UNIVERSITÀ DEGLI STUDI DELL'AQUILA

Department of Engineering and Information Science and Mathematics

Ph.D in Mathematics and Models
XXXV cycle

Thesis title

Evaluation of Non-linear Value Adjustments under multiple credit risks

SSD SECS-S/06

Ph.D Student
Ivan Gallo

Tutor
Prof. Raffaele D'Ambrosio

Supervisor
Prof. Fabio Antonelli

Coordinatore del Corso
Davide Gabrielli

Anno Accademico 2021-2022

Abstract

In this work, we consider the problem of correctly pricing financial derivatives that might be subject to multiple credit risks, such as default or liquidity risk, that gained importance after the financial crisis of $2008 - 09$.

An important achievement in the mathematical modelling of the problem was the representation of the value of such derivatives as solutions of appropriate Backward Stochastic Differential Equations (BSDE), which might be solvable in the samples cases.

When various risks are taken into account simultaneously, and correlation is admitted between the processes underlying the price formation, the picture becomes much more complex, and although the BSDE representation still applies, explicit solvability becomes impossible.

Monte Carlo simulations, usually requiring long computational times, are often the only way to get an approximation of the solution, so it might be important to develop alternative approximation techniques that require shorter computational times yet preserving accuracy. Once the BSDE's representation is developed, in a Markovian setting, the derivative's can be rewritten as a deterministic function of the state variables, which verifies a non-linear PDE.

Thus, we decided to employ a PDE discretization approach to approximate the PDE solution. By employing an adaptation of the simple method of lines, we were able to construct an approximation method that turned out to be accurate and efficient, thus producing a valid alternative to Monte Carlo simulations.

Keywords: XVA, Value Adjustments, Backward Stochastic Differential Equation, Non-linear Valuation, Credit Risk, Defaultable Claims

Contents

List of Figures

List of Tables

Introduction

The PhD thesis is mostly about how to estimate the prices of financial derivatives that are subject to different credit risks, like default or liquidity risks.

Before the financial crisis of $2008-09$, the value of a financial contract, such as an option, was determined by taking the conditional expectation of the discounted expected price under a risk-neutral measure without taking into account any credit risk.

In periods of financial trouble or crisis, some traditional financial models could no longer account for all the risks. As a matter of fact, in 2004, the Basel Committee signed the Basel II agreement regarding the capital requirements banks must meet to curb financial risks. In particular, Basel II set up the accounting standards regarding Counterparty Credit Risk (CCR), which is the risk that a counterparty might default before honouring its engagements, and it covers loans, repurchase agreements (Repo) transactions, and most importantly, over-the-counter (OTC) derivatives. In the last decade, the interest in CCR increased remarkably, and a theory of Value Adjustments was developed. The first to be introduced was the Credit Value Adjustment (CVA), which is the difference between the default-free value of a portfolio and the valuation taking into account the possibility of counterparty default, while the investor is always considered default-free. After the financial crisis in 2009, the Basel Committee issued a new version of the act called Basel III, pushing financial institutions to incorporate default risks of either party when evaluating products with cashflows in both directions.

A new measure called *Debt Value Adjustment* (DVA) was introduced as an accruement of the claim's value due to the investor's default risk.

To mitigate credit risk, parties often employ collateralization to balance their exposure to the reciprocal default event. A collateral account, underwritten by both parties, is established for this purpose, in which they deposit or withdraw assets to cover the risk of default. The collateral in this account can sometimes be rehypothecated for self-financing, which means that those who withdraw assets may use them to finance other activities.

As investments/collateralizations are often funded also by external sources, further risks are involved and further adjustments have to be made. The introduction of the Funding Value Adjustment (FVA) and Liquidity Value Adjustment (LVA), makes the pricing problem recursive and non-linear, as those quantities are closely linked to the adjusted price itself.

The first important result was to achieve a representation of this derivative's adjusted values as solutions of appropriate Backward Stochastic Differential Equations (BSDE): the literature on this matter is nowadays very extensive and it includes the works [15, 16, 21].

In some simpler cases, these BSDE's can be explicitly solved since they turn out to be linear. When taking into account multiple risks simultaneously, the problem becomes significantly more complex. More precisely, we consider the case of a European claim subject to funding, liquidity, and default risks of either party.

