ dp_\epsilon(t) &= -[Ap_\epsilon(t)dt + DF(X_\epsilon)p_\epsilon(t) - Dg(X_\epsilon(t))]dt + \kappa_\epsilon(t)\sqrt{Q}dW(t), \quad t \in [0, T], \\ p_\epsilon(T) &= -Dg_0(X_\epsilon(T)), \end{aligned} \tag{3.109}$$

Using the Itô formula applied to $|X|_2^2$, we have that $\forall \epsilon > 0$ it holds

$$\begin{aligned} |X_\epsilon(t)|_H^2 &= |x|_H^2 + 2 \int_0^t \langle AX_\epsilon(s) + F(X_\epsilon(s)) + Bu_\epsilon(s), X_\epsilon(s) \rangle_H ds + \\ &+ TrQt + 2 \int_0^t \langle X_\epsilon(s), \sqrt{Q}dW(s) \rangle_H. \end{aligned} \tag{3.110}$$

(Here and everywhere in the following we shall denote by C several positive constants independent of ϵ .)

From the fact that $\langle X_\epsilon(s), \sqrt{Q}dW(s) \rangle_H$ is a square integrable martingale, [DPZ14, Th. 3.14, Th. 4.12] and recalling the assumption $TrAQ < \infty$ we have that

$$\mathbb{E} \sup_{t \in [0, T]} \left| \int_0^t \langle X_\epsilon(s), \sqrt{Q}dW(s) \rangle_H \right| \leq C \mathbb{E} \int_0^T |X_\epsilon(t)|_H^2 dt,$$

and from the fact that A generates a strongly continuous semigroup, see, e.g. [BM08], we have that

$$\int_0^t \langle AX_\epsilon(s), X_\epsilon(s) \rangle_H ds \leq C_1 \int_0^t |X_\epsilon(s)|_V^2 ds.$$

We also have that it holds,

$$\int_0^t \langle F(X_\epsilon(s)), X_\epsilon(s) \rangle_H ds \leq C |X_\epsilon(t)|_H^2,$$

see, e.g. [ADPM11, BM08] for details. Eventually from assumption (ii) we have

$$\int_0^t \langle Bu(s), X_\epsilon(s) \rangle_H ds \leq L^{-1} \int_0^T |u_\epsilon(t)|_U^2 dt.$$

Taking then the expectation on both side of (3.110) yields

$$\mathbb{E} \sup_{t \in [0, T]} |X_\epsilon(t)|_H^2 + \mathbb{E} \int_0^T |X_\epsilon(t)|_V^2 dt \leq C_1 + C_2 \int_0^T \mathbb{E} \sup_{s \in [0, t]} |X_\epsilon(s)|_H^2 dt$$

and applying Gronwall's lemma it follows eventually that

$$\mathbb{E} \sup_{t \in [0, T]} |X_\epsilon(t)|_H^2 + \mathbb{E} \int_0^T |X_\epsilon(t)|_V^2 dt \leq C(1 + |x|_H^2). \quad (3.111)$$

In an analogous manner, applying Itô formula to $|p_\epsilon|_H^2$ by (3.109) we obtain that

$$\begin{aligned} \frac{1}{2} d|p_\epsilon(t)|_H^2 &= - \langle Ap_\epsilon(t) + DF(X_\epsilon(t))p_\epsilon(t) - Dg(X_\epsilon(t)), p_\epsilon(t) \rangle_H + \\ &= \frac{1}{2} \langle \kappa_\epsilon(t), \kappa_\epsilon(t) \rangle_H dt + \left\langle p_\epsilon(t), \kappa_\epsilon(t) \sqrt{Q} dW(t) \right\rangle_H. \end{aligned}$$

which yields after applying arguments similar to the ones above

$$\begin{aligned} \mathbb{E} \sup_{t \in [0, T]} |p_\epsilon(t)|_H^2 + \mathbb{E} \int_0^T |p_\epsilon(t)|_V^2 dt + \mathbb{E} \int_0^T |\kappa_\epsilon(t)|_H^2 dt &\leq \\ \leq C + \mathbb{E}|X_\epsilon(T)|_H^2 &\leq C, \quad \forall \epsilon > 0. \end{aligned} \quad (3.112)$$

We have that

$$\begin{aligned} \frac{\partial}{\partial t} (X_\epsilon(t) - X_\lambda(t)) &= A(X_\epsilon(t) - X_\lambda(t)) + (F(X_\epsilon(t)) - F(X_\lambda(t))) + \\ &+ BB^*(p_\epsilon(t) - p_\lambda(t)) + B(\theta_\epsilon(t) - \theta_\lambda(t)). \end{aligned} \quad (3.113)$$

In virtue of (3.112) this yields

$$\begin{aligned} \frac{1}{2} |X_\epsilon(t) - X_\lambda(t)|_H^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_V^2 ds &\leq \\ \leq \int_0^t \langle F(X_\epsilon(s)) - F(X_\lambda(s)), X_\epsilon(s) - X_\lambda(s) \rangle_H ds &+ \\ + L \int_0^t |p_\epsilon(s) - p_\lambda(s)|_H |X_\epsilon(s) - X_\lambda(s)|_H ds &+ \\ + C \int_0^t |\theta_\epsilon(s) - \theta_\lambda(s)|_U |X_\epsilon(s) - X_\lambda(s)|_H ds, \quad \forall t \in [0, T], \end{aligned}$$

where $L = \|(\partial h)^{-1}\|_{Lip}$.

We further have that, see, e.g. [ADPM11, BM08]

$$\langle F(X_\epsilon) - F(X_\lambda), X_\epsilon - X_\lambda \rangle_H \leq C |X_\epsilon - X_\lambda|_H^2,$$

which yields, for $t \in [0, T]$, applying Young inequality,

$$\begin{aligned} |X_\epsilon(t) - X_\lambda(t)|_H^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_V^2 ds &\leq \\ \leq C \left(L \int_0^t |p_\epsilon(s) - p_\lambda(s)|_H^2 ds + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_H^2 ds + \epsilon + \lambda \right). \end{aligned} \quad (3.114)$$

Applying Gronwall's lemma in (3.114), we have

$$\begin{aligned} & |X_\epsilon(t) - X_\lambda(t)|_2^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_V^2 ds \leq \\ & \leq C \left(L \int_0^T |p_\epsilon(s) - p_\lambda(s)|_2^2 ds + \epsilon + \lambda \right), \quad \forall \epsilon, \lambda > 0, t \in [0, T]. \end{aligned} \quad (3.115)$$

Similarly we get by the Itô formula

$$\begin{aligned} & |p_\epsilon(t) - p_\lambda(t)|_H^2 + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_V^2 ds + \frac{1}{2} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_H^2 ds = \\ & = |Dg_0(X_\epsilon(T)) - Dg_0(X_\lambda(T))|_H^2 + \\ & \quad + \int_t^T \langle DF(X_\epsilon(s))p_\epsilon(s) - DF(X_\lambda(s))p_\lambda(s), p_\epsilon(s) - p_\lambda(s) \rangle_H ds + \\ & \quad - \int_t^T \langle \kappa_\epsilon(s) - \kappa_\lambda(s) \rangle_{\sqrt{Q}dW(s), X_\epsilon(s) - X_\lambda(s)} \rangle_H \leq \\ & = \int_t^T \langle DF(X_\epsilon(s))(p_\epsilon(s) - p_\lambda(s)), p_\epsilon(s) - p_\lambda(s) \rangle ds + \\ & \quad + \int_t^T \langle p_\lambda(s)(DF(X_\epsilon(s)) - DF(X_\lambda(s))), p_\epsilon(s) - p_\lambda(s) \rangle_H ds + \\ & \quad + \int_t^T \langle \kappa_\epsilon(s) - \kappa_\lambda(s) \rangle_{\sqrt{Q}dW(s), X_\epsilon(s) - X_\lambda(s)} \rangle_H + \\ & \quad + |Dg_0(X_\epsilon(T)) - Dg_0(X_\lambda(T))|_H^2 \leq \\ & \leq C \left(\int_t^T (|X_\epsilon(s)|_H^2 + 1) |p_\epsilon(s) - p_\lambda(s)|_H^2 ds \right) + \\ & \quad + C \left(\int_t^T (1 + |X_\epsilon(s)|^2 + |X_\lambda(s)|^2) |X_\epsilon(s) - X_\lambda(s)|_H |p_\epsilon(s) - p_\lambda(s)|_H |p_\epsilon(s)|_H ds \right) + \\ & \quad + \int_t^T \langle \kappa_\epsilon(s) - \kappa_\lambda(s) \rangle_{\sqrt{Q}dW(s), X_\epsilon(s) - X_\lambda(s)} \rangle_H + \\ & \quad + \|Dg_0\|_{Lip} |X_\epsilon(T) - X_\lambda(T)|_H^2, \quad t \in [0, T], \mathbb{P} - a.s.. \end{aligned} \quad (3.116)$$

Exploiting again Young's inequality, and denoting for short

$$T_{\epsilon, \lambda} := (1 + |X_\epsilon|_H^2 + |X_\lambda|_H^2) |p_\epsilon|_H,$$

we get,

$$\begin{aligned} & (|X_\epsilon(s) - X_\lambda(s)|_H |p_\epsilon(s) - p_\lambda(s)|_H) T_{\epsilon, \lambda} \leq \\ & \leq C \left(|X_\epsilon - X_\lambda|_H^2 + |p_\epsilon - p_\lambda|_H^2 \right) T_{\epsilon, \lambda}. \end{aligned} \quad (3.117)$$

Substituting now (3.117) into (3.114), (3.116), we obtain \mathbb{P} -a.s.

$$\begin{aligned}
 & |X_\epsilon(t) - X_\lambda(t)|_H^2 + |p_\epsilon(t) - p_\lambda(t)|_H^2 + \int_0^t |X_\epsilon(s) - X_\lambda(s)|_V^2 ds + \\
 & + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_V^2 ds + \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_H^2 ds \leq \\
 & \leq C \left(L \int_0^t |p_\epsilon(s) - p_\lambda(s)|_H^2 ds + \epsilon + \lambda \right) + C \int_t^T |p_\epsilon(s) - p_\lambda(s)|_2^2 |X_\epsilon(s)|_H^2 ds + \\
 & + \|Dg_0\|_{Lip} |X_\epsilon(T) - X_\lambda(T)|_2^2 + \\
 & + C \int_t^T \left(|X_\epsilon(s) - X_\lambda(s)|_H^2 + |p_\epsilon(s) - p_\lambda(s)|_H^2 \right) T_{\epsilon,\lambda}(s) ds + \\
 & - \int_t^T \left\langle (\kappa_\epsilon(s) - \kappa_\lambda(s)) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \right\rangle_H, \quad \forall t \in [0, T].
 \end{aligned} \tag{3.118}$$

Exploiting thus the fact that the process $r \mapsto \int_t^r \langle (\kappa_\epsilon - \kappa_\lambda) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \rangle_2$ is a local martingale on $[t, T]$, hence by the Burkholder-Davis-Gundy inequality, see, e.g., [DPZ96, p.58], we have for all $r \in [t, T]$

$$\begin{aligned}
 & \mathbb{E} \sup_{r \in [t, T]} \left| \int_t^r \left\langle (\kappa_\epsilon(s) - \kappa_\lambda(s)) \sqrt{Q} dW(s), X_\epsilon(s) - X_\lambda(s) \right\rangle_H \right| \leq \\
 & \leq C \left(\mathbb{E} \int_0^r |\kappa_\epsilon(s) - \kappa_\lambda(s)|_H^2 |X_\epsilon(s) - X_\lambda(s)|_H^2 ds \right)^{\frac{1}{2}} \leq \\
 & \leq C \mathbb{E} \sup_{s \in [t, r]} |X_\epsilon(s) - X_\lambda(s)|_H^2 + C \mathbb{E} \int_t^r |\kappa_\epsilon(s) - \kappa_\lambda(s)|_H^2 ds.
 \end{aligned} \tag{3.119}$$

Taking then the expectation in and by (3.118), and using (3.119) we get

$$\begin{aligned}
 & \mathbb{E} \sup_{s \in [t, T]} \left(|X_\epsilon(s) - X_\lambda(s)|_H^2 + |p_\epsilon(s) - p_\lambda(s)|_H^2 \right) \\
 & + \mathbb{E} \int_0^T |X_\epsilon(s) - X_\lambda(s)|_V^2 ds + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_H^2 ds \\
 & + \mathbb{E} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_H^2 ds \leq \\
 & \leq \|Dg_0\| \mathbb{E} |X_\epsilon(T) - X_\lambda(T)|_H^2 + C \left(L \mathbb{E} \int_0^T |p_\epsilon(s) - p_\lambda(s)|_H^2 ds + \epsilon + \lambda \right) \\
 & + C \mathbb{E} \sup_{s \in [t, T]} |X_\epsilon(s) - X_\lambda(s)|_H^2 \\
 & + C \mathbb{E} \int_t^T \left(|p_\epsilon(s) - p_\lambda(s)|_H^2 + |X_\epsilon(s) - X_\lambda(s)|_H^2 \right) \left(|X_\epsilon(s)|_H^2 + T_{\epsilon,\lambda}(s) \right) ds.
 \end{aligned} \tag{3.120}$$

Taking into account estimates (3.115) and (3.116), from (3.120) we have

$$\begin{aligned}
& \mathbb{E} \sup_{s \in [t, T]} (|X_\epsilon(s) - X_\lambda(s)|_H^2 + |p_\epsilon(s) - p_\lambda(s)|_H^2) \\
& \quad + \mathbb{E} \int_0^T |X_\epsilon(s) - X_\lambda(s)|_V^2 ds + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_H^2 ds \\
& \quad + \mathbb{E} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_H^2 ds \leq \\
& \leq \tilde{C} \left(L \mathbb{E} \int_0^T |p_\epsilon(s) - p_\lambda(s)|_H^2 ds \right) \\
& \quad + \tilde{C} \left(\mathbb{E} \int_t^T |p_\epsilon(s) - p_\lambda(s)|_H^2 (|X_\epsilon(s)|_H^3 + T_{\epsilon, \lambda}(s)) ds \right) \\
& \quad + \tilde{C} \|Dg_0\|_{Lip} \mathbb{E} |X_\epsilon(T) - X_\lambda(T)|_H^2 + \tilde{C}(\epsilon + \lambda).
\end{aligned} \tag{3.121}$$

where \tilde{C} is a positive constant independent of ϵ and λ . It follows that if $\tilde{C}(LT + \|Dg_0\|_{Lip}) < 1$, then, for any $t \in [0, T]$,

$$\begin{aligned}
& \mathbb{E} \sup_{s \in [t, T]} (|X_\epsilon(s) - X_\lambda(s)|_H^2 + |p_\epsilon(s) - p_\lambda(s)|_H^2) \\
& \quad + \mathbb{E} \int_0^T |X_\epsilon(s) - X_\lambda(s)|_V^2 ds + \int_t^T |p_\epsilon(s) - p_\lambda(s)|_H^2 ds \\
& \quad + \mathbb{E} \int_t^T |\kappa_\epsilon(s) - \kappa_\lambda(s)|_H^2 ds \leq \\
& \leq C \mathbb{E} \int_t^T |p_\epsilon(s) - p_\lambda(s)|_H^2 (|X_\epsilon(s)|_H^2 + T_{\epsilon, \lambda}(s)) ds + C(\epsilon + \lambda).
\end{aligned} \tag{3.122}$$

Let us define for $j \in \mathbb{N}$

$$\Omega_j := \left\{ \omega \in \Omega : \sup_{\epsilon} \sup_{t \in [0, T]} (|X_\epsilon(t)|_H^2 + |X_\epsilon(t)|_V^2 + |p_\epsilon(t)|_H^2) dt \leq j \right\},$$

then estimates (3.111) implies that

$$\mathbb{P}(\Omega_j) \geq 1 - \frac{C}{j}, \quad \forall j \in \mathbb{N},$$

for some constant C independent of ϵ .

If we set $X_\epsilon^j := \mathbb{1}_{[\Omega_j]} X_\epsilon$, $p_\epsilon^j := \mathbb{1}_{[\Omega_j]} p_\epsilon$ and $\kappa_\epsilon^j := \mathbb{1}_{[\Omega_j]} \kappa_\epsilon$, then such quantities satisfy the system (3.109), with $\mathbb{1}_{[\Omega_j]} \sqrt{Q} dW$. The latter means that estimate (3.122) still holds in this context, so that we have

$$\begin{aligned}
& \mathbb{E} \sup_{s \in [t, T]} |X_\epsilon^j(s) - X_\lambda^j(s)|_H^2 + \sup_{s \in [t, T]} |p_\epsilon^j(s) - p_\lambda^j(s)|_H^2 \\
& \quad + \mathbb{E} \int_t^T |p_\epsilon^j(s) - p_\lambda^j(s)|_V^2 ds + \mathbb{E} \int_t^T |(\kappa_\epsilon(s) - \kappa_\lambda(s)) \chi_j|_H^2 ds \leq \\
& \leq C_j \int_t^T \mathbb{E} |p_\epsilon^j(s) - p_\lambda^j(s)|_H^2 ds + C(\epsilon + \lambda), \quad j \in \mathbb{N}.
\end{aligned} \tag{3.123}$$

By Gronwall's lemma we get, for any $t \in [0, T]$

$$\mathbb{E} \sup_{s \in [t, T]} |X_\epsilon^j(s) - X_\lambda^j(s)|_H^2 + \sup_{s \in [t, T]} |p_\epsilon^j(s) - p_\lambda^j(s)|_H^2 \leq C(\epsilon + \lambda)e^{C_j T}, \quad (3.124)$$

hence, for $\epsilon \rightarrow 0$ and all $j \in \mathbb{N}$ and all $t \in [0, T]$, we obtain

$$\begin{aligned} X_\epsilon^j &\rightarrow X^j && \text{in } L^2(\Omega_j; L^2([0, T] \times \mathcal{O}) \times L^2([0, T] \times \mathcal{O})), \\ p_\epsilon^j &\rightarrow p^j && \text{in } L^2(\Omega_j; L^2([0, T] \times \mathcal{O}) \times L^2([0, T] \times \mathcal{O})). \end{aligned} \quad (3.125)$$

Therefore for each $\omega \in \Omega$, we have that $\{X_\epsilon(t, \omega), p_\epsilon(t, \omega)\}$ are Cauchy sequences in $L^2([0, T] \times \mathcal{O})$, with respect to ϵ and by estimates (3.111) and (3.112) it follows that taking related subsequences, still denoted by ϵ , we have

$$\begin{aligned} X_\epsilon &\rightharpoonup X^* && \text{in } L^2([0, T] \times \Omega; V), \\ p_\epsilon &\rightharpoonup p^* && \text{in } L^2([0, T] \times \Omega \times \mathcal{O} \times \mathcal{O}), \\ p_\epsilon &\rightharpoonup p^* && \text{in } L^2([0, T] \times \Omega; V), \\ u_\epsilon &\rightharpoonup u^* && \text{in } L^\infty([0, T]; L^2(\Omega \times U)), \end{aligned} \quad (3.126)$$

where \rightharpoonup means weak (respectively, weak-star) convergence, so we have for $\epsilon \rightarrow 0$

$$X_\epsilon \rightarrow X^*, \quad p_\epsilon \rightarrow p^*, \text{ a.e. in } [0, T] \times \Omega \times \mathcal{O} \times \mathcal{O}. \quad (3.127)$$

We also have, since $\{I_{ion}(v_\epsilon)\}$ is bounded in $L^{\frac{4}{3}}([0, T] \times \Omega \times \mathcal{O})$, then it is weakly compact in $L^1([0, T] \times \Omega \times \mathcal{O})$ and by (3.127) we have that for a subsequence $\{\epsilon\} \rightarrow 0$,

$$I_{ion}(v_\epsilon) \rightarrow I_{ion}(v^*), \quad \text{a.e. in } [0, T] \times \Omega \times \mathcal{O},$$

which implies that

$$I_{ion}(v_\epsilon) \rightarrow I_{ion}(v^*) \quad \text{in } L^1([0, T] \times \Omega \times \mathcal{O}). \quad (3.128)$$

Then, letting $\epsilon \rightarrow 0$ we obtain

$$\begin{cases} dX^*(t) = AX^*(t)dt + F(X^*(t))dt + \sqrt{Q}dW(t) + Bu^*(t)dt, t \in [0, T], \\ X^*(0) = x, \end{cases}.$$

Taking into account that Ψ is weakly lower semicontinuous in \mathcal{U} we infer by (3.104) that

$$\Psi(u^*) = \inf \{\Psi(u); u \in \mathcal{U}\},$$

therefore (X^*, u^*) is optimal for the problem (P) and the proof of existence is therefore complete. \square

Concerning the uniqueness for the optimal pair (X^*, u^*) given by Th. 3.3.4, we have that it follows by the same argument via the maximum principle result for problem (P), namely one has the following result.

Theorem 3.3.5. *Let (X^*, u^*) be optimal in problem (P), then*

$$u^* = (\partial h)^{-1}(B^*p), \text{ a.e. } t \in [0, T], \quad (3.129)$$

where p is the solution to the backward stochastic equation (3.108).

Proof. If (X^*, u^*) is optimal for the problem (P), then by the same argument used to prove Th. 3.3.4, see (5.25), we have

$$\begin{aligned} & \mathbb{E} \int_0^T \langle Dg(X^*(t)), Z^v(t) \rangle_2 dt + \mathbb{E} \int_0^T h'(u^*(t), v(t)) dt \\ & + \mathbb{E} \langle Dg_0(X^*(T)), Z^v(T) \rangle_2 \leq 0, \quad \forall v \in \mathcal{U}, \end{aligned} \quad (3.130)$$

where Z^v is solution to equation (3.107) with X_ϵ replaced by X^* . This implies as above that (3.129) holds. \square

The uniqueness in (P). If (X^*, u^*) is optimal in (P) then it satisfies systems (3.102), (3.129) and (3.130), so that arguing as in the proof of Th. 3.3.4, the same set of estimates implies that the previous system has at most one solution if $LT + \|Dg_0\|_{Lip} < C^*$, where C^* is sufficiently small. \square

4 | Infinite dimensional analysis on networks

Section 4.1 is taken from [CDP16b] whereas Section 4.2 is taken from [CDP16d].

Abstract

In the present chapter we prove the existence and uniqueness for the solution to a stochastic reaction–diffusion equation, defined on a network, and subjected to non–local dynamic stochastic boundary conditions. The result is obtained by deriving a Gaussian-type estimate for the related leading semigroup, under rather mild regularity assumptions on the coefficients. An application of the latter to a stochastic optimal control problem on graphs, is also provided.

We thus consider a reaction-diffusion equation on a network subjected to dynamic boundary conditions, with time delayed behaviour, also allowing for multiplicative Gaussian noise perturbations. Exploiting semigroups theory, we rewrite the aforementioned stochastic problem as an abstract stochastic partial differential equation taking values in a suitable product Hilbert space, for which we prove the existence and uniqueness of a mild solution. Eventually, a stochastic optimal control application is studied.

4.1 Gaussian estimate on networks with dynamic stochastic boundary conditions

Starting with the introductory work [Lum80], where elliptic operators acting on suitable functions spaces on network have been first introduced, several works related to a wide set of physical phenomena whose dynamics are carried out on graphs, have appeared, e.g., concerning the study of heat diffusion, see, e.g. [Mug10], applications to quantum mechanics, see, e.g., [Tum06], the stochastic modelisation of neurobiological activities, particularly with respect to the analysis of the *FitzHugh-Nagumo equation*, see, e.g., [ADP11, BCP15, BMZ⁺08, CGMY03], and references therein, the quest for invariant measures, see, e.g., [ADPM13], the problem of suitable types of estimates, as in the case of the Gaussian one, see, e.g., [DPZ11], and references therein, etc.

A powerful technique often used to address aforementioned problems, consists in introducing a suitable infinite dimensional product space and then study the diffusion problem exploiting a semigroup theory approach, see [Mug14] and references therein, for a detailed analysis of the latter subject. Moreover, to what concerns standard problems of existence and uniqueness for the solution of a diffusion problem, as well as the spectral properties of related the leading semigroup, the attention has often been put on the determination of proper boundary condition for the particular diffusion problem one is interested in.

When the focus is on diffusion problems governed by a second order differential operator, then typical boundary condition are the so-call *generalized Kirchhoff conditions*, see, e.g., [Mug10]. Nevertheless, during recent years, also different type of rather general boundary conditions has been proposed. The latter is the case, e.g., of non-local boundary conditions, allowing for non-local interaction of non-adjacent vertex of the graph, see, e.g., [CGMY03, DPZ11], dynamic boundary conditions, see, e.g., [BMZ⁺08, MR07], etc.

The main goal of the present work is to generalize previously mentioned approaches in order to achieve a unified perspective. We will start from a completely general non-local diffusion problem, endowed with non-local boundary conditions which will be both dynamic and static. In such a setting, we state our main result, namely we prove a Gaussian upper bound for the semigroup generated by a proper infinitesimal generator acting on a suitable Hilbert space. We would like to underline that latter type of bound turns out to be extremely powerful when one wants to prove existence and uniqueness of a solution to a stochastic partial differential equation (SPDE), since this immediately leads the operator to be *Hilbert-Schmidt*, allowing to relax regularity assumptions on the coefficients of the SPDE.

The general approach that can be used to show the *Hilbert-Schmidt property* of the leading semigroup, typically relies on the study of its spectral properties. However it is not always possible to give a precise characterization of the semigroup eigenvalues, particularly whit respect to diffusive problems on a graph. In such a case a complete characterization of the spectrum can be obtained by considering the topological structure of the graph. Alternatively, one can try to derive a *heat kernel* which leads to prove a Gaussian upper bound for the semigroup. The latter approach will be the one we will pursue in the present paper.

The work is so structured, in Section 4.2.1, exploiting the theory of sesquilinear form, we will introduce a suitable infinite dimensional space, showing that our equation can be rewritten as an infinite dimensional problem where the differential operator generates a strongly continuous analytic semigroup, hence obtaining the well-posedness of the abstract Cauchy problem. Then, in Section 4.1.1, we will prove a Gaussian estimates for the operator, while in Section 4.2.3 a suitable stochastic multiplicative perturbation will be introduced in order to show both the existence and the uniqueness of a *mild solution*, in a suitable sense, under rather mild assumptions on the coefficients. Eventually, in Section 4.2.4, a stochastic optimal control application will be proposed.

4.1.1 General framework

Let us consider a finite connected network identified with a graph \mathbb{G} composed by a finite number $n \in \mathbb{N}$ of vertices, indicated by v_1, \dots, v_n and linked by a finite number $m \in \mathbb{N}$ of edges, indicated by e_1, \dots, e_m and assumed to be of unitary length. For the sake of readability, let us also introduce the following notations: we use Latin letters $i, j, k = 1, \dots, m$, to denote quantities related to edges, so that u_i will stand for a function on the

edge e_i , for $i = 1, \dots, m$; while we use Greek letters $\alpha, \beta, \gamma = 1, \dots, n$, to denote quantities related to vertices, so that d_α , $\alpha = 1, \dots, n$, will be the values of the unknown function evaluated at the vertex v_α , with $\alpha = 1, \dots, n$.

In order to describe the structure of the graph \mathbb{G} we will exploit the *incidence matrix* $\Phi := (\phi_{\alpha,i})_{n \times m}$, see, e.g., [Mug14], which is defined as follows: $\Phi := \Phi^+ - \Phi^-$, where the sum is intended componentwise, with $\Phi^+ = (\phi_{\alpha,i}^+)_{n \times m}$, resp. $\Phi^- = (\phi_{\alpha,i}^-)_{n \times m}$, is the *incoming incidence matrix*, resp. the *outgoing incidence matrix*. In particular, both of them have value 1, whenever the vertex v_α is the initial point, resp. the terminal point, of the edge e_i , and 0 otherwise. The latter implies that

$$\phi_{\alpha,i}^+ = \begin{cases} 1 & v_\alpha = e_i(0), \\ 0 & \text{otherwise} \end{cases}, \quad \phi_{\alpha,i}^- = \begin{cases} 1 & v_\alpha = e_i(1), \\ 0 & \text{otherwise} \end{cases}.$$

Aforementioned definition is consistent with the idea that if $|\phi_{\alpha,i}| = 1$, then we the edge e_i is called *incident* to the vertex v_α , and it remains defined the set

$$\Gamma(v_\alpha) = \{i \in \{1, \dots, m\} : |\phi_{\alpha,i}| = 1\},$$

of all the *incident edges* to the vertex v_α .

In order to consider the most general framework, we allow the dynamic of the unknown function u , defined on the network, to depend non-locally on the underlying graph \mathbb{G} , which implies to take into account non-local interactions, namely the process taking place on the edge e_i can be affected by the process that takes place on the edge e_j , $i, j = 1, \dots, m$, even if the edge e_j is not directly connected with the edge e_i .

We also introduce, see [CMN08], the *ephaptic incidence tensor*, which is defined as follows

$$\mathcal{I} := \mathcal{I}^+ - \mathcal{I}^-, \quad \mathcal{I}^+ := \Phi^+ \otimes \Phi^+, \quad \mathcal{I}^- := \Phi^- \otimes \Phi^-,$$

being \otimes the Kronecker product of two $n \times m$ matrices, defined as

$$(A \otimes B)_{\beta j}^{\alpha i} := a_{\alpha i} b_{\beta j},$$

in particular $(A \otimes B)$ is a $n^2 \times m^2$ matrix and, in our case, it is worth to mention that the matrix $(A \otimes B)$ is symmetric.

Using previous notation, in what follows we will denote by $\iota_{\beta j}^{\alpha i}$, resp. ${}^+ \iota_{\beta,j}^{\alpha,i}$, resp. ${}^- \iota_{\beta j}^{\alpha i}$, the entries of the matrix \mathcal{I} , resp. of the matrix \mathcal{I}^+ , resp. of the matrix \mathcal{I}^- .

Remark 4.1.1. We underline that the entry $\iota_{\beta j}^{\alpha i}$ represents the influence that the vertex v_β , as an endpoint of the edge e_j , plays on the vertex v_α which is an endpoint of the edge e_i .

We will thus define the *weighted incidence tensor* $\mathcal{D} = (\delta_{\beta,j}^{\alpha,i})$, $\alpha, \beta = 1, \dots, n$, $i, j = 1, \dots, m$, as follows

$$\delta_{\beta j}^{\alpha i} = c_{ij}(v_\beta) \iota_{\beta j}^{\alpha i}, \tag{4.1}$$

where the function c is a smooth enough function that we will specify later on.

Eventually, we consider two different type of boundary conditions. In particular we will assume that the vertices v_α , $\alpha = 1, \dots, n_0$, $1 \leq n_0 \leq n$, have some non-local static generalized *Kirchhoff type* conditions, whereas we equip the remaining nodes v_α , $\alpha = n_0 + 1, \dots, n$, with some non-local dynamic boundary conditions.

Let us thus consider the following diffusion problem on a finite and connected graph \mathbb{G} ,

$$\begin{cases} \dot{u}_j(t, x) = \sum_{i=1}^m (c_{ij}u'_i)'(t, x) + \sum_{i=1}^m p_{ij}u_i(t, x), & t \geq 0, x \in (0, 1), j = 1, \dots, m, \\ u_j(t, v_\alpha) = u_l(t, v_\alpha) =: d_\alpha^u(t), & t \geq 0, l, j \in \Gamma(v_\alpha), j = 1, \dots, m, \\ \sum_{\beta=1}^n b_{\alpha\beta}d_\beta^u(t) = \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{\alpha i} u'_j(t, v_\alpha), & t \geq 0, \alpha = n_0 + 1, \dots, n, \\ \dot{d}_\alpha^u(t) = - \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{\alpha i} u'_j(t, v_\beta) + \sum_{\beta=1}^n b_{\alpha\beta}d_\beta^u(t), & t \geq 0, \alpha = 1, \dots, n_0, \\ u_j(0, x) = u_j^0(x), & x \in (0, 1), j = 1, \dots, m, \\ d_i^u(0) = d_i^0, & i = 1, \dots, n_0, \end{cases} \quad (4.2)$$

where we have denoted by $\dot{u}(t, x)$ the time derivative of the unknown function u , whereas $u'(t, x)$ denotes its space-derivative.

Moreover, for $x \in [0, 1]$, $t \in [0, T]$, we defined the unknown functions $u(t, x)$ and $d^u(t)$, by

$$u(t, x) = (u_1(t, x), \dots, u_m(t, x))^T, \quad d^u(t) = (d_1^u(t), \dots, d_{n_0}^u(t), d_{n_0+1}^u(t), \dots, d_n^u(t))^T,$$

and we consider the $n \times n$ matrix $B = (b_{\alpha\beta})_{\alpha,\beta=1,\dots,n}$, defined as $B := B_1 + B_2$, B_1 being the $n \times n$ matrix defined as

$$B_1 := \begin{pmatrix} b_{1,1} & \dots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{n_0,1} & \dots & b_{n_0,n} \\ 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}, \quad (4.3)$$

while B_2 is the $n \times n$ matrix defined as

$$B_2 := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \\ b_{n_0+1,1} & \dots & b_{n_0+1,n} \\ \vdots & \ddots & \vdots \\ b_{n,1} & \dots & b_{n,n} \end{pmatrix}.$$

If not stated otherwise, we use $\langle \cdot, \cdot \rangle_m$, resp. $|\cdot|_m$, to denote the standard scalar product, resp. the related norm, in \mathbb{R}^m .

Throughout the paper we will assume the following assumptions to hold:

Hypothesis 4.1.2. **(i)** for any $i, j = 1, \dots, m$, we have that $c_{ij}(x) \in C^1(0, 1)$, also assuming that the matrix $C := (c_{ij})_{i,j=1,\dots,m}$ is positive definite, uniformly in $[0, 1]$, namely for any $x \in [0, 1]$, $\bar{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$, there exists $\lambda^C > 0$ such that

$$\langle C(x)\bar{y}, \bar{y} \rangle_m = \sum_{i,j=1}^m c_{ij}(x)y_jy_i \geq \lambda^C |\bar{y}|_m^2; \quad (4.4)$$

(ii) for any $i, j = 1, \dots, m$ we have that $p_{ij}(x) \in L^\infty(0, 1)$, also assuming that the matrix $P := (p_{ij})_{i,j=1,\dots,m}$ is negative semi-definite, uniformly in $[0, 1]$, namely for any $x \in [0, 1]$, $\bar{y} = (y_1, \dots, y_m) \in \mathbb{R}^m$, there exists $\lambda^P \geq 0$ such that

$$\langle P(x)\bar{y}, \bar{y} \rangle_m = \sum_{i,j=1}^m p_{ij}(x)y_j y_i \leq -\lambda^P |\bar{y}|_m^2; \quad (4.5)$$

The abstract setting

In what follows we introduce the abstract setting which allows us to rewrite equation (4.2) as an abstract Cauchy problem. In particular, let us first consider the following spaces

$$X^2 := (L^2([0, 1]))^m, \quad \text{resp. } \mathbb{R}^n,$$

equipped with the standard inner products, denoted by $\langle \cdot, \cdot \rangle_2$, resp. $\langle \cdot \rangle_n$, and norms denoted by $|\cdot|_2$, resp. $|\cdot|_n$. Then, we define the product Hilbert space $X^{t,x} := X^2 \times \mathbb{R}^n$, equipped with the inner product

$$\left\langle \begin{pmatrix} u \\ d^u \end{pmatrix}, \begin{pmatrix} v \\ d^v \end{pmatrix} \right\rangle_{X^{t,x}} := \sum_{j=1}^m \int_0^1 u_j(x)v_j(x)dx + \sum_{\alpha=1}^n d_\alpha^u d_\alpha^v,$$

where $u, v \in X^2$, $d^u, d^v \in \mathbb{R}^n$, with associated norm denoted by $|\cdot|_{X^{t,x}}$. Analogously, we define the Banach space

$$X^p := (L^p([0, 1]))^m, \quad \mathcal{X}^p := X^p \times \mathbb{R}^n, \quad p \in [1, \infty],$$

Remark 4.1.3. In [DPZ11, MR07] the authors consider a diffusion problem similar to the one represented by eq. (4.2), and where the boundary conditions depend on some phenomenological positive constants μ and ν . For ease of notation, we have dropped latter constants in the present work without loose of generality. In fact, our results remain valid also when previous constants are explicitly considered, since it is sufficient to consider some weighted spaces of the form

$$X_\mu^2 := \prod_{j=1}^m L^2([0, 1]; \mu_j dx), \quad \mathbb{R}_\nu^n := \prod_{\alpha=1}^n \mathbb{R} \frac{1}{\nu_\alpha}.$$

Recalling the definition of *incidence matrix* Φ given in Sec. 4.2.1, we introduce the associated *Kirchhoff operators* $\Phi_\delta^+, \Phi_\delta^- : (H^1(0, 1))^m \rightarrow \mathbb{R}^n$, which are defined as follows

$$\begin{aligned} \Phi_\delta^+ u' &:= \left(\sum_{i,j=1}^m \sum_{\alpha=1}^n +\delta_{1j}^{\alpha i} u'_i(v_1), \dots, \sum_{i,j=1}^m \sum_{\alpha=1}^n +\delta_{nj}^{\alpha i} u'_i(v_n) \right)^T, \\ \Phi_\delta^- u' &:= \left(\sum_{i,j=1}^m \sum_{\alpha=1}^n -\delta_{1j}^{\alpha i} u'_i(v_1), \dots, \sum_{i,j=1}^m \sum_{\alpha=1}^n -\delta_{nj}^{\alpha i} u'_i(v_n) \right)^T, \end{aligned}$$

where the notation $+\delta$, resp. $-\delta$, means that ι in equation (4.1) belongs to \mathcal{I}^+ , resp. \mathcal{I}^- , namely

$$+\delta_{\beta j}^{\alpha i} = \begin{cases} c_{ij}(v_\beta) \iota_{\beta j}^{\alpha i} & \text{if } \iota_{\beta j}^{\alpha i} \in \mathcal{I}^+, \\ 0 & \text{otherwise,} \end{cases}, \quad -\delta_{\beta j}^{\alpha i} = \begin{cases} c_{ij}(v_\beta) \iota_{\beta j}^{\alpha i} & \text{if } \iota_{\beta j}^{\alpha i} \in \mathcal{I}^-, \\ 0 & \text{otherwise.} \end{cases}$$

Let us then introduce the differential operator $(A, D(A))$ as

$$Au = \begin{pmatrix} (c_{1,1}u_1') + p_{1,1}u_1 & \cdots & (c_{1,m}u_1') + p_{1,m}u_m \\ \vdots & \ddots & \vdots \\ (c_{m,1}u_1') + p_{m,1}u_1 & \cdots & (c_{m,m}u_m') + p_{m,m}u_m \end{pmatrix},$$

which has domain defined as

$$D(A) = \left\{ u \in (H^2(0,1))^m : \exists d^u(t) \in \mathbb{R}^n \text{ s.t. } (\Phi^+)^T d^u(t) = u(0), \right. \\ \left. (\Phi^-)^T d^u(t) = u(1), \Phi_\delta^+ u'(0) - \Phi_\delta^- u'(1) = B_2 d^u(t) \right\}.$$

Then, we define the operator matrix

$$\mathcal{A} = \begin{pmatrix} A & 0 \\ C & B_1 \end{pmatrix}, \quad (4.6)$$

where C represents the *feedback operator* acting from $D(C) := D(A)$ to \mathbb{R}^n and defined as follows

$$Cu := \left(- \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{1i} u_j'(v_1), \dots, - \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{n_0 i} u_j'(v_{n_0}), 0, \dots, 0 \right)^T,$$

and

$$D(\mathcal{A}) = \left\{ \begin{pmatrix} u \\ d^u \end{pmatrix} \in D(A) \times \mathbb{R}^n : u_i(v_\alpha) = d_\alpha^u, \quad \forall i \in \Gamma(v_\alpha), \alpha = 1, \dots, n \right\}.$$

Exploiting previous definitions, we can rewrite equation (4.2) as the following abstract infinite dimensional equation stated on the Hilbert space $X^{t,x}$

$$\begin{cases} \dot{\mathbf{u}}(t) = \mathcal{A}\mathbf{u}(t), & t \geq 0, \\ \mathbf{u}(0) = \mathbf{u}_0, \end{cases} \quad (4.7)$$

where

$$\mathbf{u} := (u, d^u)^T = (u_1, \dots, u_m, d_1^u, \dots, d_{n_0}^u, d_{n_0+1}^u, \dots, d_n^u)^T \in X^{t,x},$$

and

$$\mathbf{u}_0 := (u_1(0, x), \dots, u_m(0, x), d_1^u(0), \dots, d_{n_0}^u(0), 0, \dots, 0)^T \in X^{t,x}.$$

Then we introduce the sesquilinear form $\mathfrak{a} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$, where the space \mathcal{V} is a suitable subspace of $X^{t,x}$, see below, defined as

$$\begin{aligned} \mathfrak{a}(\mathbf{u}, \mathbf{v}) &:= \langle Cu', v' \rangle_2 - \langle Pu, v \rangle_2 - \langle B_1 d^u, d^v \rangle_n - \langle B_2 d^u, d^v \rangle_n = \\ &= \sum_{i,j=1}^m \int_0^1 (c_{i,j}(x)u_j'(x)v_i'(x) - p_{i,j}(x)u_j(x)v_i(x)) dx - \sum_{\alpha,\beta=1}^n b_{\alpha\beta} d_\alpha^u d_\beta^v, \end{aligned} \quad (4.8)$$

for any $\mathbf{u}, \mathbf{v} \in X^{t,x}$.

In particular, the subspace \mathcal{V} , domain of the form \mathfrak{a} , is defined by the following lemma

Lemma 4.1.4. *Let us consider the linear subspace*

$$\mathcal{V} := \left\{ \begin{pmatrix} u \\ d^u \end{pmatrix} \in (H^1(0,1))^m \times \mathbb{R}^n : u_i(\mathbf{v}_\alpha) = d_\alpha^u, \forall i \in \Gamma(\mathbf{v}_\alpha), \alpha = 1, \dots, n \right\},$$

then \mathcal{V} is densely and compactly embedded in $X^{t,x}$. In particular \mathcal{V} is a Hilbert space equipped with the scalar product

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{V}} := \sum_{j=1}^m \int_0^1 (u'_j(x)v'_j(x) + u_j(x)v_j(x)) dx + \sum_{\alpha=1}^n d_\alpha^u d_\alpha^v. \quad (4.9)$$

The corresponding norm will be denoted by $|\cdot|_{\mathcal{V}}$.

Proof. See, e.g., [CMN08, Lemma 3.1] or [MR07, Lemma 3.1]. \square

Remark 4.1.5. One of the main advantages in using the theory of sesquilinear form is that, under suitable assumptions, a sesquilinear form \mathbf{a} can be uniquely associated to an infinitesimal generator of an analytic strongly continuous semigroup. In particular, if we prove that the form \mathbf{a} satisfies some regularity conditions, then we also have a corresponding regularity for the associated semigroup. In the next proposition we gather several properties satisfied by the form \mathbf{a} defined in (4.8). We would like to underline that such results have already been proved separately, and under a different setting, in different works, see, e.g., [BMZ⁺08, DPZ11, Mug10, MR07] and reference therein. Nevertheless, for the sake of completeness, we will provide for the latter a sketch of their proofs.

Proposition 4.1.6.