The default event is when the (random) parties' default times happen before the end of the contract. In general, these random times are not necessarily stopping times in terms of the filtration generated by the processes that represent the prices of the traded asset. To price the defaultable contract, we first need to extend this filtration generated a longer filtration that makes these times stopping times, by progressive enlargement. By exploiting the so-called reduced form approach for the default times, we assume that they are the first jump times of two Cox processes with stochastic positive intensities adapted to the market filtration. To make sure that the intensities are positive, they are represented by CIR processes, and the representation of the adjusted as a solution of BSDE works.

This equation depends on the so-called “close-out value", which is a portion of a contractually agreed price to be paid as partial compensation when the default of one of the parties occurs. There are fundamentally two possibilities: either the close-out value is taken as a portion of the default-free price or of the defaultable contract price. The first choice usually determines a solvable linear BSDE, the second returns a non-linear BSDE, not explicitly solvable. Our main goal is to treat the second case, while the first is the most commonly treated in the literature's. As in [2, 16], intensities are usually considered deterministic, while considering them stochastic allows for correlations among the fundamentals defining the market, which is a desirable feature to include in the model. Unfortunately, in presence of correlations, model affinity often fails and transform techniques cannot be applied.

This leaves only Monte Carlo simulations in high dimensions (see, for instance, [41]) as the only technique to obtain a numerical approximation of the solution, which, typically have very long computational times. Hence, finding alternative numerical methods with lower computational costs becomes a key issue.

When the processes dynamics are Markovian, it is possible to express the derivative's adjusted value as a deterministic function of the state variables, which satisfies a semi-linear PDE with final condition given by the product's payoff. Exploiting this representation, it is natural to investigate discretization methods of this associated PDE as an alternative numerical approach to the problem. In the thesis work, we included correlations in the model, generating a non-linear PDE in $[0, T] \times \mathbb{R}^3$ that we treated by adapting the so-called "method of lines", which approximated the spatial derivatives with finite differences, and it generates a system of ODEs at each point of the discretization grid, that can be solved by a suitable time integration method.

The spatial domain $\mathbb{R}^3_+$ is unbounded, so we needed to restrict it to an appropriate bounded rectangle. This truncation required defining appropriate boundary conditions, that were imposed by identifying, when possible, the asymptotic behaviour of the solution. We obtained by appropriately modifying the Black & Scholes formula with adjusted rates including the default intensities.

We used the Euler explicit scheme for the time integration method. To the best of our knowledge, in the literature, we could not find numerical methods covering this general case, so we employed Monte Carlo simulations to provide a benchmark to compare our results with.

The approach proved particularly advantageous in terms of efficiency (it significantly decreased the computation times), reaching an accuracy equivalent to that of the Monte Carlo method.

By taking into account that the explicit Euler method can produce serious numerical instabilities. We also developed a semi-implicit or implicit scheme, even though this choice implied longer quite computational times.

By comparing the three schemes, it turned out the explicit scheme was quite stable and it achieved a competitive accuracy in shorter computational times. Indeed, we were able to keep the so-called Courant-Friedricks-Levy (or CFL) number below the critical value to guarantee the explicit scheme stability. We were able do to so without compromising efficiency, by keeping the number of temporal nodes suitable bigger than the number of spatial nodes. Lastly, we ran a short sensitivity analysis to estimate the impact of stochastic intensities on the prices.

The work is presented in the following manner: Chapter 1 introduces the Adjusted Value Theory, showing how to incorporate various risks such as CVA, DVA, FVA, collateralization, and default into the pricing evaluation. The chapter is concluded by presenting the theoretical equation that characterizes the price. Chapter 2 discusses the theory of BSDEs and presents some results regarding the well-posedness, existence and uniqueness of solutions. Finally, under Markovian assumptions, the Feynman-Kac formula is used to derive an associated PDE from the BSDE.

In chapter 3 we specialize the theory of to our pricing equation (Chapter 1). Under those choices, the backward equation becomes non-linear, to which a 4 dimensional non-linear PDE is associated. Using this representation, it makes sense to look into discretization methods for the PDE as a different way to solve the problem numerically. In Chapter 4, we briefly describe the method of lines, and apply it to the PDE in the specific case of a European call. Lastly, we discuss some numerical results, and a sensitivity analysis is performed.

Chapter 1

Value Adjustment Theory

Since the financial crisis of $2008-09$, when many large financial institutions as Lehman Brothers went bankrupt, in the United States, United Kingdom, and Europe, the so-called counterparty risk has become extremely important. Financial contracts usually involve two parties, the investor (or buyer) and the counterparty (or seller), and they require the solvency of a future payment on one or both sides. When a counterparty becomes insolvent, i.e., it does not fulfill its contractual obligations, causing the investor's loss of money, we speak of "counterparty risk". The financial crisis showed that it is appropriate to take this aspect into account for all the financial contracts.

In the last 10 years, the global crisis made banks follow stricter rules and procedures to avoid being exposed to such risks. Indeed, Governors of the 10 most industrialised nations were already aware of those risks when they signed the so-called Basilea Agreements in 1998, where it was specified the amount of capital banks might borrow to undertake financial operations.

To keep up with the growing complexity of the financial sector, agreements were updated, and four official Basel Committee versions were released. The latest, in 2017, influenced heavily the scientific community, sparring the development of a "Value Adjustment Theory".

This theory incorporates default risks from one or both parties, but it takes also into account the lack of liquidity, and the financing demands. Historically, the first adjustment to be introduced was the so-called Credit Value Adjustment (CVA). CVA is very intuitive, and it simply adjusts the price to account for the seller's (counterparty) chance of default by discounting the default-free price. Over the years, the role of the CVA increased considerably, and its correct formulation became crucial in derivatives trading in the OTC markets, so stimulating much research in the field: we refer the reader to [20, 27, 35, 36], and [11] for a general discussion on the subject.

When both parties are defaultable, this adjustment is no longer sufficient. In fact, also the investor that enters a future might default, and this bilateral exposure to risk must be taken into account when evaluating the contract. This new adjustment is commonly referred to as Debt Value Adjustment (DVA). An example of a contract with bilateral exposure is the forward rate agreement (FRAs).

Over the years, the market and regulators implemented ways to reduce the risks. The most widespread techniques to curb the exposure to default are the collateralization and close-out netting rules.

A collateral account is required of both parties of the contract to guarantee a given level of solvency in case of default. In this way, when the default occurs, the party that is positively exposed may obtain at least a partial recovery of the loss.

It is to be kept in mind that the assets recevived as collateral during the contract lifetime are often reinvested in a rehypothecation market, making the liquidity risks of both parties more complex.

Also, close-out netting rules are often contractually agreed to define a "close-out value", paid as a partial recovery of the losses at the time of the party's default. The impact of all these requirements, on the derivative's price was analyzed in [13, 14, 18].

Finally, when a derivatives desk [1] executes an order by a client, it backs the trade by hedging it with other dealers. These operations modify the available liquidity and the relative adjustment is called Funding Value Adjustment (FVA) (see [16, 25, 26] for a detailed discussion).

As we mentioned before, in [13, 17, 27] an important achievement was the representation of the derivative's adjusted value as solution of an appropriate Backward Stochastic Differential Equation (BSDE), considering all the aspects described above. In the next section, we are going to describe the step-by-step construction of the representation BSDE obtained, including progressively the above-mentioned features.

1.1 Unilateral CVA

As mentioned before, the first value adjustment to be modelled was CVA, which is based solely on the investor's exposure at time of the counterparty's default, since the investor is considered to be default-free. For this reason, we refer to it as "*Unilateral Credit Value Adjustment*".

We refer the reader to [19], for a coherent treatment of the approach to valuation of derivatives with CVA, while some particular applications by the same authors, for a coherent treatment as approach are given in [15, 20].

Following the discussion in [13], we are going to give a first construction of this adjustment. We assume to be in a probability space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$, where $\mathbb{G} = (\mathcal{G}_t)_{t\geq 0}$ is the filtration representing the complete information flow from whole market including default. We assume to be in an arbitrage-free setting so that $\mathbb{P}$ is a risk-neutral measure already selected by some criterion. In particular, the default event is represented by the occurrence of a default time, that we denote by τ. The probability space is endowed also with a right-continuous and complete sub-filtration $\mathcal{F}_t$ representing all the observable market quantities. This filtration does not necessarily include the default event, thus we have $\mathcal{F}_t \subseteq \mathcal{F}_t \vee \mathcal{H}_t = \mathcal{G}_t$, where $\mathcal{H}_t = \sigma(\{\mathbf{1}_{\{\tau\leq s\}}, s \leq t\})$, is the smallest filtration making the random variable τ a stopping time.