(i) *If Assumptions 4.2.2 hold, then the form $\mathbf{a} : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{R}$ defined in (4.8) is:*

- *continuous, i.e. it exists $M > 0$, such that*

$$|\mathbf{a}(\mathbf{u}, \mathbf{v})| \leq M |\mathbf{u}|_{\mathcal{V}} |\mathbf{v}|_{\mathcal{V}}; \quad (4.10)$$

- *$X^{t,x}$ -elliptic, i.e. there exist $\lambda > 0$ and $\omega \in \mathbb{R}$, such that*

$$\mathbf{a}(\mathbf{u}, \mathbf{u}) \geq \lambda |\mathbf{u}|_{\mathcal{V}}^2 - \omega |\mathbf{u}|_{X^{t,x}}^2; \quad (4.11)$$

- *closed, i.e. \mathcal{V} is complete with respect to the following norm*

$$\|\mathbf{u}\|_{\mathbf{a}}^2 := \mathbf{a}(\mathbf{u}, \mathbf{u}) + \|\mathbf{u}\|_{X^{t,x}}^2; \quad (4.12)$$

(ii) *If Assumptions 4.2.2 hold and the matrix B is negative defined, i.e. there exists $\mu > 0$ such that*

$$\langle B\bar{y}, \bar{y} \rangle_n \leq -\mu |\bar{y}|_n^2, \forall \bar{y} \in \mathbb{R}^n,$$

then \mathbf{a} is coercive, namely it is $X^{t,x}$ -elliptic with $\omega = 0$, hence

$$\mathbf{a}(\mathbf{u}, \mathbf{u}) \geq \lambda |\mathbf{u}|_{\mathcal{V}}^2; \quad (4.13)$$

(iii) *If Assumptions 4.2.2 hold and the matrices C , P and B are all symmetric, then the form \mathbf{a} is symmetric as well.*

Proof. (i) To simplify notations, let us define the following quantities

$$\begin{aligned}\bar{c} &:= \min_{x \in [0,1]} \sum_{i,j=1}^m c_{i,j}(x), & \bar{C} &:= \max_{1 \leq j \leq m} \sum_{i,j=1}^m c_{i,j}(x), \\ \bar{p} &:= \min_{1 \leq j \leq m} \sum_{i,j=1}^m (1 - p_{i,j}(x)), & \bar{P} &:= \max_{1 \leq j \leq m} \sum_{i,j=1}^m (1 - p_{i,j}(x)), \\ \bar{b} &:= \min_{i,l} b_{i,l}, & \bar{B} &:= \sum_{\alpha,\beta=1}^n b_{\alpha,\beta}.\end{aligned}$$

Proceeding as in [MR07, Lemma 3.2], we have that \mathcal{V} , equipped with the inner product defined in equation (4.9), is a Hilbert space, moreover it is a closed subspace of $(H^1(0,1))^m \times \mathbb{R}^n$. From the continuous embedding of $H^1(0,1)$ into $C(0,1)$, see, e.g., [MR07, Lemma 3.2], we obtain

$$|d_i^u| \leq \max_{1 \leq j \leq m} \max_{x \in [0,1]} |u_j(x)| \leq \max_{1 \leq j \leq m} |u_j|_{H^1(0,1)} \leq \sum_{j=1}^m |u_j|_{H^1(0,1)},$$

hence the norm defined in eq. (4.9) is equivalent to

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathcal{V}} := \sum_{j=1}^m (u_j'(x)v_j'(x) + u_j(x)v_j(x))\mu_j dx, \quad (4.14)$$

and, again by [MR07, Lemma 3.2], it also follows that there exists $K > 0$ such that

$$|d_i^u| \leq K|\mathbf{u}|_{\mathcal{V}}, \quad i = 1, \dots, n,$$

then, defining

$$\tilde{c} := \min\{\bar{c}, \bar{p}\}, \quad \tilde{C} := \max\{\bar{C}, (1 - \bar{B})K^2, \bar{P}\},$$

we have that the norm generated by \mathcal{V} is equivalent to the one generated by \mathbf{a} , which, from the completeness of \mathcal{V} , implies the closure of \mathbf{a} . In what follows the Hilbert space \mathcal{V} will be equipped with the inner product (4.14) and the corresponding norm.

Concerning the continuity of \mathbf{a} , from assumptions 4.2.2, we have

$$\begin{aligned}|\mathbf{a}(\mathbf{u}, \mathbf{v})| &\leq \sum_{i,j=1}^m \int_0^1 (|c_{i,j}(x)u_i'(x)v_j'(x)| + |p_{i,j}(x)u_i(x)v_j(x)|) dx + \\ &\quad - \sum_{\alpha,\beta=1}^n b_{\alpha,\beta} |d_\alpha^u| |d_\beta^v| \leq \\ &\leq 2L \sum_{i,j=1}^m \langle u_i, v_j \rangle_{H^1((0,1);\mu_j dx)} - K^2 \bar{B} |\mathbf{u}|_{\mathcal{V}} |\mathbf{v}|_{\mathcal{V}} \leq \\ &\leq 2L \left(\sum_{j=1}^m |u_j|_{H^1((0,1);\mu_j dx)}^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^m |v_j|_{H^1((0,1);\mu_j dx)}^2 \right)^{\frac{1}{2}} + \\ &\quad - K^2 \bar{B} |\mathbf{u}|_{\mathcal{V}} |\mathbf{v}|_{\mathcal{V}} = \\ &= (2L - \bar{B}K^2) |\mathbf{u}|_{\mathcal{V}} |\mathbf{v}|_{\mathcal{V}} \leq M |\mathbf{u}|_{\mathcal{V}} |\mathbf{v}|_{\mathcal{V}}.\end{aligned}$$

where L , resp. M , is defined by $L := \max\{\bar{C}, \bar{P}\}$, resp. by $M := (2L - K^2\bar{B})$. Moreover assumptions 4.2.2 also implies that the form

$$\mathbf{a}_1 := \langle Cu', v' \rangle_2 - \langle Pu, v \rangle_2,$$

is $X^{t,x}$ -elliptic. In fact, by [Bur98, Cor. 4.11], see also [CMN08], we have, for some constant $K > 0$, that the following inequality holds

$$\max_{x \in [0,1]} u(x) \leq K \|u\|_{L^2}^{\frac{1}{2}} \|u\|_{H^1}^{\frac{1}{2}},$$

hence, introducing $\mathbf{a}_2 := -\langle Bd^u, d^u \rangle$, we can decompose \mathbf{a} as $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$, so that the claim follows from [Mug07, Lemma 2.1], see also [CMN08, Th. 2.3] and [CM08, Lemma 2.1, Cor. 2.2].

- (ii) from assumptions 4.2.2 and denoting by μ^C , resp. μ^P , the constant in equation (4.4), resp. equation (4.5), we have that

$$\begin{aligned} \mathbf{a}(\mathbf{u}, \mathbf{u}) &= \int_0^1 (\langle C(x)u'(x), u'(x) \rangle_m - \langle P(x)u(x), u(x) \rangle_m) dx - \langle Bd^u, d^u \rangle \\ &\geq \int_0^1 (\mu^C |u(x)|_m^2 + \mu^P |u(x)|_m^2) dx + \mu^B |d^u|^2 \geq \lambda \|\mathbf{u}\|_{\mathcal{V}}. \end{aligned} \quad (4.15)$$

- (iii) it immediately follows from the very definition of \mathbf{a} , see eq. (4.8). □

In force of Proposition 4.1.6, we recall the following result, see [MR07, Lemma 3.3].

Proposition 4.1.7. *The operator associated with the form \mathbf{a} defined in (4.8) is the operator $(\mathcal{A}, D(\mathcal{A}))$ defined in equation (4.45).*

Proof. See [MR07, Lemma 3.3] or [Ouh09, Prop. 1.51, Th. 1.52], □

We end the present subsection characterizing the semigroup generated by the operator $(\mathcal{A}, D(\mathcal{A}))$ defined in equation (4.45). Such result will be used later on to prove the *Gaussian bound*, see Sec. 4.1.1 below.

Proposition 4.1.8. *If assumptions 4.2.2 hold, then the operator associated with the form \mathbf{a} defined in equation (4.8), is densely defined, sectorial and resolvent compact, hence it generates an analytic and compact C_0 -semigroup $\mathcal{T}(t)$. We also have the following properties for the semigroup*

- (i) *if the matrix B is negative definite, then the semigroup is uniformly exponentially stable;*
- (ii) *if the matrices C , P and B are symmetric, then the semigroup is self-adjoint;*
- (iii) *if the matrices C and P are diagonal, and the matrix B has entries that are positive off-diagonal and it also satisfies*

$$b_{\alpha\alpha} + \sum_{\beta \neq \alpha} b_{\alpha\beta} \leq 0, \quad \text{for any } \alpha = 1, \dots, n,$$

then the semigroup is positive and \mathcal{X}^∞ -contractive in the sense of [Ouh09, Ch. 2].

Proof. The main claim follows exploiting Lemma 4.1.4, Proposition 4.1.6, Proposition 4.1.7 and [Dav90, Th. 1.2.1]. Concerning (i) the uniformly exponential stability it is enough to see that the shifted form $\lambda - \mathfrak{a}(\cdot, \cdot)$ is accretive, whereas point (ii) follows from the fact that the form \mathfrak{a} is symmetric, while point (iii) follows from [CMN08, Th. 2.3] and [Mug10, Cor. 3.4]. \square

Gaussian bounds

In what follows we state our main result concerning Gaussian estimates and, in order to achieve the result, we require assumptions stated in (4.2.2) as well as the following

Hypothesis 4.1.9. The matrices C and P are diagonal and B has entries that are positive off-diagonal and it satisfies, for any $\alpha = 1, \dots, n$,

$$b_{\alpha\alpha} + \sum_{\beta \neq \alpha} |b_{\alpha\beta}| < 0,$$

Under the current assumptions we have that the semigroup \mathcal{T} generated by \mathcal{A} is analytic, compact, positive, \mathcal{X}^∞ -contractive and uniformly exponential stable on $X^{t,x}$, see Proposition 4.1.8.

Let us also recall, see [MR07, Lemma 5.2]. the following lemma,

Lemma 4.1.10. *Let us consider a set of functions $u_j : [0, 1] \rightarrow \mathbb{R}$, $j = 1, \dots, m$, and let us then define the map $Uu : [0, m] \rightarrow \mathbb{R}$ by*

$$Uu(x) := u_j(x - j + 1), \quad \text{if } x \in (j - 1, j),$$

then the map U is a one-to-one map from $(L^2(0, 1))^m$ onto $L^2(0, m)$. Also U is an isometry if we consider $(L^2(0, 1))^m$ with the norm

$$|u|_{(L^2(0,1))^m} = \left(\sum_{j=1}^m |u_j|_{L^2(0,1)} \right)^{\frac{1}{2}}.$$

We then consider the product space $X^{t,x} := (L^2(0, 1))^m \times \mathbb{R}^n$, hence, in virtue of Lemma 4.1.10, defining $\Omega := (0, m) \times (0, n)$, and

$$\mu := dx \oplus \delta_1 \oplus \dots \oplus \delta_n,$$

where δ_{x_0} is the Dirac measure centred at x_0 , then we have that the map $U : X^{t,x} \rightarrow L^2(\Omega, \mu)$ is an isomorphism. Since we have required assumptions 4.2.2 to hold, then we know that the operator associated with the form \mathfrak{a} , see eq. (4.8), generates an analytic and compact C_0 -semigroup, which we have defined as $\mathcal{T}(t)$, moreover we have

Theorem 4.1.11. *The semigroup $\mathcal{T}(t)$, acting on the space $X^{t,x}$ and associated to \mathfrak{a} , is ultracontractive, namely there exists a constant $M > 0$ such that*

$$\|\mathcal{T}(t)\mathbf{u}\|_{\mathcal{X}^\infty} \leq Mt^{-\frac{1}{4}} \|\mathbf{u}\|_{X^{t,x}}, \quad t \in [0, T], \mathbf{u} \in X^{t,x}. \quad (4.16)$$

Proof. By the *Nash-type inequality* for weighted L^p -space, we have that there exists a constant $M_1 > 0$ such that

$$\|f\|_{L^2(\Omega, \mu)} \leq M_1 \left(\|f'\|_{L^2(\Omega, \mu)} + \|f\|_{L^2(\Omega, \mu)} \right)^{\frac{1}{3}} \|f\|_{L^2(\Omega, \mu)}^{\frac{2}{3}} \leq M_1 \|f\|_{H^1(\Omega, \mu)}^{\frac{1}{3}} \|f\|_{L^2(\Omega, \mu)}^{\frac{2}{3}},$$

hence, for $\mathbf{u} \in V_0$, we have

$$\begin{aligned} \|\mathbf{u}\|_{X^{t,x}}^2 &= \sum_{j=1}^m \|u_j\|_2^2 + \sum_{i=1}^n |d_i^u| \leq M_1^2 \sum_{j=1}^m \|u_j\|_{H^1}^{\frac{2}{3}} \|u_j\|_2^{\frac{4}{3}} + \sum_{i=1}^n |d_i^u|, \leq \\ &\leq M_1^2 \left(\sum_{j=1}^m \|u_j\|_{H^1}^2 + \sum_{i=1}^n |d_i^u| \right)^{\frac{1}{3}} \left(\sum_{j=1}^m \|u_j\|_{L^1}^2 + \sum_{i=1}^n |d_i^u| \right)^{\frac{1}{3}} \leq \\ &\leq M_2 \|\mathbf{u}\|_{V_0}^{\frac{2}{3}} \|\mathbf{u}\|_{\mathcal{X}^1}^{\frac{4}{3}}, \end{aligned}$$

and the claim follows from the equivalence between the norms $\|\cdot\|_{\mathfrak{a}}$ and $\|\cdot\|_{V_0}$, as have been shown in Prop. 4.1.6 and [Ouh09, Lemma 5.2]. \square

Moreover, Th. 4.1.11 implies the following

Corollary 4.1.12. *The semigroup $\mathcal{T}(t)$ on $X^{t,x}$ satisfies*

$$\|\mathcal{T}(t)\mathbf{u}\|_{\mathcal{X}^\infty} \leq M \left(\frac{1-t\omega}{t} \right)^{\frac{1}{4}} e^{1+t\omega} \|\mathbf{u}\|_{X^{t,x}},$$

where $\omega < 0$ is the spectral bound of the semigroup $\mathcal{T}(t)$.

Proof. The claim follows from Prop. 4.1.8, Th. 4.1.11 and [Ouh09, Lemma 6.5]. \square

Besides the ultracontractivity of $\mathcal{T}(t)$ together with Cor. 4.1.12, implies that the semigroup has an integral Kernel, see [Dav90, Lemma 2.1.2.]. More precisely let us denote by $\tilde{\mathcal{T}}(t) := U^{-1}\mathcal{T}(t)U$ the similar semigroup, see, e.g., [EN00b], acting on $L^2(\Omega, \mu)$, being U the isomorphism introduced above. Then, Lemma [Dav90, Lemma 2.1.2] gives us that the action of $\left(\tilde{\mathcal{T}}(t) \right)_{t \geq 0}$, reads as follow

$$\left(\tilde{\mathcal{T}}(t)g \right) (\cdot) = \int_{\Omega} K_t(\cdot, y)g(y)\mu(dy), \quad g \in L^2(\Omega, \mu),$$

for a suitable kernel $K_t \in L^\infty(\Omega \times \Omega)$. Besides, we can rewrite eq. (4.16) as follows

$$\|\mathcal{T}(t)\mathbf{u}\|_{\mathcal{X}^\infty} \leq e^{\kappa(t)} \|\mathbf{u}\|_{X^{t,x}}, \quad t \in [0, T], \quad \mathbf{u} \in X^{t,x},$$

where

$$\kappa(t) := \log M - \frac{1}{4} \log t.$$

Then, applying [Dav90, Th. 2.2.3], we can derive the following logarithmic Sobolev inequality

$$\int_{\Omega} \tilde{\mathbf{u}} \log \tilde{\mathbf{u}} dx \leq \epsilon \mathfrak{a}(\mathbf{u}, \mathbf{u}) + \kappa(\epsilon) \|\mathbf{u}\|_{X^{t,x}}^2 + \|\mathbf{u}\|_{X^{t,x}}^2 \log \|\mathbf{u}\|_{X^{t,x}}, \quad (4.17)$$

for any $\mathbf{u} \geq 0$, $\mathbf{u} \in V_0$ and $\epsilon > 0$, and where $\tilde{\mathbf{u}} \in L^2(\Omega, \mu)$ denotes the function isometric to \mathbf{u} under the isomorphism U . Eventually, inequality (4.17) implies the next result

Theorem 4.1.13. *The Gaussian upper bound*

$$0 \leq K_t(x, y) \leq c_\delta t^{-\frac{1}{2}} e^{-\frac{|x-y|^2}{\sigma t}}, \quad (4.18)$$

holds for the heat kernel K_t introduced above, such that it holds

$$[\mathcal{T}(t)g](x) = \int_{\Omega} K_t(x, y)g(y)\mu(dy), \quad y \in L^2(\Omega, \mu).$$

Proof. The claim follows from [Dav90, Th. 3.2.7], taking into account the logarithmic Sobolev inequality (4.17), see, e.g., [Mug10, Th. 4.8] and [DPZ11]. \square

Exploiting Th. 4.1.13 it is also possible to prove the existence of a *mild solution*, in a suitable sense, to equation (4.2) perturbed by a multiplicative Gaussian noise. Before state latter result, let us denote by $\mathcal{L}_2(X^{t,x})$ the class of *Hilbert-Schmidt* operator from $X^{t,x}$ to $X^{t,x}$, while $|\cdot|_{HS}$ denotes the standard *Hilbert-Schmidt* norm. We refer the reader to, e.g., [DPZ14, Appendix. C], for a dense *résumé* of the main properties of *Hilbert-Schmidt* operators.

Proposition 4.1.14. *Let assumptions 4.2.2-4.1.9 hold, then, for any $t > 0$, the semigroup $\mathcal{T}(t) \in \mathcal{L}_2(X^{t,x})$, moreover there exists $M > 0$ such that*

$$|\mathcal{T}(t)|_{HS} \leq Mt^{-\frac{1}{4}}.$$

Proof. Since

$$|\mathcal{T}(t)|_{HS} = |\tilde{\mathcal{T}}(t)|_{HS} = |K_t|_{L^2(\Omega \times \Omega)},$$

where K_t is the kernel defined in equation (4.18), then, by Th. 4.1.13, eq. (4.18), see also [DPZ11, Cor.2], we obtain the existence of a constant $C > 0$ such that, $\forall t \in [0, T]$, it holds

$$|\mathcal{T}(t)|_{HS}^2 = \int_{\Omega \times \Omega} |K_t(x, y)|^2 dx dy \leq C\sqrt{2\pi\sigma t^{-1}},$$

which implies the existence of a positive constant M such that, $\forall t \in [0, T]$, the following hold

$$|\mathcal{T}(t)|_{HS} \leq Mt^{-\frac{1}{4}}.$$

\square

4.1.2 The perturbed stochastic problem

In the present section we focus our attention on the problem (4.2) by perturbing it with multiplicative Gaussian noise. Let us first consider the following complete, filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, with respect to which, we state the following system

$$\begin{cases} \dot{u}_j(t, x) = (c_j u_j')'(t, x) + p_i u_i(t, x) + g_j(t, x, u_j(t, x)) \dot{W}_j^1(t, x), \\ \text{for } t \geq 0, x \in (0, 1), j = 1, \dots, m, \\ u_j(t, v_\alpha) = u_l(t, v_\alpha) =: d_\alpha^u(t), \quad t \geq 0, l, j \in \Gamma(v_\alpha), j = 1, \dots, m, \\ \sum_{\beta=1}^n b_{\alpha\beta} d_\beta^u(t) = \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{\alpha i} u_j'(t, v_\alpha), \quad t \geq 0, \alpha = n_0 + 1, \dots, n, \\ \dot{d}_\alpha^u(t) = - \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{\alpha i} u_j'(t, v_\beta) + \sum_{\beta=1}^n b_{\alpha\beta} d_\beta^u(t) + \tilde{g}_\alpha(t, d_\alpha^u(t)) \dot{W}_\alpha^2(t, v_\alpha), \\ \text{for } t \geq 0, \alpha = 1, \dots, n_0, \\ u_j(0, x) = u_j^0(x), \quad x \in (0, 1), j = 1, \dots, m, \\ d_i^u(0) = d_i^0, \quad i = 1, \dots, n_0, \end{cases} \quad (4.19)$$

where, for every $(j, \alpha) \in \{1, \dots, m\} \times \{1, \dots, n_0\}$, W_j^1 and W_α^2 are independent Wiener processes adapted to \mathcal{F}_t- , while \dot{W} is the formal time derivative. In particular, for every $j = 1, \dots, m$, W_j^1 is a space time Wiener process with values in $L^2(0, 1)$. Then, we denote by $W^1 := (W_1^1, \dots, W_m^1)$, a space time Wiener process with values in the product space $X^2 := (L^2(0, 1))^m$. Analogously, for every $\alpha = 1, \dots, n$, W_α^2 is a space time Wiener process taking values in \mathbb{R} , hence we denote by $W^2 := (W_1^2, \dots, W_n^2)$ the standard Wiener process with values in \mathbb{R}^n . Consequently, $W := (W^1, W^2)$ indicates the standard space time Wiener process with values in $X^{t,x} := X^2 \times \mathbb{R}^n$, being $(\mathcal{F}_t)_{t \in [0, T]}$ the natural filtration generated by W , augmented by all \mathbb{P} -null sets of \mathcal{F}_T .

Besides assumptions 4.2.2 and 4.1.9 we will also assume the following to hold.

Hypothesis 4.1.15.

- (i) For every $j = 1, \dots, m$, the functions $g_j : [0, T] \times [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$, are measurable, bounded and uniformly Lipschitz in the third component, namely there exist constants $C_j > 0$ and K_j such that, for any $(t, x, y_1) \in [0, T] \times [0, 1] \times \mathbb{R}$ and $(t, x, y_2) \in [0, T] \times [0, 1] \times \mathbb{R}$, the following holds

$$|g_j(t, x, y_1)| \leq C_j, \quad |g_j(t, x, y_1) - g_j(t, x, y_2)| \leq K_j |y_1 - y_2|;$$

- (ii) For every $\alpha = 1, \dots, n_0$, the functions $\tilde{g}_\alpha : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$, are measurable, bounded and uniformly Lipschitz with respect to the second component, namely there exist constants $C_\alpha > 0$ and K_α such that, for any $(t, y_1) \in [0, T] \times \mathbb{R}$ and $(t, y_2) \in [0, T] \times \mathbb{R}$, the following holds

$$|\tilde{g}_\alpha(t, y_1)| \leq C_\alpha, \quad |\tilde{g}_\alpha(t, y_1) - \tilde{g}_\alpha(t, y_2)| \leq K_\alpha |y_1 - y_2|.$$

With the help of the notations just introduced, see also Sec. 5.2, the problem (4.56) can be rewritten as an abstract infinite dimensional Cauchy problem of the form

$$\begin{cases} d\mathbf{u}(t) = \mathcal{A}\mathbf{u}(t)dt + G(t, \mathbf{u}(t))dW(t), & t \geq 0, \\ \mathbf{u}(0) = \mathbf{u}_0 \in X^{t,x}, \end{cases} \quad (4.20)$$

where \mathcal{A} is the operator introduced in (4.45), while $G : [0, T] \times X^{t,x} \rightarrow \mathcal{L}(X^{t,x})$, $\mathcal{L}(X^{t,x})$ being the space of linear and bounded operator from $X^{t,x}$ to $X^{t,x}$ equipped with standard operator norm $|\cdot|_{\mathcal{L}}$, is defined as

$$G(t, \mathbf{u})\mathbf{v} = (\sigma_1(t, u)v, \sigma_2(t, y)z)^T, \quad \mathbf{u} = (u, y), \quad \mathbf{v} = (v, z) \in X^{t,x}, \quad (4.21)$$

with

$$\begin{aligned} (\sigma_1(t, u)v)(x) &= (g_1(t, x, u_1(t, x)), \dots, g_m(t, x, u_m(t, x)))^T, \\ \sigma_2(t, y)z &= (\tilde{g}_1(t, y_1)z_1, \dots, \tilde{g}_{n_0}(t, y_{n_0})z_{n_0}, 0, \dots, 0)^T. \end{aligned}$$

It is worth to mention that, in order to guarantee the existence and uniqueness of a *mild solution* to equation (4.57), in a suitable sense to be introduced in a while, we have to require the stronger property that $G : [0, T] \times X^{t,x} \rightarrow \mathcal{L}_2(X^{t,x})$, where $\mathcal{L}_2(X^{t,x})$ is the space of *Hilbert-Schmidt* operator from $X^{t,x}$ into itself equipped with standard *Hilbert-Schmidt norm* denoted by $|\cdot|_{HS}$, see, e.g., [DPZ14, Appendix C]. Nevertheless, by Prop. 4.1.14, we can show that the semigroup $\mathcal{T}(t)$ is *Hilbert-Schmidt*, and that to have a unique solution in a *mild sense* we can weaken the condition on G requiring it to take values in $\mathcal{L}(X^{t,x})$.

The aforementioned *mild solution* to equation (4.57), is intended in the following sense

Definition 4.1.16. We will say that \mathbf{u} is a *mild solution* to equation (4.57), if it is a mean square continuous $X^{t,x}$ -valued process adapted to the filtration generated by W , such that for any $t \geq 0$ we have that $\mathbf{u} \in L^2(\Omega, C([0, T]; X^{t,x}))$, and it holds

$$\mathbf{u}(t) = \mathcal{T}(t)\mathbf{u}_0 + \int_0^t \mathcal{T}(t-s)G(s, \mathbf{u}(s))dW(s), \quad t \geq 0. \quad (4.22)$$

We thus have the following.

Proposition 4.1.17. *Let assumptions 4.2.2-4.1.9-4.2.8 hold, then the map $G : [0, T] \times X^{t,x} \rightarrow \mathcal{L}(X^{t,x})$ defined in eq. (4.58) satisfies:*

(i) *for any $\mathbf{u} \in X^{t,x}$, the map $G(\cdot, \cdot)\mathbf{u} : [0, T] \times X^{t,x} \rightarrow X^{t,x}$ is measurable;*

(ii) *$\mathcal{T}(t)G(s, \mathbf{u}) \in \mathcal{L}_2(X^{t,x})$, for any $t > 0$, $s \in [0, T]$ and $\mathbf{u} \in X^{t,x}$;*

(iii) *for any $t > 0$, $s \in [0, T]$, $\mathbf{u}, \mathbf{v} \in X^{t,x}$, and for some constant $M > 0$, it holds*

$$|\mathcal{T}(t)G(s, \mathbf{u})|_{HS} \leq Mt^{-\frac{1}{4}}(1 + |\mathbf{u}|_{X^{t,x}}), \quad (4.23)$$

$$|\mathcal{T}(t)G(s, \mathbf{u}) - \mathcal{T}(t)G(s, \mathbf{v})|_{HS} \leq Mt^{-\frac{1}{4}}|\mathbf{u} - \mathbf{v}|_{X^{t,x}}, \quad (4.24)$$

$$|G(s, \mathbf{u})|_{\mathcal{L}} \leq M(1 + |\mathbf{u}|_{X^{t,x}}). \quad (4.25)$$

Proof. Point (i) immediately follows from assumptions 4.2.8, whereas (ii) follows from equation (4.63). Concerning point (iii), we have that eq. (4.63) immediately follows from assumptions 4.2.8. To derive eq. (4.64), we first denote by $\{\phi_k\}_{k \in \mathbb{N}}$ an orthonormal basis in $X^{t,x}$. Then, denoting in what follows by $M > 0$ several different constants, and exploiting assumptions 4.2.8, we have

$$\begin{aligned} |\mathcal{T}(t)G(s, \mathbf{u})|_{HS}^2 &= \sum_{j,k \in \mathbb{N}} \langle \mathcal{T}(t)G(s, \mathbf{u})\phi_j, \phi_k \rangle_{X^{t,x}}^2 = \\ &= \sum_{j,k \in \mathbb{N}} \langle G(s, \mathbf{u})\phi_j, \mathcal{T}(t)\phi_k \rangle_{X^{t,x}}^2 \leq |G(s, \mathbf{u})|_{\mathcal{L}}^2 |\mathcal{T}(t)|_{HS}^2 \leq \\ &\leq M(1 + |\mathbf{u}|_{X^{t,x}}^2) |\mathcal{T}(t)|_{HS}^2 \leq Mt^{-\frac{1}{4}}(1 + |\mathbf{u}|_{X^{t,x}}), \end{aligned} \quad (4.26)$$

where the last inequality follows from Prop. 4.1.14, hence, proceeding as for eq. (4.66), we obtain eq. (4.64). \square

Theorem 4.1.18. *Let assumptions 4.2.2-4.1.9-4.2.8 hold, then there exists a unique mild solution in the sense of Def. 4.2.16.*

Proof. The result can be derived exploiting [DPZ96, Th. 5.3.1], together with Prop. 4.2.11, see also [DPZ11]. \square

Existence and uniqueness for the non-linear equation

In what follows we generalize eq. (4.56), and consequently the abstract Cauchy problem (4.57), taking into account a non-linear Lipschitz perturbation. The notation is as in previous sections. In particular we consider the following non-linear stochastic boundary value problem

$$\left\{ \begin{array}{l} \dot{u}_j(t, x) = (c_j u'_j)'(t, x) + p_j u_j(t, x) + f_j(t, x, u_j(t, x)) + g_j(t, x, u_j(t, x)) \dot{W}_j^1(t, x), \\ \quad \text{for } t \geq 0, x \in (0, 1), j = 1, \dots, m, \\ u_j(t, v_\alpha) = u_l(t, v_\alpha) =: d_\alpha^u(t), \quad t \geq 0, l, j \in \Gamma(v_\alpha), j = 1, \dots, m, \\ \sum_{\beta=1}^n b_{\alpha\beta} d_\beta^u(t) = \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{\alpha i} u'_j(t, v_\alpha), \quad t \geq 0, \alpha = n_0 + 1, \dots, n, \\ \dot{d}_\alpha^u(t) = - \sum_{i,j=1}^m \sum_{\beta=1}^n \delta_{\beta j}^{\alpha i} u'_j(t, v_\beta) + \sum_{\beta=1}^n b_{\alpha\beta} d_\beta^u(t) + \tilde{g}_\alpha(t, d_\alpha^u(t)) \dot{W}_\alpha^2(t, v_\alpha), \\ \quad \text{for } t \geq 0, \alpha = 1, \dots, n_0, \\ u_j(0, x) = u_j^0(x), \quad x \in (0, 1), j = 1, \dots, m, \\ d_i^u(0) = d_i^0, \quad i = 1, \dots, n_0. \end{array} \right. \quad (4.27)$$

Besides the assumptions 4.2.2-4.1.9-4.2.8, we also require that

Hypothesis 4.1.19. For every $j = 1, \dots, m$, the functions $f_j : [0, T] \times [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$, are measurable, bounded and uniformly Lipschitz continuous with respect to the third component, namely there exist constants $C_j > 0$ and K_j , such that, for any $(t, x, y_1) \in [0, T] \times [0, 1] \times \mathbb{R}$ and $(t, x, y_2) \in [0, T] \times [0, 1] \times \mathbb{R}$, it holds

$$|f_j(t, x, y_1)| \leq C_j, \quad |f_j(t, x, y_1) - f_j(t, x, y_2)| \leq K_j |y_1 - y_2|.$$

Analogously to what has been made in Sec. 4.2.3, we reformulate eq. (4.68) as follows

$$\left\{ \begin{array}{l} d\mathbf{u}(t) = [\mathcal{A}\mathbf{u}(t) + F(t, \mathbf{u}(t))] dt + G(t, \mathbf{u}(t)) dW(t), \quad t \geq 0, \\ \mathbf{u}(0) = \mathbf{u}_0 \in X^{t,x}, \end{array} \right. \quad (4.28)$$

moreover we define $F : [0, T] \times X^{t,x} \rightarrow X^{t,x}$, such that

$$F(t, \mathbf{u}) := (f(t, u), 0)^T, \quad \mathbf{u} = (u, y) \in X^{t,x} := X^2 \times \mathbb{R}^n, \quad (4.29)$$

with

$$(f(t, u))(x) := (f_1(t, x, u_1(t, x)), \dots, f_m(t, x, u_m(t, x)))^T.$$

Then, we can state the following result for the existence and uniqueness of a *mild solution* to the eq. (4.69)

Theorem 4.1.20. *Let assumptions 4.2.2-4.1.9-4.2.8-4.2.13 hold, then there exists a unique mild solution to eq. (4.69) in the sense of Def. 4.2.16.*

Proof. It is enough to show that the map F defined in eq. (4.70) is Lipschitz continuous on the space $X^{t,x}$. In fact, from assumptions 4.2.13, it holds

$$|F(t, \mathbf{u}) - F(t, \mathbf{v})|_{X^{t,x}} = |f(t, u) - f(t, v)|_{X^2} \leq K |u - v|_{X^2}. \quad (4.30)$$

Then, exploiting eq. (4.71) together with Prop.4.2.11, the existence of a unique mild solution is a direct application of [DPZ96, Th. 5.3.1], see also [DPZ11]. \square

Remark 4.1.21. A result similar to Th.4.2.14 can be also proved under the assumption of F to be only a function of polynomial growth at infinity, see, e.g., [BM08].

4.1.3 Application to stochastic optimal control

In the present section, in the light of previously obtained results, we consider an optimal control problem related to a general nonlinear control system, written in the following form

$$\begin{cases} du(t)^z = [Au^z(t) + F(t, \mathbf{u}^z) + G(t, \mathbf{u}^u(t))R(t, \mathbf{u}(t), z(t))] dt \\ \quad \quad \quad + G(t, \mathbf{u}^z(t))dW(t), \quad t \in [t_0, T], \\ \mathbf{u}^z(t_0) = \mathbf{u}_0 \in X^{t,x}, \end{cases} \quad (4.31)$$

where z denotes the control and the subscript \mathbf{u}^z denotes the dependence of the process $\mathbf{u} \in X^{t,x}$ from the control z . In particular, we analyse the system (4.72) following the approach given in [FT05], searching for its weak solutions, see, e.g., [FS06].

Let us fix $t_0 \geq 0$ and $\mathbf{u}_0 \in X^{t,x}$, then an *admissible control system* (ACS) is given by $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}, (W(t))_{t \geq 0}, z)$, where

- $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ is a complete probability space;
- $(\mathcal{F}_t)_{t \geq 0}$ is a filtration, in the aforementioned probability space, satisfying the usual assumptions;
- $(W(t))_{t \geq 0}$ is a \mathcal{F}_t -adapted Wiener process with values in $X^{t,x}$;
- z is a process taking values in the space Z , predictable with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$, and such that $z(t) \in Z$ \mathbb{P} -a.s., for almost any $t \in [t_0, T]$, Z being a suitable domain of Z .

To each ACS we associate the mild solution of the abstract equation (4.72) $\mathbf{u}^z \in C([t_0, T]; L^2(\Omega; X^{t,x}))$, and we introduce the following cost functional

$$J(t_0, \mathbf{u}_0, z) := \mathbb{E} \int_{t_0}^T l(t, \mathbf{u}^z(t), z(t)) dt + \mathbb{E} \varphi(\mathbf{u}^z(T)), \quad (4.32)$$

where the function l , resp. the function φ , denotes the *running cost*, resp. the *terminal cost*. Then, the main goal is to chose a control z belonging to a given set of admissible controls, and such that it minimizes the cost functional (4.74). If such a control z exists, it will be called optimal control.

In what follows, besides the assumptions 4.2.2-4.1.9-4.2.84.2.13, we will also require the following to hold

Hypothesis 4.1.22. (i) the map $R : [0, T] \times X^{t,x} \times Z \rightarrow X^{t,x}$ is measurable and, for some $C_R > 0$, it satisfies

$$\begin{aligned} |R(t, \mathbf{u}, z) - R(t, \mathbf{u}, z)|_{X^{t,x}} &\leq C_R(1 + |\mathbf{u}|_{X^{t,x}} + |\mathbf{v}|_{X^{t,x}})^m |\mathbf{u} - \mathbf{v}|_{X^{t,x}}, \\ |R(t, \mathbf{u}, z)|_{X^{t,x}} &\leq C_R; \end{aligned}$$

(ii) the map $l : [0, T] \times X^{t,x} \times \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$ is measurable and, for some $C_l > 0$ and $C \geq 0$, it satisfies

$$\begin{aligned} |R(t, \mathbf{u}, z) - R(t, \mathbf{v}, z)| &\leq C_l(1 + |\mathbf{u}|_{X^{t,x}} + |\mathbf{v}|_{X^{t,x}})^m |\mathbf{u} - \mathbf{v}|_{X^{t,x}}, \\ |R(t, 0, z)|_{X^{t,x}} &\geq -C, \\ \inf_{z \in \mathcal{Z}} l(t, 0, z) &\leq C_l; \end{aligned}$$

(iii) for some $C_\varphi > 0$ and $m \geq 0$, the map $\varphi : X^{t,x} \rightarrow \mathbb{R}$ satisfies

$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v})| \leq C_\varphi(1 + |\mathbf{u}|_{X^{t,x}} + |\mathbf{v}|_{X^{t,x}})^m |\mathbf{u} - \mathbf{v}|_{X^{t,x}}.$$

Following [FT05], if we let assumptions 4.2.2-4.1.9-4.2.84.2.13-4.2.15 to hold, then an ACS can be constructed as follows: first we arbitrarily chose the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ and W as above, then we consider the uncontrolled problem

$$\begin{cases} d\mathbf{u}(t) = [\mathcal{A}\mathbf{u}(t) + F(t, \mathbf{u})] dt + G(t, \mathbf{u}(t))dW(t), & t \geq 0, \\ \mathbf{u}(0) = \mathbf{u}_0 \in X^{t,x}, \end{cases} \quad (4.33)$$

under above assumptions. Then, by Th. 4.2.14, we have the existence for a unique mild solution to eq. (4.75). Moreover, by the boundedness of R and applying the *Girsanov theorem*, we obtain that, for any fixed $\zeta \in \mathcal{Z}$, there exists a probability measure \mathbb{P}^ζ such that

$$W^\zeta(t) := W(t) - \int_{t_0 \wedge t}^{t \wedge T} R(s, \mathbf{u}(s), \zeta) ds,$$

is a Wiener process, so that, for any $t \in [0, T]$, $\mathbf{u}, \mathbf{v} \in X^{t,x}$, we can classically define the Hamiltonian function associated to the problem (4.75), as follows

$$\begin{aligned} \psi(t, \mathbf{u}, \mathbf{v}) &= - \inf_{z \in \mathcal{Z}} \{l(t, \mathbf{u}, z) + \mathbf{v}R(t, \mathbf{u}, z)\}, \\ \Gamma(t, \mathbf{u}, \mathbf{v}) &= \{z \in \mathcal{Z} : \psi(t, \mathbf{u}, \mathbf{v}) + l(t, \mathbf{u}, z) + \mathbf{v}R(t, \mathbf{u}, z) = 0\}, \end{aligned} \quad (4.34)$$

where we note that $\Gamma(t, \mathbf{u}, w)$ is a (possibly empty) subset of \mathcal{Z} , and the function ψ satisfies assumptions 4.2.15. In the present setting we can apply [FT05, Th. 5.1] which allows us to write the *Hamilton-Jacobi-Bellman* (HJB) equation for the problem (4.72)-(4.74), as follows

$$\begin{cases} \frac{\partial w(t, \mathbf{u})}{\partial t} + \mathcal{L}_t w(t, \mathbf{u}) = \psi(t, \mathbf{u}, \nabla w(t, \mathbf{u})G(t, \mathbf{u})), \\ w(T, \mathbf{u}) = \varphi(\mathbf{u}), \end{cases} \quad (4.35)$$

where

$$\mathcal{L}_t w(\mathbf{u}) := \frac{1}{2} Tr [G(t, \mathbf{u})G(t, \mathbf{u})^* \nabla^2 w(\mathbf{u})] + \langle \mathcal{A}\mathbf{u}, \nabla w(\mathbf{u}) \rangle_{X^{t,x}},$$

is the infinitesimal generator associated to the eq. (4.72), Tr denotes the trace, G^* is the adjoint of G and ∇ is a suitable notion of gradient to be introduced in a while.

In particular, see, e.g., [FT05, Def. 5.1], w is said to be a *mild solution in the sense of generalized gradient*, or simply *mild solution*, according to the following definition

Definition 4.1.23. We say that a function $w : [0, T] \times X^{t,x} \rightarrow \mathbb{R}$ is a mild solution to equation (4.77) if the following hold:

- (i) there exist $C > 0$ and $m \geq 0$, such that for any $t \in [0, T]$, and for any $\mathbf{u}, \mathbf{v} \in X^{t,x}$, it holds

$$\begin{aligned} |w(t, \mathbf{u}) - w(t, \mathbf{v})| &\leq C(1 + |\mathbf{u}|_{X^{t,x}} + |\mathbf{v}|_{X^{t,x}})^m |\mathbf{u} - \mathbf{v}|_{X^{t,x}}, \\ |w(t, 0)| &\leq C; \end{aligned}$$

- (ii) for any $0 \leq t \leq T$, $\mathbf{u} \in X^{t,x}$, we have that

$$w(t, \mathbf{u}) = P_{t,T}\varphi(\mathbf{u}) - \int_t^T P_{t,s}\psi(s, \cdot, w(s, \cdot), \rho(s, \cdot))(\mathbf{u})ds,$$

where ρ is an arbitrary element of the *generalized directional gradient* $\nabla^G w$ defined in [FT05], while $P_{t,T}$ is the Markov semigroup generated by the forward process (4.72).

In particular we would like to underline that, thanks to the approach developed in [FT05], we do not need to require any differentiability properties for the functions F , G and w . In fact, the notion of gradient appearing in equation (4.77) is to be intended in a weak sense, which is exactly the notion of the *generalized directional gradient* we have reminded before, see [FT05]. In particular, the latter means that if w is regular enough, then ∇w coincides with the standard notion of gradient, namely, with respect to the present case, it coincides with the *Fréchet derivative*, resp. with the *Gâteaux derivative*, if we assume w to be *Fréchet differentiable*, resp. to be *Gâteaux differentiable*.

We thus have the following result.

Proposition 4.1.24. *Let us consider the optimal control problem (4.72)-(4.74), then the associated HJB equation is represented by eq. (4.77). Moreover, if assumptions 4.2.2-4.1.9-4.2.8-4.2.13-4.2.15 hold, then we have that the HJB equation (4.77) admits a unique mild solution in the sense of definition 4.2.16.*

Proof. The proof immediately follows from [FT05, Th. 5.1]. □

As a direct consequence of Proposition 4.2.18, we have the following

Theorem 4.1.25. *Let assumptions 4.2.2-4.1.9-4.2.8-4.2.13-4.2.15 hold, w be a mild solution to the HJB equation (4.77) and ρ is an element of the generalized directional gradient $\nabla^G w$. Then, for all ACS, we have $J(t_0, \mathbf{u}_0, z) \geq w(t_0, \mathbf{u}_0)$, and the equality holds if and only if the following feedback law is verified by z and \mathbf{u}^z*

$$z(t) = \Gamma(t, \mathbf{u}^z(t), G(t, \rho(t, \mathbf{u}^z(t))), \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T]. \quad (4.36)$$

Moreover, if there exists a measurable function $\gamma : [0, T] \times X^{t,x} \times X^{t,x} \rightarrow \mathcal{Z}$ with

$$\gamma(t, \mathbf{u}, \mathbf{v}) \in \Gamma(t, \mathbf{u}, \mathbf{v}), \quad t \in [0, T], \quad \mathbf{u}, \mathbf{v} \in X^{t,x},$$

then there exists at least one ACS for which

$$\bar{z}(t)\gamma(t, \mathbf{u}^{\bar{z}}(t), \rho(t, \mathbf{u}^{\bar{z}}(t))), \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T].$$

Eventually, we have that $\mathbf{u}^{\bar{z}}$ is a mild solution of equation (4.72).

Proof. See [FT05, Th. 7.2]. □

Example 4.1.1 (The heat equation with controlled stochastic boundary conditions on a graph). In what follows we give an example concerning the heat equation defined on a graph \mathbb{G} , as it has been defined in Sec. 4.2.1. On every nodes of \mathbb{G} we assume local controlled dynamic boundary conditions. Hence, according with the setting introduced above, we have m nodes and $n_0 = n$ nodes equipped with dynamic boundary conditions. We also assume to do not have any noise on the heat equation, whereas we assume the boundary condition to be perturbed by an additive Wiener process. Then, we are considering a system of the following form

$$\begin{cases} \dot{u}_j(t, x) = (c_j u'_j)'(t, x), & t \geq 0, x \in (0, 1), j = 1, \dots, m, \\ u_j(t, v_\alpha) = u_l(t, v_\alpha) =: d_\alpha^u(t), & t \geq 0, l, j \in \Gamma(v_\alpha), j = 1, \dots, m, \\ \dot{d}_\alpha^u(t) = -\sum_{j=1}^m \phi_{\alpha, j} c_j(v_\alpha) u'_j(t, v_\alpha) + b_\alpha d_\alpha^u(t) + \tilde{g}_\alpha(t) (z(t) + \dot{W}_\alpha^2(t)), & t \geq 0, \alpha = 1, \dots, n, \\ u_j(0, x) = u_j^0(x), & x \in (0, 1), j = 1, \dots, m, \\ d_i^u(0) = d_i^0, & i = 1, \dots, n. \end{cases} \quad (4.37)$$

Miming what we have done during previous section, we rewrite (4.79) as an abstract Cauchy problem on the Hilbert space $X^{t,x}$, obtaining

$$\begin{cases} d\mathbf{u}(t)^z = \mathcal{A}\mathbf{u}^z(t)dt + G(t, \mathbf{u}^z(t)) (Rz(t) + dW(t)), & t \in [t_0, T], \\ \mathbf{u}^z(t_0) = \mathbf{u}_0 \in X^{t,x}, \end{cases} \quad (4.38)$$

where $R : \mathbb{R}^n \rightarrow X^{t,x}$ is the immersion of the boundary space \mathbb{R}^n into the product space $X^{t,x} := X^2 \times \mathbb{R}^n$. In the present setting the control z takes values in \mathbb{R}^n , and \mathcal{Z} is a subset of \mathbb{R}^n . Then, if we consider a cost functional of the form (4.74), we have, by Prop. 4.2.18 and Theorem 4.2.20, the existence of at least one ACS for the HJB equation (4.77) which is associated to the stochastic control problem (4.80)-(4.74). Moreover, we can derive the following

Theorem 4.1.26. *Let assumptions 4.2.2-4.1.9-4.2.8-4.2.13-4.2.15 hold, and let w be a mild solution to the HJB equation (4.77), and ρ be an element of the generalized directional gradient $\nabla^G w$. Then, for all ACS, we have $J(t_0, \mathbf{u}_0, z) \geq w(t_0, \mathbf{u}_0)$, and the equality holds if and only if the following feedback law is verified by z and \mathbf{u}^z*

$$z(t) = \Gamma(t, \mathbf{u}^z(t), G(t, \rho(t, \mathbf{u}^z(t)))) , \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T]. \quad (4.39)$$

Besides, if there exists a measurable function $\gamma : [0, T] \times X^{t,x} \times X^{t,x} \rightarrow \mathcal{Z}$, with

$$\gamma(t, \mathbf{u}, \mathbf{v}) \in \Gamma(t, \mathbf{u}, \mathbf{v}), \quad t \in [0, T], \mathbf{u}, \mathbf{v} \in X^{t,x},$$

then there exists at least one ACS such that

$$\bar{z}(t) \gamma(t, \mathbf{u}^{\bar{z}}(t), \rho(t, \mathbf{u}^{\bar{z}}(t))), \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T].$$

Eventually, we have that $\mathbf{u}^{\bar{z}}$ is a mild solution to the eq. (4.72).

4.1.4 Conclusions

In the present paper, we have generalized previously obtained results concerning different evolution problems on networks, by taking into account a diffusion problem on a graph

which has been endowed with non-local boundary static and dynamic conditions, and also considering a stochastic perturbation. We would like to underline that assumptions we made throughout the paper, could be relaxed taking into account the particular geometry of the graph, as it can be constructed according with the peculiarities of the concrete problem in which one is interested.

A second possible generalization of the results presented here, consists in considering time-non-local boundary conditions. The latter, leads to a problem that, as it is standard when dealing with delay equations, can be studied by introducing a suitable path space, with its associated corresponding operator. The price to pay regards the regularity of the leading operator, which is no longer analytic. This implies that the Gaussian estimate, obtained in the present work, does not hold, hence the *Hilbert-Schmidt* property of the semigroup has to be proved with different techniques.

4.2 Stochastic reaction-diffusion equations on networks with dynamic time-delayed boundary conditions

Recent years have seen an increasing attention to the study of diffusion problems on networks, especially in connection with the theory of stochastic processes. In fact, there is a broad area of possible applications where the mathematical use of graphs and random dynamics stated on them, play a crucial role, as in the case, e.g., of quantum mechanics, see, e.g. [Tum06], the books [GZ04, Kle09] and references therein; in neurobiology, as an example concerning the study of stochastic system of the *FitzHugh-Nagumo* type, see, e.g., [ADP11, All10, BCP15, BMZ⁺08, CGMY03]; or in finance, see, e.g., [HN13] and references therein, particularly in the light of numerical applications.

Concerning the aforementioned ambit, a possible approach which has shown to be particularly useful, is to introduce a suitable infinite dimensional space of functions that takes into account the underlying graph domain and then tackle the diffusion problem exploiting both functional analytic tools and infinite dimensional analysis. This technique had led to a systematic study of Stochastic Partial Differential Equations (SPDEs) on networks, showing that it is in general possible to rewrite a diffusion problem defined on a network in a general abstract form, see, e.g., [BMZ⁺08, CGMY03, CM08, DPZ11], and the monograph [Mug14] for a detailed introduction to the subject.

One of the main issues that appears in rewriting the initial problem into an operatorial abstract setting, is to chose the right boundary conditions (BC), that the diffusion problem has to satisfy. In order to overcome the latter, a systematic study of abstract SPDE equipped with different possible BC has been carried up during last years. The typical conditions when one is to deal with diffusion problems governed by a second order differential operator are the so-call *generalized Kirchhoff conditions*, see, e.g., [Mug10]. Nevertheless rather recently, different types of general BC has been proposed, such as non-local BC, allowing for non-local interaction of non-adjacent vertex of the graph, see, e.g., [CGMY03, DPZ11], or dynamic BC, see, e.g., [BMZ⁺08, MR07], or also mixed type BC, allowing for both static and dynamic non-local boundary conditions, see, e.g., [CDP16b].

In the present work we consider a new type of non-local BC. In fact, in any of the aforementioned works, only non-local spatial BC have been considered, while we will focus our attention on boundary conditions which are non-local in time. We refer to [IW06b, IW06a, IP10, Web05], and references therein, for concrete applications that can be potentially stud-

ied in the light of the approach that we develop in our work.

In particular, our study exploits the theory of delay equations, see, e.g., [BP04, BP01], so that we will lift the time-delayed boundary conditions to have values in a suitable infinite dimensional path space, showing that the corresponding differential operator does in fact generate a strongly continuous semigroup on an appropriate space of paths.

The work is so structured: in Sec. 4.2.1 we will introduce the setting and the main notations; in Sec. 4.2.2, exploiting the theory of delay operators, we will introduce the infinite dimensional product space we will work in, also showing that we can rewrite our equation as an infinite dimensional problem where the differential operator generates a strongly continuous semigroup, this immediately lead to the wellposedness of the abstract Cauchy problem; in Sec. 4.2.3 we will introduce a stochastic multiplicative perturbation of Brownian type, showing the existence and uniqueness of a *mild solution*, in a suitable sense, under rather mild assumptions on the coefficients; eventually, in Sec. 4.2.4, we provide an application of the developed theory to a stochastic optimal control problem.

4.2.1 General framework

Let us consider a finite, connected network identified with a finite graph composed by $n \in \mathbb{N}$ vertices v_1, \dots, v_n , and by $m \in \mathbb{N}$ edges e_1, \dots, e_m which are assumed to be normalized on the interval $[0, 1]$. Moreover, we will assume that on the nodes v_1, \dots, v_n of \mathbb{G} are endowed with dynamic boundary conditions to be specified later on.

We would like to recall that in [CM08, CDP16b, DPZ11], it has been considered a diffusion problem, stated on a finite graph, where the boundary conditions exhibit non-local behaviour, namely what happens on a given node also depends on the state of the remaining nodes, even without a direct connection. In the present work, we will consider a different type of non-local condition, studying a diffusion on a finite graph where the boundary conditions, at a given time, are affected by the present value of the state equation on each nodes, as well as by the past values of the underlying dynamic.

In particular we exploiting the semigroup theory, see, e.g. [EN00b] for a detailed introduction to semigroup theory and [Mug14] to what concerns its application on networks, to show how to rephrase our main problem as an abstract Cauchy problem, so that the well-posedness of the solution will be linked to the fact that a certain matrix operator generates a C_0 -semigroup on a suitable, infinite dimensional, space.

In what follows we will employ the following notation: we will use the Latin letter $i, j, k = 1, \dots, m$, $m \in \mathbb{N}^+$, to denote the edges, hence u_i it will be a function on the edge e_i , $i = 1, \dots, m$; while we will use Greek letters $\alpha, \beta, \gamma = 1, \dots, n$, $n \in \mathbb{N}^+$, to denote the vertexes, consequently d_α it will be a function evaluated at the node v_α , $\alpha = 1, \dots, n$.

To describe the graph structure we use the so-called *incidence matrix* $\Phi = (\phi_{\alpha,i})_{(n+1) \times m}$, defined as $\Phi := \Phi^+ - \Phi^-$, where $\Phi^+ = (\phi_{\alpha,i}^+)_{(n+1) \times m}$, resp. $\Phi^- = (\phi_{\alpha,i}^-)_{(n+1) \times m}$, is the *incoming incidence matrix*, resp. the *outgoing incidence matrix*. Let us note that $\phi_{\alpha,i}^+$, resp. $\phi_{\alpha,i}^-$, takes value 1 whenever the vertex v_α is the initial point, resp. the terminal point, of the edge e_i , and 0 otherwise, that is it holds

$$\phi_{\alpha,i}^+ = \begin{cases} 1 & v_\alpha = e_i(0), \\ 0 & \text{otherwise} \end{cases}, \quad \phi_{\alpha,i}^- = \begin{cases} 1 & v_\alpha = e_i(1), \\ 0 & \text{otherwise} \end{cases},$$

moreover, if $|\phi_{\alpha,i}| = 1$, the edge e_i is called *incident* to the vertex v_α and accordingly,

we define

$$\Gamma(v_\alpha) = \{i \in \{1, \dots, m\} : |\phi_{\alpha i}| = 1\},$$

as the set of incident edges to the vertex v_α .

Taking into consideration the above introduced notations, we state the following diffusion problem on the finite and connected graph \mathbb{G}

$$\begin{cases} \dot{u}_j(t, x) = (c_j u_j')'(t, x), & t \geq 0, x \in (0, 1), j = 1, \dots, m, \\ u_j(t, v_\alpha) = u_l(t, v_\alpha) =: d^\alpha(t), & t \geq 0, l, j \in \Gamma(v_\alpha), j = 1, \dots, m, \\ \dot{d}^\alpha(t) = -\sum_{j=1}^m \phi_{j\alpha} u_j'(t, v_\alpha) + b_\alpha d^\alpha(t) + \int_{-r}^0 d^\alpha(t + \theta) \mu(d\theta), & t \geq 0, \alpha = 1, \dots, n, \\ u_j(0, x) = u_j^0(x), & x \in (0, 1), j = 1, \dots, m, \\ d^\alpha(0) = d_\alpha^0, & \alpha = 1, \dots, n, \\ d^\alpha(\theta) = \eta_\alpha^0(\theta), & \theta \in [-r, 0], \alpha = 1, \dots, n. \end{cases} \quad (4.40)$$

where $\mu \in \mathcal{M}([-r, 0])$ and $\mathcal{M}([-r, 0])$ is the set of Borel measure on $[-r, 0]$, being $r > 0$ a finite constant. Before state the main assumptions concerning the terms appearing in (4.40), let us make the following

Remark 4.2.1. We would like to underline that the approach we are going to develop can be generalized, exploiting the same techniques, to the case where only $0 < n_0 < n$ nodes have dynamics conditions, whereas the remaining $n - n_0$ nodes exhibit standard *Kirchhoff type* conditions. Since our interest mainly concerns the study of dynamic boundary conditions, and to consider a mixed boundary type conditions does not affect neither the approach nor the final result, for the sake of simplicity we will assume that all the n nodes composing the graph are endowed with dynamic boundary conditions.

With respect to the definition of the terms we have introduced in (4.40), in order to consider the diffusion problem on \mathbb{G} , we assume the following to hold

Hypothesis 4.2.2. (i) for any $j = 1, \dots, m$, the function $c_j \in C^1([0, 1])$, while $c(x) > 0$ for a.a. $x \in [0, 1]$;

(ii) for any $\alpha = 1, \dots, n$, we have that $b_\alpha \leq 0$, moreover there exists at least one $\alpha \in \{1, \dots, n\}$, such that $b_\alpha < 0$.

The typical approach concerning the study of delay differential equations consists in lifting the underlying process, which originally takes values in a finite dimensional space, to a suitable infinite dimensional path space, usually the space of square integrable Lebesgue functions or the space of continuous functions.

In particular, we consider the following Hilbert spaces

$$\begin{aligned} X^2 &:= (L^2([0, 1]))^m, & Z^2 &:= L^2([-r, 0]; \mathbb{R}^n), \\ X^{t,x} &:= X^2 \times \mathbb{R}^n, & \mathcal{E}^2 &:= X^{t,x} \times Z^2, \end{aligned}$$

equipped with the standard graph norms and scalar products. Since we are interested in applying the aforementioned *lifting* procedure to rewrite the dynamic of the \mathbb{R}^n -valued process d as it takes values in an infinite dimensional space, we introduce the notion of

segment. In particular, we consider the process $d : [-r, T] \rightarrow \mathbb{R}^n$, and, for any $t \geq 0$, we define the *segment* as

$$d_t : [-r, 0] \rightarrow \mathbb{R}^n, \quad [-r, 0] \ni \theta \mapsto d_t(\theta) := d(t + \theta) \in \mathbb{R}^n. \quad (4.41)$$

As it is standard in dealing with delay equation, we denote by $d(t)$ the present \mathbb{R}^n -value of the process d , whereas d_t stands for the *segment* of the process d , i.e. $d_t = (d(t + \theta))_{\theta \in [-r, 0]}$. More precisely, we have

$$\begin{aligned} u(t) &:= (u_1(t), \dots, u_m(t))^T \in X^2, \\ d(t) &:= (d^1(t), \dots, d^n(t))^T \in \mathbb{R}^n, \\ d_t &:= (d_t^1, \dots, d_t^n)^T \in Z^2. \end{aligned}$$

Exploiting latter notations, we can rewrite the system (4.40), as follows

$$\begin{cases} \dot{u}(t) = A_m u(t), & t \in [0, T], \\ \dot{d}(t) = C u(t) + \Phi d_t + \tilde{B} d(t), & t \in [0, T], \\ \dot{d}_t = A_\theta d_t, & t \in [0, T], \\ Lu(t) = d(t), \\ u(0) = u_0 \in X^2, \quad d_0 = \eta \in Z^2, \quad d(0) = d^0 \in \mathbb{R}^n, \end{cases} \quad (4.42)$$

where A_m is the differential operator defined by

$$A_m u(t, x) = \begin{pmatrix} \frac{\partial}{\partial x} (c_j(x) \frac{\partial}{\partial x} u_1(t, x)) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \frac{\partial}{\partial x} (c_m(x) \frac{\partial}{\partial x} u_m(t, x)) \end{pmatrix},$$

and such that $A_m : D(A_m) \subset X^2 \rightarrow X^2$, with domain

$$D(A) := \left\{ u \in (H^2([0, 1]))^m : \exists d \in \mathbb{R}^n : Lu = d \right\},$$

where $L : (H^1([0, 1]))^m \rightarrow \mathbb{R}^n$ is the following *boundary evaluation operator*

$$Lu(t, x) := (d^1(t), \dots, d^n(t))^T, \quad d^\alpha(t) := u_j(t, v_\alpha), \quad j \in \Gamma(v_\alpha).$$

We underline that the operator $(A, D(A))$ just defined, generates a C_0 -semigroup on the space X^2 , see, e.g., [BMZ⁺08, DPZ11, Mug10]. Moreover, in writing system (4.42), we also made use of the so-called *feedback operator* $C : D(A) \rightarrow \mathbb{R}^n$, which is defined as follows

$$C u(t, x) := \left(- \sum_{j=1}^m \phi_{j1} u_j'(t, v_1), \dots, - \sum_{j=1}^m \phi_{jn} u_j'(t, v_n) \right)^T,$$

furthermore, we have set B to be the following $n \times n$ diagonal matrix

$$B = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & b_n \end{pmatrix},$$

where b_α , $\alpha = 1, \dots, n$, satisfy assumptions 4.2.2; also the operator

$$\Phi : H^1([-r, 0]; \mathbb{R}^n) \rightarrow \mathbb{R}^n, \quad (4.43)$$

is a bounded linear operator, commonly indicated as the *delay operator*. Eventually, we have denoted by $A_\theta : D(A_\theta) \subset Z^2 \rightarrow Z^2$, the linear differential operator defined by

$$A_\theta \eta := \frac{\partial}{\partial \theta} \eta(\theta), \quad D(A_\theta) = \{\eta \in H^1([-r, 0]; \mathbb{R}^n) : \eta(0) = d^0\},$$

where the derivative $\frac{\partial}{\partial \theta}$ has to be intended as the weak distributional derivative in Z^2 .

Remark 4.2.3. A particular case of the setting introduced above is given by choosing the so-called *continuous delay operator* $\Phi d_t = \int_{-r}^0 d^u(t+\theta)\mu(d\theta)$, which ensures that (4.40) satisfies the aforementioned assumptions. Another possible choice is represented by the *discrete delay operator* $\Phi d_t = d^u(t-r)$, which is obtained by the previous one taking $\mu = \delta_{-r}$, where δ_{-r} is the Dirac delta centered at $-r$. In what follows we do not specify the particular form of the *delay operator*, in order to prove our results in the general case of a bounded linear operator Φ .

Summing up the previously introduced notation, we can rewrite equation (4.42) more compactly, namely

$$\begin{cases} \dot{\mathbf{u}}(t) = \mathcal{A}\mathbf{u}(t), & t \in [0, T], \\ \mathbf{u}(0) = \mathbf{u}_0 \in \mathcal{E}^2, \end{cases} \quad (4.44)$$

where $\mathbf{u}(t) := (u(t), d(t), d_t)^T$, $\mathbf{u}_0 := (u_0, d^0, \eta) \in \mathcal{E}^2$, and the operator \mathcal{A} is defined as

$$\mathcal{A} := \begin{pmatrix} A_m & 0 & 0 \\ C & B & \Phi \\ 0 & 0 & A_\theta \end{pmatrix}, \quad (4.45)$$

with domain $D(\mathcal{A}) := D(A_m) \times D(A_\theta)$. We will show later that the matrix operator $(\mathcal{A}, D(\mathcal{A}))$ in equation (4.45), generates a C_0 -semigroup on the Hilbert space \mathcal{E}^2 , which implies the wellposedness as well as the uniqueness of the solution, in a suitable sense, for the equation (4.44).

4.2.2 On the infinitesimal generator

The present section will be mainly dedicated to the study of the operator defined in equation (4.45), aiming at proving that it generates a C_0 -semigroup. For the sake of completeness, we recall that the operator \mathcal{A} generates a strongly continuous semigroup in the case that no delay on the boundary is taken into account. In fact, according to the notation introduced within section 4.2.1, if we consider the operator

$$A_a := \begin{pmatrix} A_m & 0 \\ C & B \end{pmatrix}, \quad (4.46)$$

with domain

$$D(A_a) := \{\mathbf{u} = (u, d) \in X^{t,x} : u \in D(A_m), u_j(v_\alpha) = d^\alpha \quad j \in \Gamma(v_\alpha)\}, \quad (4.47)$$

then we have the following result.

Proposition 4.2.4. *Let assumptions 4.2.2 hold true, then the operator $(A_{\mathbf{a}}, D(A_{\mathbf{a}}))$ is self-adjoint, dissipative and has compact resolvent. In particular $A_{\mathbf{a}}$ generates an analytic C_0 -semigroup of contractions on the Hilbert space $X^{t,x}$. Moreover, the semigroup $(T_{\mathbf{a}}(t))_{t \geq 0}$, generated by $A_{\mathbf{a}}$, is uniformly exponentially stable.*

Proof. A proof of the claim can be found in [BMZ⁺08, Prop. 2.4], as well as in [MR07, Cor 3.4], nevertheless we give a sketch of it to better clarify the type of methods involved. We consider the sesquilinear form $\mathbf{a} : V_{\mathbf{a}} \times V_{\mathbf{a}} \rightarrow \mathbb{R}$, defined, for any $\mathbf{u} = (u, d)$, $\mathbf{v} = (v, h) \in X^{t,x}$, by

$$\mathbf{a}(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^m \int_0^1 c_j(x) u'_j(x) v'_j(x) dx + \sum_{\alpha=1}^n b_{\alpha} d^{\alpha} h^{\alpha}. \quad (4.48)$$

and with dense domain $V_{\mathbf{a}} \subset X^{t,x}$ defined as follows

$$V_{\mathbf{a}} := \left\{ \mathbf{u} = (u, d) \in X^{t,x} : u \in (H^1(0, 1))^m, \right. \\ \left. u_j(v_{\alpha}) = d^{\alpha}, \quad \alpha = 1, \dots, n, j \in \Gamma(v_{\alpha}) \right\}.$$

Exploiting [MR07, Lemma 3.2], it can be shown that the form \mathbf{a} is symmetric, closed, continuous and positive, then, by [MR07, Lemma 3.3], it is associated to the operator $(A_{\mathbf{a}}, D(A_{\mathbf{a}}))$, and the result follows by using classical results on sesquilinear forms, see, e.g., [Ouh09]. \square

Using the operator defined in (4.46)–(4.47), and exploiting a well known perturbation result, it is possible to show that the operator $(\mathcal{A}, D(\mathcal{A}))$ generates a C_0 -semigroup. We will first prove that the diagonal operator defined as

$$\mathcal{A}_0 := \begin{pmatrix} A_{\mathbf{a}} & 0 \\ 0 & A_{\theta} \end{pmatrix}, \quad D(\mathcal{A}_0) = D(\mathcal{A}), \quad (4.49)$$

generates a C_0 -semigroup on the Hilbert space \mathcal{E}^2 .

Theorem 4.2.5. *Let assumptions 4.2.2 hold true, then the matrix operator $(\mathcal{A}_0, D(\mathcal{A}_0))$, defined in equation (4.49), generates a C_0 -semigroup given by*

$$\mathcal{T}_0(t) = \left(\begin{array}{c|cc} T_{\mathbf{a}}(t) & 0 & \\ \hline 0 & T_t & T_0(t) \end{array} \right), \quad (4.50)$$

where $T_{\mathbf{a}}$ is the C_0 -semigroup generated by $(A_{\mathbf{a}}, D(A_{\mathbf{a}}))$, see equations (4.46)–(4.47), $T_0(t)$ is the nilpotent left-shift semigroup

$$(T_0(t)\eta)(\theta) := \begin{cases} \eta(t+\theta) & t+\theta \leq 0, \\ 0 & t+\theta > 0, \end{cases}, \quad \eta \in Z^2, \quad (4.51)$$

and $T_t : \mathbb{R}^n \rightarrow Z^2$ is defined by

$$(T_t d)(\theta) := \begin{cases} e^{(t+\theta)B} d & -t < \theta \leq 0, \\ 0 & -r \leq \theta \leq -t, \end{cases}, \quad d \in \mathbb{R}^n, \quad (4.52)$$

$e^{(t+\theta)B}$ being the semigroup generated by the finite dimensional $n \times n$ matrix B , as follows

$$e^{tB} := \sum_{i=0}^{\infty} \frac{(tB)^i}{i!}.$$

Proof. From the strong continuity of T_a and $T_0(t)$ and exploiting the equation (4.52), we have that the semigroup $\mathcal{T}_0(t)$, see equation (4.50), is strongly continuous. Hence, we can compute the resolvent for the semigroup (4.50), showing that the corresponding generator is given by (4.49). To what concerns the resolvent of the operator \mathcal{A}_0 , namely $R(\lambda, \mathcal{A}_0)$, we thus have

$$R(\lambda, \mathcal{A}_0)\mathbf{X} = \int_0^{\infty} e^{-\lambda t} \mathcal{T}_0(t)\mathbf{X} dt, \quad \lambda \in \mathbb{C}, \quad \mathbf{X} \in \mathcal{E}^2.$$

Let us take $\mathbf{u} := (u, d) \in D(A_a)$ and $\eta \in H^1([-r, 0]; \mathbb{R}^n)$, such that the following holds

$$(\lambda - A_a)(u, d)^T = (v, d^v)^T, \quad (v, h)^T \in X^{t,x}, \quad (4.53)$$

$$\lambda\eta - \eta' = \zeta, \quad \eta(0) = d, \quad \zeta \in Z^2, \quad (4.54)$$

then a solution to equation (4.54) is given by

$$\eta(\theta) = e^{\lambda\theta} \left(d + \int_{\theta}^0 e^{-\lambda t} \zeta(t) dt \right).$$

Moreover, if we indicate with A_{θ}^0 the infinitesimal generator of the nilpotent left shift, namely

$$A_{\theta}^0 \eta = \eta' \quad D(A_{\theta}^0) = \{\eta \in H^1([-r, 0]; \mathbb{R}^n) : \eta(0) = 0\},$$

we have that its resolvent is given by

$$(R(\lambda, A_{\theta}^0)\zeta)(\theta) = e^{\lambda\theta} \int_{\theta}^0 e^{-\lambda t} \zeta(t) dt,$$

see, e.g., [EN00b], therefore, taking $\mathbf{Y} = (v, h, \zeta)^T$, the resolvent for \mathcal{A}_0 reads as follows

$$\begin{aligned} R(\lambda, \mathcal{A}_0)\mathbf{Y} &= (R(\lambda, A_a)(v, h), e^{\lambda\theta} R(\lambda, B)h + R(\lambda, A_{\theta}^0)\zeta)^T = \\ &= \left(\begin{array}{c|c} R(\lambda, A_a) & \begin{matrix} 0 \\ 0 \end{matrix} \\ \hline 0 & e^{\lambda\theta} R(\lambda, B) \quad | \quad R(\lambda, A_{\theta}^0)(t) \end{array} \right) \mathbf{Y}. \end{aligned}$$

Summing up, the result follows noticing that

$$\begin{aligned} \int_0^{\infty} e^{-\lambda t} (T_t d)(\theta) dt &= \int_{-\theta}^{\infty} e^{-\lambda t} e^{(t+\theta)B} d(t) dt = \\ &= e^{\lambda\theta} \int_0^{\infty} e^{(t+\theta)B} d(t) dt = e^{\lambda\theta} R(\lambda, B), \end{aligned}$$

so that, we have

$$R(\lambda, \mathcal{A}_0) = \int_0^{\infty} e^{-\lambda t} \mathcal{T}_0(t) dt,$$

which implies that the semigroup $(\mathcal{T}_0(t))_{t \geq 0}$, defined in equation (4.50), is generated by $(\mathcal{A}_0, D(\mathcal{A}_0))$ in (4.49). \square

In what follows we prove that the matrix operator $(\mathcal{A}, D(\mathcal{A}))$ (4.45) generates a C_0 -semigroup on the Hilbert space \mathcal{E}^2 , exploiting a perturbation approach. In particular, we exploit firstly the *Miyadera-Voigt perturbation theorem*, see, e.g., [EN00b, Cor. III.3.16], which states the following

Theorem 4.2.6. *Let $(G, D(G))$ be the generator of a strongly continuous semigroup $(S(t))_{t \geq 0}$, defined on a Banach space X , and let $K \in \mathcal{L}((D(G), \|\cdot\|_G); X)$. Assume that there exist constants $t_0 > 0$ and $0 \leq q \leq 1$, such that*

$$\int_0^{t_0} \|KS(t)x\| dt \leq q\|x\|, \quad \forall x \in D(G). \quad (4.55)$$

Then $(G + K, D(G))$ generates a strongly continuous semigroup $(U(t))_{t \geq 0}$ on X , which satisfies

$$U(t)x = S(t)x + \int_0^t S(t-s)KU(s)x ds,$$

and

$$\int_0^{t_0} \|KU(t)x\| dt \leq \frac{q}{1-q}\|x\|, \quad \forall x \in D(G), t \geq 0.$$

Let us now to consider the operator matrix

$$\mathcal{A}_1 := \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \Phi \\ 0 & 0 & 0 \end{pmatrix} \in \mathcal{L}(D(\mathcal{A}_0), \mathcal{E}^2),$$

where Φ is the delay operator defined in equation (4.43). Exploiting Theorem 4.2.6 we show that, under a suitable assumption on Φ , the matrix operator $\mathcal{A} = \mathcal{A}_0 + \mathcal{A}_1$ generates a C_0 -semigroup on \mathcal{E}^2 .

Theorem 4.2.7. *Let assumptions 4.2.2 hold true, then the operator $(\mathcal{A}, D(\mathcal{A}))$ defined in equation (4.45), generates a strongly continuous semigroup.*

Proof. The result follows applying the *Miyadera-Voigt perturbation theorem* 4.2.6, together with the assumption for the *delay operator* Φ to be bounded, see equation (4.43), therefore the perturbation operator \mathcal{A}_1 is bounded. In fact, from the boundness of Φ , we have that, for $\mathbf{X} = (u, d, \eta)^T$, it holds

$$\int_0^{t_0} \|\mathcal{A}_1 \mathcal{T}_0(t) \mathbf{X}\| dt = \int_0^{t_0} \|\Phi(T_t d + T_0(t) \eta)\| dt \leq M \|T_t d + T_0(t) \eta\|,$$

and equation (4.55) is satisfied for a sufficiently small t_0 , from which the result. \square

A typical example of *delay operator* is represented by the bounded and linear operator $\Phi : C([-r, 0]; \mathbb{R}^n) \rightarrow \mathbb{R}^n$, defined by

$$\Phi(\eta) = \int_{-r}^0 \eta(\theta) \mu(d\theta),$$

where μ is of bounded variation. In fact, since $H^1([-r, 0]; \mathbb{R}^n)$ is continuously embedded in $C([-r, 0]; \mathbb{R}^n)$, then $\Phi : H^1([-r, 0]; \mathbb{R}^n) \rightarrow \mathbb{R}^n$ is bounded.

We note that Theorem 4.2.7 holds for more general type of *delay operators*, namely taking into consideration weaker assumptions on its definition. In fact, by the result contained in [BP04, Th. 1.17], we have that $(\mathcal{A}, D(\mathcal{A}))$, defined in equation (4.45), generates a strongly continuous semigroup for a general unbounded operator Φ , provided that there exist $t_0 > 0$ and $0 < q < 1$ such that

$$\int_0^{t_0} \|\Phi(S_t \mathbf{u} + T_0(t)\eta)\| dt \leq q \|(\mathbf{u}, \eta)\|.$$

4.2.3 The perturbed stochastic problem

In the present section we study the system defined in (4.40) perturbed by a multiplicative Gaussian noise. We will carry out our analysis with respect to the following standard, complete and filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, then we define the following system

$$\begin{cases} \dot{u}_j(t, x) = (c_j u_j')'(t, x) + g_j(t, x, u_j(t, x)) \dot{W}_j^1(t, x), & t \geq 0, x \in (0, 1), j = 1, \dots, m, \\ u_j(t, v_\alpha) = u_l(t, v_\alpha) =: d^\alpha(t), & t \geq 0, l, j \in \Gamma(v_i), j = 1, \dots, m, \\ \dot{d}^\alpha(t) = -\sum_{j=1}^m \phi_{j\alpha} u_j'(t, v_\alpha) + b_\alpha d^\alpha(t) + \int_{-r}^0 d^\alpha(t+\theta) \mu(d\theta) + \tilde{g}_\alpha(t, d^\alpha(t), d_t^\alpha) \dot{W}_\alpha^2(t, v_\alpha), & t \geq 0, \alpha = 1, \dots, n, \\ u_j(0, x) = u_j^0(x), & x \in (0, 1), j = 1, \dots, m, \\ d^\alpha(0) = d_\alpha^0, & \alpha = 1, \dots, n, \\ d^\alpha(\theta) = \eta_\alpha^0(\theta), & \theta \in [-r, 0], \alpha = 1, \dots, n. \end{cases} \quad (4.56)$$

where W_j^1 and W_α^2 , $j = 1, \dots, m$, $\alpha = 1, \dots, n_0$, are independent \mathcal{F}_t -adapted space time Wiener processes to be specified in a while, and \dot{W} indicates the *formal* time derivative. In particular W_j^1 , $j = 1, \dots, m$, is a space time Wiener process taking values in $L^2(0, 1)$, consequently we denote by $W^1 = (W_1^1, \dots, W_m^1)$ a space time Wiener process with values in $X^2 := (L^2(0, 1))^m$. Similarly, we have that each W_α^2 , $\alpha = 1, \dots, n$, is a space time Wiener process with values in \mathbb{R} , so that we denote by $W^2 = (W_1^2, \dots, W_n^2)$ the standard Wiener process with values in \mathbb{R}^n . Eventually, we indicate by $W := (W^1, W^2)$ a standard space time Wiener process with values in $X^{t,x} := X^2 \times \mathbb{R}^n$.

In what follows we require both the assumptions stated in 4.2.2, as well as the following *Hypothesis* 4.2.8. (i) The functions

$$g_j : [0, T] \times [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}, \quad j = 1, \dots, m,$$

are measurable, bounded and uniformly Lipschitz with respect to the third component, namely there exist $C_j > 0$ and $K_j > 0$, such that, for any $(t, x, y_1) \in [0, T] \times [0, 1] \times \mathbb{R}$ and $(t, x, y_2) \in [0, T] \times [0, 1] \times \mathbb{R}$, it holds

$$|g_j(t, x, y_1)| \leq C_j, \quad |g_j(t, x, y_1) - g_j(t, x, y_2)| \leq K_j |y_1 - y_2|;$$

(ii) The functions

$$\tilde{g}_\alpha : [0, T] \times \mathbb{R} \times Z^2 \rightarrow \mathbb{R}, \quad \alpha = 1, \dots, n_0,$$

are measurable, bounded and uniformly Lipschitz with respect to the second component, namely there exist $C_\alpha > 0$ and $K_\alpha > 0$, such that, for any $(t, u, \eta) \in [0, T] \times \mathbb{R} \times Z^2$ and $(t, v, \zeta) \in [0, T] \times \mathbb{R} \times Z^2$, it holds

$$|\tilde{g}_\alpha(t, u, \eta)| \leq C_\alpha, \quad |\tilde{g}_\alpha(t, u, \eta) - \tilde{g}_\alpha(t, v, \zeta)| \leq K_\alpha(|u - v|_n + |\eta - \zeta|_{Z^2}).$$

Using previously introduced notations, the problem in (4.56) can be rewritten as the following abstract infinite dimensional Cauchy problem

$$\begin{cases} d\mathbf{X}(t) = \mathcal{A}\mathbf{X}(t)dt + G(t, \mathbf{X}(t))dW(t), & t \geq 0, \\ \mathbf{u}(0) = \mathbf{u}_0 \in \mathcal{E}^2, \end{cases} \quad (4.57)$$

where \mathcal{A} is the operator introduced in (4.45), the map G is defined as the following application

$$G : [0, T] \times \mathcal{E}^2 \rightarrow \mathcal{L}(X^{t,x}; \mathcal{E}^2),$$

being $\mathcal{L}(X^{t,x}; \mathcal{E}^2)$ the space of linear and bounded operator from $X^{t,x}$ to \mathcal{E}^2 , equipped with standard norm $|\cdot|_{\mathcal{L}}$, other terms are intended such as they have been defined within Sec. 4.2.2, and $W = (W^1, W^2)$ is a $X^{t,x}$ -valued standard Brownian motion.

In particular, if $\mathbf{X} = (\mathbf{u}, \eta)^T = (u, y, \eta)$, and $\mathbf{Y} = (\mathbf{v}, \eta)^T = (v, z, \eta)$, then G is defined as

$$G(t, \mathbf{X})\mathbf{Y} = (\sigma_1(t, u)v, \sigma_2(t, y, \eta)z, 0)^T, \quad (4.58)$$

with

$$\begin{aligned} (\sigma_1(t, u)v)(x) &= (g_1(t, x, u_1(t, x)), \dots, g_m(t, x, u_m(t, x)))^T, \\ \sigma_2(t, y, \eta)z &= (\tilde{g}_1(t, y_1, \eta)z_1, \dots, \tilde{g}_n(t, y_n, \eta)z_n)^T. \end{aligned}$$

Our next step concerns how to obtain a *mild solution* to equation (4.57), namely a solution defined in the following sense

Definition 4.2.9. We will say that \mathbf{X} is *mild solution* to equation (4.57) if it is a mean square continuous \mathcal{E}^2 -valued process, adapted to the filtration generated by W , such that, for any $t \geq 0$, we have that $\mathbf{X} \in L^2(\Omega, C([0, T]; \mathcal{E}^2))$ and it holds

$$\mathbf{X}(t) = \mathcal{T}(t)\mathbf{X}_0 + \int_0^t \mathcal{T}(t-s)G(s, \mathbf{X}(s))dW(s), \quad t \geq 0. \quad (4.59)$$

In general, in order to guarantee the existence and uniqueness of a mild solution to equation (4.57), we have to require that

$$G : [0, T] \times \mathcal{E}^2 \rightarrow \mathcal{L}_2(X^{t,x}; \mathcal{E}^2),$$

being $\mathcal{L}_2(X^{t,x}; \mathcal{E}^2)$ the space of *Hilbert-Schmidt* operator from $X^{t,x}$ to \mathcal{E}^2 equipped with its standard norm denoted as $|\cdot|_{HS}$, see, e.g., [DPZ14, Appendix C]. Nevertheless, when dealing with a diffusion problem where the leading term is a second order differential operator, it is enough to require that G takes value in $\mathcal{L}(X^{t,x}; \mathcal{E}^2)$ since, in this particular case, the map G inherits the needed regularity from the analytic semigroup generated by the second order differential operator. On the other hand, if we consider a delay operator then, due to the presence of the first order differential operator A_θ , the operator \mathcal{A} , defined in equation

(4.45), does not generate an analytic semigroup on the space \mathcal{E}^2 . The latter suggests that it seems reasonable to require G to take values in $\mathcal{L}_2(X^{t,x}; \mathcal{E}^2)$, in order to have both existence and uniqueness for a solution to equation (4.57). In what follows, we will show that, since $A_{\mathbf{a}}$ generates an analytic semigroup, and exploiting the particular form for G in equation (4.58), we have that $\mathcal{T}(t)G(s, \mathbf{X})$ belongs to $\mathcal{L}_2(X^{t,x}; \mathcal{E}^2)$, hence, by assumptions 4.2.8 on the functions g and \tilde{g} , the existence and uniqueness of a mild solution to equation (4.57) follows. In particular, we have the following result

Proposition 4.2.10. *Let assumptions 4.2.2–4.2.8 hold true, then we have that $\mathcal{T}(t)G(s, \mathbf{X}) \in \mathcal{L}_2(X^{t,x}; \mathcal{E}^2)$, namely*

$$\sum_{i=1}^{\infty} \|\mathcal{T}(t)G(s, \mathbf{X})\tilde{\phi}_i\|_{\mathcal{E}^2}^2 < \infty, \quad (4.60)$$

where $\{\tilde{\phi}_i\}_{i=1}^{\infty}$ is an orthonormal basis in $X^{t,x}$.

Proof. Let $\{\tilde{\phi}_i\}_{i=1}^{\infty}$, resp. $\{\phi_i\}_{i=1}^{\infty}$, resp. $\{e_i\}_{i=1}^n$, be an orthonormal basis in $X^{t,x}$, resp. in X^2 , resp. in \mathbb{R}^n . Exploiting the explicit form of the semigroups \mathcal{T} and G , see equation (4.58), we have that

$$\begin{aligned} \sum_{i=1}^{\infty} \|\mathcal{T}(t)G(s, \mathbf{X})\tilde{\phi}_i\|_{\mathcal{E}^2}^2 &= \sum_{i=1}^{\infty} \|T_{\mathbf{a}}(t) (\sigma_1(s, u), \sigma_2(s, d^u, \eta))^T \tilde{\phi}_i\|_{X^{t,x}}^2 + \\ &+ \sum_{i=1}^n \|T_t \sigma_2(s, d^u, \eta) e_i\|_{Z^2}^2. \end{aligned} \quad (4.61)$$

The first sum on the right hand side of equation (4.61) converges because, see [BCM08, Th. 8], $T_{\mathbf{a}}$ is indeed a *Hilbert-Schmidt* operator. Concerning the second term, i.e. $\sum_{i=1}^n \|T_t \sigma_2(s, d^u, \eta) e_i\|_{Z^2}^2$, we have that, since the matrix B is diagonal, $e^{tB} = \text{diag}(e^{tb_1}, \dots, e^{tb_n})$, see, e.g., [EN00b]. Therefore, exploiting the particular form of T_t in equation (4.52), we have that the following holds for any e_i

$$(T_t e_i) = \begin{cases} (0, \dots, 0, e^{(t+\theta)b_i}, 0, \dots, 0) & , -t < \theta < 0, \\ 0 & , -r \leq \theta \leq -t, \end{cases} \quad (4.62)$$

hence, by assumptions 4.2.8, we also obtain

$$\|T_t \sigma_2(s, d^u, \eta) e_i\|_{Z^2}^2 = \int_{-t}^0 e^{2(t+\theta)b_i} \sigma_2(s, d^u, \eta) d\theta < \infty,$$

which implies that the second sum on the right hand side of (4.61) is finite, and the claim is proved. \square

The next result will be later used in order to show the existence and uniqueness of a mild solution to equation (4.57).