In the market, some assets (including a bond) are traded, and an investor forms an investing

[1] The term "derivative desk" typically refers to a specific department within a financial institution, such as a bank or investment firm, that is responsible for trading and managing financial derivatives.

strategy with finite maturity T generating a cashflow during the lifetime of the contract. This cashflow might consider or not the default possibility and we denote by $\Pi^D(t,T), \Pi(t,T)$ the discounted cashflows between t and T respectively generated by the defaultable and the default-free portfolio.

We set V_t as the Net Present Value, $V_t = \mathbb{E}_t^{\mathcal{G}}\big[\Pi(t,T)\big]$, $\mathbb{E}^{\mathcal{G}}$ will denote the conditional expectation $\mathcal{G}_t$. Thus V_t is the default-free price. We assume the investor's prospective, and we want to assess the contract value subject a counterparty risk, that we indicate by $\widetilde{V}_t$. If there is no default by the counterparty during the lifetime of the contract, the on price is given by the final condition.

If, on the other hand, the counterparty goes into default before the maturity, the defaultable price is made up of two parts: the discounted cashflow until the default time and a part depending upon the contract value at default at t.

If V_τ is negative, the investor has to pay the value due to the counterparty, even if the counterparty goes bankrupt. As a result, the investor is solely responsible for paying it. If V_τ is positive, the investor has the right to be paid by the counterparty, but due to the counterparty's default, the investor receives only a recovery fraction R of V_τ, called the Recovery Rate. It is a constant in $[0,1]$, and it is usually agreed contractually at the initial time, possibly because a certain amount of wealth was secured by a collateralization requirement. Then we have

$$\widetilde{V}_t := \mathbb{E}_t^{\mathcal{G}}\Big[\Pi^D(t,T)\Big] = \mathbb{E}_t^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau\leq T\}}\big(\Pi(t,\tau) + \mathrm{e}^{-\int_t^\tau r_s ds}(RV_\tau^+ + V_\tau^-)\big)\Big], \quad (1.1)$$

where we denoted by $\{r_t\}_t$ the interest rate process determining the money market account, and we set $V_\tau^+ = max\big(V_\tau, 0\big)$, $V_\tau^- = min(V_\tau, 0)$.

Theorem 1.1.1 (***General counterparty risk pricing formula*** [19]**).** *The price of a financial contract at time $t \geq 0$ under counterparty risk with default time $\tau > t$ is*

$$\mathbb{E}_t^{\mathcal{G}}\Big[\Pi^D(t,T)\Big] = \mathbb{E}_t^{\mathcal{G}}\Big[\Pi(t,T) - LGD\mathbf{1}_{\{t<\tau\leq T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+\Big], \quad (1.2)$$

with $LGD = 1 - R$ (Loss Given Default) and the recovery fraction $0 \leq R \leq 1$ is assumed to be deterministic.

Proof. Since

$$\Pi(t,T) = \mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau\leq T\}}\Pi(t,T),$$

we can write the random variable inside the expectation in (1.2), as

$$\begin{aligned}
&\mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau\leq T\}}\Pi(t,T) - LGD\mathbf{1}_{\{t<\tau\leq T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+ \\
=&\mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau\leq T\}}\Pi(t,T) + (R-1)\mathbf{1}_{\{t<\tau\leq T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+ \\
=&\mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau\leq T\}}\Pi(t,T) + R\mathbf{1}_{\{t<\tau\leq T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+ - \mathbf{1}_{\{t<\tau\leq T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+.
\end{aligned}$$

We consider the conditional expectation with respect to $\mathcal{G}_\tau$,

$$\begin{aligned}&\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathbf{1}_{\{\tau>T\}}\Pi(t,T)+\mathbf{1}_{\{t<\tau\le T\}}\Pi(t,T)+R\mathbf{1}_{\{t<\tau\le T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau-\mathbf{1}_{\{t<\tau\le T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau\Big]\\ =&\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathbf{1}_{\{\tau>T\}}\Pi(t,T)\Big]+\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathbf{1}_{\{t<\tau\le T\}}\Pi(t,T)+\mathbf{1}_{\{t<\tau\le T\}}\big(R\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau-\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau\big)\Big].\end{aligned}$$

We note that

$$\begin{aligned}\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathbf{1}_{\{\tau>T\}}\Pi(t,T)\Big]&=\mathbf{1}_{\{\tau>T\}}\Pi(t,T),\\ \mathbf{1}_{\{t<\tau\le T\}}\Pi(t,T)&=\mathbf{1}_{\{t<\tau\le T\}}\big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T)\big).\end{aligned}$$