Proposition 4.2.11. *Let assumptions 4.2.2–4.2.8 hold true, then the map $G : [0, T] \times \mathcal{E}^2 \rightarrow \mathcal{L}(X^{t,x}, \mathcal{E}^2)$, defined in equation (4.58), satisfies:*

(i) *for any $\mathbf{u} \in X^{t,x}$ the map $G(\cdot, \cdot)\mathbf{u} : [0, T] \times \mathcal{E}^2 \rightarrow \mathcal{E}^2$, is measurable;*

(ii) for any $t > 0$, $s \in [0, T]$, $\mathbf{X}, \mathbf{Y} \in \mathcal{E}^2$, there exists a constant $M > 0$, such that

$$|\mathcal{T}(t)G(s, \mathbf{X})|_{HS} \leq Mt^{-\frac{1}{4}}(1 + |\mathbf{X}|_{\mathcal{E}^2}), \quad (4.63)$$

$$|\mathcal{T}(t)G(s, \mathbf{X}) - \mathcal{T}(t)G(s, \mathbf{Y})|_{HS} \leq Mt^{-\frac{1}{4}}|\mathbf{X} - \mathbf{Y}|_{\mathcal{E}^2}, \quad (4.64)$$

$$|G(s, \mathbf{X})|_{\mathcal{L}} \leq M(1 + |\mathbf{X}|_{\mathcal{E}^2}). \quad (4.65)$$

Proof. Point (i) and (4.65) in point (ii), immediately follow from assumptions 4.2.8. To prove the inequality (4.63) in point (ii), we proceed analogously as in the proof of Prop. 4.2.10. In particular, let $\{\phi_k\}_{k \in \mathbb{N}}$, resp. $\{e_k\}_{k=1}^n$, resp. $\{\tilde{\phi}_k\}_{k \in \mathbb{N}}$, resp. $\{\psi_k\}_{k \in \mathbb{N}}$, resp. $\{\tilde{\psi}_k\}_{k \in \mathbb{N}}$, be an orthonormal basis in X^2 , resp. in \mathbb{R}^n , resp. a basis in $X^{t,x}$, resp. an orthonormal basis in Z^2 , resp. a basis in \mathcal{E}^2 , and, for the sake of clarity, let us denote with $M > 0$ different constants throughout what follows. Then, exploiting assumptions 4.2.8, together with the particular form of G given in equation(4.58), we have

$$\begin{aligned} |\mathcal{T}(t)G(s, \mathbf{X})|_{HS}^2 &= \sum_{j,k \in \mathbb{N}} \left\langle \mathcal{T}(t)G(s, \mathbf{X})\tilde{\phi}_j, \tilde{\psi}_k \right\rangle_{\mathcal{E}^2}^2 = \\ &= \sum_{j,k \in \mathbb{N}} \left\langle T_{\mathbf{a}}(t)(\sigma_1(s, u), \sigma_2(s, d^u, \eta))\tilde{\phi}_j, \tilde{\phi}_k \right\rangle_{X^{t,x}}^2 + \\ &+ \sum_{j=1}^n \sum_{k \in \mathbb{N}} \left\langle T_t \sigma_2(s, d^u, \eta)e_j, \psi_k \right\rangle_{Z^2}. \end{aligned} \quad (4.66)$$

Since $T_{\mathbf{a}}$ is self-adjoint and by [BCM08, Prop. 10], we have that

$$\begin{aligned} &\sum_{j,k \in \mathbb{N}} \left\langle T_{\mathbf{a}}(t)(\sigma_1(s, u), \sigma_2(s, d^u, \eta))\tilde{\phi}_j, \tilde{\phi}_k \right\rangle_{X^{t,x}}^2 = \\ &= \sum_{j,k \in \mathbb{N}} \left\langle (\sigma_1(s, u), \sigma_2(s, d^u, \eta))\tilde{\phi}_j, T_{\mathbf{a}}(t)\tilde{\phi}_k \right\rangle_{X^{t,x}}^2 \leq \\ &\leq \|(\sigma_1(s, u), \sigma_2(s, d^u, \eta))\|_{\mathcal{L}(X^{t,x})}^2 |T_{\mathbf{a}}(t)|_{\mathcal{L}_2(X^{t,x})}^2 \leq |G(s, \mathbf{X})|_{\mathcal{L}(X^{t,x}; \mathcal{E}^2)}^2 |T_{\mathbf{a}}(t)|_{\mathcal{L}_2(X^{t,x})}^2 \leq \\ &\leq Mt^{-\frac{1}{2}}(1 + |\mathbf{X}|_{\mathcal{E}^2}^2). \end{aligned} \quad (4.67)$$

Moreover, because \mathbb{R}^n is finite dimensional and $\mathcal{L}_2(\mathbb{R}^n; Z^2) = \mathcal{L}(\mathbb{R}^n; Z^2)$, we immediately have that the following holds

$$|\mathcal{T}(t)G(s, \mathbf{X})|_{HS} \leq Mt^{-\frac{1}{4}}(1 + |\mathbf{X}|_{\mathcal{E}^2}),$$

and equation (4.63) is thus proved. Finally, the proof of the inequality (4.64) in (ii) proceeds the same way as the latter one. \square

Summing up previous results, we are now in position to state the following

Theorem 4.2.12. *Let assumptions 4.2.2–4.2.8 hold true, then there exists a unique mild solution, in the sense of Definition 4.2.16, to equation (4.57).*

Proof. The result follows by [DPZ96, Th. 5.3.1], see also [DPZ11], together with propositions 4.2.10 and 4.2.11. \square

Existence and uniqueness for the non-linear equation

The present subsection is devoted to the generalisation of the existence and uniqueness of a mild solution, see Th. 4.2.12, to the abstract formulation, see eq. (4.57), of the problem stated by eq. (4.56). In particular we shall consider the addition of a non-linear Lipschitz perturbation. The notation used in what follows is as in previous sections.

We will thus focus on the following non-linear stochastic dynamic boundary value problem

$$\left\{ \begin{array}{l} \dot{u}_j(t, x) = \left(c_j u_j' \right)' (t, x) + f_j(t, x, u_j(t, x)) + g_j(t, x, u_j(t, x)) \dot{W}_j^1(t, x), \\ \quad t \geq 0, x \in (0, 1), j = 1, \dots, m, \\ u_j(t, v_\alpha) = u_l(t, v_\alpha) =: d^\alpha(t), \quad t \geq 0, l, j \in \Gamma(v_i), j = 1, \dots, m, \\ \dot{d}^\alpha(t) = -\sum_{j=1}^m \phi_{j\alpha} u_j'(t, v_\alpha) + b_\alpha d^\alpha(t) + \int_{-r}^0 d^\alpha(t+\theta) \mu(d\theta) + \tilde{g}_\alpha(t, d^\alpha(t), d_t^\alpha) \dot{W}_\alpha^2(t, v_\alpha), \\ \quad t \geq 0, \alpha = 1, \dots, n, \\ u_j(0, x) = u_j^0(x), \quad x \in (0, 1), j = 1, \dots, m, \\ d^\alpha(0) = d_\alpha^0, \quad \alpha = 1, \dots, n, \\ d^\alpha(\theta) = \eta_\alpha^0(\theta), \quad \theta \in [-r, 0], \alpha = 1, \dots, n. \end{array} \right. \quad (4.68)$$

In what follows, besides assumptions 4.2.2 and 4.2.8 and in order to deal with functions f_j appearing in eq. (4.68), we also require the following

Hypothesis 4.2.13. The functions

$$f_j : [0, T] \times [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}, \quad j = 1, \dots, m,$$

are measurable mappings, bounded and uniformly Lipschitz continuous with respect to the third component, namely, for $j = 1, \dots, m$, there exist positive constants C_j and K_j , such that, for any $(t, x, y_1) \in [0, T] \times [0, 1] \times \mathbb{R}$ and any $(t, x, y_2) \in [0, T] \times [0, 1] \times \mathbb{R}$, it holds

$$|f_j(t, x, y_1)| \leq C_j, \quad |f_j(t, x, y_1) - f_j(t, x, y_2)| \leq K_j |y_1 - y_2|.$$

Proceeding similarly to what is seen in Sec. 4.2.3, we reformulate equation (4.68) as an abstract Cauchy problem as follows

$$\left\{ \begin{array}{l} d\mathbf{X}(t) = [\mathcal{A}\mathbf{X}(t) + F(t, \mathbf{X})] dt + G(t, \mathbf{X}(t)) dW(t), \quad t \geq 0, \\ \mathbf{X}(0) = \mathbf{X}_0 \in \mathcal{E}^2, \end{array} \right. \quad (4.69)$$

where $F : [0, T] \times \mathcal{E}^2 \rightarrow \mathcal{E}^2$, and such that

$$F(t, \mathbf{X}) = (f(t, u), 0, 0)^T, \quad \text{being } \mathbf{X} = (u, y, \eta) \in \mathcal{E}^2, \quad (4.70)$$

with

$$(f(t, u))(x) = (f_1(t, x, u_1(t, x)), \dots, f_m(t, x, u_m(t, x)))^T.$$

The following result provides the existence and uniqueness of a *mild solution* to equation (4.69).

Theorem 4.2.14. *Let assumptions 4.2.2, 4.2.8 and 4.2.13, hold true. Then, there exists a unique mild solution, in the sense of the Definition 4.2.16, to equation (4.69).*

Proof. It is enough to show that the map F defined in equation (4.70) is Lipschitz continuous on the Hilbert space \mathcal{E}^2 . In fact, assumptions 4.2.13 implies that

$$|F(t, \mathbf{X}) - F(t, \mathbf{Y})|_{\mathcal{E}^2} = |f(t, u) - f(t, v)|_{X^2} \leq K|u - v|_{X^2} \leq |\mathbf{X} - \mathbf{Y}|_{\mathcal{E}^2}, \quad (4.71)$$

for any $\mathbf{X} = (u, y, \eta)^T$ and any $\mathbf{Y} = (v, z, \zeta)^T \in \mathcal{E}^2$. Then, exploiting equation (4.71), together with Proposition 4.2.11, the existence of a unique mild solution is a direct application of [DPZ96, Th. 5.3.1], see also [DPZ11]. \square

4.2.4 Application to stochastic optimal control

The present section is mainly devoted to the study and characterization of the stochastic optimal control associated to a general non-linear system of the form

$$\begin{cases} d\mathbf{X}^z(t) = [A\mathbf{X}^z(t) + F(t, \mathbf{X}^z) + G(t, \mathbf{X}^z(t))R(t, \mathbf{X}^z(t), z(t))] dt + G(t, \mathbf{X}^z(t))dW(t), \\ \mathbf{X}^z(t_0) = \mathbf{X}_0 \in \mathcal{E}^2, \end{cases} \quad (4.72)$$

where, besides having used the notations defined along previous sections, we denote by z the control, while we use the notation \mathbf{X}^z , to indicate the explicit dependence of the process $\mathbf{X} \in \mathcal{E}^2$, from the control z . In what follows we exploit the results contained in [FT05], where a general characterization of stochastic optimal control problem in infinite dimension is given by mean of a forward-backward-SDE approach. Therefore, the control problem defined by equation (4.72), is to be understood in the weak sense, see also, e.g., [DPZ11, FS06]. In particular, we aim at finding a control z , within a given set of admissible controls, such that it minimizes the following cost functional

$$J(t_0, \mathbf{X}_0, z) = \mathbb{E} \int_{t_0}^T l(t, \mathbf{X}^z(t), z(t)) dt + \mathbb{E}\varphi(\mathbf{X}^z(T)). \quad (4.73)$$

As stated in [FT05], we first fix $t_0 \geq 0$ and $\mathbf{X}_0 \in \mathcal{E}^2$, then an *Admissible Control System* (ACS) is given by $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P}, (W(t))_{t \geq 0}, z)$, where

- $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ is a complete probability space, where the filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfies the usual assumptions;
- $(W(t))_{t \geq 0}$ is a \mathcal{F}_t -adapted Wiener process taking values in \mathcal{E}^2 ;
- z is a process taking values in the space Z , predictable with respect to the filtration $(\mathcal{F}_t)_{t \geq 0}$, and such that $z(t) \in Z$ \mathbb{P} -a.s., for almost any $t \in [t_0, T]$, being Z a suitable domain of Z .

To each ACS, we associate the mild solution $\mathbf{X}^z \in C([t_0, T]; L^2(\Omega; \mathcal{E}^2))$ to the abstract equation (4.72). Consequently, we can introduce the functional cost

$$J(t_0, \mathbf{X}_0, z) = \mathbb{E} \int_{t_0}^T l(t, \mathbf{X}^z(t), z(t)) dt + \mathbb{E}\varphi(\mathbf{X}^z(T)), \quad (4.74)$$

where the function l , resp. φ , denotes the *running cost*, resp. the *terminal cost*. Our goal is to minimize the functional J over all admissible control system. If a minimizing ACS for the functional J exists, then it is called optimal control.

Throughout this section we will make use of the assumptions 4.2.2, 4.2.8, and 4.2.13, moreover we will also assume the following

Hypothesis 4.2.15. (i) the map $R : [0, T] \times \mathcal{E}^2 \times \mathcal{Z} \rightarrow \mathcal{E}^2$ is measurable and it satisfies

$$\begin{aligned} |R(t, \mathbf{X}, z) - R(t, \mathbf{Y}, z)|_{\mathcal{E}^2} &\leq C_R(1 + |\mathbf{X}|_{\mathcal{E}^2} + |\mathbf{Y}|_{\mathcal{E}^2})^m |\mathbf{X} - \mathbf{Y}|_{\mathcal{E}^2}, \\ |R(t, \mathbf{X}, z)|_{\mathcal{E}^2} &\leq C_R; \end{aligned}$$

for some $C_R > 0$;

(ii) the map $l : [0, T] \times \mathcal{E}^2 \times \mathcal{Z} \rightarrow \mathbb{R} \cup \{+\infty\}$ is measurable and it satisfies

$$\begin{aligned} |l(t, \mathbf{X}, z) - l(t, \mathbf{Y}, z)| &\leq C_l(1 + |\mathbf{X}|_{\mathcal{E}^2} + |\mathbf{Y}|_{\mathcal{E}^2})^m |\mathbf{X} - \mathbf{Y}|_{\mathcal{E}^2}, \\ |l(t, 0, z)|_{\mathcal{E}^2} &\geq -C, \\ \inf_{z \in \mathcal{Z}} l(t, 0, z) &\leq C_l; \end{aligned}$$

for some $C > 0$ and $C_l \geq 0$;

(iii) the map $\varphi : \mathcal{E}^2 \rightarrow \mathbb{R}$ satisfies

$$|\varphi(\mathbf{X}) - \varphi(\mathbf{Y})| \leq C_\varphi(1 + |\mathbf{X}|_{\mathcal{E}^2} + |\mathbf{Y}|_{\mathcal{E}^2})^m |\mathbf{X} - \mathbf{Y}|_{\mathcal{E}^2}.$$

for some $C_\varphi > 0$ and $m \geq 0$.

Under assumptions 4.2.2, 4.2.8, 4.2.13, and 4.2.15, we can construct, see [FT05], an ACS as follows. Let us arbitrarily chose the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$, and W as above. Then, we consider the uncontrolled problem

$$\begin{cases} d\mathbf{X}(t) = [\mathbf{A}\mathbf{X}(t) + F(t, \mathbf{X})] dt + G(t, \mathbf{X}(t))dW(t), & t \geq 0, \\ \mathbf{X}(0) = \mathbf{X}_0 \in \mathcal{E}^2. \end{cases} \quad (4.75)$$

Exploiting Theorem 4.2.14, we have that there exists a unique mild solution to equation (4.75). Moreover, by the boundedness of R and applying the *Girsanov theorem*, we have that, $\forall \zeta \in \mathcal{Z}$, there exists a probability measure \mathbb{P}^ζ , such that

$$W^\zeta(t) := W(t) - \int_{t_0 \wedge t}^{t \wedge T} R(s, \mathbf{X}(s), \zeta) ds,$$

is a Wiener process. Consequently, $\forall t \in [0, T]$, and $\forall (\mathbf{X}, \mathbf{Y}) \in \mathcal{E}^2 \times \mathcal{E}^2$, we define the Hamiltonian function related to the aforementioned problem, as follows

$$\begin{aligned} \psi(t, \mathbf{X}, \mathbf{Y}) &:= - \inf_{z \in \mathcal{Z}} \{l(t, \mathbf{X}, z) + \mathbf{Y}R(t, \mathbf{X}, z)\}, \\ \Gamma(t, \mathbf{X}, \mathbf{Y}) &:= \{z \in \mathcal{Z} : \psi(t, \mathbf{X}, \mathbf{Y}) + l(t, \mathbf{X}, z) + \mathbf{v}R(t, \mathbf{X}, z) = 0\}, \end{aligned} \quad (4.76)$$

where we would underline that the set $\Gamma(t, \mathbf{X}, w)$ is a (possibly empty) subset of \mathcal{Z} , while the function ψ satisfies assumptions 4.2.15.

Within the present setting, we can apply [FT05, Th. 5.1] to write the *Hamilton-Jacobi-Bellman* (HJB) equation associated to the problem stated by (4.72) together with (4.74). In particular, we have

$$\begin{cases} \frac{\partial w(t, \mathbf{X})}{\partial t} + \mathcal{L}_t w(t, \mathbf{X}) = \psi(t, \mathbf{X}, \nabla w(t, \mathbf{X})G(t, \mathbf{X})), \\ w(T, \mathbf{X}) = \varphi(\mathbf{X}), \end{cases} \quad (4.77)$$

where

$$\mathcal{L}_t w(\mathbf{X}) := \frac{1}{2} Tr [G(t, \mathbf{X})G(t, \mathbf{X})^* \nabla^2 w(\mathbf{X})] + \langle A\mathbf{X}, \nabla w(\mathbf{X}) \rangle_{\mathcal{E}^2},$$

is the infinitesimal generator of the equation (4.72), while Tr stands for the *trace*, and G^* is the adjoint of G .

In what follows we exploit the following definition, see, e.g., [FT05, Def. 5.1].

Definition 4.2.16. A function $u : [0, T] \times X^{t,x} \rightarrow \mathbb{R}$ is defined to be a mild solution in the sense of generalized gradient, to equation (4.77) if the following hold:

(i) there exists $C > 0$ and $m \geq 0$ such that for any $t \in [0, T]$ and any $\mathbf{u}, \mathbf{v} \in X^{t,x}$ it holds

$$\begin{aligned} |w(t, \mathbf{X}) - w(t, \mathbf{Y})| &\leq C(1 + |\mathbf{X}|_{\mathcal{E}^2} + |\mathbf{Y}|_{\mathcal{E}^2})^m |\mathbf{X} - \mathbf{Y}|_{\mathcal{E}^2}, \\ |w(t, 0)| &\leq C; \end{aligned}$$

(ii) for any $0 \leq t \leq T$ and $\mathbf{X} \in \mathcal{E}^2$, we have that

$$w(t, \mathbf{X}) = P_{t,T} \varphi(\mathbf{X}) - \int_t^T P_{t,s} \psi(s, \cdot, w(s, \cdot), \rho(s, \cdot))(\mathbf{X}) ds,$$

where ρ is an arbitrary element of the *generalized directional gradient* $\nabla^G w$, as it has been defined in [FT05], while $P_{t,T}$ is the Markov semigroup generated by the forward process (4.72).

Remark 4.2.17. We would like to underline that, following the approach developed in [FT05], we do not need to require any differentiability properties for the function F , G and w . In fact, the notion of *gradient* appearing in equation (4.77), is to be intend in a weak sense, namely in terms of the *generalized directional gradient*. In fact, in [FT05] the authors show that, if w is regular enough, then ∇w coincides with the standard notion of gradient. The latter implies that, in the present case, the *generalized directional gradient* coincides with the *Fréchet derivative*, resp. with the *Gâteaux derivative*, if we assume w to be *Fréchet differentiable*, resp. to be *Gâteaux differentiable*.

In the light of Definition 4.2.16 and Remark 4.2.17, we have the following.

Proposition 4.2.18. *Let us consider the optimal control problem defined by (4.72) and (4.74), then the equation (4.77) provides the associated HJB problem. Moreover, if assumptions 4.2.2, 4.2.8, 4.2.13, and 4.2.15 hold true, then we have that the HJB equation (4.77) admits a unique mild solution, in the sense of the definition 4.2.16.*

Proof. The proof immediately follows exploiting [FT05, Th. 5.1]. □

As a direct consequence of Proposition 4.2.18, we provide a *synthesis* of the optimal control problem, by the following

Theorem 4.2.19. *Let assumptions 4.2.2, 4.2.8, 4.2.13, and 4.2.15 hold true. Let w be a mild solution to the HJB equation (4.77), and chose ρ to be an element of the generalized directional gradient $\nabla^G w$. Then, for all ACS, we have that $J(t_0, \mathbf{X}_0, z) \geq w(t_0, \mathbf{X}_0)$, and the equality holds if and only if the following feedback law is verified by z and \mathbf{u}^z*

$$z(t) = \Gamma(t, \mathbf{X}^z(t), G(t, \rho(t, \mathbf{X}^z(t))), \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T]. \quad (4.78)$$

then there also exists at least one ACS, such that

$$\bar{z}(t)\gamma(t, \mathbf{X}^z(t), \rho(t, \mathbf{X}^z(t))), \quad \mathbb{P} - a.s. \text{ for a.a. } t \in [t_0, T].$$

Eventually, we have that $\mathbf{X}^{\bar{z}}$ is a mild solution to equation (4.72).

Part III

Rigorous asymptotic expansions

5 | Small noise asymptotic expansion

The present chapter is taken from [ASG16, CDP15e].

Abstract

In this chapter we study the small noise asymptotic expansions for certain classes of local volatility models arising in finance. We provide explicit expressions for the involved coefficients as well as accurate estimates on the remainders. Moreover, we perform a detailed numerical analysis, with accuracy comparisons, of the obtained results by mean of the standard Monte Carlo technique as well as exploiting the polynomial Chaos Expansion approach.

We also present rigorous small noise expansion results for a Lévy perturbed Vasicek model. Estimates for the remainders as well as an application to ZCB pricing are also provided.

5.1 Asymptotic expansion for some local volatility models arising in finance

In the present paper we shall provide small noise asymptotic expansions for some local volatility models (LVMs) arising in finance. Our approach is based on the rigorous results on asymptotic expansions for solutions of finite dimensional SDE's obtained in [AS15] (following the approach proposed in [Gar85, Sec.6.2]); extensions to a class of SPDE's and infinite dimensional SDE's were presented in [ADPM11, ADPMS14a, ADPMS14b]. In particular we consider underlyings whose behaviour is characterized by a stochastic volatility term of *small* amplitude ϵ with respect to which we perform a formal, based on [Gar85, Sec. 6.2], resp. an asymptotic, based on [AS15], expansion. The latter implies that the equation characterizing the particular LVM of interest is approximated by a finite recursive system of a number N of linear equations with random coefficients. We then exploit the solutions of the latter system to provide a formal, resp. an asymptotic, approximation of smooth functions of the original solution for the particular LVM of interest. In a similar way we derive the corresponding approximation for the expected value of the related option price in the *risk neutral setting*. Errors estimates and explicit expressions for the involved approximations

are also provided for some specific cases, together with a detailed numerical analysis.

We would like to recall that LVMs are commonly used to analyse options markets where the underlying volatility strongly depends on the level of the underlying itself. Let us mention that although time-homogeneous local volatilities are supposedly inconsistent with the dynamics of the equity index implied volatility surface, see, e.g., [MH14], some authors, see, e.g., [Cré04], claim that such models provide the best average hedge for equity index options.

Let us also note that, particularly during recent years, different asymptotic expansions approaches to other particular problems in mathematical finance have been developed, see, e.g., [AL13, BGM09, CDP15e, FPS00, FT12, GHL⁺12, Gul12, Lüt04, TMH80, UY04, Yos03], see also [AHK12, IPW⁺09, PR⁺05] for applications to other areas.

Eventually we shall provide rigorous small noise expansion results for the Lévy perturbed Vasicek model. Our analysis is based on [AS15], in the setting proposed in [Gar85, Sec.6.2]. Let us underline that during recent years a wide range of small noise expansion techniques have been developed, particularly with respect to the *so called* Local Volatility Models (LVMs), see, e.g., [AL13, FPS00, Lüt04, UY04]. LVMs are commonly used to analyse options markets where the underlying volatility strongly depends on the level of the underlying itself, therefore LVMs are also widely accepted as tools to model interest-rate derivatives as is the case for the Vasicek model. The paper is organized as follows: in Sect. 5.2 we present the approach developed in [AS15] is presented and then applied,

The present paper is organized as follows: in Sect. 5.2 the basic general asymptotic expansions approach, based on [AS15] is presented. Then, in Sec. 5.4 we apply the aforementioned results to important examples in financial mathematics. In particular in Sec. 5.4 we study a perturbation up to the first order around the Black-Scholes model as well as a correction with jumps for the case of a generic smooth volatility function f . We then give more detailed results for the case of an exponential volatility function f , in 5.3.1 with Brownian motion driving, in 5.3.2 with an additional jump term. In 5.3.3 we shall present detailed corresponding results for the case of a polynomial volatility function f , in 5.3.4 we treat the case of corrections for f being a polynomial and the noise containing jumps. In Sect. 5.3.5 we apply the results of Sect. 5.4 to provide corrections up to first order of European call options in Black-Scholes models, with stochastic interest rate keeping Brownian motion as a driving process. Corresponding results for the case where an additional Poisson driving process is added are presented in 5.3.5, also in Sect. 5.4, to provide order corrections to both the Vasicek model and its zero coupon bond price. To validate the expansion we have performed, we present their numerical implementations obtained by exploiting the *Polynomial Chaos Expansion* approach, see the Appendix, as well as the standard *Monte Carlo* technique, also providing a detailed comparison between the twos in terms of accuracy.

5.2 The asymptotic expansion

5.2.1 The general setting

Let us consider the following stochastic differential equation (SDE), indexed by a parameter $\epsilon \geq 0$

$$\begin{cases} dX_t^\epsilon = \mu^\epsilon(X_t^\epsilon) dt + \sigma^\epsilon(X_t^\epsilon) dL_t, \\ X^\epsilon(0) = x_0^\epsilon \in \mathbb{R}, \quad t \in [0, \infty) \end{cases}; \quad (5.1)$$

where L_t , $t \in [0, \infty)$, is a Lévy process of jump diffusion type, subject to some restrictions which will be specified later on and $\mu^\epsilon : \mathbb{R}^d \rightarrow \mathbb{R}$, $\sigma^\epsilon : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ are Borel measurable functions for any $\epsilon \geq 0$ satisfying some additional technical conditions in order to have existence and uniqueness of strong solutions of (5.1), e.g., locally Lipschitz and sublinear growth at infinity, see, e.g., [App09, Lud73, MR15a, GS79, IPW+09, Sch04]. If the Lévy process L_t has a jump component, then X_t^ϵ in eq. (5.1) has to be understood as $X_{t-}^\epsilon := \lim_{s \uparrow t} X_s^\epsilon$, see, e.g., [MR15a] for details.

Hypothesis 5.2.1. Let us assume that:

- (i) $\mu^\epsilon, \sigma^\epsilon \in C^{k+1}(\mathbb{R})$ in the space variable, for any fixed value $\epsilon \geq 0$ and for all $k \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$;
- (ii) the maps $\epsilon \mapsto \alpha^\epsilon(x)$, where $\alpha = \mu, \sigma$, are in $C^M(I)$ in ϵ , for some $M \in \mathbb{N}$, for every fixed $x \in \mathbb{R}$ and where $I := [0, \epsilon_0]$, $\epsilon_0 > 0$.

Our goal is to show that under hypothesis 5.2.1 and some further conditions on μ^ϵ and σ^ϵ (needed for the construction of the coefficients X_t^i , $i = 0, 1, \dots, N$ appearing in (5.2) to be discussed below), a solution X_t^ϵ of equation (5.1) can be represented as a power series with respect to the parameter ϵ , namely

$$X_t^\epsilon = X_t^0 + \epsilon X_t^1 + \epsilon^2 X_t^2 + \dots + \epsilon^N X_t^N + R_N(t, \epsilon), \quad (5.2)$$

where $X^i : [0, \infty) \rightarrow \mathbb{R}$, $i = 0, \dots, N$, are continuous functions, while $|R_N(t, \epsilon)| \leq C_N(t)\epsilon^{N+1}$, $\forall N \in \mathbb{N}$ and $\epsilon \geq 0$, for some $C_N(t)$ independent of ϵ , but in general dependent of randomness, through $X_t^0, X_t^1, \dots, X_t^N$. The functions X_t^i are determined recursively as solutions of random differential equations in terms of the X_t^j , $j \leq i-1$, $\forall i \in \{1, \dots, N\}$.

Before giving the proof of the validity of the expression in eq. (5.2), let us recall the following result, see, e.g. [GM10].

Lemma 5.2.2. *Let f be a real (resp. complex) valued function in $C^{M+1}(B(x_0, r))$, $r > 0$, $x_0 \in \mathbb{R}$ for some $M \in \mathbb{N}_0$, where $(B(x_0, r))$ denotes the ball of center x_0 and radius r .*

Then for any $x \in B(x_0, r)$ the following Taylor expansion formula holds

$$f(x) = \sum_{p=0}^M \frac{D^p f(x_0)}{p!} (x - x_0)^p + R_M(D^{M+1} f(x_0, x)),$$

with $D^p f(x_0) := D^p f(x)|_{x=x_0}$ the p -th derivative at x_0 and

$$R_M(f^{(M+1)}(x_0, x)) := \frac{(M+1)(x-x_0)^{M+1}}{(M+1)!} \int_0^1 (1-s)^M D^{M+1} f(x_0 + s(x-x_0)) ds.$$

Moreover setting

$$C_M(x_0, x) := \frac{M+1}{(M+1)!} \int_0^1 (1-s)^M D^{M+1} f(x_0 + s(x-x_0)) ds,$$

we have

$$|C_M(x_0, x)| \leq \frac{M+1}{(M+1)!} \int_0^1 (1-s)^M \sup_{x \in B(x_0, r)} |D^{M+1} f(x_0 + s(x-x_0))| ds =: \tilde{C}_M(x_0) < +\infty \quad (5.3)$$

and also

$$|R_M(f^{(M+1)}(x_0, x))| \leq |C_M(x_0)||x - x_0|^{M+1} \leq \tilde{C}_M(x_0)|x - x_0|^{M+1}, \quad M \in \mathbb{N}_0.$$

With previous lemma in mind, let us then consider a function $f : \mathbb{R}^+ \times \mathbb{R} \rightarrow \mathbb{R}$, and $f_\epsilon(x) := f(\epsilon, x)$, $\epsilon \geq 0$, $x \in \mathbb{R}$. If we then suppose that for any fixed $x \in \mathbb{R}$, f is of class $C^{K+1}(I)$ in ϵ for some $K \in \mathbb{N}_0$, $I = [0, \epsilon_0]$, $\epsilon_0 > 0$, we can write the Taylor expansion of f around $\epsilon = 0$, w.r.t. $\epsilon \in I$ for any fixed $x \in \mathbb{R}$, as follows

$$f_\epsilon(x) = \sum_{j=0}^K f_j(x) \epsilon^j + R_K^{f_\epsilon}(\epsilon, x), \quad (5.4)$$

where f_j is the j -th coefficient in the expansion provided by Lemma 5.2.2, while $\sup_x |R_K^{f_\epsilon}(\epsilon, x)| \leq C_{K,f} \epsilon^{K+1}$ for some $C_{K,f} > 0$, independent of ϵ . Assume in addition that $x \mapsto f_j(x)$ are in C^{M+1} , $j = 0, \dots, K$, for some $M \in \mathbb{N}_0$, then, applying Lemma 5.2.2 to the function f_j around $x = x_0$, we obtain

$$f_\epsilon(x) = \sum_{j=0}^K \epsilon^j \left[\sum_{\gamma=0}^M \frac{D^\gamma f_j(x_0)}{\gamma!} (x - x_0)^\gamma + R_M(f_j^{(M+1)}(x_0, x)) \right] + R_K^{f_\epsilon}(\epsilon, x), \quad (5.5)$$

with $R_M(f_j^{(M+1)}(x_0, x))$ estimated as in Lemma 5.2.2 (with f_j replacing f) and $R_K^{f_\epsilon}(\epsilon, x)$ as in (5.4).

Let us now take $x = x(\epsilon)$ assuming $\epsilon \mapsto x(\epsilon)$ in C^{N+1} , with $0 \leq \epsilon \leq \epsilon_0$, $0 < \epsilon_0 < 1$ and $x(0) = x_0 \in \mathbb{R}$. Then by Lemma 5.2.2

$$x(\epsilon) = \sum_{j=0}^N \epsilon^j x_j + R_N^x(\epsilon), \quad N \in \mathbb{N}_0, \quad x_j \in \mathbb{R}, j = 0, 1, \dots, N, \quad (5.6)$$

with f replaced by x , M replaced by N , x by ϵ , x_0 by 0, $D^{M+1}(f(\cdot))$ by $f^{(M+1)}(\cdot)$ and $R_M(f^{(M+1)}(x_0, x))$ by $R_N^x(\epsilon)$. In particular

$$|R_N^x(\epsilon)| \leq \tilde{C}_N(0) \epsilon^{N+1}, \quad (5.7)$$

with $\tilde{C}_N(0) \epsilon^{N+1}$ independent of ϵ .

Plugging (5.6) into (5.5) we get

$$\begin{aligned} f_\epsilon(x(\epsilon)) &= \sum_{j=0}^K \epsilon^j \left[\sum_{\gamma=0}^M \frac{D^\gamma f_j(x_0)}{\gamma!} (x(\epsilon) - x_0)^\gamma + R_M(f_j^{(M+1)}(x_0, x(\epsilon))) \right] + R_K^{f_\epsilon}(\epsilon, x(\epsilon)) = \\ &= \sum_{j=0}^K \epsilon^j \left[\sum_{\gamma \leq M} \frac{D^\gamma f_j(x_0)}{\gamma!} \left(\sum_{k=1}^N \epsilon^k x_k + R_N^x(\epsilon) \right)^\gamma + R_M(f_j^{(M+1)}(x_0, x(\epsilon))) \right] \\ &\quad + R_K^{f_\epsilon}(\epsilon, x(\epsilon)). \end{aligned} \quad (5.8)$$

The estimates on R_M , $R_K^{f_\epsilon}$ and R_N^x have been given above in Lemma 5.2.2, resp. after (5.4), resp. (5.7).

By Newton's formula we have that, $\forall \gamma \in \mathbb{N}_0$, the following holds

$$\left(\sum_{j=1}^N \epsilon^j x_j + R_N^x(\epsilon) \right)^\gamma = \sum_{*}^{\gamma} \frac{\gamma!}{\gamma_1! \dots \gamma_{N+1}!} \epsilon^{\gamma_1 + 2\gamma_2 + \dots + N\gamma_N} x_1^{\gamma_1} \dots x_N^{\gamma_N} (R_N^x(\epsilon))^{\gamma_{N+1}}, \quad (5.9)$$

where we have used the notation

$$\sum_{*}^{\gamma} = \sum_{\substack{\gamma_1, \dots, \gamma_{N+1} = 0 \\ \gamma_1 + 2\gamma_2 + \dots + N\gamma_N + \gamma_{N+1} = \gamma}}^{\gamma};$$

hence using (5.9) to rewrite (5.8) we obtain the following.

Lemma 5.2.3. *If, for $0 \leq \epsilon < \epsilon_0$, $\epsilon \mapsto x(\epsilon)$ is in $C^{N+1}(I)$, $I = [0, \epsilon_0]$, and $\epsilon \mapsto f_\epsilon(y)$ is $C^{K+1}(\mathbb{R})$ in $\epsilon \in I$ and for any $y \in \mathbb{R}$, $y \mapsto f_\epsilon(y)$ is in C^{M+1} , the following expansion in powers of ϵ holds:*

$$\begin{aligned} f_\epsilon(x(\epsilon)) &= \sum_{j=0}^K \epsilon^j \left[\sum_{\gamma=0}^M \frac{D^\gamma f_j(x_0)}{\gamma!} \sum_{*}^{\gamma} \frac{\gamma!}{\gamma_1! \dots \gamma_{N+1}!} \epsilon^{\gamma_1 + 2\gamma_2 + \dots + N\gamma_N} x_1^{\gamma_1} \dots x_N^{\gamma_N} (R_N^x(\epsilon))^{\gamma_{N+1}} \right. \\ &\quad \left. + R_M \left(f_j^{(M+1)}(x_0, x(\epsilon)) \right) \right] + R_K^{f_\epsilon}(\epsilon, x(\epsilon)), \end{aligned} \quad (5.10)$$

The estimates for the remainders are as follow

$$\begin{aligned} |R_N^x(\epsilon)| &\leq \tilde{C}_N(0) \epsilon^{N+1}, \\ R_M \left(f_j^{(M+1)}(x_0, x(\epsilon)) \right) &\leq \tilde{C}_M(x_0) |x - x_0|^{M+1}, \\ \sup_{x, \epsilon} |R_K^{f_\epsilon}(\epsilon, x)| &\leq C_{K,f}, \end{aligned}$$

with $\tilde{C}_N(0)$, $\tilde{C}_M(x_0)$ and $C_{K,f}$ independent of ϵ .

Taking eq. (5.10) into account, we can group all the terms with the same power $k \in \mathbb{N}_0$ of ϵ . Calling $[f_\epsilon(x(\epsilon))]_k$ the coefficient of ϵ^k , and using $k = j + \gamma$ with $j = 0, \dots, K$, $\gamma_1 + 2\gamma_2 + \dots + N\gamma_N = \gamma$ with $\gamma = 0, \dots, M$, we have the following, see, [AS15].

Proposition 5.2.4. *Let $x(\epsilon)$ be as in (5.6) let f_ϵ as in (5.4) with $f_j \in C^{M+1}$, $j = 0, \dots, K$. Then*

$$f_\epsilon(x(\epsilon)) = \sum_{k=0}^{K+M} \epsilon^k [f_\epsilon(x(\epsilon))]_k + R_{K+M}(\epsilon),$$

with $|R_{K+M}(\epsilon)| \leq C_{K+M} \epsilon^{K+M+1}$, for some constant $C_{K+M} \geq 0$, independent of ϵ , $0 \leq \epsilon \leq \epsilon_0$, and coefficients $[f_\epsilon(x(\epsilon))]_k$ defined by

$$\begin{aligned} [f_\epsilon(x(\epsilon))]_0 &= f_0(x_0); \\ [f_\epsilon(x(\epsilon))]_1 &= Df_0(x_0)x_1 + f_1(x_0); \\ [f_\epsilon(x(\epsilon))]_2 &= Df_0(x_0)x_2 + \frac{1}{2}D^2f_0(x_0)x_1^2 + Df_1(x_0)x_1 + f_2(x_0); \\ [f_\epsilon(x(\epsilon))]_3 &= Df_0(x_0)x_3 + \frac{1}{6}D^3f_0(x_0)x_1^3 + Df_1(x_0)x_2 + Df_2(x_0)x_1 + D^2f_1(x_0)x_1^2 + f_3(x_0). \end{aligned}$$

The general case has the following form

$$[f_\epsilon(x(\epsilon))]_k = Df_0(x_0)x_k + \frac{1}{k!}D^k f_0(x_0)x_1^k + f_k(x_0) + B_k^f(x_0, x_1, \dots, x_{k-1}), \quad k = 1, \dots, K+M \quad (5.11)$$

where B_k^f is a real function depending on $(x_0, x_1, \dots, x_{k-1})$ only.

5.2.2 The asymptotic character of the expansion of the solution X_t^ϵ of the SDE in powers of ϵ

Theorem 5.2.5. *Let us assume that the coefficients α^ϵ , $\alpha = \mu, \sigma$, of the stochastic differential equation (5.1) are in $C^{K\alpha}(I)$ as functions of ϵ , $\epsilon \in I = [0, \epsilon_0]$, $\epsilon_0 > 0$, and in $C^{M\alpha}(\mathbb{R})$ as functions of x . Let us also assume that α^ϵ are such that there exists a solution X_t^ϵ in the probabilistic strong, resp. weak sense of (5.1) and that the recursive system of random differential equations*

$$dX_t^j = [\mu^\epsilon(X_t^\epsilon)]_j dt + [\sigma^\epsilon(X_t^\epsilon)]_j dL_t, \quad j = 0, 1, \dots, N, \quad t \geq 0,$$

has a unique solution.

Then there exists a sequence $\epsilon_n \in (0, \epsilon_0]$, $\epsilon_0 > 0$, $\epsilon_n \downarrow 0$ as $n \rightarrow \infty$ such that $X_t^{\epsilon_n}$ has an asymptotic expansion in powers of ϵ_n , up to order N , in the following sense:

$$X_t^{\epsilon_n} = X_t^0 + \epsilon_n X_t^1 + \dots + \epsilon_n^N X_t^N + R_N(\epsilon_n, t),$$

with

$$\text{st-lim}_{\epsilon_n \downarrow 0} \frac{\sup_{s \in [0, t]} |R_N(\epsilon_n, s)|}{\epsilon_n^{N+1}} \leq C_{N+1},$$

for some deterministic $C_{N+1} \geq 0$, independent of $\epsilon \in I$, where *st-lim* stands for the limit in probability.

Proof. We proceed by slightly modifying the proof in [AS15] since we have to take care of the presence of the explicit dependence on ϵ of the drift coefficient.

We shall use the fact that

$$T_N(\epsilon, t) := \frac{\left[X_t^\epsilon - \sum_{j=0}^N \epsilon^j X_t^j \right]}{\epsilon^{N+1}}, \quad \epsilon \in (0, \epsilon_0],$$

satisfies a random differential equation of the form

$$\epsilon^{N+1} dT_N(\epsilon, t) = A_{N+1}^{\mu^\epsilon}(X_t^0, \dots, X_t^N, R^N(t, \epsilon)) dt + A_{N+1}^{\sigma^\epsilon}(X_t^0, \dots, X_t^N, R^N(t, \epsilon)) dL_t,$$

with coefficients $A_{N+1}^{\alpha^\epsilon}$, $\alpha = \mu, \sigma$ given by

$$A_{N+1}^{\alpha^\epsilon}(y_0, y_1, \dots, y_N, y) = \left[\alpha^\epsilon \left(\sum_{j=0}^N \epsilon^j y_j + \epsilon^{N+1} y \right) - \sum_{j=0}^N \epsilon^j \alpha_j(y_0, y_1, \dots, y_N) \right],$$

with α_j , $j = 0, 1, \dots, N$ the expansion coefficients of α^ϵ in powers of $\epsilon \in I$.

By Taylor's theorem one proves

$$\frac{1}{\epsilon^{N+1}} \sup_{s \in [0, t]} |A_{N+1}^{\alpha^\epsilon}(X_s^0, \dots, X_s^N, R_s^N(\epsilon))| \leq C_{N+1}, \quad \epsilon \in (0, \epsilon_0],$$

for some $C_{N+1} \geq 0$, independent of ϵ , $0 \leq \epsilon \leq \epsilon_0$.

From this one deduces that one can find a sequence $\epsilon_n \rightarrow 0$ as $n \rightarrow \infty$ s.t.

$$\text{st-lim}_{\substack{\epsilon_n \downarrow 0 \\ n \rightarrow \infty}} \frac{1}{\epsilon_n^{N+1}} \sup_{s \in [0, t]} |A_{N+1}^{\alpha \epsilon_n}(X_s^0, \dots, X_s^N, R_s^N(\epsilon_n))|$$

exists and it is bounded by C_{N+1} .

Under some assumptions on μ^ϵ , σ^ϵ and L it follows then from a theorem by Skorohod, on the continuous dependence of solutions of SDE's on the coefficients, see, e.g. [GS79], that

$$\text{st-lim}_{\substack{\epsilon_n \downarrow 0 \\ n \rightarrow \infty}} \sup_{s \in [0, t]} |T_N(\epsilon_n, s)|$$

exists and it is bounded by C_{N+1} , which proves the result.

See [AS15] for more details. □

Remark 5.2.6. It can be seen that in general the k -th equation for X_t^k in Th. 5.2.5 is a nonhomogeneous linear equation in X_t^k , but with random coefficients depending on X_t^0, \dots, X_t^{k-1} and with a random inhomogeneity depending on X_t^k . Thus it has the general form

$$\begin{aligned} dX_t^k = & f_k(X_t^0, \dots, X_t^{k-1}) X_t^k dt + g_k(X_t^0, \dots, X_t^{k-1}) dt \\ & + \tilde{g}_k(X_t^0) dL_t + h_k(X_t^0, \dots, X_t^{k-1}) X_t^k dL_t, \end{aligned} \tag{5.12}$$

for some continuous functions f_k, g_k, \tilde{g}_k and h_k .

Let us now look at particular cases.

Example 5.2.1. Let $\mu^\epsilon = (a + \epsilon b)x$ and $\sigma^\epsilon = (\sigma_0 + \epsilon \sigma_1)x$ with a, b, σ_0 and σ_1 some real constants. Applying Proposition 5.2.4 we get

$$\begin{aligned} X_t^0 &= x_0 + \int_0^t a X_s^0 ds + \int_0^t \sigma_0 X_s^0 dL_s, \\ X_t^1 &= \int_0^t a X_s^1 ds + \int_0^t b X_s^0 ds + \int_0^t \sigma_1 X_s^0 dL_t + \int_0^t \sigma_0 X_s^1 dL_t, \\ X_t^k &= \int_0^t a X_s^k ds + \int_0^t b X_s^{k-1} ds + \int_0^t \sigma_1 X_s^{k-1} dL_s + \int_0^t \sigma_0 X_t^k dL_t, k \geq 2. \end{aligned} \tag{5.13}$$

If we consider the special case of $\mu^\epsilon(x) = ax + b$, independent of ϵ , $\sigma^\epsilon(x) = cx + \epsilon \tilde{d}x$, for some real constants a, b, c and \tilde{d} , independent of ϵ , and where the Lévy process is taken to be a standard Brownian motion, $L_t = W_t$, then by eq. (5.11) we have that X_t^k satisfies a linear equation with constant coefficients for any $k \in \mathbb{N}$, thus applying standard results, see, e.g., [Lud73, Gar88], an explicit solution for X_t^k can be retrieved.

Let us describe this in the case where we have a set of K coupled linear stochastic equations with random coefficients of the form

$$\begin{cases} dX_t^k = [A^k(t)X_t^k + f^k(t)] dt + [B^k(t)X_t^k + g(t)] dW_t, \\ X_0^k = x_0^k \in \mathbb{R}, \quad t \geq 0 \end{cases} \tag{5.14}$$

where, for any $k = 1, \dots, K$, all the functions A^k, B^k, f^k and g are assumed to be Lipschitz and with at most linear growth. A solution of equation (5.14) is then given by

$$X_t^k = \sum_{k=0}^K \Phi_k(t) \left[\int_0^t \Phi_k^{-1}(s) (f^k(s) - B^k(s)g^k(s)) ds + \int_0^t \Phi_k^{-1}(s)g^k(s)dW_s \right] \quad (5.15)$$

where $\Phi_k(t)$ is the fundamental $K \times K$ matrix solution of the corresponding homogeneous equation, i.e. it is the solution of the problem

$$\begin{cases} d\Phi_k(t) = A^k(t)\Phi_k(t)dt + B^k(t)\Phi_k(t)dW_t, \\ \Phi_k(0) = I, t \geq 0, \end{cases}, \quad (5.16)$$

being I the unit $K \times K$ matrix.

Remark 5.2.7. In the case where $K = 1$ we have that Φ reduces to a scalar and is given by

$$\Phi(t) = \exp \left\{ \int_0^t \left(A(s) - \frac{1}{2}B^2(s) \right) ds + \int_0^t B(s)dW_s \right\}.$$

Still in the case $K = 1$ but with a more general noise, i.e. W_t in eq. (5.14) replaced by a Lévy process composed by a Brownian motion plus W_t a jump component expressed by \tilde{N} , eq. (5.16) is replaced by

$$\begin{cases} d\Phi(t) = A(t)\Phi(t)dt + B(t)\Phi(t)dW_t + \int_{\mathbb{R}_0} \Phi(t_-)C(t, x)\tilde{N}(dt, dx), \\ \Phi(0) = I, t \geq 0 \end{cases}. \quad (5.17)$$

with A, B and C Lipschitz and with at most linear growth, and where $\tilde{N}(dt, dx)$ is a Poisson compensated random measure to be understood in the following sense: $\tilde{N}(t, A) := N(t, A) - t\nu(A)$ for all $A \in \mathcal{B}(\mathbb{R}_0)$, $0 \notin \bar{A}$, with \bar{A} the closure of A , N being a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_0$ and $\nu(A) := \mathbb{E}(N(1, A))$, while $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$ and $\int_{\mathbb{R}_0} (|x|^2 \wedge 1)\nu(dx) < \infty$, ν is the Lévy measure to \tilde{N} , see, e.g. [App09, IPW⁺09, MR15a].

Denoting then eq. (5.17) for short as

$$d\Phi(t) = \Phi(t_-)dX(t), \quad (5.18)$$

with

$$dX(t) = A(t)dt + B(t)dW_t + \int_{\mathbb{R}_0} C(t, x)\tilde{N}(dt, dx),$$

we have then that the solution to eq. (5.18) is explicitly given by

$$\begin{aligned} \Phi(t) = \exp \left\{ 1 + \int_0^t \left(A(s) - \frac{1}{2}B^2(s) \right) ds + \int_0^t B(s)dW_s \right. \\ \left. + \int_{\mathbb{R}_0} C(s, x)\tilde{N}(ds, dx) \right\} \prod_{0 < s \leq t} (1 + \Delta X_s) e^{-\Delta X_s}, \end{aligned} \quad (5.19)$$

where $\Delta X(s) := X_s - X_{s-}$ is the jump at time $s \in (0, t]$. The stochastic process (5.19) is called *Doléans-Dade exponential* (or stochastic exponential) and it is usually denoted by $\Phi(t) = \mathcal{E}(X_t)$. The *Doléans-Dade exponential* has a wide use in finance since it is the natural extension to the Lévy case of the standard geometric Brownian motion, see, e.g., [Lud73, Gar85] for a more extensive treatment of the fundamental solution of the homogeneous equation for system of linear SDE's and [App09] for more details on the *Doléans-Dade exponential*.

5.3 Corrections around the Black-Scholes price (with Brownian, resp. Brownian plus jumps)

We shall study an asset S_t^ϵ evolving according to the particular stochastic differential equation (SDE) governing the Black-Scholes (BS) model, with the possible addition of some driving term determined by a compound Poisson process, see, e.g. [BS73, Sch04], resp. [BGM09, Mer76]. Our aim is to apply the theory developed in Sec. 5.2 in order to give corrections around the price given by the BS model for an option with terminal payoff Φ written on the underlying S_t^ϵ (Φ is a given real valued function assumed here to be sufficiently smooth). In particular, if we consider the return process defined as $X_t^\epsilon := \log S_t^\epsilon$ (S_t^ϵ being supposed to be strictly positive, at least almost surely) we have that the price $P(t, T)$ at time t of the option with final payoff Φ with maturity time T , $0 \leq t \leq T$, is given by

$$P(t, T) = \mathbb{E}^{\mathbb{Q}} \left[e^{r(T-t)} \Phi(X_T) \middle| \mathcal{F}_t \right], \tag{5.20}$$

where \mathbb{Q} is a relevant equivalent martingale measure, called in financial application *risk-neutral measure*, $\mathbb{E}^{\mathbb{Q}}[\cdot]$ the corresponding conditional expectation given the σ -algebra \mathcal{F}_t at time t associated with the underlying Brownian motion, $r > 0$ is the constant interest rate. We refer to, e.g., [BS73, DF01, CIJR85, Fil01, KK99, Sch04] for a general introduction to option pricing.

From Theorem 5.2.5 and using Lemma 5.2.3 we deduce that $\Phi(X_t^\epsilon)$ has an asymptotic expansion in powers of $\epsilon \in [0, \epsilon_0)$, $\epsilon_0 > 0$, in the sense of Theorem 5.2.5, of the form

$$\Phi(X_t^\epsilon) = \sum_{k=0}^H \epsilon^k [\Phi(X_t^\epsilon)]_k + R_H(\epsilon, t), \tag{5.21}$$

with

$$\sup_{s \in [0, t]} |R_H(\epsilon, s)| \leq C_{H+1}(t) \epsilon^{H+1},$$

for any $H \in \mathbb{N}$ and the coefficients can be computed from the expansions coefficients of X_t^ϵ , as discussed in section 5.2.

More concretely we will deal with two particular cases. In the first case we have an asset S^ϵ evolving according to a geometric Brownian motion with a small perturbation in the diffusion. Namely the asset evolves, in a risk neutral setting, according to

$$\begin{cases} dS_t^\epsilon = S_t^\epsilon [(\sigma_0 + \epsilon \sigma_1 \bar{f}(S_t^\epsilon)) dW_t] , \\ s_0 = s_0, t \geq 0, \end{cases}, \tag{5.22}$$

where $\sigma_0 \neq 0$ and σ_1 are real constants, $s_0 > 0$, and W_t is a \mathbb{Q} Brownian motion adapted to the filtration $(\mathcal{F}_t)_t$, $\bar{f}(S_t^\epsilon) := f(X_t^\epsilon)$ with f a given smooth function on \mathbb{R} . In particular the existence and uniqueness of a strong solution to equation (5.22) follows under the general assumption of $\bar{f} \in C^1$ from [McK69, Problem 3.3.2]. We have assumed σ_0 and σ_1 to be time independent for the sake of simplicity. The generalization to time dependent functions is an immediate generalization with no complication in the results developed in what follows.

Suppose, for all $t \geq 0$, $S_t^\epsilon > 0$ a.s. (which is the case if ϵ is sufficiently small). Applying Itô's lemma to $X_t^\epsilon := \log S_t^\epsilon$, we end up with the following evolution for X_t^ϵ , the return of

the asset price

$$X_t^\epsilon = x_0 - \int_0^t \left[\frac{\sigma_0^2}{2} + \epsilon \sigma_0 \sigma_1 f(X_s^\epsilon) + \epsilon^2 \frac{\sigma_1^2 f(X_s^\epsilon)^2}{2} \right] ds + \int_0^t [\sigma_0 + \epsilon \sigma_1 f(X_s^\epsilon)] dW_s, \quad (5.23)$$

where we have set $x_0 := \log s_0$.

Applying the results obtained in Sec. 5.2 and expanding eq. (5.23) to the second order in ϵ we get

$$\begin{aligned} X_t^0 &= x_0 - \frac{\sigma_0^2}{2}t + \sigma_0 W_t, & \text{with law } \mathcal{N}(x_0 + \mu t, \sigma_0^2 t), \\ X_t^1 &= - \int_0^t \sigma_0 \sigma_1 f(X_s^0) ds + \int_0^t \sigma_1 f(X_s^0) dW_s, \\ X_t^2 &= - \int_0^t \left(\frac{\sigma_1^2 f(X_s^0)^2}{2} + 2\sigma_0 \sigma_1 f'(X_s^0) X_s^1 \right) ds + \int_0^t \sigma_1 f'(X_s^0) X_s^1 dW_s, \end{aligned} \quad (5.24)$$

where $\mathcal{N}\left(-\frac{\sigma_0^2}{2}t, \sigma_0^2 t\right)$ denotes the Gaussian distribution of mean μt and variance $\sigma_0^2 t$, f' the derivative of f .

The second model we will deal with, following [Mer76, BGM09], is the previous one with an addition of a small compound Poisson process

$$Z_t = \sum_{i=1}^{N_t} J_i,$$

with N_t a standard Poisson process with intensity $\lambda > 0$ and $(J_i)_{i=1, \dots, N_t}$ being independent normally distributed random variables, namely such that

$$J_i \text{ has law } \mathcal{N}(\gamma, \delta^2).$$

We thus have that the Lévy measure $\nu(dz)$ of Z reads as

$$\nu(dz) = \frac{\lambda}{\sqrt{2\pi\delta}} e^{-\frac{(z-\gamma)^2}{2\delta^2}} dz, \quad z \in \mathbb{R},$$

and the cumulant function of Z is

$$\kappa(\zeta) = \lambda \left(e^{\gamma\zeta + \frac{\delta^2\zeta^2}{2}} - 1 \right).$$

In particular we assume the asset S^ϵ to evolve according to a geometric Lévy process with a small perturbation in the diffusion. Namely the asset evolves, in a risk neutral setting, according to

$$\begin{cases} dS_t^\epsilon = S_t^\epsilon \left[(\sigma_0 + \epsilon \sigma_1 \bar{f}(S_t^\epsilon)) dW_t + \epsilon \sum_{i=1}^{N_t} J_i \right], \\ S_0^\epsilon = s_0 > 0, t \geq 0, \end{cases} \quad (5.25)$$

Again the existence and uniqueness of a strong solution to equation (5.25) can be obtained with arguments similar to the ones used in [McK69, Problem 3.3.2] together with the properties of $\sum_{i=1}^{N_t} J_i$.

Proceeding as above, and applying Itô's lemma to $X_t^\epsilon := \log S_t^\epsilon$, we have that the log-return process X_t^ϵ evolves according to

$$\begin{aligned}
 X_t^\epsilon = & x_0 - \int_0^t \left[\frac{\sigma_0^2}{2} + \epsilon \sigma_0 \sigma_1 f(X_s^\epsilon) + \epsilon^2 \frac{\sigma_1^2 f(X_s^\epsilon)^2}{2} \right] ds + \epsilon \lambda t \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \\
 & + \int_0^t (\sigma_0 + \epsilon \sigma_1 f(X_s^\epsilon)) dW_s + \epsilon \sum_{i=1}^{N_t} J_i,
 \end{aligned} \tag{5.26}$$

for $\epsilon \in I = [0, \epsilon_0]$, $\epsilon_0 > 0$.

In the present case it is more tricky to deal with the risk neutral probability measure \mathbb{Q} . Under suitable assumptions on the coefficients and noise one can assure the existence (but not necessarily the uniqueness) of an equivalent probability measure \mathbb{Q} . We will assume the process (5.26) to evolve under a risk-neutral measure \mathbb{Q} , see, e.g. [App09].

In particular we will use two specific forms for the function f , that is an exponential function and a polynomial function. The former is of special interest for its general application to integral transforms, such as Fourier or Laplace transforms, see, e.g. Section 5.3.1, Remark 5.3.2. The latter mimics a polynomial volatility process (these type of processes have been widely used in finance since they can be easily implemented, see, e.g. [CFR13] and reference therein).

5.3.1 A correction given by an exponential function

Let us consider the first model described by equations (5.23) and (5.24), i.e. an asset S^ϵ evolving according to a geometric Brownian motion under the unique risk neutral probability measure \mathbb{Q} , recalling that $X_t^\epsilon = \log S_t^\epsilon$. Let us first look at the particular case $f(x) = e^{\alpha x}$, for some $\alpha \in \mathbb{R}$. We take into account the particular case of an exponential function due to the fact that it can be easily extended to the much more general case where the function f can be written as a Fourier transform or a Laplace transform of some bounded measure on the real line, as it will be further discussed in Rem. 5.3.2 below. We then get the following proposition.

Proposition 5.3.1. *Let us consider the SDE (5.23) in the particular case where $f(x) = e^{\alpha x}$, for some $\alpha \in \mathbb{R}_0$, $\sigma_0 \in \mathbb{R}_0$.*

Then the following expansion $X_t^\epsilon = X_t^0 + \epsilon X_t^1 + \epsilon^2 X_t^2 + R_2(\epsilon, t)$ holds, where the coefficients are given by

$$\begin{aligned}
 X_t^0 &= x_0 + \mu t + \sigma_0 W_t, \quad \text{with law } \mathcal{N}(x_0 + \mu t, \sigma_0^2 t); \\
 X_t^1 &= \int_0^t K_\alpha e^{\alpha X_s^0} ds + \frac{\sigma_1}{\alpha \sigma_0} \left(e^{\alpha X_t^0} - 1 \right); \\
 X_t^2 &= C_\alpha^1 \int_0^t e^{2\alpha X_s^0} ds + C_\alpha^2 e^{\alpha X_t^0} \int_0^t e^{\alpha X_s^0} ds + C_\alpha^3 \int_0^t e^{\alpha X_s^0} ds \\
 &\quad + C_\alpha^4 \int_0^t e^{\alpha X_s^0} \int_0^s e^{\alpha X_r^0} dr ds + C_\alpha^5 e^{2\alpha X_t^0} + C_\alpha^6 e^{\alpha X_t^0} + C_\alpha^7,
 \end{aligned} \tag{5.27}$$

with

$$\begin{aligned} K_\alpha &:= -\sigma_1\left(\frac{\mu}{\sigma_0} + \frac{\alpha\sigma_0}{2} + \sigma_0\right), \quad C_\alpha^1 := -\sigma_1^2\left(\frac{5}{2} + \frac{\mu}{\sigma_0^2} + \alpha + \frac{K_\alpha}{\sigma_0\sigma_1}\right), \quad C_\alpha^2 := K_\alpha\frac{\sigma_1}{\sigma_0}, \\ C_\alpha^3 &:= \sigma_1^2\left(\frac{\mu}{\sigma_0^2} + \frac{\alpha}{2} + 2\right), \quad C_\alpha^4 := -K_\alpha\sigma_1\alpha\left(2\sigma_0 + \frac{\mu}{\sigma_0} + \frac{\alpha\sigma_0}{2}\right), \\ C_\alpha^5 &:= \frac{\sigma_1^2}{2\alpha\sigma_0^2}, \quad C_\alpha^6 := -\frac{\sigma_1^2}{\alpha\sigma_0^2}, \quad C_\alpha^7 := \frac{\sigma_1^2}{2\alpha\sigma_0^2}, \quad \mu = -\frac{\sigma_0^2}{2}. \end{aligned}$$

Furthermore $R_2(\epsilon, t)$ satisfies the bound

$$\text{st-lim}_{\epsilon_n \downarrow 0} \frac{\sup_{s \in [0, t]} |R_2(\epsilon, s)|}{\epsilon_n^3} \leq C_3,$$

for some subsequence $\epsilon_n \downarrow 0$ and with some constant $C_3 \geq 0$.

Proof. The proof consists in a repeated application of the Itô formula and the stochastic Fubini theorem.

In fact substituting $f(x) = e^{\alpha x}$ into system (5.24) we immediately obtain

$$\begin{aligned} X_t^0 &= x_0\mu t + \sigma_0 W_t, \quad \text{with law } \mathcal{N}(x_0 + \mu t, \sigma_0^2 t); \\ X_t^1 &= -\int_0^t \sigma_0\sigma_1 e^{\alpha X_s^0} ds + \int_0^t \sigma_1 e^{\alpha X_s^0} dW_s; \\ X_t^2 &= -\int_0^t \left(\frac{\sigma_1^2}{2} e^{2\alpha X_s^0} + 2\sigma_0\sigma_1\alpha e^{\alpha X_s^0} X_s^1\right) ds + \int_0^t \sigma_1\alpha e^{\alpha X_s^0} X_s^1 dW_s. \end{aligned} \tag{5.28}$$

To compute X_t^1 we apply Itô's lemma to the function $g(X_t^0) = e^{\alpha X_t^0}$ to get

$$e^{\alpha X_t^0} = 1 + \int_0^t (e^{\alpha X_s^0} \alpha \mu + \frac{\alpha^2}{2} \sigma_0^2 e^{\alpha X_s^0}) ds + \int_0^t e^{\alpha X_s^0} \alpha \sigma_0 dW_s. \tag{5.29}$$

Expressing the latter integral involving dW_s by the other terms in eq. (5.29) and substituting it in the stochastic integral of X_t^1 in the system (5.28) we get the result for X_t^1 in eq. (5.27).

In order to derive the expression for X_t^2 we use again Itô's lemma, in particular eq. (5.29), getting from (5.28)

$$\begin{aligned} X_t^2 &= -\int_0^t \left(\frac{\sigma_1^2}{2} e^{2\alpha X_s^0} + 2\sigma_0\sigma_1\alpha e^{\alpha X_s^0} X_s^1\right) ds + \int_0^t \alpha\sigma_1 e^{\alpha X_s^0} X_s^1 dW_s = \\ &= -\int_0^t \sigma_1^2 \left(2\alpha + \frac{1}{2}\right) e^{2\alpha X_s^0} ds + \int_0^t 2\alpha\sigma_1^2 e^{\alpha X_s^0} ds - \int_0^t \int_0^s 2K_\alpha\sigma_1\sigma_0\alpha e^{\alpha X_s^0} e^{\alpha X_r^0} dr ds \\ &+ \underbrace{\int_0^t \frac{\sigma_1^2\alpha}{\sigma_0} e^{2\alpha X_s^0} dW_s}_{(1)} - \underbrace{\int_0^t \frac{\alpha\sigma_1^2}{\sigma_0} e^{\alpha X_s^0} dW_s}_{(2)} + \underbrace{\int_0^t K_\alpha\alpha\sigma_1 e^{\alpha X_s^0} \int_0^s e^{\alpha X_r^0} dr dW_s}_{(3)}. \end{aligned}$$

For the terms (1) and (2) we use eq. (5.29), resp. Itô's lemma applied to the function $g(X_t^0) = e^{2\alpha X_t^0}$, as before to replace the stochastic integral by an integral against Lebesgue

measure. In order to treat the term (3) we use the stochastic Fubini theorem, see, e.g. Th. 6.2 in [Fil01], to get

$$(3) = \frac{K_\alpha \sigma_1}{\sigma_0} \int_0^t \int_0^s \alpha \sigma_0 e^{\alpha X_s^0} e^{\alpha X_r^0} dr dW_s = \frac{K_\alpha \sigma_1}{\sigma_0} \int_0^t e^{\alpha X_r^0} \int_r^t \alpha \sigma_0 e^{\alpha X_s^0} dW_s dr.$$

Using the expression for the integral in dW_s coming from (5.29) we then get

$$\begin{aligned} (3) &= \frac{K_\alpha \sigma_1}{\sigma_0} \int_0^t e^{\alpha X_r^0} \int_r^t \alpha \sigma_0 e^{\alpha X_s^0} dW_s dr = \\ &= \frac{K_\alpha \sigma_1}{\sigma_0} e^{\alpha X_t^0} \int_0^t e^{\alpha X_s^0} ds - \frac{K_\alpha \sigma_1}{\sigma_0} \int_0^t e^{2\alpha X_s^0} ds - \frac{K_\alpha \sigma_1}{\sigma_0} \left(\alpha \mu + \frac{\alpha^2 \sigma_0^2}{2} \right) \times \\ &\quad \times \int_0^t \int_0^s e^{\alpha X_s^0} e^{\alpha X_r^0} dr ds. \end{aligned}$$

Substituting now everything into the original system (5.28), rearranging and grouping the integrals of the same type we get the desired result in (5.27).

The estimate on the remainder is a consequence of Theorem 5.2.5. □

Remark 5.3.2. The particular choice of $f(x) = e^{\alpha x}$ can easily be extended to any real function which can be written as a Fourier transform, resp. Laplace transform, $f(x) = \int_{\mathbb{R}_0} e^{ixy} \lambda(d\alpha)$, resp. $f(x) = \int_{\mathbb{R}_0} e^{\alpha x} \lambda(d\alpha)$, of some positive measure λ on \mathbb{R}_0 (e.g. a symmetric probability measure) resp. which has finite Laplace transform. Formula (5.27) holds with $K_\alpha e^{\alpha X_\tau^0}$ replaced by $\int_{\mathbb{R}_0} K_\alpha e^{i\alpha X_\tau^0} \lambda(d\alpha)$, resp. $\int_{\mathbb{R}_0} K_\alpha e^{\alpha X_\tau^0} \lambda(d\alpha)$, which are finite if, e.g. $\int_{\mathbb{R}_0} |K_\alpha| \lambda(d\alpha) < \infty$, resp. λ has, e.g., compact support. In fact eq. (5.29) gets replaced by

$$\begin{aligned} \int_{\mathbb{R}_0} e^{\alpha X_t^0} \lambda(d\alpha) &= 1 + \int_{\mathbb{R}} \left[\int_0^t \left(e^{\alpha X_s^0} \alpha \mu + \frac{\alpha^2}{2} \sigma_0^2 e^{\alpha X_s^0} \right) ds \right] \lambda(d\alpha) \\ &\quad + \int_{\mathbb{R}} \left[\int_0^t e^{\alpha X_s^0} \alpha \sigma_0 dW_s \right] \lambda(d\alpha). \end{aligned} \tag{5.30}$$

By repeating the steps used before and exploiting again the Stochastic Fubini's theorem we get the statements in Prop. 5.3.1 extended to these more general cases.

If we assume the payoff function $x \mapsto \Phi(x)$ to be smooth, $x \in \mathbb{R}_+$, we can expand $\Phi(X_t^\epsilon)$ in powers of ϵ using the formulae in Prop. 5.2.4. Then, exploiting eq. (5.86) with $H = 1$, i.e. stopping at the first order, we get

$$\Phi(X_t^\epsilon) = \Phi(X_t^0) + \epsilon \Phi'(X_t^0) X_t^1 + R_1(\epsilon, t), \tag{5.31}$$

with $\sup_{s \in [0, t]} |R_1(\epsilon, s)| \leq \tilde{C}(s) \epsilon^2$, for some \tilde{C} independent of ϵ (Φ' is the derivative of Φ).

Calling Φ_1 the terms on the r.h.s. in eq. (5.31) minus the reminder term $R_1(\epsilon, t)$ we get that the corresponding corrected fair price $Pr^1(0; T)$, up to the first order in ϵ , of an option written on the underlying $S_t^\epsilon := e^{X_t^0}$ at time $t = 0$ with maturity T , reads as follow

$$\begin{aligned} Pr^1(0; T) &= e^{-rT} \mathbb{E}^\mathbb{Q} [\Phi_1(X_T^\epsilon)] = e^{-rT} \mathbb{E}^\mathbb{Q} [\Phi(X_T^0) + \epsilon \Phi'(X_T^0) X_T^1] = \\ &= Pr_{BS} + \epsilon e^{-rT} \mathbb{E}^\mathbb{Q} [\Phi'(X_T^0) X_T^1], \end{aligned} \tag{5.32}$$

where Pr_{BS} stands for the standard B-S price with underlying $S_t^0 := e^{X_t^0}$, see, e.g [BS73].

This formula yields thus, for a smooth payoff function, the corrected price up to the first order, with an error term related to the "full price" and bounded in modulus by $C_2\epsilon^2$ for a constant $C_2 \geq 0$ independent of ϵ .

Remark 5.3.3. It is worth to recall that the payoff function usually fails to be smooth such as in the case of European call options where $\Phi(x) = (e^x - K)^+$, $K > 0$ being the strike price. The latter payoff function always presents a point of non differentiability at $e^X = K$. Anyhow we can, as generalized functions, consider a smoothed version of the payoff function, namely $\Phi_h := \Phi * \rho_h$, with ρ_h some smooth kernel s.t. $\Phi_h \rightarrow \Phi$ as $h \rightarrow \infty$ in distributional sense. With the smoothed payoff function Φ_h , eq. (5.32) is well defined. In particular the first derivative appearing in eq. (5.32) is given by a regularized version of $\mathbb{1}_{[x > \ln K]}(x)$. Heuristically, interchanging the limits involved in the expansion with the removing of regularization we can look at $Pr^1(0, T)$ as given by (5.32) also in the case of the payoff function $\Phi(x) = (e^x - K)^+$, $x \in \mathbb{R}$, as approximation of the price, with $\Phi'(x) = \mathbb{1}_{[x > \ln K]}(x)$ given as above. This lead us to the following.

Proposition 5.3.4. *Let us consider the particular case of an European call option Φ with payoff given by $\Phi(X_T^\epsilon) = \max\{e^{X_T^\epsilon} - K, 0\} =: (e^{X_T^\epsilon} - K)_+$, K being the strike price. Then the approximated price up to the first order, $Pr^1(0; T)$, in the sense of Remark 5.3.3, is explicitly given by*

$$Pr^1(0; T) = P_{BS} + \epsilon \mathcal{K}_1 s_0^{\alpha+1} I_1(s, T, \alpha) - \epsilon \mathcal{K}_2 s_0 N(d_1) + \epsilon \mathcal{K}_3 s_0^{\alpha+1} N(d(2\alpha + 1)), \quad (5.33)$$

with $N(x)$ the cumulative function of the standard Gaussian distribution and

$$d(\alpha) = \frac{1}{\sigma_0 \sqrt{T}} \left(\log \frac{s_0}{K} + \left(r - \frac{\sigma_0^2}{2} \alpha \right) T \right), \quad d_1 := d(1), \quad d_2 := (d_1 + \sigma_0 \sqrt{T}),$$

$$\mathcal{K}_1 = K_\alpha e^{-\frac{\sigma_0^2}{2} T}, \quad \mathcal{K}_2 = \frac{\sigma}{\alpha \sigma_0}, \quad \mathcal{K}_3 = \frac{\sigma_1}{\alpha \sigma_0} e^{\frac{\sigma_0^2}{2} T \alpha(\alpha+1) + \alpha r T},$$

$$I_1(s, T, \alpha) = \int_0^T e^{\alpha \mu s} \int_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{\{x+y > \sqrt{T} d_2\}} e^{\sigma_0 x} e^{(1+\alpha)\sigma_0 y} \phi(x, 0, T-s) \phi(y, 0, s) dx dy ds,$$

where we have denoted by $\phi(x; \mu, \sigma)$ the density function of the normal distribution with mean μ and variance σ .

Proof. Given the exponential function $f(x) = e^{\alpha x}$, where $\alpha \in \mathbb{R}$, the approximated price up to the first order, $Pr^1(0; T)$ of an European call option with payoff function $\Phi(X_T^\epsilon) = (e^{X_T^\epsilon} - K)_+$ is

$$\begin{aligned} Pr^1(0; T) &= P_{BS} + \epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} [\Phi'(X_T^0) X_T^1] = \\ &= P_{BS} + \epsilon e^{-rT} \left\{ \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{[X_0^T > \ln(K)]} e^{X_T^0} \int_0^T K_\alpha e^{\alpha X_s^0} ds \right] + \right. \\ &\quad \left. - \mathcal{K}_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{[X_0^T > \ln(K)]} e^{X_T^0} \right] + \mathcal{K}_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{[X_0^T > \ln(K)]} e^{X_T^0} e^{\alpha X_T^0} \right] \right\}, \end{aligned} \quad (5.34)$$

where P_{BS} is the standard B-S price with underlying $S_t^0 = e^{X_t^0}$.

Let us first compute the integral

$$\epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{[X_T^0 > \ln(K)]} e^{X_T^0} \int_0^T K_{\alpha} e^{\alpha X_s^0} ds \right]$$

By means of Fubini Theorem, we can exchange the expectation with respect to the integration in time so that we obtain

$$\epsilon e^{-rT} K_{\alpha} \int_0^T \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{[X_t^0 > \ln(K)]} e^{X_t^0} e^{\alpha X_s^0} \right] ds . \quad (5.35)$$

From the definition of X_T^0 and X_s^0 , for every fixed $0 < s < T$, we have

$$\begin{aligned} X_T^0 &= x_0 + \mu T + \sigma_0 W_T , \\ X_s^0 &= x_0 + \mu s + \sigma_0 W_s , \end{aligned}$$

are two correlated random variables, by means of the Wiener processes involved. By algebraic manipulation let us define $W_T = W_T - W_s + W_s$, where $X := W_T - W_s$ is $\mathcal{N}(0, T - s)$ independent with respect to W_s . Then $X_T^0 = x_0 + \mu T + \sigma_0 X + \sigma_0 W_s$ and (5.76) becomes

$$\begin{aligned} \epsilon e^{-rT} K_{\alpha} \int_0^T \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{\{\sigma_0 X + \sigma_0 W_s > \ln(\frac{K}{s_0}) - \mu T\}} e^{(1+\alpha)x_0 + \mu T} e^{\alpha \mu s} e^{\sigma_0 X} e^{(1+\alpha)\sigma_0 W_s} \right] ds = \\ = \epsilon e^{-rT} K_{\alpha} s_0^{(1+\alpha)} e^{rT} e^{-\frac{\sigma_0^2}{2}T} \int_0^T e^{\alpha \mu s} \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{\{\sigma_0 X + \sigma_0 W_s > \ln(\frac{K}{s_0}) - \mu T\}} e^{\sigma_0 X} e^{(1+\alpha)\sigma_0 W_s} \right] ds . \end{aligned}$$

The expectation with respect to the risk-neutral measure can be exchanged with the time integration. Moreover by exploiting the independence of X and W_s , we get the final result

$$\begin{aligned} \epsilon K_{\alpha} s_0^{(1+\alpha)} e^{-\frac{\sigma_0^2}{2}T} \int_0^T e^{\alpha \mu s} \int_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{\{x+y > -\sqrt{T}d_2\}} e^{\sigma_0 x} e^{(1+\alpha)\sigma_0 y} \phi(x, 0, T-s) \phi(y, 0, s) dx dy ds = \\ = \epsilon s_0^{(1+\alpha)} \mathcal{K}_1 \int_0^T e^{\alpha \mu s} \int_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{\{x+y > -\sqrt{T}d_2\}} e^{\sigma_0 x} e^{(1+\alpha)\sigma_0 y} \phi(x, 0, T-s) \phi(y, 0, s) dx dy ds . \end{aligned}$$

Then we have from the definition of X_T^0

$$\begin{aligned} \mathbb{E} \left[\mathbb{1}_{[X_T^0 > \ln(K)]} e^{X_T^0} \right] &= \int_{x > -d_2} e^{x_0 + \mu T + \sigma_0 \sqrt{T}x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\ &= s_0 e^{rT} e^{-\frac{\sigma_0^2}{2}T} \int_{x > -d_2} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x}{\sqrt{2}} - \frac{\sigma_0 \sqrt{T}}{\sqrt{2}}\right)^2} e^{\frac{\sigma_0^2 T}{2}} dx = \\ &= s_0 e^{rT} \int_{x > -d_2} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x}{\sqrt{2}} - \frac{\sigma_0 \sqrt{T}}{\sqrt{2}}\right)^2} dx . \end{aligned} \quad (5.36)$$

By setting $y = x - \sigma_0 \sqrt{T}$, the integral in (5.79) reads as

$$\mathbb{E} \left[\mathbb{1}_{[X_T^0 > \ln(K)]} e^{X_T^0} \right] = s_0 e^{rT} \int_{y > -d_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = s_0 e^{rT} N(d_1) . \quad (5.37)$$

Eventually by multiplying by $-\epsilon e^{-rT} \mathcal{K}_2$, we obtain

$$-\epsilon e^{-rT} \mathcal{K}_2 \mathbb{E} \left[\mathbb{1}_{[X_0^T > \ln(K)]} e^{X_T^0} \right] = -\epsilon \mathcal{K}_2 s_0 N(d_1) \quad (5.38)$$

Let us now compute the last term in the bracket $\{ \}$ in (5.34). We have

$$\begin{aligned} \mathcal{K}_2 \mathbb{E} \left[\mathbb{1}_{[X_0^T > \ln(K)]} e^{X_T^0} e^{\alpha X_T^0} \right] &= \mathcal{K}_2 \int_{x_0 + \mu T + \sigma_0 \sqrt{T} x > \ln(K)} e^{(1+\alpha)(x_0 + \mu T + \sigma_0 \sqrt{T} x)} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\ &= \mathcal{K}_2 \int_{x > -d_2} e^{(1+\alpha)(x_0 + \mu T)} e^{(1+\alpha)\sigma_0 \sqrt{T} x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\ &= \mathcal{K}_2 s_0^{(1+\alpha)} e^{(1+\alpha)rT} e^{-(1+\alpha)\frac{\sigma_0^2}{2}T} \int_{x > -d_2} e^{(1+\alpha)\sigma_0 \sqrt{T} x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \end{aligned} \quad (5.39)$$

The integrand function can be recast as

$$\frac{1}{\sqrt{2\pi}} e^{(1+\alpha)\sigma_0 \sqrt{T} x} e^{-\frac{x^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x}{\sqrt{2}} - \frac{(1+\alpha)\sigma_0 \sqrt{T}}{\sqrt{2}}\right)^2} e^{\frac{\sigma_0^2}{2}(1+\alpha)^2 T}.$$

By the change of variable $x \mapsto y = x - (1+\alpha)\sigma_0 \sqrt{T}$, the domain of integration becomes

$$\begin{aligned} y > -d_2 - (1+\alpha)\sigma_0 T &= -\frac{1}{\sigma_0 \sqrt{T}} \left(\ln \left(\frac{K}{s_0} \right) - rT + \sigma_0^2/2 - (1+\alpha)\sigma_0^2 T \right) = \\ &= -\frac{1}{\sigma_0 \sqrt{T}} \left(\ln \left(\frac{K}{s_0} \right) + rT + \frac{\sigma_0^2}{2}(2\alpha + 1)T \right) = \\ &= -d(2\alpha + 1). \end{aligned}$$

Therefore (5.39) becomes

$$\begin{aligned} \mathcal{K}_2 \mathbb{E} \left[\mathbb{1}_{[X_0^T > \ln(K)]} e^{X_T^0} e^{\alpha X_T^0} \right] &= \mathcal{K}_2 s_0^{(1+\alpha)} e^{(1+\alpha)rT} e^{-(1+\alpha)\frac{\sigma_0^2}{2}T} e^{\frac{\sigma_0^2}{2}(1+\alpha)^2 T} \int_{y > -d(2\alpha+1)} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \\ &= \mathcal{K}_2 s_0^{(1+\alpha)} e^{(1+\alpha)rT} e^{\alpha(1+\alpha)\frac{\sigma_0^2}{2}T} N(d(2\alpha + 1)) \end{aligned}$$

Eventually by multiplying by ϵe^{-rT} we get

$$\mathcal{K}_2 \mathbb{E} \left[\mathbb{1}_{[X_0^T > \ln(K)]} e^{X_T^0} e^{\alpha X_T^0} \right] = \mathcal{K}_2 s_0^{(1+\alpha)} e^{\alpha r T} e^{\alpha(1+\alpha)\frac{\sigma_0^2}{2}T} N(d(2\alpha + 1)) = \epsilon \mathcal{K}_3 s_0^{(1+\alpha)} N(d(2\alpha + 1))$$

□

By Prop. 5.3.4 we have that the explicit computation of the corrected fair price is reduced to a numerical evaluation of a deterministic integral, which might be more efficient than directly simulating the random variables involved.

Remark 5.3.5. We could have also considered the second order perturbation $Pr^2(0; T)$ around the BS price. This is given by

$$Pr^2(0; T) = Pr^1(0; T) + \epsilon^2 e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\Phi(X_T^0)' X_T^2 \right] + e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\Phi(X_T^0)'' (X_T^1)^2 \right],$$

with Pr^1 the up to first order price in eq. (5.32). For the particular case of a European call option we have that $\Phi'' = \delta(X - \log K)e^X + \mathbb{1}_{[X > \log K]}e^X$, with δ the Dirac measure. Thus the correction up to the second order of the BS price for a European call option reads

$$\begin{aligned} Pr^2(0; T) = & Pr^1 + \epsilon^2 \mathcal{K}_4 s_0^{2\alpha+1} I_1(s, T, 2\alpha) + \epsilon^2 \mathcal{K}_5 s_0^{2\alpha+1} I_2(s, T) + \epsilon^2 \mathcal{K}_6 s_0^{\alpha+1} I_1(s, T, \alpha) + \\ & + \epsilon^2 \mathcal{K}_7 s_0^{2\alpha+1} I_3(r, s, T) + \epsilon^2 \mathcal{K}_8 s_0^{2\alpha+1} N(d(-3 - 4\alpha)) + \\ & + \epsilon^2 \mathcal{K}_9 s_0^{\alpha+1} N(d(-1 - 2\alpha)) + \epsilon^2 \mathcal{K}_{10} s_0 N(d(1)), \end{aligned} \tag{5.40}$$

with Pr^1 as in eq. (5.33), the notations as in Prop. 5.3.4 and

$$\begin{aligned} \mathcal{K}_4 &= \left(C_\alpha^1 + 2K_\alpha \frac{\sigma_1}{\alpha\sigma_0} \right) e^{-\frac{\sigma_0^2}{2}T}, \quad \mathcal{K}_5 = C_\alpha^2 e^{\alpha r T - \frac{\sigma_0^2}{2}(\alpha+1)T}, \\ \mathcal{K}_6 &= (C_\alpha^3 + 2K_\alpha \frac{\sigma_1}{\alpha\sigma_0}) e^{-\frac{\sigma_0^2}{2}T}, \quad \mathcal{K}_7 = (C_\alpha^4 + 2K_\alpha^2) e^{-\frac{\sigma_0^2}{2}T}, \\ \mathcal{K}_8 &= C_\alpha^5 e^{\frac{\sigma_0^2}{2}T\alpha(2\alpha+1)+2\alpha r T}, \quad \mathcal{K}_9 = (C_\alpha^6 + \frac{\sigma_1}{\alpha\sigma_0}) e^{\frac{\sigma_0^2}{2}T\alpha(\alpha+1)+\alpha r T}, \\ \mathcal{K}_{10} &= (C_\alpha^7 - \frac{\sigma_1}{\alpha\sigma_0}), \\ I_2(s, T) &= \int_0^T \int_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{[x+y > -\sqrt{T}d(1)]} e^{\alpha\mu s + (2\alpha+1)\sigma_0 y + (\alpha+1)\sigma_0 x} \phi(x; y, T-s) \phi(y; 0, s) dx dy ds, \\ I_3(r, s, T) &= \int_0^T \int_0^s \int_{\mathbb{R} \times \mathbb{R} \times \mathbb{R}} \mathbb{1}_{[x+y+z > -\sqrt{T}d(1)]} e^{\alpha\mu s + \alpha\mu r + \sigma_0 x + (1+\alpha)\sigma_0 y + (1+2\alpha)\sigma_0 z} \times \\ &\quad \times \phi(x; y, T-s) \phi(y; z, s-r) \phi(z; 0, r) dx dy dz dr ds, \end{aligned}$$

Numerical results concerning the pricing formula in Prop. 5.3.4.

We will now use the techniques introduced in the Appendix, which are based on the *multi-element Polynomial Chaos Expansion* (PCE) approach, to show the accuracy of the above derived approximated pricing formula in Proposition 5.3.4.

In what follows we will numerically compute the first order correction of the price of an European call option, whose payoff function is $(e^{X_T^c} - K)_+$. In particular we focus our attention on the second summand of

$$Pr_1(0; T) = Pr_{BS} + \epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} [\Phi'(X_T^0) X_T^1], \tag{5.41}$$

where Pr_{BS} denotes the usual B-S price with underlying $S_t^0 = e^{X_t^0}$. Also, X_T^0 and X_T^1 are defined as in Prop. 5.3.1.

The expectation is detected by means of the standard Monte Carlo method, using 10000 independent realization, and by mean of the multi-element PCE, see the Appendix for a brief introduction to the latter. Indeed, the random variable of interest is

$$\mathbb{1}_{\{X_0^T(\omega) > \ln(K)\}} \exp(X_T^0) X_T^1.$$

For both methods we will use the available analytical expression of X_T^0 and X_T^1 , depending on the function $f(x)$. In what follows $D := \{X_0^T(\omega) > \ln(K)\}$.

In particular exploiting the linearity of the expectation and the definition of the two random variables involved, (5.41) becomes

$$\mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \int_0^T K_{\alpha} e^{\alpha X_s^0} ds \right] + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} e^{\alpha X_T^0} \right] - K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \right] \tag{5.42}$$

Then we perform a *multi-element* PCE-approximation of each random variable in (5.42), setting the degree of the approximation, see the Appendix and in particular eq. (5.78), to be $p = 15$, since the degree of precision reached for such approximation is enough. Moreover for higher degree the computational cost increase as well as numerical fluctuation, due to implementation by means of NISP toolbox for scilab, becomes relevant for Multi-element approximation. It is worth to mention that *multi-element* PCE is nothing else that a PCE focused on D . Moreover the global statistics are a scaled by means of the weight w of the element D . See 5.3.5 for further details.

The numerical values of the parameters are gathered in Table 5.1.