So we have

$$\begin{aligned}&\mathbf{1}_{\{\tau>T\}}\Pi(t,T)+\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathbf{1}_{\{t<\tau\le T\}}\big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T)+R\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau-\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau\big)\Big]\\ =&\mathbf{1}_{\{\tau>T\}}\Pi(t,T)+\mathbf{1}_{\{t<\tau\le T\}}\Pi(t,\tau)+\mathbf{1}_{\{t<\tau\le T\}}\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T)-\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau\\ &+R\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau\Big].\end{aligned}$$

Notice that $V^+_\tau=\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]^+$, hence

$$\begin{aligned}&\mathbf{1}_{\{\tau>T\}}\Pi(t,T)+\mathbf{1}_{\{t<\tau\le T\}}\Pi(t,\tau)+\mathbf{1}_{\{t<\tau\le T\}}\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T)-\mathrm{e}^{-\int_t^\tau r_s ds}\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]^+\Big]\\ &+\mathbf{1}_{\{t<\tau\le T\}}\mathbb{E}^{\mathcal{G}}_\tau\Big[R\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau\Big],\end{aligned}$$

where

$$\begin{aligned}&\mathbb{E}^{\mathcal{G}}_\tau\Big[\mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T)-\mathrm{e}^{-\int_t^\tau r_s ds}\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]^+\Big]\\ =&\mathrm{e}^{-\int_t^\tau r_s ds}\mathbb{E}^{\mathcal{G}}_\tau\Big[\Pi(\tau,T)\Big]-\mathrm{e}^{-\int_t^\tau r_s ds}\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]^+\\ =&\mathrm{e}^{-\int_t^\tau r_s ds}\left[\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]-\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]^+\right]\\ =&\begin{cases}0, & if\quad \mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]>0\\ \mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big], & otherwise,\end{cases}\end{aligned}$$

hence

$$\mathrm{e}^{-\int_t^\tau r_s ds}\left[\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]-\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]^+\right]=\mathrm{e}^{-\int_t^\tau r_s ds}\mathbb{E}^{\mathcal{G}}_\tau\big[\Pi(\tau,T)\big]^-.$$

Finally,

$$\begin{aligned}&\mathbf{1}_{\{\tau>T\}}\Pi(t,T)+\mathbf{1}_{\{t<\tau\le T\}}\big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}V^-_\tau+R\mathrm{e}^{-\int_t^\tau r_s ds}V^+_\tau\big)\\ =&\mathbf{1}_{\{\tau>T\}}\Pi(t,T)+\mathbf{1}_{\{t<\tau\le T\}}\big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}(RV^+_\tau+V^-_\tau)\big)=\Pi^D(t,T).\end{aligned}$$

Conditioning the obtained result with respect to the information available at t, and using the fact that $\mathbb{E}^{\mathcal{G}}_t\big[\mathbb{E}^{\mathcal{G}}_\tau[\cdot]\big]=\mathbb{E}^{\mathcal{G}}_t\big[\big]$ due to $t<\tau$, we obtain (1.2). □

The value of a defaultable claim is clearly the value of the corresponding default-free claim minus an option part evaluated at τ, which represents the discount with respect to the default-free price

$$CVA = \mathbb{E}_t^{\mathcal{G}}\left[LGD\mathbf{1}_{\{t<\tau\leq T\}}\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+\right] > 0. \tag{1.3}$$

In particular, we can write

$$\widetilde{V}_t = V_t - CVA.$$

Let's provide a simple example to better understand this concept and provide an initial approximation.

In the derivatives market, there are products known as options.

An option is a derivative financial contract that bestows upon its holder the right, though not the obligation, to engage in the purchase or sale of a specified quantity of an underlying asset at a predetermined price, known as the strike price, on a future date. The underlying asset encompasses a wide range of financial instruments, commodities, indices, or currencies. When an individual purchases an option, she effectively transfers the potential downside risk to the counterparty. Conversely, the seller of the option assumes the obligation to fulfill the contractual terms by trading the underlying asset at the agreed-upon price, regardless of whether it leads to a loss relative to the prevailing market price.

It is essential to recognize that the buyer and seller positions within an option contract are inherently asymmetric. To compensate the seller for undertaking the risk, they receive an initial payment, typically referred to as the premium, which reflects the value of the option. The determination of this value relies on pertinent market data and must be sufficiently equitable to incentivize both parties to enter into the contract.