Parameters	α	r	σ_1	K	T
Values	0.1	0.03	0.15	100	0.5

Table 5.1: Numerical values of the parameters employed in further computations

The fair price is numerically determined for the set of spot prices $s_0 = \in \{90, 100, 110\}$ and volatility value $\sigma_0 \in \{15\%, 25\%, 35\%\}$.

The PCE computation will be compared with standard Monte-Carlo simulation for the integrals and expansions in (5.42). The number of independent realizations is set as 10000. Moreover as benchmark we use the results presented in Proposition 3.1. These data are collected in Tables 5.2, 5.3, 5.4.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	12.38180	12.37737	12.36010	2.22240	2.22195	2.23204
	Error		4.4374e-03	2.3950e-01		4.4374e-04	2.3995e-02
	Time		0.0580	0.3200		0.0530	0.2890
$\sigma_0 = 25\%$	Results	14.09613	14.08919	14.14155	4.31567	4.31498	4.28696
	Error		6.9451e-03	1.7882e-01		6.9451e-04	1.7755e-02
	Time		0.0530	0.3060		0.0690	0.4130
$\sigma_0 = 35\%$	Results	15.08779	15.07774	15.30850	6.58042	6.57941	6.57030
	Error		1.0044e-02	1.4500e-01		1.0044e-03	1.4255e-02
	Time		0.0690	0.3460		0.0630	0.3420

Table 5.2: Numerical values for PCE and MC estimation of equation 25, for $s_0 = 90$, $\alpha = 0.1$, $\sigma_1 = 0.15$, $r = 0.03$ and $T = 0.5$.

5.3.2 A correction given by an exponential function and jumps

In what follows we extend the results in Sec. 5.3.1 to the second model in Sec. 5.4. In particular we will consider a correction up to the first order around the BS price (for a

5.3 Corrections around the Black-Scholes price (with Brownian, resp. Brownian plus jumps) **247**

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	39.39600	39.38877	38.97870	8.42541	8.42468	8.46801
	Error		7.2374e-03	3.2398e-01		7.2374e-04	3.2376e-02
	Time		0.0610	0.3180		0.3160	0.3260
$\sigma_0 = 25\%$	Results	28.38116	28.37206	28.57793	9.82235	9.82144	9.82024
	Error		9.0927e-03	2.1197e-01		9.0927e-04	2.1097e-02
	Time		0.0520	0.3000		0.0590	0.2860
$\sigma_0 = 35\%$	Results	25.56320	25.55082	25.60074	12.03973	12.03850	12.03580
	Error		1.2374e-02	1.6429e-01		1.2374e-03	1.6466e-02
	Time		0.0530	0.3190		0.0550	0.2940

Table 5.3: Numerical values for PCE and MC estimation of equation 25, for $s_0 = 100$, $\alpha = 0.1$, $\sigma_1 = 0.15$, $r = 0.03$ and $T = 0.5$.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	69.70042	69.69460	69.68928	18.07600	18.07542	18.08538
	Error		5.8153e-03	2.6109e-01		5.8153e-04	2.5932e-02
	Time		0.0560	0.4000		0.0700	0.3330
$\sigma_0 = 25\%$	Results	45.21665	45.20739	45.04317	17.54193	17.54100	17.52920
	Error		9.2595e-03	2.0951e-01		9.2595e-04	2.0818e-02
	Time		0.0690	0.3460		0.0550	0.3120
$\sigma_0 = 35\%$	Results	37.87932	37.86590	37.32253	19.09570	19.09436	19.11644
	Error		1.3416e-02	1.7161e-01		1.3416e-03	1.7169e-02
	Time		0.0530	0.3510		0.0660	0.3240

Table 5.4: Numerical values for PCE and MC estimation of equation 25, for $s_0 = 110$, $\alpha = 0.1$, $\sigma_1 = 0.15$, $r = 0.03$ and $T = 0.5$.

European call option) where both diffusive and jump perturbations are taken into account. We consider an asset whose return evolves according to eq. (5.26) and consider as before the particular case where $f(x) = e^{\alpha x}$, $\alpha \in \mathbb{R}_0$. Carrying out the asymptotic expansion in powers of ϵ , $0 \leq \epsilon \leq \epsilon_0$, and stopping it at the second order we get the following proposition:

Proposition 5.3.6. *Let us assume X_t^ϵ evolves according to eq. (5.26) with $f(x) = e^{\alpha x}$, for some $\alpha \in \mathbb{R}$, then we have the asymptotic expansion up to the second order in powers of ϵ , $0 \leq \epsilon \leq \epsilon_0$, $X_t^\epsilon = X_t^0 + \epsilon X_t^1 + \epsilon^2 X_t^2 + R_2(\epsilon, t)$, where the coefficients are given by*

$$\begin{aligned}
X_t^0 &= x_0 + \mu t + \sigma_0 W_t, \quad \text{with law } \mathcal{N}(x_0 + \mu t, \sigma_0^2 t); \\
X_t^1 &= \int_0^t K_\alpha e^{\alpha X_s^0} ds + \frac{\sigma_1}{\alpha \sigma_0} (e^{\alpha X_t^0} - 1) + \lambda t (e^{\gamma + \frac{\delta^2}{2}} - 1) + \sum_{i=1}^{N_t} J_i; \\
X_t^2 &= C_\alpha^1 \int_0^t e^{2\alpha X_s^0} ds + C_\alpha^2 e^{\alpha X_t^0} \int_0^t e^{\alpha X_s^0} ds + C_\alpha^3 \int_0^t e^{\alpha X_s^0} ds + C_\alpha^4 \int_0^t e^{\alpha X_s^0} \int_0^s e^{\alpha X_r^0} dr ds \\
&\quad + C_\alpha^5 e^{2\alpha X_t^0} + C_\alpha^6 e^{\alpha X_t^0} + C_\alpha^7 + C_\alpha^8 \lambda (e^{\gamma + \frac{\delta^2}{2}} - 1) \nu(dx) \int_0^t s e^{\alpha X_s^0} ds \\
&\quad - t e^{\alpha X_t^0} \lambda (e^{\gamma + \frac{\delta^2}{2}} - 1) + \frac{\sigma_1}{\sigma_0} \lambda (e^{\gamma + \frac{\delta^2}{2}} - 1) \int_0^t e^{\alpha X_s^0} ds \\
&\quad + C_\alpha^9 \int_0^t \sum_{i=1}^{N_s} J_i e^{\alpha X_s^0} ds + \frac{\sigma_1}{\sigma_0} e^{\alpha X_t^0} \sum_{i=1}^{N_t} J_i - \frac{\sigma_1}{\sigma_0} \sum_{i=1}^{N_t} J_i \int_0^t e^{\alpha X_s^0} ds,
\end{aligned} \tag{5.43}$$

with the constants as in Prop. 5.3.1 and

$$C_\alpha^8 = \frac{\sigma_1}{\sigma_0} \alpha \mu + \frac{\sigma_0 \sigma_1}{2} \alpha^2 - 2\sigma_0 \sigma_1 \alpha, \quad C_\alpha^9 = 2\sigma_0 \sigma_1 \alpha - \frac{\sigma_1}{\sigma_0} \alpha \mu - \frac{\sigma_0 \sigma_1}{2} \alpha^2.$$

Proof. The proof follows from Prop. 5.3.1 just taking into account the presence of the Poisson random measure terms and applying Itô's lemma, together with the stochastic Fubini theorem. \square

Remark 5.3.7. Just as in the Remark 5.3.2 it is easy to extend Prop. 5.3.6 and formula (5.32) to the case where $f(x) = e^{\alpha x}$ is replaced by $\int_{\mathbb{R}_0} e^{i\alpha x} \lambda(d\alpha)$, resp. $\int_{\mathbb{R}_0} e^{\alpha x} \lambda(d\alpha)$, with assumptions corresponding to those in Remark 5.3.2.

Proposition 5.3.8. *Let us consider the model described by (5.26) in the particular case of an European call option Φ with payoff given by $\Phi(X_T^\varepsilon) = (e^{X_T^\varepsilon} - K)_+$. Then the approximated price up to the first order $Pr_\nu^1(0; T)$, in the sense explained in Remark 5.3.3, is explicitly given by*

$$Pr_\nu^1(0; T) = Pr^1 + \epsilon T s_0 N(d_1) (e^{\gamma + \frac{\delta^2}{2}} - 1) + \epsilon T s_0 N(d_1) \delta \lambda,$$

where Pr^1 is the corrected fair price up to the first order as given in eq. (5.33) (the notations are as Prop. 5.3.4).

Proof. The proof is analogous of the proof of Prop. 5.3.4 adding the jump process. The claim follows then from the independence of the jump process and of the Brownian motion together with the fact that $\mathbb{E} \sum_{i=1}^{N_t} J_i = \delta T \lambda$. \square

Numerical results concerning the pricing formula in Prop. 5.3.6

We consider numerically the model discussed in Prop. 5.3.8, assuming that the J_i are independent and normally distributed random variable

$$J_i \sim \mathcal{N}(\gamma, \delta^2) \quad \gamma = 0.05, \quad \delta = 0.02,$$

and $\lambda = 2$. In particular we are aiming at numerically computing the expectations in the second summand of (5.41), which in the present case reads

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \int_0^T K_\alpha e^{\alpha X_s^0} ds \right] + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} e^{\alpha X_T^0} \right] - K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \right] \\ + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \lambda T \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \right] + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \sum_{i=1}^{N_T} J_i \right]. \end{aligned} \tag{5.44}$$

By means of independence of the jumps and $\mathbb{E}_t \left[\sum_{i=1}^{N_T} J_i \right] = \lambda T \delta$, we get

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \int_0^T K_\alpha e^{\alpha X_s^0} ds \right] + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} e^{\alpha X_T^0} \right] - K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \right] \\ K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \lambda T \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \right] + K_2 \lambda T \delta \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \right]. \end{aligned} \tag{5.45}$$

We are going to compute (5.63) by *multi-element* PCE-approximations.

The other parameters entering the model are taken from Table 5.1 and the three spot price considered are $s_0 \in \{90, 100, 110\}$. The results are presented in Tables 5.5, 5.6, 5.7.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	12.51812	12.51387	12.56922	2.23603	2.23560	2.21394
	Error		4.2567e-03	2.4285e-01		4.2567e-04	2.4069e-02
	Time		0.0830	0.5250		0.0860	0.5160
$\sigma_0 = 25\%$	Results	14.31171	14.30550	14.15258	4.33723	4.33661	4.34650
	Error		6.2145e-03	1.8182e-01		6.2145e-04	1.8295e-02
	Time		0.0850	0.5050		0.0880	0.5190
$\sigma_0 = 35\%$	Results	15.34622	15.33806	15.39336	6.60626	6.60544	6.61219
	Error		8.1646e-03	1.4758e-01		8.1646e-04	1.4754e-02
	Time		0.0880	0.5630		0.0990	0.5080

Table 5.5: Numerical values for PCE and MC estimation of equation 25, for $s_0 = 90$, $\alpha = 0.1$, $\sigma_1 = 0.15$, $r = 0.03$, $\lambda = 2$, $\gamma = 0.05$, $\delta = 0.02$ and $T = 0.5$.

5.3.3 A correction given by a polynomial function

Let us consider eq. (5.23) with f a polynomial correction, namely $f(x) = \sum_{i=0}^N \alpha_i x^i$, with $\alpha_i \in \mathbb{R}$ and $N \in \mathbb{N}_0$. We then get the following proposition.

Proposition 5.3.9. *Let us consider the case of the B-S model corrected by a non-linear term given by (5.23) with $f(x) = \sum_{i=0}^N \alpha_i x^i$, for some $\alpha_i \in \mathbb{R}$, then the expansion coefficients for*

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	39.80797	39.80128	39.84062	8.46660	8.46593	8.48244
	Error		6.6879e-03	3.2715e-01		6.6879e-04	3.2848e-02
	Time		0.0870	0.5490		0.0860	0.5270
$\sigma_0 = 25\%$	Results	28.78634	28.77863	28.54522	9.86287	9.86209	9.86723
	Error		7.7114e-03	2.1566e-01		7.7114e-04	2.1567e-02
	Time		0.0900	0.5370		0.0910	0.6180
$\sigma_0 = 35\%$	Results	25.96991	25.96051	26.21060	12.08041	12.07947	12.05190
	Error		9.3989e-03	1.6859e-01		9.3989e-04	1.6726e-02
	Time		0.1070	0.5240		0.0920	0.5180

Table 5.6: Numerical values for PCE and MC estimation of equation 25, for $s_0 = 100$, $\alpha = 0.1$, $\sigma_1 = 0.15$, $r = 0.03$, $\lambda = 2$, $\gamma = 0.05$, $\delta = 0.02$ and $T = 0.5$.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	70.37793	70.37303	70.57558	18.14375	18.14326	18.20012
	Error		4.9058e-03	2.6286e-01		4.9058e-04	2.6111e-02
	Time		0.0850	0.4990		0.0850	0.5210
$\sigma_0 = 25\%$	Results	45.81367	45.80646	45.98197	17.60163	17.60091	17.59774
	Error		7.2116e-03	2.1243e-01		7.2116e-04	2.1281e-02
	Time		0.0950	0.5270		0.0910	0.5180
$\sigma_0 = 35\%$	Results	38.43779	38.42848	38.04608	19.15155	19.15062	19.13383
	Error		9.3058e-03	1.7688e-01		9.3058e-04	1.7692e-02
	Time		0.0910	0.5370		0.0990	0.5430

Table 5.7: Numerical values for PCE and MC estimation of equation 25, for $s_0 = 110$, $\alpha = 0.1$, $\sigma_0 = 0.15$, $r = 0.03$, $\lambda = 2$, $\gamma = 0.05$, $\delta = 0.02$ and $T = 0.5$.

the solution X_t^ϵ of (5.23) up to the second order are given by the system

$$\begin{aligned}
 X_t^0 &= x_0 + \mu t + \sigma_0 W_t, \quad \text{with law } \mathcal{N}(x_0 + \mu t, \sigma_0^2 t); \\
 X_t^1 &= \sum_{i=1}^N \tilde{K}_i (X_t^0)^{i+1} - \sum_{i=0}^N \int_0^t K_i (X_s^0)^i ds + \sigma_1 \alpha_0 W_t; \\
 X_t^2 &= \sum_{k=1}^{2N+1} C_k^1 (X_t^0)^k - \sum_{k=1}^{2N+1} \int_0^t C_k^2 (X_s^0)^k ds + \sum_{i=1}^N \sum_{j=0}^N \int_0^t \int_0^s C_{i,j}^3 (X_s^0)^{i-1} (X_r^0)^j dr ds \\
 &\quad + \sum_{i=1}^N \sum_{j=0}^N (X_t^0)^i \int_0^s C_{i,j}^4 (X_r^0)^j dr.
 \end{aligned} \tag{5.46}$$

where the constants are given by

$$K_i = \begin{cases} \sigma_0 \sigma_1 \alpha_i + \frac{\sigma_1}{\sigma_0} \mu \alpha_i + \frac{\sigma_0 \sigma_1}{2} \alpha_{i+1} (i+1), & i \neq 0, i \neq N, \\ \sigma_0 \sigma_1 \alpha_0 + \frac{\sigma_0 \sigma_1 \alpha_1}{2}, & i = 0, \\ \sigma_0 \sigma_1 \alpha_N + \frac{\sigma_1}{\sigma_0} \mu \alpha_N, & i = N, \end{cases} \quad \tilde{K}_i = \frac{\sigma_1}{\sigma_0} \frac{\alpha_i}{(i+1)},$$

$$C_k^1 = \gamma_k^1 + \gamma_k^2 + \gamma_k^3,$$

where

$$\begin{aligned} \gamma_k^1 &= \begin{cases} \sum_{k=i+j+1} \mu i \alpha_i + \frac{\sigma_1}{\sigma_0} - \frac{\sigma_0}{2}(i+j+1), & k \neq 1, k \neq 2N, \\ \frac{\sigma_0}{2}, & k = 0, \\ \mu \frac{\sigma_1}{\sigma_0} N \alpha_N (\sigma_0 \sigma_1 \alpha_N + \frac{\sigma_1}{\sigma_0} \mu \alpha_N), & k = 2N, \end{cases} \\ \gamma_k^2 &= \begin{cases} \left(\frac{(-1)^k + 1}{2} \frac{\sigma_1^2}{2} \right) \alpha_k^2, & \text{if } 1 \leq k \leq N, \\ 0, & \text{otherwise} \end{cases} \\ \gamma_k^3 &= \sum_{i+j=k-1} 2\sigma_0 \sigma_1 \alpha_i i \tilde{K}_j, \\ C_{i,j}^3 &= - \begin{cases} \frac{\sigma_1}{\sigma_0} \alpha_1 K_0, & \text{if } i = 1, j = 0, \\ \frac{\sigma_1}{\sigma_0} i \alpha_i K_j + \frac{\sigma_0 \sigma_1}{2} i \alpha_i K_j (i-1), & \text{otherwise.} \end{cases} \\ C_{i,j}^4 &= \frac{\sigma_1}{\sigma_0} \alpha_i K_j, \end{aligned}$$

Proof. The proof consists in a series of applications of Itô's formula and stochastic Fubini theorem, see, e.g. [Fil01] Th. 6.2. In fact, substituting $f(x) = \sum_{i=0}^N \alpha_i x^i$ into system (5.24) we obtain

$$\begin{aligned} X_t^0 &= x_0 + \mu t + \sigma_0 W_t, \quad \text{with law } \mathcal{N}(x_0 + \mu t, \sigma_0^2 t); \\ X_t^1 &= - \int_0^t \sigma_0 \sigma_1 \left(\sum_{i=0}^N \alpha_i (X_s^0)^i \right) ds + \int_0^t \sigma_1 \left(\sum_{i=0}^N \alpha_i (X_s^0)^i \right) dW_s; \\ X_t^2 &= - \int_0^t \frac{\sigma_1^2}{2} \left(\sum_{i=0}^N \alpha_i (X_s^0)^i \right)^2 + 2\sigma_1 \left(\sum_{i=0}^N \alpha_i (X_s^0)^i \right)' X_s^1 ds \\ &\quad + \int_0^t \sigma_1 \left(\sum_{i=0}^N \alpha_i (X_s^0)^i \right)' X_s^1 dW_s. \end{aligned} \tag{5.47}$$

To compute X_t^1 obtaining eq. (5.46) we apply Itô's lemma to the function $g(X_t^0) = \alpha_{i+1} (X_t^0)^{i+1}$ to get

$$(X_t^0)^{i+1} = \int_0^t \left(\mu(i+1)(X_s^0)^i + \frac{1}{2} \sigma_0^2 i(i+1)(X_s^0)^{i-1} \right) ds + \int_0^t (X_s^0)^i (i+1) \sigma_0 dW_s. \tag{5.48}$$

Then, summing up we obtain

$$\sum_{i=1}^N \int_0^t (X_s^0)^i (i+1) \sigma_0 dW_s = \sum_{i=1}^N (X_s^0)^{i+1} - \sum_{i=1}^N \int_0^t \left(\mu(i+1)(X_s^0)^i + \frac{1}{2} \sigma_0^2 i(i+1)(X_s^0)^{i-1} \right) ds. \tag{5.49}$$

Substituting now eq. (5.49) into X^1 in eq. (5.47) we obtain the following

$$\begin{aligned} X_t^1 &= \sum_{i=1}^N \frac{\sigma_1}{\sigma_0} \frac{\alpha_i}{(i+1)} (X_t^0)^{i+1} - \sum_{i=1}^N \int_0^t \sigma_0 \sigma_1 \alpha_i (X_s^0)^i - \sum_{i=1}^N \int_0^t \mu(i+1) \frac{\sigma_1 \alpha_i}{\sigma_0 (i+1)} (X_s^0)^i + \\ &\quad - \sum_{i=1}^N \int_0^t \frac{1}{2} \sigma_0^2 i(i+1) \frac{\sigma_1 \alpha_i}{\sigma_0 (i+1)} (X_s^0)^{i-1} ds, \end{aligned}$$

and rearranging the terms we then get the desired result in (5.46) for X_t^1 .

Substituting the expression of X_t^1 into X_t^2 we obtain

$$\begin{aligned} X_t^2 &= - \sum_{i=1}^N \int_0^t \frac{\sigma_1^2}{2} \alpha_i^2 (X_s^0)^{2i} ds - \sum_{i,j=1}^N \int_0^t 2\sigma_0 \sigma_1 \alpha_i i \tilde{K}_j (X_s^0)^{i-1} (X_s^0)^{j+1} ds = \\ &= \sum_{j=0}^N \sum_{i=1}^N \int_0^t \int_0^s 2\sigma_0 \sigma_1 \alpha_i i K_j (X_s^0)^{i-1} (X_r^0)^j dr ds + \sum_{i,j=1}^N \int_0^t \sigma_1 \alpha_i i K_j (X_s^0)^{i-1} (X_s^0)^{j+1} dW_s \\ &\quad - \sum_{j=0}^N \sum_{i=1}^N \int_0^t \int_0^s \sigma_1 \alpha_i i K_j (X_s^0)^{i-1} (X_r^0)^j dr dW_s. \end{aligned}$$

Exploiting again the stochastic Fubini theorem, from eq. (5.49) and grouping the terms with the same powers we obtain (5.46). \square

Proposition 5.3.10. *Let us consider the particular case of $N = 1$, i.e. a linear perturbation, namely $f(x) = \alpha_0 + \alpha_1 x$, $\alpha_i \in \mathbb{R}$, $i = 0, 1$. Then the terms up to the first order in equation (5.46) read*

$$\begin{aligned} X_t^0 &= x_0 + \mu t + \sigma_0 W_t, \\ X_t^1 &= \beta_1 t + \beta_2 t^2 + \beta_3 W_t + \beta_4 W_t^2 + \beta_5 t W_t - \int_0^t \beta_6 W_s ds, \end{aligned} \quad (5.50)$$

with

$$\begin{aligned} \beta_1 &= -\sigma_0 \sigma_1 \alpha_0 - \sigma_0 \sigma_1 \alpha_1 x_0 - \frac{\sigma_0 \sigma_1 \alpha_1}{2}, \quad \beta_2 = -\frac{\sigma_0 \sigma_1 \alpha_1 \mu}{2}, \quad \beta_3 = \alpha_1 \sigma_0 + x_0 \sigma_1 \alpha_1, \\ \beta_4 &= \frac{\sigma_0 \sigma_1 \alpha_1}{2}, \quad \beta_5 = \sigma_1 \alpha_1 \mu, \quad \beta_6 = \sigma_1 \alpha_1 \mu + \sigma_0^2 \sigma_1 \alpha_1. \end{aligned}$$

The first order correction (in the sense discussed in Remark 5.3.3), of the price of an European call option Φ with payoff given by $\Phi(X_T^\epsilon) = (e^{X_T^\epsilon} - K)_+$ is explicitly given by

$$\begin{aligned} Pr^1(0; T) &= P_{BS} + \epsilon s_0 (\beta_1 + \sigma_0 \beta_3 + \beta_4) T N(d_1) + \epsilon s_0 (\beta_2 + \sigma_0^2 \beta_4) T^2 N(d_1) \\ &\quad + \epsilon s_0 (\beta_3 + 2\sigma_0 \beta_4 T + T \beta_5) \sqrt{T} \phi(-d_1) - \epsilon s_0 \beta_4 T d_1 \phi(d_1) + \\ &\quad + \epsilon s_0 T^2 \beta_5 \sigma_0 T^2 N(d_1) - \epsilon s_0 e^{+\frac{\sigma_0^2}{2} T} \beta_6 I(s, T), \end{aligned} \quad (5.51)$$

where the notation is as in Prop. 5.3.4 and we have denoted for short by $\phi(x)$ the density function of the standard Gaussian law and we have set

$$I(s, T) = \int_0^T \int_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{[x+y > -\sqrt{T} d_1]} e^{\sigma_0(x+y)} y \phi(x; 0, T-s) \phi(y; 0, s) dx dy ds.$$

Proof. Let us consider the linear function $f(x) = \alpha_0 + \alpha_1 x$, where $\alpha_0, \alpha_1 \in \mathbb{R}$. The approximated price up to the first order, $Pr^1(0; T)$ of an European call option with payoff function $\Phi(X_T^\epsilon) = (e^{X_T^\epsilon} - K)_+$ is

$$Pr^1(0; T) = P_{BS} + \epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} [\Phi(X_T^0) X_T^1] \quad (5.52)$$

where P_{BS} is the standard B-S price with underlying $s_0(t) = e^{X_t^0}$.

In particular we have that X_T^0 and X_T^1 are defined as

$$X_T^0 = x_0 + \mu T + \sigma_0 W_T \quad (5.53)$$

$$X_T^1 = \beta_1 T + \beta_2 T^2 + \beta_3 W_T + \beta_4 W_T^2 + \beta_5 T W_T - \beta_6 \int_0^T W_s ds. \quad (5.54)$$

By linearity of the expectation, (5.52) becomes, collecting the terms with coefficients β_3 and β_5 ,

$$\begin{aligned} Pr^1(0; T) = P_{BS} + \epsilon e^{-rT} & \left\{ \mathbb{E}^{\mathbb{Q}} \left[\beta_1 T \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] + \mathbb{E}^{\mathbb{Q}} \left[\beta_2 T^2 \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] + \right. \\ & + \mathbb{E}^{\mathbb{Q}} \left[\beta_{3,5}^T W_T \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] + \mathbb{E}^{\mathbb{Q}} \left[\beta_4 W_T^2 \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] + \\ & \left. + \mathbb{E}^{\mathbb{Q}} \left[\beta_6 \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \int_0^T W_s ds \right] \right\}, \end{aligned} \quad (5.55)$$

with $\beta_{3,5}^T := \beta_3 + T\beta_5$.

From the definition of X_T^0 we have that

$$\epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\beta_1 T \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] = \epsilon T \beta_1 s_0 N(d_1),$$

and as above we have

$$\epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\beta_2 T^2 \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] = \epsilon T^2 \beta_2 s_0 N(d_1)$$

Concerning the third term in (5.55), we have that,

$$\begin{aligned} \beta_{3,5}^T \mathbb{E}^{\mathbb{Q}} \left[W_T \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] & = \beta_{3,5}^T s_0 e^{rT} e^{-\frac{\sigma_0^2}{2} T} \sqrt{T} \int_{x > -d_2} e^{\sigma_0 \sqrt{T} x} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\ & = \beta_{3,5}^T s_0 e^{rT} e^{-\frac{\sigma_0^2}{2} T} \sqrt{T} \int_{x > -d_2} x \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x}{\sqrt{2}} - \frac{\sigma_0 \sqrt{T}}{\sqrt{2}}\right)^2} e^{\frac{\sigma_0^2}{2} T} dx, \end{aligned}$$

and by setting $y = x - \sigma_0 \sqrt{T}$, we get that the r.h.s. is given by

$$\begin{aligned} \beta_{3,5}^T s_0 e^{rT} \sqrt{T} \int_{y > -d_1} (\sigma_0 \sqrt{T} + y) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy & = \beta_{3,5}^T T s_0 e^{rT} \sigma_0 N(d_1) - \beta_{3,5}^T \sqrt{T} s_0 e^{rT} \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \right]_{-d_1}^{+\infty} = \\ & = \beta_{3,5}^T T s_0 e^{rT} \sigma_0 N(d_1) + \beta_{3,5}^T \sqrt{T} s_0 e^{rT} \phi(-d_1, 0, 1). \end{aligned}$$

Hence the third term in (5.55) reads

$$\epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\beta_3 W_T \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] = \epsilon \beta_3 T \sigma_0 s_0 N(d_1) + \epsilon \beta_3 s_0 \sqrt{T} \phi(-d_1, 0, 1).$$

Exploiting the definition of X_T^0 occurring in the fourth term in (5.55), as well as similar algebraic computation as in the previous previous section, we get

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\beta_4 W_T^2 \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_T^0} \right] & = \beta_4 s_0 e^{rT} e^{-\frac{\sigma_0^2}{2} T} \int_{x > -d_2} T x^2 e^{\sigma_0 \sqrt{T} x} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \\ & = \beta_4 s_0 e^{rT} T \int_{y > -d_1} (y + \sigma_0 \sqrt{T})^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \end{aligned}$$

Developing the square and using the linearity property of the integral we get that the r.h.s. is equal to

$$\begin{aligned} & \int_{y>-d_1} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy + \int_{y>-d_1} 2\sigma_0\sqrt{T}y \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy + \int_{y>-d_1} \sigma_0^2 T \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \\ & = \int_{y>-d_1} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy + 2\sigma_0\sqrt{T}\phi(-d_1, 0, 1) + \sigma_0^2 TN(d_1). \end{aligned}$$

The first term is computed using integration by parts,

$$\int_{y>-d_1} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = -d_1\phi(d_1) + N(d_1)$$

therefore

$$\epsilon e^{-rT} \mathbb{E}^{\mathbb{Q}} \left[\beta_4 W_T^2 \mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_0^T} \right] = \epsilon \beta_4 s_0 T \left(-d_1\phi(d_1) + N(d_1) + 2\sigma_0\sqrt{T}\phi(-d_1, 0, 1) + \sigma_0^2 TN(d_1) \right).$$

To compute the fifth in (5.55) term we use Fubini theorem to exchange the expectation with the integral with respect to time, getting

$$\beta_6 \int_0^T \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{\{X_0^T > \ln(K)\}} e^{X_0^T} W_s ds \right]. \quad (5.56)$$

For every fixed $s \in [0, T]$, W_s and W_T , the latter is included in X_0^T by its very definition, are Gaussian random variable jointly distributed. Therefore exploiting basic properties of Brownian motion we can recast them by means of a sum of independent random variable, namely

$$\begin{aligned} W_s &= Y \sim \mathcal{N}(0, s), \\ W_T &= W_T - W_s + W_s = X + Y. \end{aligned}$$

In particular $X \sim \mathcal{N}(0, T-s)$ and it is independent with respect to Y . Thus (5.56) reads

$$\begin{aligned} \beta_6 \int_0^T \int_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{\{x+y > -\sqrt{T}d_2\}} e^{x_0 + \mu T} e^{\sigma_0(x+y)y} \frac{1}{\sqrt{2\pi}} e^{\frac{x^2}{2(T-s)}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2s}} ds = \\ = \beta_6 s_0 e^{rT} e^{-\frac{\sigma_0^2}{2}T} \int_0^T \int_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{\{x+y > -\sqrt{T}d_2\}} e^{\sigma_0(x+y)y} \phi(x; 0; T-s) \phi(y, 0, s) dx dy ds, \end{aligned}$$

and the claim follows. □

Numerical results concerning the pricing formula in Prop. 5.3.10

Let us consider the case of the B-S model corrected by a linear term given as in Prop. 5.3.10 by $f(x) = \alpha_0 + \alpha_1 x$. We compute the first order correction of the price of an European call option with $\Phi(X_T^\epsilon) = (e^{X_T^\epsilon} - K)_+$ as payoff function, according to Prop. 5.3.10.

Our aim is computing the expectation in (5.41) in the present case. By the very definition of X_0^T and X_T^T and the form of Φ' , it reads as

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_1 T \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_2 T^2 \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_3 W_T \right] \\ & + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_4 W_T^2 \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_5 T W_T \right] - \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_6 \int_0^T W_s ds \right] \end{aligned} \tag{5.57}$$

Each random variable in the brackets is approximated by means of a *multi-element* PCE of degree $p = 15$ and respectively by means of standard Monte Carlo methods, using $N = 10000$ independent simulations of the random variable involved.

The accuracy of *PCE* is represented by its absolute error, using as benchmark the analytical value coming from (5.51). Due to the *Law of Large Numbers*, the accuracy of MC-estimation of (5.32) is provided by its standard error (SE_{MC}). Upon considering $N = 10000$ realizations (Y_j) of the random variable $Y := \Phi'(X_T^0)X_T^1$ inside the expectation in the r.h.s. of equation (5.32), let us compute

$$SE_{MC} = \frac{\hat{\sigma}}{\sqrt{N}} \tag{5.58}$$

where $\hat{\sigma}^2 = \frac{1}{N-1} \sum_{j=1}^N (Y_j - \mu_{MC})^2$ and $\mu_{MC} = \frac{1}{N} \sum_{j=1}^N Y_j$.

The numerical values of the parameters involved are collected in Table 5.5

Parameters	α_0	α_1	r	σ_1	K	T
Values	0.3	0.5	0.03	0.1	100	0.5

Table 5.8: Numerical values of the parameters employed in further computations

The computations are made setting the parameters as in Table 5.8 and for a set of volatility values $\sigma_0 \in \{15\%, 25\%, 35\%\}$ and for a set of increasing spot price $s_0 \in \{90, 100, 110\}$.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	1.45057	1.45049	1.44774	1.12927	1.12927	1.12870
	Error		7.9315e-05	8.2194e-03		7.9315e-06	8.2548e-04
$\sigma_0 = 25\%$	Results	3.82504	3.82488	3.83225	3.28856	3.28855	3.28849
	Error		1.5990e-04	1.1922e-02		1.5990e-05	1.1596e-03
$\sigma_0 = 35\%$	Results	6.44932	6.44905	6.44766	5.71657	5.71654	5.71482
	Error		2.7379e-04	1.5876e-02		2.7379e-05	1.5413e-03

Table 5.9: Numerical values for PCE and MC estimation of equation (5.41), $s_0 = 90$, $\alpha_0 = 0.3$, $\alpha_1 = 0.5$, $\sigma_1 = 0.10$, $r = 0.03$, $K = 100$ and $T = 0.5$.

5.3.4 A correction given by a polynomial function and jumps

In the present section we generalize the results obtained in the previous subsection 5.3.3 adding a compensated Poisson random measure. In particular let us assume that the normal

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	5.53650	5.53637	5.53931	5.03945	5.03944	5.04061
	Error		1.2763e-04	9.2641e-03		1.2763e-05	9.3834e-04
$\sigma_0 = 25\%$	Results	8.50577	8.50556	8.52655	7.83481	7.83479	7.83473
	Error		2.0666e-04	1.3297e-02		2.0666e-05	1.3194e-03
$\sigma_0 = 35\%$	Results	11.50225	11.50192	11.49891	10.63364	10.63361	10.63396
	Error		3.3318e-04	1.7426e-02		3.3318e-05	1.7601e-03

Table 5.10: Numerical values for PCE and MC estimation of equation (5.41), $s_0 = 100$, $\alpha_0 = 0.3$, $\alpha_1 = 0.5$, $\sigma_1 = 0.10$, $r = 0.03$, $K = 100$ and $T = 0.5$.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	12.70933	12.70924	12.72127	12.37689	12.37688	12.37642
	Error		9.7072e-05	1.2457e-02		9.7072e-06	1.2373e-03
$\sigma_0 = 25\%$	Results	15.16320	15.16299	15.16028	14.53658	14.53656	14.53696
	Error		2.0462e-04	1.5486e-02		2.0462e-05	1.5493e-03
$\sigma_0 = 35\%$	Results	17.99767	17.99732	18.01788	17.10754	17.10750	17.10605
	Error		3.5427e-04	2.0558e-02		3.5427e-05	1.9645e-03

Table 5.11: Numerical values for PCE and MC estimation of equation (5.41), $s_0 = 110$, $\alpha_0 = 0.3$, $\alpha_1 = 0.5$, $\sigma_1 = 0.10$, $r = 0.03$, $K = 100$ and $T = 0.5$.

return of the asset price evolves according to eq. (5.26) with a polynomial f . Then we have the following proposition.

Proposition 5.3.11. *Let us consider the case of the B-S model with added compensated Poisson noise and corrected by a non-linear term given by (5.26) with $f(x) = \sum_{i=0}^N \alpha_i x^i$, for some $\alpha_i \in \mathbb{R}$, then the expansion coefficients for the solution X_t^ϵ of (5.26) up to the second order are given by the system*

$$\begin{aligned}
X_t^0 &= x_0 + \mu t + \sigma_0 W_t, \quad \text{with law } \mathcal{N}(x_0 + \mu t, \sigma_0^2 t); \\
X_t^1 &= \sum_{i=1}^N \tilde{K}_i (X_t^0)^{i+1} - \sum_{i=0}^N \int_0^t K_i (X_s^0)^i ds + \sigma_1 \alpha_0 W_t - \lambda t \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \sum_{i=1}^{N_t} J_i; \\
X_t^2 &= \sum_{k=1}^{2N+1} C_k^1 (X_t^0)^k - \sum_{k=1}^{2N+1} \int_0^t C_k^2 (X_s^0)^k ds + \sum_{i=1}^N \sum_{j=0}^N \int_0^t \int_0^s C_{i,j}^3 (X_s^0)^{i-1} (X_r^0)^j dr ds \\
&\quad + \sum_{i=1}^N \sum_{j=0}^N (X_t^0)^i \int_0^s C_{i,j}^4 (X_r^0)^j dr + \sum_{i=0}^{N-1} C_i^5 \lambda \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \int_0^t s (X_s^0)^i ds \\
&\quad - \alpha_{i+1} \sigma_1 t (X_t^0)^i \lambda \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \int_0^t \sigma_1 \alpha_1 W_s ds \lambda \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) - \sigma_1 t \alpha_1 W_t \lambda \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \\
&\quad + \sum_{i=2}^N \alpha_i \sigma_1 \lambda \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \int_0^t (X_s^0)^i ds + \sigma_1 \alpha_1 W_t + \sum_{i=2}^N \sigma_1 \alpha_i (X_s^0)^i \sum_{i=1}^{N_t} J_i \\
&\quad - \sum_{i=2}^N \sigma_1 \alpha_i \int_0^t \int_{\mathbb{R}_0} (X_s^0)^i ds \sum_{i=1}^{N_t} J_i + \sum_{i=0}^{N-1} C_i^5 \int_0^t \sum_{i=1}^{N_s} J_i (X_s^0)^i ds
\end{aligned} \tag{5.59}$$

where the constants are as in Prop. 5.3.9 and

$$C_i^5 = \begin{cases} \sigma_0^2 \sigma_1 \alpha_2 + 2\sigma_0 \sigma_1 \alpha_1, & i = 0, \\ \sigma_1 \mu \alpha_{i+1} (i+1) + \frac{\sigma_0^2}{2} (i+2)(i+1) + 4\sigma_0 \sigma_1 \alpha_{i+1}, & i \neq 0, i \neq N, \\ \alpha_N \sigma_1 N \mu + 2\sigma_0 \sigma_1 N \alpha_{N+1}, & i = N, \end{cases} \quad \tilde{K}_i = \frac{\sigma_1}{\sigma_0} \frac{\alpha_i}{(i+1)},$$

Proof. The proof is analogous to the one in Prop. 5.3.9 taking into account the compensated Poisson random measure terms and applying Itô's lemma together with the stochastic Fubini theorem. \square

Proposition 5.3.12. *Let us consider the particular case of $N = 1$, i.e. a linear perturbation, namely $f(x) = \alpha_0 + \alpha_1 x$ in Prop. 5.3.11. Then the terms up to the first order in equation (5.59) read*

$$\begin{aligned} X_t^0 &= x_0 + \mu t + \sigma_0 W_t, \\ X_t^1 &= \beta_1 t + \beta_2 t^2 + \beta_3 W_t + \beta_4 W_t^2 + \beta_5 t W_t - \int_0^t \beta_6 W_s ds - \lambda t \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \sum_{i=1}^{N_t} J_i; \end{aligned} \quad (5.60)$$

with the constants being as in Prop. 5.3.10.

Also, the first order correction of the price of an European call option Φ with payoff given by $\Phi(X_T^\epsilon) = (e^{X_T^\epsilon} - K)_+$ (in the sense of Remark 5.3.3) is explicitly given by

$$Pr^1(0; T) = Pr^1 + \epsilon T s_0 N(d(1)) \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \epsilon T s_0 N(d(1)) \delta \lambda, \quad (5.61)$$

where Pr^1 is the corrected fair price up to the first order as given in eq. (5.51) and the notations are as above.

Proof. The proof follows as in Prop. 5.3.10. \square

Numerical results concerning the pricing formula in Prop. 5.3.11

The J_i are assumed to be independent and normally distributed random variables

$$J_i \sim \mathcal{N}(\gamma, \delta^2), \quad \text{for all } i \in \{1, 2, \dots, N_T\}, \quad \gamma = 0.05, \quad \delta = 0.02,$$

and $\lambda = 2$. In particular we are aiming at computing the expectation in (5.41) for the model described in Prop. 5.3.11. In the present case we have that this expectation is equal to

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_1 T \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_2 T^2 \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_3 W_T \right] \\ & + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_4 W_T^2 \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_5 T W_T \right] - \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_6 \int_0^T W_s ds \right] \\ & + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \lambda T \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \right] + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \sum_{i=1}^{N_T} J_i \right]. \end{aligned} \quad (5.62)$$

By means of the independence of the jumps and $\mathbb{E} \left[\sum_{i=1}^{N_T} J_i \right] = \delta \lambda T$, we can rewrite (5.62) as

$$\begin{aligned} & \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_1 T \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_2 T^2 \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_3 W_T \right] \\ & + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_4 W_T^2 \right] + \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_5 T W_T \right] - \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_0^T} \beta_6 \int_0^T W_s ds \right] \\ & + K_2 \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \lambda T \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) \right] + K_2 \lambda T \delta \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_D e^{X_T^0} \right]. \end{aligned} \tag{5.63}$$

We shall then compute *multi-element* PCE-approximations for this expression.

The parameters are taken from Table 5.8 and the three spot prices, resp, volatilities, considered are $s_0 \in \{90, 100, 110\}$, resp. $\sigma_0 \in \{15\%, 25\%, 35\%\}$.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	1.45057	1.45049	1.44774	1.12927	1.12927	1.12870
	Error		7.9315e-05	8.2194e-03		7.9315e-06	8.2548e-04
$\sigma_0 = 25\%$	Results	3.82504	3.82488	3.83225	3.28856	3.28855	3.28849
	Error		1.5990e-04	1.1922e-02		1.5990e-05	1.1596e-03
$\sigma_0 = 35\%$	Results	6.44932	6.44905	6.44766	5.71657	5.71654	5.71482
	Error		2.7379e-04	1.5876e-02		2.7379e-05	1.5413e-03

Table 5.12: Numerical values for PCE and MC estimation of equation (5.41), $s_0 = 90$, $\alpha_0 = 0.3$, $\alpha_1 = 0.5$, $\sigma_1 = 0.10$, $r = 0.03$, $K = 100, \lambda = 2$, $\gamma = 0.05$, $\delta = 0.02$ and $T = 0.5$.

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	5.53650	5.53637	5.53931	5.03945	5.03944	5.04061
	Error		1.2763e-04	9.2641e-03		1.2763e-05	9.3834e-04
$\sigma_0 = 25\%$	Results	8.50577	8.50556	8.52655	7.83481	7.83479	7.83473
	Error		2.0666e-04	1.3297e-02		2.0666e-05	1.3194e-03
$\sigma_0 = 35\%$	Results	11.50225	11.50192	11.49891	10.63364	10.63361	10.63396
	Error		3.3318e-04	1.7426e-02		3.3318e-05	1.7601e-03

Table 5.13: Numerical values for PCE and MC estimation of equation (5.41), $s_0 = 100$, $\alpha_0 = 0.3$, $\alpha_1 = 0.5$, $\sigma_1 = 0.10$, $r = 0.03$, $K = 100, \lambda = 2$, $\gamma = 0.05$, $\delta = 0.02$ and $T = 0.5$.

5.3.5 The Black-Scholes model with stochastic interest rate

We will consider in the present section an asset S_t^ϵ evolving, under the risk neutral probability \mathbb{Q} , according to the SDE

$$\begin{cases} dS_t = S_t r_t^\epsilon dt + S_t \sigma_0 dW_t^1, \\ s_0 = s_0 \in \mathbb{R}, \end{cases}, \tag{5.64}$$

		$\epsilon = 0.1$			$\epsilon = 0.01$		
		Analytical	PCE	standard MC	Analytical	PCE	standard MC
$\sigma_0 = 15\%$	Results	12.70933	12.70924	12.72127	12.37689	12.37688	12.37642
	Error		9.7072e-05	1.2457e-02		9.7072e-06	1.2373e-03
$\sigma_0 = 25\%$	Results	15.16320	15.16299	15.16028	14.53658	14.53656	14.53696
	Error		2.0462e-04	1.5486e-02		2.0462e-05	1.5493e-03
$\sigma_0 = 35\%$	Results	17.99767	17.99732	18.01788	17.10754	17.10750	17.10605
	Error		3.5427e-04	2.0558e-02		3.5427e-05	1.9645e-03

Table 5.14: Numerical values for PCE and MC estimation of equation (5.41), $s_0 = 110$, $\alpha_0 = 0.3$, $\alpha_1 = 0.5$, $\sigma_1 = 0.10$, $r = 0.03$, $K = 100$, $\lambda = 2$, $\gamma = 0.05$, $\delta = 0.02$ and $T = 0.5$.

with $\sigma_0 \in \mathbb{R}_+$ and r^ϵ is a stochastic interest rate evolving according to

$$\begin{cases} dr_t^\epsilon = b(t, r_t^\epsilon)dt + \epsilon\nu(t, r^\epsilon)dW_t^2, \\ r_0 = r_0 \in \mathbb{R}, \end{cases}, \quad (5.65)$$

with $\epsilon > 0$ a constant, and b and ν real functions satisfying some regularity and growth assumptions guaranteeing existence and uniqueness of the solution of (5.65). Let us assume further that W^1 and W^2 are two standard Brownian motions with correlation

$$d\langle W^1 W^2 \rangle_t = \bar{\rho}dt,$$

for some constant $\bar{\rho} > 0$.

We will refer to [Sch04] for an introduction on pricing under stochastic interest rate.

We will assume in what follows that $b(t, r) = \kappa(\theta - r)$, for some real constants κ and θ , to take into account the mean reverting property usually shown by interest rates. Furthermore we choose $\nu(t, r) = \bar{\nu}$ or $\nu(t, r) = \bar{\nu}\sqrt{r}$, for some constants $\bar{\nu}$, $r \in \mathbb{R}_+$.

Remark 5.3.13. Although $\nu(t, r_t^\epsilon) = \bar{\nu}\sqrt{r_t^\epsilon}$ does not satisfy the smoothness assumptions we assumed up to now, our expansion works, at least as formal power series expansion, due the existence and uniqueness result for such a type of diffusion term proven in [CIJR85].

We will from now on consider the normal return of the asset S_t , namely we will consider the process $X_t^\epsilon := \log S_t^\epsilon$ evolving according to

$$\begin{cases} dX_t^\epsilon = \left(r_t^\epsilon - \frac{\sigma_0^2}{2} \right) dt + \sigma_0 dW_t^1, \\ X_0^\epsilon = x_0 \in \mathbb{R}, \end{cases}. \quad (5.66)$$

Carrying out the expansion developed in Th. 5.2.5 we have the following result.

Proposition 5.3.14. *Let us consider the normal return process X_t^ϵ , $t \in [0, T]$, evolving according to eq. (5.66) and a stochastic interest rate r_t^ϵ evolving according to the SDE (5.65) with $b(s, r) = \kappa(\theta - r)$, κ and $\theta \in \mathbb{R}$, and correlation given by $d\langle W^1 W^2 \rangle_t = \bar{\rho}dt$, for some constant $\bar{\rho} > 0$.*

Then the first order heuristic expansion for the stochastic interest rate r_t^ϵ reads

$$\begin{aligned} r_t^0 &= r_0 + \int_0^t \kappa(\theta - r_s^0) ds, \\ r_t^1 &= - \int_0^t e^{-\kappa(t-s)} \nu(s, r_s^0) dW_s^2 \quad \text{with law } \mathcal{N}\left(0, Q_t^1\right), \end{aligned} \quad (5.67)$$

with

$$Q_t^{r^1} := \int_0^t e^{-2\kappa(t-s)} \nu^2(s, r_s^0) ds.$$

Furthermore, the first order heuristic expansion for the normal return X_t^ϵ is

$$\begin{aligned} X_t^0 &= x_0 + \int_0^t \left(r_s^0 - \frac{\sigma_0^2}{2} \right) ds + \int_0^t \sigma_0 dW_t^1, \\ X_t^1 &= \int_0^t r_s^1 ds. \end{aligned} \tag{5.68}$$

Proof. The proof immediately follows applying Th. 5.2.5. In particular we have that r_t^1 evolves according to

$$r_t^1 = - \int_0^t \kappa r_s^1 ds + \int_0^t \nu(s, r_s^0) dW_s^2.$$

Applying Itô's lemma to the function $re^{\kappa t}$, noticing that r_t^0 is a deterministic process, we find the desired solution. The distribution of r_t^1 follows simply noticing that r_t^1 is an integral of a Brownian motion with deterministic integrand. \square

Our aim is to price an option written on the underlying $e^{X_t^\epsilon}$ with final payoff given by $\Phi(X_T)$ under the stochastic interest rate r^ϵ . In particular we aim at finding the value $Pr(0; T)$ given by

$$Pr(0; T) = \mathbb{E} e^{-\int_0^T r_s^\epsilon ds} \Phi(X_T^\epsilon), \tag{5.69}$$

where the expectation is taken with respect to the joint measure generated by the two correlated Brownian motions W^1 and W^2 .

In order to give an analytic expression for (5.69), we make transformations in order to replace W^2 in r_t^1 by a random variable independent of W_T^1 .

Let us first notice that, being r^1 a Gaussian random variable, also X^1 is normally distributed. In particular we have that X_T^1 has law $\mathcal{N}\left(0, Q_T^{X^1}\right)$ with

$$Q_T^{X^1} := \int_0^T \left(\int_u^T e^{\kappa s} ds \right)^2 e^{-2\kappa u} \nu(u, r_u^0)^2 du.$$

We can further compute $\Theta_T := Cov(\sigma_0 W_T^1, X_T^1)$ as

$$\Theta_T = \sigma_0 \bar{\rho} \int_0^T \int_u^T e^{\kappa s} ds e^{-\kappa u} \nu(u, r_u^0) du.$$

Exploiting now the properties of the Gaussian distribution, following [KK99], we can rewrite

$$\int_0^T r_s^1 ds = \frac{\Theta_T}{\sigma_0 T} W_T^1 + \sqrt{\Lambda_T} Z, \tag{5.70}$$

with $\Lambda_T := Q_T^{X^1} - \frac{\Theta_T^2}{\sigma_0^2 T}$ and Z with law $\mathcal{N}(0, 1)$ and independent from W_T^1 .

Using this we can now prove the following.

Proposition 5.3.15. *Let us consider the particular case of an European call option with payoff $\Phi = (e^{X_T^\epsilon} - K)_+$. The first order correction (in the sense of Remark 5.3.3) to the fair price of a contingent claim Φ given in eq. (5.69) is given by*

$$Pr^1(0; T) = P_{BS} + \epsilon e^{-\int_0^T r_s^0 ds} K \frac{\Theta}{\sigma_0 \sqrt{T}} \phi(-d(1)), \quad (5.71)$$

with ϕ the density function of the standard Gaussian law and $d(1)$ as in Prop. 5.3.4.

Proof. Applying Prop. 5.2.4 to the function $x \mapsto G(x) := e^{-X}$, with $X = \int_0^T r_s^\epsilon ds$ we get

$$\begin{aligned} e^{-\int_0^T r_s^\epsilon ds} &= e^{-\int_0^T r_s^0 ds} - \epsilon e^{-\int_0^T r_s^0 ds} \int_0^T r_s^1 ds + R_1^G(\epsilon, T) = \\ &= e^{-\int_0^T r_s^0 ds} \left(1 - \epsilon \int_0^T r_s^1 ds \right) + R_1^G(\epsilon, T). \end{aligned}$$

The proof follows then taking into account Prop. 5.2.4 applied to the function Φ .

In particular we have that

$$\begin{aligned} Pr^1(0; T) &= \mathbb{E} e^{-\int_0^T r_s^0 ds} \left(1 - \epsilon \int_0^T r_s^1 ds \right) \Phi(X_T^0) + \epsilon \mathbb{E} e^{-\int_0^T r_s^0 ds} \left(1 - \epsilon \int_0^T r_s^1 ds \right) \Phi'(X_T^0) X_T^1 = \\ &= \mathbb{E} e^{-\int_0^T r_s^0 ds} \Phi(X_T^0) - \epsilon \mathbb{E} e^{-\int_0^T r_s^0 ds} \int_0^T r_s^1 ds \Phi(X_T^0) \\ &\quad + \epsilon \mathbb{E} e^{-\int_0^T r_s^0 ds} \Phi'(X_T^0) X_T^1 - \epsilon^2 \mathbb{E} e^{-\int_0^T r_s^0 ds} \int_0^T r_s^1 ds \Phi'(X_T^0) X_T^1 + R_1(\epsilon, T) = \\ &= \mathbb{E} e^{-\int_0^T r_s^0 ds} \Phi(X_T^0) - \epsilon \mathbb{E} e^{-\int_0^T r_s^0 ds} \int_0^T r_s^1 ds \Phi(X_T^0) + \epsilon \mathbb{E} e^{-\int_0^T r_s^0 ds} \Phi'(X_T^0) X_T^1, \end{aligned} \quad (5.72)$$

where the last equality follows from incorporating the term with ϵ^2 into $R_1(\epsilon, T)$.

From the fact that $\Phi = (e^{X_T^0} - K) \mathbb{1}_{[W_T^1 > -\sqrt{T}d(1)]}$ and that $\Phi' = e^{X_T^0} \mathbb{1}_{[W_T^1 > -\sqrt{T}d(1)]}$ we eventually have that

$$Pr^1(0; T) = \mathbb{E} e^{-\int_0^T r_s^0 ds} \Phi(X_T^0) + \epsilon \mathbb{E} e^{-\int_0^T r_s^0 ds} K X_T^1 \mathbb{1}_{[W_T^1 > -\sqrt{T}d(1)]}.$$

Using now (5.70) and exploiting the independence of W_T^1 and Z and the fact that Z has zero mean, and computing explicitly the expectation we get the claim. \square

Remark 5.3.16. So opposite to previous cases discussed here, the expansion underlying our correction in Prop. 5.3.15 is only heuristic, not asymptotic.

The Black-Scholes model with stochastic interest rate and jumps

Let the general setting be as in Sec. 5.3.5, in particular let us consider an asset S_t^ϵ under the risk neutral probability \mathbb{Q} whose return $X_t^\epsilon := \log S_t^\epsilon$ evolves according to

$$\begin{cases} dX_t^\epsilon = \left(r_t^\epsilon - \frac{\sigma_0^2}{2} \right) dt - \epsilon \lambda t \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \sigma_0 dW_t^1 + \epsilon \sum_{k=1}^{N_t} J_k, \\ X_0^\epsilon = x_0, \end{cases} \quad (5.73)$$

with $\sigma_0 \in \mathbb{R}_+$ and r^ϵ is a stochastic interest rate evolving according to eq. (5.65) and \tilde{N} is as in eq. (5.26).

Apply the expansion developed in Th. 5.2.5 (in this case understood as a heuristic one) we get the following result.

Proposition 5.3.17. *Let us consider the normal return process X_t^ϵ evolving according to eq. (5.73) where the stochastic interest rate r_t^ϵ evolves according to the SDE (5.65), with $\nu(t, r)$ satisfying standard condition assuming existence and uniqueness of solutions, and $b(t, r) = \kappa(\theta - r)$ as before in this section. Let us assume further that the jump process is independent of W_t^i , $i = 1, 2$ whereas W^1 in (5.73) and W^2 in (5.67) are correlated as follows*

$$d\langle W^1 W^2 \rangle_t = \bar{\rho} dt,$$

for some constant $\bar{\rho} > 0$. Then the heuristic expansion up to the first order for the stochastic interest rate r_t^ϵ is given by (5.67).

Furthermore the heuristic expansion up to the first order for the return X_t^ϵ is

$$\begin{aligned} X_t^0 &= x_0 + \int_0^t \left(r_s^0 - \frac{\sigma_0^2}{2} \right) ds, \\ X_t^1 &= \int_0^t r_s^1 ds - \lambda t \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \sum_{k=1}^{N_t} J_k. \end{aligned} \tag{5.74}$$

Proof. The proof is similar to the one of Prop. 5.3.14 taking into account the presence of the Poisson random measure. \square

Proposition 5.3.18. *Let us consider the particular case of an European call option with payoff $\Phi(X_T^\epsilon) = (e^{X_T^\epsilon} - K)_+$. The first order correction (both heuristic and in the sense of Remark 5.3.3) to the fair price of the contingent claim Φ is given by*

$$Pr_v^1(0; T) = Pr^1(0; T) + \epsilon T e^{-\int_0^T r_s^0 ds} N(d(1)) \lambda \left(e^{\gamma + \frac{\delta^2}{2}} - 1 \right) + \epsilon T e^{-\int_0^T r_s^0 ds} N(d(1)) \delta \lambda, \tag{5.75}$$

where $Pr^1(0; T)$ is the price in Prop. 5.3.15 and N as in Prop. 5.3.4.

Proof. The proof is similar to the one in Prop. 5.3.15. Proceeding then as in Prop. 5.3.15, taking into account the presence of the jump terms, we get the claim. \square

Appendix

Polynomial Chaos Expansion

In the present section we briefly recall the main characteristics of the *Polynomial Chaos Expansion* (PCE) approach, we refer the interested reader to, e.g., [DPPB15b, EMSU12a, EMSU12b, PT11] and references therein, for a detailed introduction to such a method. The PCE method allows to approximate a random variable as a linear combination of orthogonal polynomials in order to compute its statistics with low computational effort. Before entering into details, we would like to underline that the PCE is, in fact, a generalisation of the original Wiener Chaos decomposition, see [Wie38].

Let us consider a standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the Hilbert space of real-valued random variables $L^2(\Omega, \mathcal{F}, \mathbb{P})$, i.e. of real-valued random variables X defined on $(\Omega, \mathcal{F}, \mathbb{P})$, such that

$$\mathbb{E}[X^2] = \int_{\Omega} (X(\omega))^2 \mathbb{P}(d\omega) < +\infty.$$

Moreover $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is equipped with the standard scalar product

$$\mathbb{E}[XY] = \langle X, Y \rangle_{\mathbb{P}} = \int_{\Omega} X(\omega)Y(\omega)\mathbb{P}(d\omega) ,$$

and corresponding norm

$$\|X\|_{\mathbb{P}}^2 = \mathbb{E}[X^2] .$$

The related convergence will be always referred to as *mean square convergence* or *strong convergence*.

Among the elements of $L^2(\Omega, \mathcal{F}, \mathbb{P})$, there is the class of *basic random variables*, which is used to decompose quantity of interest like, e.g., the solution of a Stochastic Differential Equation (SDE) at finite time $T > 0$. We notice that not all the functions $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ can be used to perform such a decomposition since they have to satisfy, see, e.g., [EMSU12b, Section 3], at least the following two properties

- ξ has finite moments of all orders
- the distribution function $F_{\xi}(x) := \mathbb{P}(\xi \leq x)$, $x \in \mathbb{R}$, of the basic random variables is absolutely continuous, with a probability density function (pdf) denoted by f_{ξ} .

Let us denote by $\sigma(\xi)$ the σ -algebra generated by the basic random variable ξ , hence $\sigma(\xi) \subset \mathcal{F}$. If we want to polynomially decompose a given random variable Y in terms of ξ , then Y has to be, at least, measurable with respect to the σ -algebra $\sigma(\xi)$. Exploiting the Doob-Dynkin Lemma, see, e.g., [Kal06, Lemma 1.13], we have that Y is $\sigma(\xi)$ -measurable if for some Borel measurable function $g : \mathbb{R} \rightarrow \mathbb{R}$, $Y = g(\xi)$. In what follows, without loss of generality, we restrict ourselves to consider the decomposition in $L^2(\Omega, \sigma(\xi), \mathbb{P})$. The basic random variable ξ is assumed to determine a class of orthogonal polynomials $\{\Psi_i(\xi)\}_{i \in \mathbb{N}}$, which is called the *generalized polynomial chaos* (gPC) basis. We underline that their orthogonality properties is detected by means of the measure induced by ξ in the image space $(D, \mathcal{B}(D))$, where $D \subset \mathbb{R}$ is the range of ξ and where $\mathcal{B}(D) \subset \mathcal{B}(\mathbb{R})$ denotes the Borel σ -algebra associated with D . For each $i, j \in \mathbb{N}$, we have

$$\langle \Psi_i, \Psi_j \rangle_{\mathbb{P}} = \int_{\Omega} \Psi_i(\xi(\omega)) \Psi_j(\xi(\omega)) d\mathbb{P}(\omega) = \int_D \Psi_i(x)\Psi_j(x)f_{\xi}(x)dx . \quad (5.76)$$

If ξ has law $\mathcal{N}(0, 1/2)$, namely the centred normal distribution of variance $\frac{1}{2}$, then the related set $\{\Psi_i(x)\}_{i \in \mathbb{N}}$ is represented by the family of non-normalized Hermite polynomials defined on the whole real line, namely $D = \mathbb{R}$, and

$$\begin{cases} \Psi_0(x) &= 1 \\ \Psi_1(x) &= 2x \\ \Psi_2(x) &= 4x^2 - 2 \\ &\vdots \end{cases} \quad (5.77)$$

Figure 5.1 provides the graph of the first six orthonormal polynomials, achieved by scaling each Ψ_i in (5.77) by its norm in $L^2(\Omega, \sigma(\xi), \mathbb{P})$, namely, $\forall i \in \mathbb{N}$, Ψ_i is divided by $\|\Psi_i\|_{\mathbb{P}} := \sqrt{2^i i!}$.

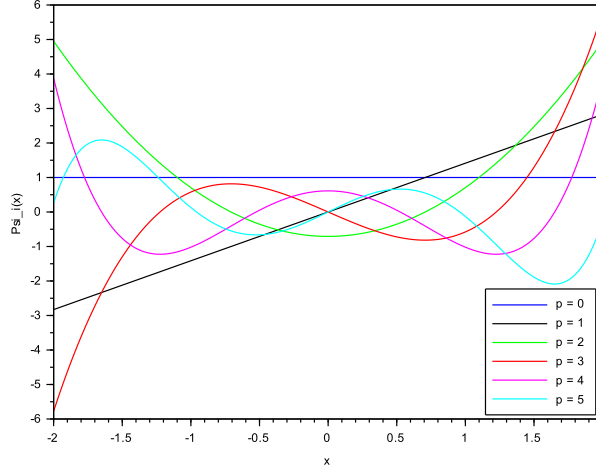


Figure 5.1: The Hermite normalized polynomials up to degree 5

Latter polynomials Ψ_i constitute a maximal system in $L^2(\Omega, \sigma(\xi), \mathbb{P})$, therefore every random variable $Y \in L^2(\Omega, \sigma(\xi), \mathbb{P})$ can be approximated as follows

$$Y^{(p)} = \sum_{i=0}^p c_i \Psi_i(\xi), \quad (5.78)$$

for suitable coefficients c_i which depend on the random variable Y , see, e.g., [EMSU12b, Section 3.1]. We refer to eq. (5.78) as the *truncated PCE*, at degree p , of Y . Exploiting previous definitions, taking $i \in \{0, \dots, p\}$, and considering the orthogonality property of the polynomials $\{\Psi_i(\xi)\}_{i \in \mathbb{N}}$, we have

$$c_i = \frac{1}{\|\Psi_i\|_{\mathbb{P}}^2} \langle Y, \Psi_i \rangle_{\mathbb{P}} = \frac{1}{\|\Psi_i\|_{\mathbb{P}}^2} \langle g, \Psi_i \rangle_{\mathbb{P}}, \quad (5.79)$$

and, since $Y = g(\xi)$, we also obtain

$$\langle Y, \Psi \rangle_{\mathbb{P}} = \langle g, \Psi \rangle_{\mathbb{P}} = \int_{\Omega} g(\xi(\omega)) \Psi_i(\xi(\omega)) dP(\omega) = \int_{\mathbb{R}} g(x) \Psi_i(x) f_{\xi}(x) dx. \quad (5.80)$$

The convergence rate of the PCE-approximation (5.78) in $L^2(\Omega, \sigma(\xi), \mathbb{P})$ norm is strictly linked to the magnitude of the coefficients of the decomposition. Indeed, by the Parseval identity, we have

$$\|Y\|_{\mathbb{P}}^2 = \sum_{i=0}^{+\infty} c_i^2 \|\Psi_i\|_{\mathbb{P}}^2,$$

furthermore, using the orthogonality property of the Hermite polynomials in $L^2(\Omega, \sigma(\xi), \mathbb{P})$, the norm of (5.78) is given by

$$\|Y^{(p)}\|_{\mathbb{P}}^2 = \sum_{i=0}^p c_i^2 \|\Psi_i\|_{\mathbb{P}}^2.$$

Exploiting the fundamental properties of the orthogonal projections in Hilbert space, see, e.g., [Rud87, Theorem 4.11], we can estimate the mean square error as

$$\|Y - Y^{(N)}\|_{\mathbb{P}}^2 = \|Y\|_{\mathbb{P}}^2 - \|Y^{(N)}\|_{\mathbb{P}}^2 = \sum_{i=N+1}^{+\infty} c_i^2 \|\Psi_i\|_{\mathbb{P}}^2, \quad (5.81)$$

thus the rate of convergence depends on the coefficients. In particular the PCE of $Y^{(p)}$ approximates the Y -statistics in terms of the c_i coefficients appearing in eq. (5.78), e.g. the first two centred moments are determined by

$$\mathbb{E} \left[Y^{(p)} \right] = c_0, \quad (5.82)$$

$$\text{Var} \left[Y^{(p)} \right] = \sum_{i=1}^p c_i^2 \|\Psi_i\|_{\mathbb{P}}^2. \quad (5.83)$$

Multi-element decomposition

Concerning the application of the PCE method to the approximation of quantities as in the case of European call options, we have implemented a method called multi-element generalized polynomial chaos (ME-gPC) method, see, e.g., [PT11], and references therein. Without entering into technical details, let us mention that it is an extension of the PCE approach which can be applied to arbitrary probability measures. In particular, see, e.g., [WK05, WK06], the ME-gPC approach can be effectively used to numerically solve S(P)DEs, by decomposing the random inputs, e.g. the Brownian motion, into smaller elements. Each of the latter is then used to define a new random variable, with respect to a conditional probability density function, and a set of orthogonal polynomials defined in terms of the aforementioned random variable. Then, the procedure we have recalled in the Appendix, is applied element-by-element and, thanks to the convergence of the method, the final result is achieved rearranging, in a suitable way, the ones obtained for each term.

5.4 Small noise expansion for the Lévy perturbed Vasicek model

The Vasicek model (together with the CIR model) is one of the most used short rate models. It assumes that the interest rate under the risk neutral measure \mathbb{Q} evolves according to a mean reverting Ornstein-Uhlenbeck process with constant coefficients, see, e.g. [DF01] for details. In particular the interest rate r_t is the solution of the following linear stochastic equation

$$\begin{cases} dr_t = \kappa [\theta - r_t] dt + \sigma dW_t, \\ r_0 = r_0, \end{cases} \quad (5.84)$$

with κ , θ , σ and r_0 some positive constants. The price of a pure discounted bond, better known as *zero-coupon bond* (ZCB), in the Vasicek model can be explicitly computed, see, e.g. [DF01] as

$$\text{ZCB}(t; T) = \mathbb{E}_t \left[e^{-\int_t^T r_s ds} \right] = A(t; T) e^{-B(t; T) r_t}, \quad (5.85)$$

$$A(t; T) := \exp \left\{ \left(\theta - \frac{\sigma^2}{2\kappa^2} \right) (B(t; T) - T + t) - \frac{\sigma^2}{4\kappa} B(t; T)^2 \right\},$$

$$B(t; T) := \frac{1}{\kappa} \left(1 - e^{-\kappa(T-t)} \right).$$

The price of an option with payoff $\Phi(r_T)$ written on the interest rate r_t is given by

$$ZBO(t; T) = \mathbb{E}_t \left[e^{-\int_t^T r_s ds} \Phi(r_T) \right].$$

In the particular case of an European call/put option, as the one introduced in the previous BS model, the formula can be explicitly computed, see, e.g. [DF01] Sec. 3.2.1.

From Theorem 5.2.5 we deduce that $\Phi(r_t^\epsilon)$ has an asymptotic expansion in powers of ϵ of the form

$$\Phi(r_t^\epsilon) = \sum_{k=0}^H \epsilon^k [\Phi(r_t^\epsilon)]_k + R_H(\epsilon, t), \quad (5.86)$$

with

$$\sup_{s \in [0, t]} |R_H(\epsilon, s)| \leq C_{H+1}(t) \epsilon^{H+1},$$

and the coefficients can be computed from the expansions coefficients of r_t^ϵ , as discussed in section 5.2, where also the Taylor coefficients of Φ are treated.

5.4.1 The Vasicek model: a first order correction

Applying the results in Sec. 5.2, let us then consider the following perturbed Vasicek model

$$\begin{cases} dr_t^\epsilon = \kappa [\theta - r_t^\epsilon] dt + (\sigma_0 + \epsilon \sigma_1 f(r_t^\epsilon)) dW_t, \\ r_0^\epsilon = r_0, \end{cases}, \quad (5.87)$$

with σ_0 and σ_1 some positive constants, f a smooth real valued function, $0 \leq \epsilon \leq \epsilon_0$.

Let us now consider the particular case $f(r) = e^{\alpha r}$, for some $\alpha \in \mathbb{R}$, then we get the following proposition.

Proposition 5.4.1. *For the particular case where $f(r) = e^{\alpha r}$, for some $\alpha \in \mathbb{R}$, we have that r_t^ϵ can be written a power series, namely*

$$r_t^\epsilon = r_t^0 + \epsilon r_t^1 + R_1(\epsilon, t),$$

where the expansion coefficients read as

$$\begin{aligned} r_t^0 &= r_0 e^{-\kappa t} + \theta (1 - e^{-\kappa t}) + \sigma_0 \int_0^t e^{\kappa(t-s)} dW_s, \quad \text{with law } \mathcal{N}(\mu_t, Q_t), \\ r_t^1 &= \frac{\sigma_1}{\alpha \sigma_0} (e^{\alpha r_t^0} - e^{\alpha r_0}) + \int_0^t C_1^\alpha e^{-\kappa(t-s)} e^{\alpha r_s^0} ds + \int_0^t C_2^\alpha e^{-\kappa(t-s)} r_s^0 e^{\alpha r_s^0}, \end{aligned} \quad (5.88)$$

with

$$\begin{aligned} \mu_t &= r_0 e^{-\kappa t} + \theta (1 - e^{-\kappa t}), \quad Q_t = \frac{\sigma_0^2}{2\kappa} (1 - e^{-2\kappa t}), \\ C_1^\alpha &= -\frac{\sigma_1}{\alpha \sigma_0} \left(\kappa + \kappa \theta \alpha + \frac{1}{2} \alpha^2 \sigma_0^2 \right), \quad C_2^\alpha = \frac{\sigma_1}{\alpha \sigma_0} \kappa \alpha. \end{aligned}$$

Proof. Applying Th. 5.2.5 we have that expanding eq. (5.87) up to the first order we get

$$\begin{aligned} r_t^0 &= r_0 + \int_0^t \kappa [\theta - r_s^0] ds + \int_0^t \sigma_0 dW_s, \\ r_t^1 &= - \int_0^t \kappa r_s^1 ds + \int_0^t \sigma_1 e^{\alpha r_s^0} dW_s, \end{aligned} \quad (5.89)$$

An application of Itô's lemma to $g(s, r) = e^{\kappa t} r_t^0$ gives us that

$$r_t^0 = r_0 e^{-\kappa t} + \theta (1 - e^{-\kappa t}) + \sigma_0 \int_0^t e^{\kappa(t-s)} W_s, \text{ with law } \mathcal{N}(\mu_t, Q_t).$$

Computing r_t^1 , in the same manner, we have that applying Itô's lemma to $g(s, r) = e^{\kappa t} r_t^1$ it follows

$$r_t^1 = \sigma_1 \int_0^t e^{-\kappa(t-s)} e^{\alpha r_s^0} dW_s,$$

Applying again Itô's lemma to the function $h(r) = e^{\alpha r} e^{\kappa s}$ we get

$$e^{\alpha r_t^0} - e^{\alpha r_0^0} = \int_0^t (\gamma - \kappa \alpha r_s^0) e^{\alpha r_s^0} ds + \int_0^t \alpha \sigma_0 e^{\alpha r_s^0} dW_s, \quad (5.90)$$

with $\gamma := \kappa + \alpha \kappa \theta + \frac{\alpha^2 \sigma_0^2}{2}$. The expression for r_t^1 thus follows applying eq. (5.90) and solving for integral w.r.t. the Brownian motion. \square

Remark 5.4.2. With the same argument we can derive also the second correction term r_t^2 . In fact applying Th. 5.2.5 we have that

$$r_t^2 = - \int_0^t \kappa r_s^2 ds + \int_0^t \sigma_1 \alpha e^{\alpha r_s^0} dW_s = \sigma_1 \alpha \int_0^t e^{-\kappa(t-s)} e^{\alpha r_s^0} dW_s.$$

The particular choice of $f(r) = e^{\alpha r}$ can easily be extended to any real function which can be written as a Fourier transform, resp. Laplace transform, $f(r) = \int_{\mathbb{R}_0} e^{ir\alpha} \lambda(d\alpha)$, resp. $f(r) = \int_{\mathbb{R}_0} e^{\alpha r} \lambda(d\alpha)$ of some positive measure λ on \mathbb{R}_0 (e.g. a probability measure) resp. which has finite Laplace transform. Formulae (5.92) holds with $K_\alpha e^{\alpha r_\tau^0}$ replaced by $\int_{\mathbb{R}_0} K_\alpha e^{i\alpha r_\tau^0} \lambda(d\alpha)$, resp. $\int_{\mathbb{R}_0} K_\alpha e^{\alpha r_\tau^0} \lambda(d\alpha)$, which are finite if, e.g. $\int_{\mathbb{R}_0} |K_\alpha| \lambda(d\alpha) < \infty$, resp. λ has compact support. In fact eq. (5.90) gets replaced by

$$\begin{aligned} \int_{\mathbb{R}_0} e^{\alpha r_t^0} \lambda(d\alpha) &= 1 + \int_{\mathbb{R}} \left[\int_0^t \left(e^{\alpha r_s^0} \alpha \mu + \frac{\alpha^2}{2} \sigma_0^2 e^{\alpha r_s^0} \right) ds \right] \lambda(d\alpha) \\ &\quad + \int_{\mathbb{R}} \left[\int_0^t e^{\alpha r_s^0} \alpha \sigma_0 dW_s \right] \lambda(d\alpha). \end{aligned}$$

By repeating the steps used before we get the statements in Prop. 5.4.3 extended to these more general cases.

5.4.2 The Vasicek model: a first order correction with jumps

In the present section we will deal with the previous model with an addition of a small perturbed Poisson compensated measure $\tilde{\mathcal{N}}$. In particular we will assume $\tilde{N}(t, A) := N(t, A) - t\nu(A)$ for all $A \in \mathcal{B}(\mathbb{R}, 0)$, $0 \notin \bar{A}$, N being a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}_0$ and $\nu(A) = \mathbb{E}(N(1, A))$, while $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$ and $\int_{\mathbb{R}_0} (|x|^2 \wedge 1) \nu(dx) < \infty$. Eventually the Poisson random measure is assumed to be independent of the Brownian motion W_t . We refer to [App09] for details on Levy processes.

Under previous conditions let us assume we are given an interest rate r_t^ϵ evolving according to the SDE

$$\begin{cases} dr_t = \kappa[\theta - r_t] dt + (\sigma_0 + \epsilon\sigma_1 f(r_t^\epsilon)) dW_t + \epsilon \int_0^t \int_{\mathbb{R}_0} x \tilde{N}(ds, dx), \\ r_0 = r_0, \end{cases} \quad (5.91)$$

with the notation as previously introduced.

Let us again consider the particular case $f(r) = e^{\alpha r}$, for some $\alpha \in \mathbb{R}$, then we get the following proposition.

Proposition 5.4.3. *For the particular case where $f(r) = e^{\alpha r}$, for some $\alpha \in \mathbb{R}_0$, we have that r_t^ϵ can be written a power series, namely*

$$r_t^\epsilon = r_t^0 + \epsilon r_t^1 + R_1(\epsilon, t),$$

where the expansion coefficients read as

$$\begin{aligned} r_t^0 &= r_0 e^{-\kappa t} + \theta (1 - e^{-\kappa t}) + \sigma_0 \int_0^t e^{\kappa(t-s)} dW_s, \quad \text{with law } \mathcal{N}(\mu_t, Q_t), \\ r_t^1 &= \frac{\sigma_1}{\alpha\sigma_0} (e^{\alpha r_t^0} - e^{\alpha r_0}) + \int_0^t C_\alpha^1 e^{-\kappa(t-s)} e^{\alpha r_s^0} ds + \int_0^t C_\alpha^2 e^{-\kappa(t-s)} r_s^0 e^{\alpha r_s^0} \\ &\quad + \int_0^t \int_{\mathbb{R}_0} x \tilde{N}(ds, dx), \end{aligned} \quad (5.92)$$

with constants as in Prop. 5.4.4.

Proof. The proof is completely analogous to the one of Prop. 5.4.4 just taking into account the Poisson random measure. \square

5.4.3 Application to pricing

Expanding the payoff function Φ , assumed to be smooth, according to eq. (5.86) we have that the first order correction to the fair price of an option written on the underlying r^ϵ is given by

$$\begin{aligned} ZBO^1(0; T) &= \mathbb{E} \left[e^{-\int_0^T r_s^\epsilon ds} \Phi(r_T^\epsilon) \right] = \mathbb{E} \left[e^{-\int_0^T r_s^0 ds} \left(1 - \epsilon \int_0^T r_s^1 ds \right) \Phi(r_T^0) \right] \\ &\quad + \mathbb{E} \left[e^{-\int_0^T r_s^0 ds} \left(1 - \epsilon \int_0^T r_s^1 ds \right) \Phi'(r_T^0) \right]. \end{aligned} \quad (5.93)$$

Proposition 5.4.4. *The first order corrected fair price of an option written on the underlying r_t^ϵ reads as*

$$\begin{aligned} ZBO^1(0; T) &= \mathbb{E} \left[e^{-\int_0^T r_s^\epsilon ds} \Phi(r_T^\epsilon) \right] = ZBO + \epsilon \mathbb{E} \left[e^{-\int_0^t r_s^0 ds} \int_0^T r_s^1 ds \Phi(r_T^0) \right] \\ &\quad + \epsilon \mathbb{E} \left[e^{-\int_0^t r_s^0 ds} \Phi'(r_T^0) \right]. \end{aligned} \quad (5.94)$$

Proof. Expanding the r_s^ϵ in a converging power series we have that

$$e^{-\int_0^t r_s^\epsilon ds} = e^{-\int_0^t r_s^0 ds - \epsilon \int_0^t r_s^1 ds + R_1(\epsilon, t)} = e^{-\int_0^t r_s^0 ds} \left(1 - \epsilon \int_0^t r_s^1 ds + \tilde{R}_1(\epsilon, t) \right),$$

where r_t^0, r_t^1 are given in Prop. 5.4.3 for the particular case of $f(x) = e^{\alpha x}$.

The expansion in eq. (5.94) follows applying Th. 5.2.5 to the payoff function Φ . \square

If we consider the particular case of pricing a *zero coupon bond* (ZCB), namely we consider a terminal payoff $\Phi = 1$, we have the following.

Proposition 5.4.5. *The first order corrected fair price of an option written on the underlying r_t^ϵ reads as*

$$ZCB^1(0; T) = ZCB + \epsilon \mathbb{E} \left[e^{-\int_0^t r_s^0 ds} \int_0^T r_s^1 ds \right], \quad (5.95)$$

where ZCB is the price given in eq. (5.85).

Proof. The claim follows by Prop. 5.4.4, with $\Phi = 1$. \square

Bibliography

- [ADEH99] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.
- [ADP11] Sergio Albeverio and Luca Di Persio. Some stochastic dynamical models in neurobiology: recent developments. *European Communications in Mathematical and Theoretical Biology*, 14:44–53, 2011.
- [ADPM11] Sergio Albeverio, Luca Di Persio, and Elisa Mastrogiacomo. Small noise asymptotic expansions for stochastic pde’s, i. the case of a dissipative polynomially bounded nonlinearity. *Tohoku Mathematical Journal, Second Series*, 63(4):877–898, 2011.
- [ADPM13] Sergio Albeverio, Luca Di Persio, and Elisa Mastrogiacomo. Invariant measures for stochastic differential equations on networks. In *Proc. Sympos. Pure Math*, volume 87, pages 1–33, 2013.
- [ADPMS14a] Sergio Albeverio, Luca Di Persio, Elisa Mastrogiacomo, and Boubaker Smii. A class of lévy driven sdes and their explicit invariant measures. *arXiv preprint arXiv:1407.3943*, 2014.
- [ADPMS14b] Sergio Albeverio, Luca Di Persio, Elisa Mastrogiacomo, and Boubaker Smii. Invariant measures for sdes driven by lévy noise. a case study for dissipative nonlinear drift in infinite dimension. *submitted*, 2014.
- [AF90] Walter G Aiello and HI Freedman. A time-delay model of single-species growth with stage structure. *Mathematical biosciences*, 101(2):139–153, 1990.
- [AHK12] Sergio Albeverio, Astrid Hilbert, and Vassily Kolokoltsov. Uniform asymptotic bounds for the heat kernel and the trace of a stochastic geodesic flow. *Stochastics An International Journal of Probability and Stochastic Processes*, 84(2-3):315–333, 2012.
- [AHMP07] Mercedes Arriojas, Yaozhong Hu, Salah-Eldin Mohammed, and Gyula Pap. A delayed Black and Scholes formula. *Stoch. Anal. Appl.*, 25(2):471–492, 2007.
- [AL13] Leif Andersen and Alexander Lipton. Asymptotics for exponential lévy processes and their volatility smile: survey and new results. *International Journal of Theoretical and Applied Finance*, 16(01):1350001, 2013.

- [All10] Linda JS Allen. *An introduction to stochastic processes with applications to biology*. CRC Press, 2010.
- [App09] David Applebaum. *Lévy Processes and Stochastic Calculus*, volume 116 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2009.
- [AR] Nacira Agram and Elin Engen Røse. Optimal control of forward-backward mean-field stochastic delayed systems. *arXiv:1412.5291*.
- [AR01] Søren Asmussen and Jan Rosiński. Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Probab.*, 38(2):482–493, 2001.
- [Are02] Wolfgang Arendt. Semigroups and evolution equations: functional calculus, regularity and kernel estimates. *Handbook of Differential Equations: Evolutionary Equations*, 1:1–85, 2002.
- [AS99] Sergio Albeverio and Victoria Steblovskaya. Asymptotics of infinite-dimensional integrals with respect to smooth measures i. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 2(04):529–556, 1999.
- [AS14] Erdinç Akyıldırım and Halil Mete Soner. A brief history of mathematics in finance. *Borsa Istanbul Review*, 14(1):57–63, 2014.
- [AS15] Sergio Albeverio and Boubaker Smii. Asymptotic expansions for sde’s with small multiplicative noise. *Stochastic processes and their applications*, 125(3):1009–1031, 2015.
- [ASG16] Di Persio Luca Albeverio Sergio, Cordoni Francesco and Pellegrini Gregorio. Asymptotic expansion for some local volatility models arising in finance. *submitted*, 2016.
- [Bac00] Louis Bachelier. *Théorie de la spéculation*. Gauthier-Villars, 1900.
- [Bar92] Viorel Barbu. *Analysis and control of nonlinear infinite dimensional systems*. Elsevier, 1992.
- [Bar10] Viorel Barbu. *Nonlinear differential equations of monotone types in Banach spaces*. Springer Science & Business Media, 2010.
- [Bat96] David S Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of financial studies*, 9(1):69–107, 1996.
- [BBP97] Guy Barles, Rainer Buckdahn, and Etienne Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics: An International Journal of Probability and Stochastic Processes*, 60(1-2):57–83, 1997.
- [BCDNR16] David R. Baños, Francesco Cordoni, Di Persio Luca Di Nunno, Giulia, and Elin E. Røse. Stochastic systems with memory and jumps. *arXiv preprint arXiv:1603.00272*, 2016.

- [BCM08] Stefano Bonaccorsi, Fulvia Confortola, and Elisa Mastrogiacono. Optimal control of stochastic differential equations with dynamical boundary conditions. *Journal of Mathematical Analysis and Applications*, 344(2):667–681, 2008.
- [BCP15] Viorel Barbu, Francesco Cordini, and Luca Di Persio. Optimal control of stochastic fitzhugh–nagumo equation. *International Journal of Control*, pages 1–11, 2015.
- [BDNK11] Fred Espen Benth, Giulia Di Nunno, and Asma Khedher. Robustness of option prices and their deltas in markets modelled by jump-diffusions. *Commun. Stoch. Anal.*, 5(2):285–307, 2011.
- [BEK07] Pauline M Barrieu and Nicole El Karoui. Pricing, hedging and optimally designing derivatives via minimization of risk measures. 2007.
- [BFP15] Damiano Brigo, Marco Francischello, and Andrea Pallavicini. Invariance, existence and uniqueness of solutions of nonlinear valuation pdes and fbsdes inclusive of credit risk, collateral and funding costs. *Collateral and Funding Costs (June 1, 2015)*, 2015.
- [BGM09] Eric Benhamou, Emmanuel Gobet, and Mohammed Miri. Smart expansion and fast calibration for jump diffusions. *Finance and Stochastics*, 13(4):563–589, 2009.
- [BI99] Viorel Barbu and Mimmo Iannelli. Optimal control of population dynamics. *Journal of optimization theory and applications*, 102(1):1–14, 1999.
- [Bil68] Patrick Billingsley. *Convergence of Probability Measures*. John Wiley & Sons Inc., New York, 1968.
- [Bis73] Jean-Michel Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.
- [BJ00] Nadine Bellamy and Monique Jeanblanc. Incompleteness of markets driven by a mixed diffusion. *Finance and Stochastics*, 4(2):209–222, 2000.
- [Bjö04] Tomas Björk. *Arbitrage theory in continuous time*. Oxford university press, 2004.
- [BJPR09] Tomasz R Bielecki, Monique Jeanblanc-Picqué, and Marek Rutkowski. *Credit risk modeling*. Osaka University Press Osaka, 2009.
- [BK14] Tobias Breiten and Karl Kunisch. Riccati-based feedback control of the monodomain equations with the fitzhugh–nagumo model. *SIAM Journal on Control and Optimization*, 52(6):4057–4081, 2014.
- [BKR97] Tomas Björk, Yuri Kabanov, and Wolfgang Runggaldier. Bond market structure in the presence of marked point processes. *Mathematical Finance*, 7(2):211–239, 1997.
- [BM94] Alan Brace and Marek Musiela. A multifactor gauss markov implementation of heath, jarro, and morton. *Mathematical Finance*, 4(3):259–283, 1994.

- [BM⁺97] Alan Brace, Marek Musiela, et al. The market model of interest rate dynamics. *Mathematical finance*, 7(2):127–155, 1997.
- [BM08] Stefano Bonaccorsi and Elisa Mastrogiacomo. Analysis of the stochastic fitzhugh–nagumo system. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 11(03):427–446, 2008.
- [BMB05] Fred Espen Benth and Thilo Meyer-Brandis. The density process of the minimal entropy martingale measure in a stochastic volatility model with jumps. *Finance and Stochastics*, 9(4):563–575, 2005.
- [BMP13] Damiano Brigo, Massimo Morini, and Andrea Pallavicini. *Counterparty credit risk, collateral and funding: with pricing cases for all asset classes*. John Wiley & Sons, 2013.
- [BMZ⁺08] Stefano Bonaccorsi, Carlo Marinelli, Giacomo Ziglio, et al. Stochastic fitzhugh–nagumo equations on networks with impulsive noise. *Electron. J. Probab*, 13:1362–1379, 2008.
- [BN97] Ole E Barndorff-Nielsen. Processes of normal inverse gaussian type. *Finance and stochastics*, 2(1):41–68, 1997.
- [BNNS⁺02] Ole E Barndorff-Nielsen, Elisa Nicolato, Neil Shephard, et al. Some recent developments in stochastic volatility modelling. *Quantitative Finance*, 2(1):11–23, 2002.
- [BP01] András Bátkai and Susanna Piazzera. Semigroups and linear partial differential equations with delay. *Journal of mathematical analysis and applications*, 264(1):1–20, 2001.
- [BP04] András Bátkai and Susanna Piazzera. Semigroups for delay equations. *Interplay Between IC02-semigroups and PDEs*, 2004.
- [BP12] Viorel Barbu and Teodor Precupanu. *Convexity and optimization in Banach spaces*. Springer Science & Business Media, 2012.
- [BP13] Damiano Brigo and Andrea Pallavicini. Ccps, central clearing, csa, credit collateral and funding costs valuation faq: Re-hypothecation, cva, closeout, netting, wwr, gap-risk, initial and variation margins, multiple discount curves, fva? *Central Clearing, CSA, Credit Collateral and Funding Costs Valuation FAQ: Re-Hypothecation, CVA, Closeout, Netting, WWR, Gap-Risk, Initial and Variation Margins, Multiple Discount Curves, FVA*, 2013.
- [BP14] Damiano Brigo and Andrea Pallavicini. Ccp cleared or bilateral csa trades with initial/variation margins under credit, funding and wrong-way risks: A unified valuation approach. *Variation Margins Under Credit, Funding and Wrong-Way Risks: A Unified Valuation Approach (January 16, 2014)*, 2014.
- [BPP11] Damiano Brigo, Andrea Pallavicini, and Vasileios Papatheodorou. Arbitrage-free valuation of bilateral counterparty risk for interest-rate products: impact of volatilities and correlations. *International Journal of Theoretical and Applied Finance*, 14(06):773–802, 2011.

- [BR] Elena Bandini and Francesco Russo. Weak Dirichlet processes with jumps. *arXiv:1512.06236*.
- [BR11] Viorel Barbu and Michael Röckner. On a random scaled porous media equation. *Journal of Differential Equations*, 251(9):2494–2514, 2011.
- [BR13] Tomasz R Bielecki and Marek Rutkowski. *Credit risk: modeling, valuation and hedging*. Springer Science & Business Media, 2013.
- [Bre94] Karl Wilhelm Breitung. Asymptotic approximations for probability integrals. 1994.
- [BS73] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973.
- [BTR07] Ian Boutle, Richard HS Taylor, and Rudolf A Römer. El niño and the delayed action oscillator. *American Journal of Physics*, 75(1):15–24, 2007.
- [Bur98] Victor I Burenkov. *Sobolev spaces on domains*. Springer, 1998.
- [BZ05] S Bonaccorsi and G Ziglio. A semigroup approach to stochastic dynamical boundary value problems. In *Systems, control, modeling and optimization*, pages 55–65. Springer, 2005.
- [CDGR] Andrea Cosso, Cristina Di Girolami, and Francesco Russo. Calculus via regularizations in Banach spaces and Kolmogorov-type path-dependent equations. *arXiv:1411.8000*.
- [CDGR11] Rosanna Coviello, Cristina Di Girolami, and Francesco Russo. On stochastic calculus related to financial assets without semimartingales. *Bulletin des Sciences mathématiques*, 135(6):733–774, 2011.
- [CDP13] Francesco Cordoni and Luca Di Persio. Transition density for the cir process by lie symmetries and application to zcb pricing. *International Journal of Pure and Applied Mathematics*, 88(2):239–246, 2013.
- [CDP14a] Francesco Cordoni and Luca Di Persio. Asymptotic expansion for the characteristic function of a multiscale stochastic volatility model. *International Journal of Applied Mathematics*, 27(3):211–223, 2014.
- [CDP14b] Francesco Cordoni and Luca Di Persio. Backward stochastic differential equations approach to hedging, option pricing, and insurance problems. *International Journal of Stochastic Analysis*, 2014, 2014.
- [CDP14c] Francesco Cordoni and Luca Di Persio. Lie symmetry approach to the cev model. *International Journal of Differential Equations and Applications*, 13(3), 2014.
- [CDP15a] Francesco Cordoni and Luca Di Persio. Invariant measure for the vasicek interest rate model in the heath–jarrow–morton–musiela framework. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 18(03):1550022, 2015.

- [CDP15b] Francesco Cordonì and Luca Di Persio. A note on a representation theorem for path-dependent pìde with lévy noise. (*preprint*), 2015.
- [CDP15c] Francesco Cordonì and Luca Di Persio. Pricing pìde for path-dependent options in some stochastic volatility models with jumps. (*preprint*), 2015.
- [CDP15d] Francesco Cordonì and Luca Di Persio. Small noise asymptotic expansion for a infinite dimensional stochastic reaction-diffusion forced van der pol equation. *International Journal of Mathematical Models and Method in Applied Sciences*, 9, 2015.
- [CDP15e] Francesco Cordonì and Luca Di Persio. Small noise expansion for the lévy perturbed vasicek model. *International Journal of Pure and Applied Mathematics*, 98(2):291, 2015.
- [CDP16a] Francesco Cordonì and Luca Di Persio. A bsde with delayed generator approach to pricing under counterparty risk and collateralization. *International Journal of Stochastic Analysis*, 2016, 2016.
- [CDP16b] Francesco Cordonì and Luca Di Persio. Gaussian estimates on networks with dynamic stochastic boundary conditions. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 2016.
- [CDP16c] Francesco Cordonì and Luca Di Persio. Optimal control for the stochastic fitzhugh-nagumo model with recovery variable. *submitted*, 2016.
- [CDP16d] Francesco Cordonì and Luca Di Persio. A stochastic dynamic boundary problems with time delayed boundary conditions on a network. *submitted*, 2016.
- [CDPMZ16] Francesco Cordonì, Luca Di Persio, Lucian Maticiuc, and Adrian Zălinescu. A stochastic approach to path-dependent nonlinear kolmogorov equations via bsdes with time-delayed generators and applications to finance. *arXiv preprint arXiv:1602.05793*, 2016.
- [CDPO16] Francesco Cordonì, Luca Di Persio, and Immacolata Oliva. A nonlinear kolmogorov equation for stochastic functional delay differential equations with jumps. *arXiv preprint arXiv:1602.03851*, 2016.
- [CF10] Rama Cont and David-Antoine Fournié. Change of variable formulas for non-anticipative functionals on path space. *Journal of Functional Analysis*, 259(4):1043–1072, 2010.
- [CF13] Rama Cont and David-Antoine Fournié. Functional Itô calculus and stochastic integral representation of martingales. *The Annals of Probability*, 41(1):109–133, 2013.
- [CFR13] Peter Carr, Travis Fisher, and Johannes Ruf. Why are quadratic normal volatility models analytically tractable? *SIAM Journal on Financial Mathematics*, 4(1):185–202, 2013.
- [CGMY03] Peter Carr, Hélyette Geman, Dilip B Madan, and Marc Yor. Stochastic volatility for lévy processes. *Mathematical Finance*, 13(3):345–382, 2003.

- [CIJR85] John C Cox, Jonathan E Ingersoll Jr, and Stephen A Ross. A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407, 1985.
- [CLMM09] Thierry Crestaux, Olivier Le Maître, and Jean-Marc Martinez. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161–1172, 2009.
- [CM78] Anna Chojnowska-Michalik. Representation theorem for general stochastic delay equations. *Acad. Pol. Sci., Ser. Sci. Math. Astron. Phys*, 26:635–642, 1978.
- [CM96] Jakša Cvitanić and Jin Ma. Hedging options for a large investor and forward-backward sde's. *The annals of applied probability*, 6(2):370–398, 1996.
- [CM08] Stefano Cardanobile and Delio Mugnolo. Qualitative properties of coupled parabolic systems of evolution equations. *Annali della Scuola Normale Superiore di Pisa—Classe di Scienze*, 7(2):287–312, 2008.
- [CM11] Fulvia Confortola and Elisa Mastrogiacomo. Optimal control for stochastic heat equation with memory. *arXiv preprint arXiv:1112.0701*, 2011.
- [CMN08] Stefano Cardanobile, Delio Mugnolo, and Robin Nittka. Well-posedness and symmetries of strongly coupled network equations. *Journal of Physics A: Mathematical and Theoretical*, 41(5):055102, 2008.
- [Con05] Rama Cont. Modeling term structure dynamics: an infinite dimensional approach. *International Journal of Theoretical and Applied Finance*, 8(03):357–380, 2005.
- [CRa] Andrea Cosso and Francesco Russo. Functional Itô versus Banach space stochastic calculus and strict solutions of semilinear path-dependent equations. *arXiv:1505.02926*.
- [CRb] Andrea Cosso and Francesco Russo. A regularization approach to functional Itô calculus and strong-viscosity solutions to path-dependent PDEs. *arXiv:1401.5034*.
- [CR07] Serge Cohen and Jan Rosiński. Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes. *Bernoulli*, 13(1):195–210, 2007.
- [CR15] Andrea Cosso and Francesco Russo. Strong-viscosity solutions: semilinear parabolic pdes and path-dependent pdes. *arXiv preprint arXiv:1505.02927*, 2015.
- [Cré04] Stéphane Crépey. Delta-hedging vega risk? *Quantitative Finance*, 4(5):559–579, 2004.
- [Cré11] Stéphane Crépey. A bsde approach to counterparty risk under funding constraints. Available at grozny.maths.univ-evry.fr/pages/perso/crepey, 2011.

- [Cré15a] Stéphane Crépey. Bilateral counterparty risk under funding constraints—part i: Pricing. *Mathematical Finance*, 25(1):1–22, 2015.
- [Cré15b] Stéphane Crépey. Bilateral counterparty risk under funding constraints—part ii: Cva. *Mathematical Finance*, 25(1):23–50, 2015.
- [CRT13] Eduardo Casas, Christopher Ryll, and Fredi Tröltzsch. Sparse optimal control of the schlögl and fitzhugh–nagumo systems. *Computational Methods in Applied Mathematics*, 13(4):415–442, 2013.
- [CT07] René Carmona and Michael R Tehranchi. *Interest rate models: an infinite dimensional stochastic analysis perspective*. Springer Science & Business Media, 2007.
- [CY99] Mou-Hsiung Chang and Roger K. Youree. The European option with hereditary price structures: basic theory. *Appl. Math. Comput.*, 102(2-3):279–296, 1999.
- [CY07] Mou-Hsiung Chang and Roger K. Youree. Infinite-dimensional Black-Scholes equation with hereditary structure. *Appl. Math. Optim.*, 56(3):395–424, 2007.
- [Dav90] Edward Brian Davies. *Heat kernels and spectral theory*, volume 92. Cambridge University Press, 1990.
- [Dav92] Edward Brian Davies. Heat kernel bounds, conservation of probability and the feller property. *Journal d’Analyse Mathématique*, 58(1):99–119, 1992.
- [Del10] Łukasz Delong. Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management. *arXiv preprint arXiv:1005.4417*, 2010.
- [Del12] Łukasz Delong. Bsdés with time-delayed generators of a moving average type with applications to non-monotone preferences. *Stochastic Models*, 28(2):281–315, 2012.
- [Del13] Łukasz Delong. *Backward stochastic differential equations with jumps and their actuarial and financial applications*. Springer, 2013.
- [DF01] Brigo Damiano and Mercurio Fabio. Interest rate models: Theory and practice. *With Smile, Inflation and Credit*, page 1007, 2001.
- [DGR12] Cristina Di Girolami and Francesco Russo. Generalized covariation and extended Fukushima decomposition for Banach space-valued processes: applications to windows of Dirichlet processes. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 15(02), 2012.
- [DGR14] Cristina Di Girolami and Francesco Russo. Generalized covariation for Banach space valued processes, Itô formula and applications. *Osaka J. Math.*, 51(3):729–783, 2014.

- [DI10a] Łukasz Delong and Peter Imkeller. On malliavin’s differentiability of bsdes with time delayed generators driven by brownian motions and poisson random measures. *Stochastic Processes and their Applications*, 120(9):1748–1775, 2010.
- [DI⁺10b] Łukasz Delong, Peter Imkeller, et al. Backward stochastic differential equations with time delayed generators—results and counterexamples. *The Annals of Applied Probability*, 20(4):1512–1536, 2010.
- [Die97] Hans Werner Diehl. The theory of boundary critical phenomena. *International Journal of Modern Physics B*, 11(30):3503–3523, 1997.
- [DIJR96] Philip H Dybvig, Jonathan E Ingersoll Jr, and Stephen A Ross. Long forward and zero-coupon rates can never fall. *Journal of Business*, pages 1–25, 1996.
- [DMØR] Kristina R Dahl, Salah-Eldin A Mohammed, Bernt Øksendal, and Elin Røse. Optimal control of systems with noisy memory and BSDEs with Malliavin derivatives. *arXiv:1403.4034*.
- [DMØR14] Kristina R Dahl, Salah-Eldin A Mohammed, Bernt Øksendal, and Elin Røse. Optimal control of systems with noisy memory and bsdes with malliavin derivatives. *arXiv preprint arXiv:1403.4034*, 2014.
- [DNØP⁺09] Giulia Di Nunno, Bernt Karsten Øksendal, Frank Proske, Bernt Øksendal, and Bernt Øksendal. *Malliavin calculus for Lévy processes with applications to finance*. Springer, 2009.
- [Doe00] Wolfgang Doeblin. Sur l’équation de kolmogoroff, pli cachetée l’académie des sciences, édité par b. Bru et M. Yor, numéro spécial des Comptes rendus de l’académie des sciences, Paris, 331, 2000.
- [DP97] Richard WR Darling and Etienne Pardoux. Backwards sde with random terminal time and applications to semilinear elliptic pde. *The Annals of Probability*, 25(3):1135–1159, 1997.
- [DP06] Giuseppe Da Prato. *An introduction to infinite-dimensional analysis*. Springer Science & Business Media, 2006.
- [DP12] Giuseppe Da Prato. *Kolmogorov equations for stochastic PDEs*. Birkhäuser, 2012.
- [DP14] Giuseppe Da Prato. *Introduction to stochastic analysis and Malliavin calculus*, volume 13. Springer, 2014.
- [DPPB15a] Luca Di Persio, Gregorio Pellegrini, and Michele Bonollo. A computational spectral approach to interest rate models. *Available at SSRN 2635012*, 2015.
- [DPPB15b] Luca Di Persio, Gregorio Pellegrini, and Michele Bonollo. Polynomial chaos expansion approach to interest rate models. *Journal of Probability and Statistics*, 2015, 2015.
- [DPZ96] Giuseppe Da Prato and Jerzy Zabczyk. *Ergodicity for infinite dimensional systems*, volume 229. Cambridge University Press, 1996.

- [DPZ02] Giuseppe Da Prato and Jerzy Zabczyk. *Second order partial differential equations in Hilbert spaces*, volume 293. Cambridge University Press, 2002.
- [DPZ11] Luca Di Persio and Giacomo Ziglio. Gaussian estimates on networks with applications to optimal control. 2011.
- [DPZ14] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic equations in infinite dimensions*, volume 152. Cambridge university press, 2014.
- [DS88] Nelson Dunford and Jacob T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.
- [Dup09] Bruno Dupire. Functional Itô calculus. *Bloomberg Portfolio Research paper*, 2009.
- [DVGLW95] Odo Diekmann, Stephan A Van Gils, SM Verduyn Lunel, and Hans-Otto Walther. *Delay equations: functional-, complex-, and nonlinear analysis*. Springer-Verlag, 1995.
- [Ein05] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 322(8):549–560, 1905.
- [EJY00] Robert J Elliott, Monique Jeanblanc, and Marc Yor. On models of default risk. *Mathematical Finance*, 10(2):179–195, 2000.
- [Eke74] Ivar Ekeland. On the variational principle. *Journal of Mathematical Analysis and Applications*, 47(2):324–353, 1974.
- [EKKP⁺97] Nicole El Karoui, Christophe Kapoudjian, Etienne Pardoux, Shige Peng, and Marie Claire Quenez. Reflected solutions of backward sde’s, and related obstacle problems for pde’s. *the Annals of Probability*, pages 702–737, 1997.
- [EKPQ97] Nicole El Karoui, Shige Peng, and Marie Claire Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.
- [EKPQ01] Nicole El Karoui, S Peng, and MC Quenez. A dynamic maximum principle for the optimization of recursive utilities under constraints. *Annals of Applied Probability*, pages 664–693, 2001.
- [EKT⁺14] Ibrahim Ekren, Christian Keller, Nizar Touzi, Jianfeng Zhang, et al. On viscosity solutions of path dependent pdes. *The Annals of Probability*, 42(1):204–236, 2014.
- [EMSU12a] Oliver G Ernst, Antje Mugler, Hans-Jörg Starkloff, and Elisabeth Ullmann. On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(02):317–339, 2012.
- [EMSU12b] Oliver G Ernst, Antje Mugler, Hans-Jörg Starkloff, and Elisabeth Ullmann. On the convergence of generalized polynomial chaos expansions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 46(02):317–339, 2012.

- [EN00a] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194. Springer Science & Business Media, 2000.
- [EN00b] Klaus-Jochen Engel and Rainer Nagel. *One-parameter semigroups for linear evolution equations*, volume 194. Springer Science & Business Media, 2000.
- [Eng99] Klaus-J Engel. Spectral theory and generator property for one-sided coupled operator matrices. In *Semigroup Forum*, volume 58, pages 267–295. Springer, 1999.
- [ETZ12a] Ibrahim Ekren, Nizar Touzi, and Jianfeng Zhang. Viscosity solutions of fully nonlinear parabolic path dependent pdes: Part i. *arXiv preprint arXiv:1210.0006v3*, 2012.
- [ETZ12b] Ibrahim Ekren, Nizar Touzi, and Jianfeng Zhang. Viscosity solutions of fully nonlinear parabolic path dependent pdes: Part ii. *arXiv preprint arXiv:1210.0007*, 2012.
- [FGG⁺15] Peter K Friz, Jim Gatheral, Archil Gulisashvili, Antoine Jacquier, and Josef Teichmann. *Large Deviations and Asymptotic Methods in Finance*, volume 110. Springer, 2015.
- [Fil01] Damir Filipovic. *Consistency problems for Heath-Jarrow-Morton interest rate models*. Springer Science & Business Media, 2001.
- [Fit61] Richard FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal*, 1(6):445, 1961.
- [FMT10] Marco Fuhrman, Federica Masiero, and Gianmario Tessitore. Stochastic equations with delay: optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations. *SIAM Journal on Control and Optimization*, 48(7):4624–4651, 2010.
- [FPS00] Jean-Pierre Fouque, George Papanicolaou, and K Ronnie Sircar. *Derivatives in financial markets with stochastic volatility*. Cambridge University Press, 2000.
- [FRZ15] Franco Flandoli, Francesco Russo, and Giovanni Zanco. Infinite-dimensional calculus under weak spatial regularity of the processes. *arXiv preprint arXiv:1511.05744*, 2015.
- [FS06] Wendell H Fleming and Halil Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.
- [FT02] Marco Fuhrman and Gianmario Tessitore. Nonlinear kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control. *Annals of probability*, pages 1397–1465, 2002.
- [FT05] Marco Fuhrman and Gianmario Tessitore. Generalized directional gradients, backward stochastic differential equations and mild solutions of semilinear parabolic equations. *Applied Mathematics and Optimization*, 51(3):279–332, 2005.

- [FT12] Masaaki Fujii and Akihiko Takahashi. Perturbative expansion of fbsde in an incomplete market with stochastic volatility. *The Quarterly Journal of Finance*, 2(03):1250015, 2012.
- [FZ] Franco Flandoli and Giovanni Zanco. An infinite-dimensional approach to path-dependent Kolmogorov's equations. *arXiv:1312.6165*.
- [G⁺06] Gisèle Ruiz Goldstein et al. Derivation and physical interpretation of general boundary conditions. *Advances in Differential Equations*, 11(4):457–480, 2006.
- [Gar85] Crispin W Gardiner. Handbook of stochastic methods for physics, chemistry and the natural sciences, vol. 13 of. *Springer Series in Synergetics*, 1985.
- [Gar88] Thomas C Gard. *Introduction to stochastic differential equations*. M. Dekker, 1988.
- [GHL⁺12] Jim Gatheral, Elton P Hsu, Peter Laurence, Cheng Ouyang, and Tai-Ho Wang. Asymptotics of implied volatility in local volatility models. *Mathematical Finance*, 22(4):591–620, 2012.
- [Gia06] Emanuela Rosazza Gianin. Risk measures via g-expectations. *Insurance: Mathematics and Economics*, 39(1):19–34, 2006.
- [GM10] Mariano Giaquinta and Giuseppe Modica. *Mathematical analysis: an introduction to functions of several variables*. Springer Science & Business Media, 2010.
- [GR15] Zorana Grbac and Wolfgang J Runggaldier. Interest rate modeling: Post-crisis challenges and approaches, 2015.
- [GS79] Iosif Il'Ich Gihman and Anatoliĭ Vladimirovich Skorohod. *Stochastic differential equations*. Springer, 1979.
- [Gul12] Archil Gulisashvili. *Analytically tractable stochastic stock price models*. Springer Science & Business Media, 2012.
- [GZ04] Crispin Gardiner and Peter Zoller. *Quantum noise: a handbook of Markovian and non-Markovian quantum stochastic methods with applications to quantum optics*, volume 56. Springer Science & Business Media, 2004.
- [H§] Friedrich Hubalek and Sgarra §. Esscher transforms and the minimal entropy martingale measure for exponential lévy models.
- [Hen14] Marc Henrard. *Interest Rate Modelling in the Multi-Curve Framework: Foundations, Evolution and Implementation*. Palgrave Macmillan, 2014.
- [HH52] Alan L Hodgkin and Andrew F Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952.

- [HJM92] David Heath, Robert Jarrow, and Andrew Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*, pages 77–105, 1992.
- [HKRS14] Friedrich Hubalek, Martin Keller-Ressel, and Carlo Sgarra. Geometric asian option pricing in general affine stochastic volatility models with jumps. *arXiv preprint arXiv:1407.2514*, 2014.
- [HN13] Ali Hirsu and Salih N Neftci. *An introduction to the mathematics of financial derivatives*. Academic Press, 2013.
- [HS09] Friedrich Hubalek and Carlo Sgarra. On the esscher transforms and other equivalent martingale measures for barndorff-nielsen and shephard stochastic volatility models with jumps. *Stochastic Processes and their Applications*, 119(7):2137–2157, 2009.
- [HS11] Friedrich Hubalek and Carlo Sgarra. On the explicit evaluation of the geometric asian options in stochastic volatility models with jumps. *Journal of Computational and Applied Mathematics*, 235(11):3355–3365, 2011.
- [HW06] Patrick S Hagan and Graeme West. Interpolation methods for curve construction. *Applied Mathematical Finance*, 13(2):89–129, 2006.
- [HWBY07] Agnes Handwerk, Harrie Willems, Bernard Bru, and Marc Yor. *Wolfgang Doeblin: a mathematician rediscovered*. Springer, 2007.
- [IKS03] AF Ivanov, YI Kazmerchuk, and AV Swishchuk. Theory, stochastic stability and applications of stochastic delay differential equations: a survey of results. *Differential Equations Dynam. Systems*, 11(1-2):55–115, 2003.
- [IP10] Gennaro Infante and Paolamaria Pietramala. A third order boundary value problem subject to nonlinear boundary conditions. *Mathematica Bohemica*, 135(2):113–121, 2010.
- [IPW⁺09] Peter Imkeller, Ilya Pavlyukevich, Torsten Wetzl, et al. First exit times for lévy-driven diffusions with exponentially light jumps. *The Annals of Probability*, 37(2):530–564, 2009.
- [Itō51] Kiyosi Itō. *On stochastic differential equations*, volume 4. American Mathematical Soc., 1951.
- [IW06a] Gennaro Infante and Jeff R.L. Webb. Loss of positivity in a nonlinear scalar heat equation. *Nonlinear Differential Equations and Applications NoDEA*, 13(2):249–261, 2006.
- [IW06b] Gennaro Infante and Jeff R.L. Webb. Nonlinear non-local boundary-value problems and perturbed hammerstein integral equations. *Proceedings of the Edinburgh Mathematical Society (Series 2)*, 49(03):637–656, 2006.
- [JK15] Svante Janson and Sten Kaijser. Higher moments of Banach space valued random variables. *Mem. Amer. Math. Soc.*, 238(1127):vii+110, 2015.

- [JR02] Monique Jeanblanc and Marek Rutkowski. Default risk and hazard process. In *Mathematical Finance—Bachelier Congress 2000*, pages 281–312. Springer, 2002.
- [JS13] Jean Jacod and Albert Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.
- [K⁺05] Fima C Klebaner et al. *Introduction to stochastic calculus with applications*, volume 57. World Scientific, 2005.
- [Kai04] Cornelia Kaiser. Integrated semigroups and linear partial differential equations with delay. *Journal of mathematical analysis and applications*, 292(2):328–339, 2004.
- [Kal06] Olav Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 2006.
- [Kel14] Christian Keller. Viscosity solutions of path-dependent integro-differential equations. *arXiv preprint arXiv:1412.8495*, 2014.
- [Khe12] Asma Khedher. Computation of the delta in multidimensional jump-diffusion setting with applications to stochastic volatility models. *Stoch. Anal. Appl.*, 30(3):403–425, 2012.
- [KK99] Yong-Jin Kim and Naoto Kunitomo. Pricing options under stochastic interest rates: a new approach. *Asia-Pacific Financial Markets*, 6(1):49–70, 1999.
- [Kle09] Hagen Kleinert. *Path integrals in quantum mechanics, statistics, polymer physics, and financial markets*. World Scientific, 2009.
- [Kol56] Andrey N Kolmogorov. *Foundations of the Theory of Probability*. Chelsea, New York, 1933. 1956.
- [KP07] Uwe Kuchler and Eckhard Platen. *Time delay and noise explaining cyclical fluctuations in prices of commodities*. University of Technology, Sydney, 2007.
- [KS12] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.
- [KSW04] Yuriy I Kazmerchuk, Anatoly V Swishchuk, and JH Wu. Black-scholes formula revisited: securities markets with delayed response. *preprint*, 2004.
- [KSW05a] Yuriy Kazmerchuk, Anatoliy Swishchuk, and Jianhong Wu. A continuous-time garch model for stochastic volatility with delay. *Canadian Applied Mathematics Quarterly*, 13(2):123–149, 2005.
- [KSW05b] Yuriy Kazmerchuk, Anatoliy Swishchuk, and Jianhong Wu. A continuous-time Garch model for stochastic volatility with delay. *Can. Appl. Math. Q.*, 13(2):123–149, 2005.
- [KSW07a] Yuriy Kazmerchuk, Anatoliy Swishchuk, and Jianhong Wu. The pricing of options for securities markets with delayed response. *Mathematics and Computers in Simulation*, 75(3):69–79, 2007.

- [KSW07b] Yuriy Kazmerchuk, Anatoliy Swishchuk, and Jianhong Wu. The pricing of options for securities markets with delayed response. *Math. Comput. Simulation*, 75(3-4):69–79, 2007.
- [KT⁺03] Naoto Kunitomo, Akihiko Takahashi, et al. On validity of the asymptotic expansion approach in contingent claim analysis. *The Annals of Applied Probability*, 13(3):914–952, 2003.
- [KT11] Jonas Kiessling and Raúl Tempone. Diffusion approximation of Lévy processes with a view towards finance. *Monte Carlo Methods Appl.*, 17(1):11–45, 2011.
- [Kua93] Yang Kuang. *Delay differential equations with applications in population dynamics*, volume 191 of *Mathematics in Science and Engineering*. Academic Press, Inc., Boston, MA, 1993.
- [Kun04] Hiroshi Kunita. Stochastic differential equations based on Lévy processes and stochastic flows of diffeomorphisms. In *Real and stochastic analysis*, Trends Math., pages 305–373. Birkhäuser Boston, Boston, MA, 2004.
- [KW13] Karl Kunisch and Marcus Wagner. Optimal control of the bidomain system (iii): Existence of minimizers and first-order optimality conditions. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(04):1077–1106, 2013.
- [KY00] Shigeo Kusuoka and Nakahiro Yoshida. Malliavin calculus, geometric mixing, and expansion of diffusion functionals. *Probability Theory and Related Fields*, 116(4):457–484, 2000.
- [Lév25] Paul Lévy. *Calcul des probabilités*, volume 9. Gauthier-Villars Paris, 1925.
- [LMK10] Olivier P Le Maître and Omar M Knio. *Introduction: Uncertainty Quantification and Propagation*. Springer, 2010.
- [LMS11] Vincent JG Leijdekker, MRH Mandjes, and PJC Spreij. Sample-path large deviations in credit risk. *Journal of applied mathematics*, 2011, 2011.
- [Lor12] Matthew Lorig. Local lévy models and their volatility smile. *ArXiv preprint*, 2012.
- [LSUV02] Jorge A León, Josep L Solé, Frederic Utzet, and Josep Vives. On lévy processes, malliavin calculus and market models with jumps. *Finance and Stochastics*, 6(2):197–225, 2002.
- [Lud73] Arnold Ludwing. *Stochastic differential equations: Theory and applications*, 1973.
- [Lum80] Günter Lumer. Connecting of local operators and evolution equations on networks. In *Potential Theory Copenhagen 1979*, pages 219–234. Springer, 1980.
- [Lüt04] Eva Lütkebohmert. An asymptotic expansion for a black–scholes type model. *Bulletin des sciences mathématiques*, 128(8):661–685, 2004.

- [LXYW14] Jian Liu, Jihong Xiao, Lizhao Yan, and Fenghua Wen. Valuing catastrophe bonds involving credit risks. *Mathematical Problems in Engineering*, 2014, 2014.
- [Mar52] Harry M. Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.
- [Mar68] Harry M. Markowitz. *Portfolio selection: efficient diversification of investments*, volume 16. Yale university press, 1968.
- [McK69] Henry P McKean. *Stochastic integrals*, volume 353. American Mathematical Soc., 1969.
- [Mer74] Robert C. Merton. On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2):449–470, 1974.
- [Mer76] Robert C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.
- [MH14] Benoit Mandelbrot and Richard L Hudson. *The Misbehavior of Markets: A fractal view of financial turbulence*. Basic books, 2014.
- [Moh84] Salah-Eldin A. Mohammed. *Stochastic Functional Differential Equations*, volume 99 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.
- [Moh98] Salah-Eldin A. Mohammed. Stochastic differential systems with memory: theory, examples and applications. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, volume 42 of *Progr. Probab.*, pages 1–77. Birkhäuser Boston, Boston, MA, 1998.
- [MP14] Nicola Moreni and Andrea Pallavicini. Parsimonious hjm modelling for multiple yield curve dynamics. *Quantitative Finance*, 14(2):199–210, 2014.
- [MR07] Delio Mugnolo and Silvia Romanelli. Dynamic and generalized wentzell node conditions for network equations. *Mathematical methods in the applied sciences*, 30(6):681–706, 2007.
- [MR10] Lucian Maticiuc and Aurel Răşcanu. A stochastic approach to a multivalued dirichlet–neumann problem. *Stochastic Processes and their Applications*, 120(6):777–800, 2010.
- [MR14] Laura Morino and Wolfgang J Runggaldier. On multicurve models for the term structure. In *Nonlinear Economic Dynamics and Financial Modelling*, pages 275–290. Springer, 2014.
- [MR15a] Vidyadhar Mandrekar and Barbara Rüdiger. Stochastic integration in banach spaces. *VII Springer, Berlin*, 2015.
- [MR⁺15b] Lucian Maticiuc, Aurel Răşcanu, et al. Backward stochastic variational inequalities on random interval. *Bernoulli*, 21(2):1166–1199, 2015.

- [MS90] Dilip B Madan and Eugene Seneta. The variance gamma (vg) model for share market returns. *Journal of business*, pages 511–524, 1990.
- [MTU04] Ryosuke Matsuoka, Akihiko Takahashi, and Yoshihiko Uchida. A new computational scheme for computing greeks by the asymptotic expansion approach. *Asia-Pacific Financial Markets*, 11(4):393–430, 2004.
- [Mug07] Delio Mugnolo. A variational approach to strongly damped wave equations. In *Functional analysis and evolution equations*, pages 503–514. Springer, 2007.
- [Mug10] Delio Mugnolo. Gaussian estimates for a heat equation on a network. *arXiv preprint arXiv:1005.2070*, 2010.
- [Mug14] Delio Mugnolo. *Semigroup methods for evolution equations on networks*. Springer, 2014.
- [Mus93] Marek Musiela. Stochastic pdes and term structure models. *Journées Internationales de Finance, IGR-AFFI, La Baule*, 1993.
- [MZ10] Marek Musiela and Thaleia Zariphopoulou. Stochastic partial differential equations and portfolio choice. In *Contemporary Quantitative Finance*, pages 195–216. Springer, 2010.
- [NAY62] Jinichi Nagumo, Suguru Arimoto, and Shuji Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.
- [NS87] Charles R Nelson and Andrew F Siegel. Parsimonious modeling of yield curves. *Journal of business*, pages 473–489, 1987.
- [NV03] Elisa Nicolato and Emmanouil Venardos. Option pricing in stochastic volatility models of the ornstein-uhlenbeck type. *Mathematical finance*, 13(4):445–466, 2003.
- [oBSfIS00] Basle Committee on Banking Supervision and Bank for International Settlements. *Principles for the management of credit risk*. Bank for International Settlements, 2000.
- [ØS05] Bernt Karsten Øksendal and Agnes Sulem. *Applied stochastic control of jump diffusions*, volume 498. Springer, 2005.
- [Ouh09] El-Maati Ouhabaz. *Analysis of Heat Equations on Domains.(LMS-31)*. Princeton University Press, 2009.
- [PA10] Sivakumar Pitchaiah and Antonios Armaou. Output feedback control of the fitzhugh-nagumo equation using adaptive model reduction. In *Decision and Control (CDC), 2010 49th IEEE Conference on*, pages 864–869. IEEE, 2010.
- [Par67] Kalyanapuram R. Parthasarathy. *Probability measures on metric spaces*. Probability and Mathematical Statistics, No. 3. Academic Press, Inc., New York-London, 1967.

- [PB13] Andrea Pallavicini and Damiano Brigo. Interest-rate modelling in collateralized markets: multiple curves, credit-liquidity effects, ceps. *Credit-Liquidity Effects, CCPs (April 4, 2013)*, 2013.
- [Pen91] Shige Peng. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics Stochastics Rep*, 37(1-2):61–74, 1991.
- [Pen97] Shige Peng. Backward sde and related g-expectation. *Pitman research notes in mathematics series*, pages 141–160, 1997.
- [Pen04] Shige Peng. Nonlinear expectations, nonlinear evaluations and risk measures. In *Stochastic methods in finance*, pages 165–253. Springer, 2004.
- [Pen11a] Shige Peng. Backward stochastic differential equation, nonlinear expectation and their applications. In *Proceedings of the International Congress of Mathematicians*, volume 1, pages 393–432, 2011.
- [Pen11b] Shige Peng. Note on viscosity solution of path-dependent pde and g-martingales. *arXiv preprint arXiv:1106.1144*, 2011.
- [Pet06] Evangelia Petrou. *Malliavin calculus in Lévy spaces and applications in finance*. University of London, Imperial College London, 2006.
- [Phi90] Protter Philip. *Stochastic integration and differential equations*, 1990.
- [PP90] Etienne Pardoux and Shige Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.
- [PP92] Etienne Pardoux and Shige Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, pages 200–217. Springer, 1992.
- [PR⁺05] Szymon Peszat, Francesco Russo, et al. Large-noise asymptotics for one-dimensional diffusions. *Bernoulli*, 11(2):247–262, 2005.
- [Pro05] Philip E. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005.
- [PT10] Andrea Pallavicini and Marco Tarenghi. Interest-rate modeling with multiple yield curves. *Available at SSRN 1629688*, 2010.
- [PT11] Giovanni Peccati and Murad Taqqu. *Wiener Chaos: Moments, Cumulants and Diagrams: A survey with Computer Implementation*, volume 1. Springer Science & Business Media, 2011.
- [PW11] Shige Peng and Falei Wang. Bsde, path-dependent pde and nonlinear feynman-kac formula. *arXiv preprint arXiv:1108.4317*, 2011.
- [PZ98] Étienne Pardoux and Shuguang Zhang. Generalized bsdes and nonlinear neumann boundary value problems. *Probability Theory and Related Fields*, 110(4):535–558, 1998.

- [PZ07] Szymon Peszat and Jerzy Zabczyk. *Stochastic partial differential equations with Lévy noise: An evolution equation approach*, volume 113. Cambridge University Press, 2007.
- [R⁺91] Mark Rubinstein et al. Exotic options. Technical report, University of California at Berkeley, 1991.
- [Rei02] Markus Reiß. *Nonparametric estimation for stochastic delay differential equations*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, 2002.
- [Rud87] Walter Rudin. *Real and complex analysis*. Tata McGraw-Hill Education, 1987.
- [RV93] Francesco Russo and Pierre Vallois. Forward, backward and symmetric stochastic integration. *Probability theory and related fields*, 97(3):403–421, 1993.
- [RV95] Francesco Russo and Pierre Vallois. The generalized covariation process and Itô formula. *Stochastic Process. Appl.*, 59(1):81–104, 1995.
- [RV96a] Francesco Russo and Pierre Vallois. Itô formula for C^1 -functions of semimartingales. *Probab. Theory Related Fields*, 104(1):27–41, 1996.
- [RV96b] Francesco Russo and Pierre Vallois. Ito formula for C^1 -functions of semimartingales. *Probability theory and related fields*, 104(1):27–41, 1996.
- [RvG06] Riedle Markus Reiß, Markus and Onno van Gaans. Delay differential equations driven by Lévy processes: stationarity and Feller properties. *Stochastic Process. Appl.*, 116(10):1409–1432, 2006.
- [Sam73] Paul A. Samuelson. Mathematics of speculative price. *Siam Review*, 15(1):1–42, 1973.
- [Sat99] Ken-iti Sato. *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.
- [Sch04] Steven E Schreve. Stochastic calculus for finance ii, 2004.
- [SN88] Andrew F Siegel and Charles R Nelson. Long-term behavior of yield curves. *Journal of Financial and Quantitative Analysis*, 23(01):105–110, 1988.
- [ST08] Walter Schachermayer and Josef Teichmann. How close are the option pricing formulas of bachelier and black–merton–sholes? *Mathematical Finance*, 18(1):155–170, 2008.
- [SUV07] Josep Lluís Solé, Frederic Utzet, and Josep Vives. Canonical lévy process and malliavin calculus. *Stochastic processes and their Applications*, 117(2):165–187, 2007.
- [Swi05] Anatoliy Swishchuk. Modeling and pricing of variance swaps for stochastic volatilities with delay. *Wilmott Magazine*, 19(September):63–73, 2005.

- [Swi07] Anatoliy Swishchuk. Change of time method in mathematical finance. *Canad. Appl. Math. Quart.*, 15(3):299–336, 2007.
- [Swi13] Anatoliy Swishchuk. *Modeling and Pricing of Swaps for Financial and Energy Markets with Stochastic Volatilities*. World Scientific. World Scientific Publishing Co. Pte. Ltd. , Singapore., 2013.
- [SX11] Anatoliy Swishchuk and Li Xu. Pricing variance swaps for stochastic volatilities with delay and jumps. *International Journal of Stochastic Analysis*, 2011, 2011.
- [Tap13] Stefan Tappe. Compact embeddings for spaces of forward rate curves. In *Abstract and Applied Analysis*, volume 2013. Hindawi Publishing Corporation, 2013.
- [Teh05] Michael Tehranchi. A note on invariant measures for hjm models. *Finance and Stochastics*, 9(3):389–398, 2005.
- [Tes96] Gianmario Tessitore. Existence, uniqueness and space regularity of the adapted solutions of a backward spde. *Stochastic Analysis and Applications*, 14(4):461–486, 1996.
- [TMH80] Yutaka Takahashi, Hideo Miyahara, and Toshiharu Hasegawa. An approximation method for open restricted queueing networks. *Operations Research*, 28(3-part-i):594–602, 1980.
- [Tuc92] Henry C Tuckwell. Random perturbations of the reduced fitzhugh-nagumo equation. *Physica Scripta*, 46(6):481, 1992.
- [Tum06] Roderich Tumulka. The analogue of bohm–bell processes on a graph. *Physics Letters A*, 348(3):126–134, 2006.
- [UY04] Masayuki Uchida and Nakahiro Yoshida. Asymptotic expansion for small diffusions applied to option pricing. *Statistical inference for stochastic processes*, 7(3):189–223, 2004.
- [Var95] Tiziano Vargiolu. *Ph.D. Thesis*. PhD thesis, University of Padova, 1995. <http://www.math.unipd.it/~vargiolu/tesi/>.
- [Var99] Tiziano Vargiolu. Invariant measures for the musielà equation with deterministic diffusion term. *Finance and Stochastics*, 3(4):483–492, 1999.
- [Wag97] Daniel F Waggoner. Spline methods for extracting interest rate curves from coupon bond prices. *Federal Reserve Bank of Atlanta Working Paper*, pages 97–10, 1997.
- [WDH93] Paul Wilmott, Jeff Dewynne, and Sam Howison. *Option pricing: mathematical models and computation*. Oxford financial press, 1993.
- [Web05] Jeff R.L. Webb. Optimal constants in a nonlocal boundary value problem. *Nonlinear Analysis: Theory, Methods & Applications*, 63(5):672–685, 2005.

- [Wie38] Norbert Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.
- [WK05] Xiaoliang Wan and George Em Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *Journal of Computational Physics*, 209(2):617–642, 2005.
- [WK06] Xiaoliang Wan and George Em Karniadakis. Beyond wiener–askey expansions: handling arbitrary pdfs. *Journal of Scientific Computing*, 27(1-3):455–464, 2006.
- [WYZW14] Aiyin Wang, Ls Yong, Weili Zeng, and Yang Wang. The optimal analysis of default probability for a credit risk model. In *Abstract and Applied Analysis*, volume 2014. Hindawi Publishing Corporation, 2014.
- [YM05] Feng Yan and Salah Mohammed. A stochastic calculus for systems with memory. *Stoch. Anal. Appl.*, 23(3):613–657, 2005.
- [Yos03] Nakahiro Yoshida. Conditional expansions and their applications. *Stochastic processes and their applications*, 107(1):53–81, 2003.
- [Ză112] Adrian Zălinescu. Second order hamilton–jacobi–bellman equations with an unbounded operator. *Nonlinear Analysis: Theory, Methods & Applications*, 75(13):4784–4797, 2012.
- [Zei95] Eberhard Zeidler. *Applied Functional Analysis. Main Principles and Their Applications*. Springer-Verlag 1995. New York, third edition, 1995.
- [Zig08] Giacomo Ziglio. Spdes with random dynamic boundary conditions. 2008.