There exist two primary classifications of options: calls and puts. A call option grants the buyer the right to purchase the underlying asset at the strike price, while a put option grants the buyer the right to sell it. Both buyers and sellers have the flexibility to choose which side of the option contract they wish to engage in.

Additionally, options can be further categorized as European or American. European options can only be exercised upon reaching the maturity date, whereas American options offer the flexibility to be exercised at any time during the contract's lifespan.

Let's consider an investor who purchases a one-year $T = 1$ European call option on an underlying asset with an initial value of $S = 100$, a strike price $K = 90$, a deterministic risk-free interest rate $r = 0.05$, and an underlying asset volatility $\sigma = 0.4$. We assume that the investor is default-free, while the counterparty may default, and the loss given default (LGD) is 0.5.

The default event, denoted as τ, can occur at any point during the contract's lifetime $[0, T]$.

We aim at understanding how the option value changes when considering counterparty credit risk compared to the case without default risk.

By simulating 10^4 paths of the underlying asset S, we apply a simple Monte Carlo method to obtain two sets of results.[2]

In Figure 1.1, the blue curve represents the option value when not considering counterparty risk, while the red curve represents the value of the call option when the counterparty may default.

We observe that both curves converge to the same value at maturity, as no default event has occurred in the defaultable case, resulting in identical valuations.

However, during the contract's lifetime, we can see how the possibility of counterparty default decreases the value of the contract. Initially, the value of the default-free contract is 20.9094, while considering the counterparty's default risk reduces the contract value to 20.5963, resulting in an approximation of the credit value adjustment of 0.3132.

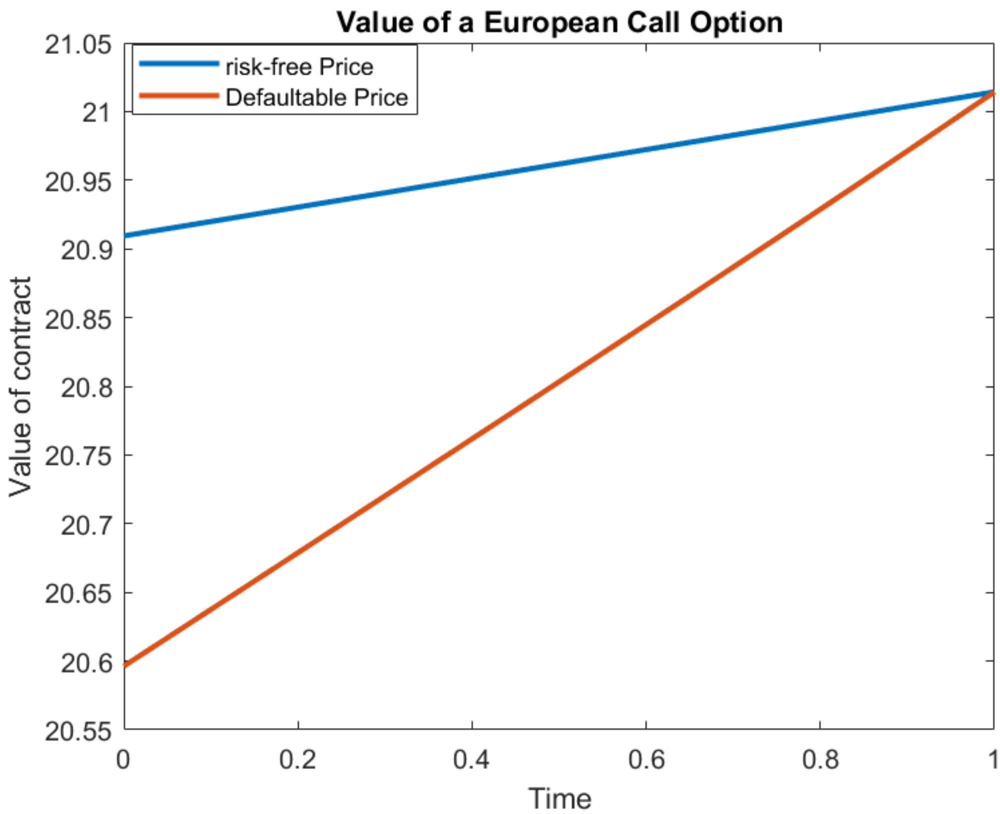

Figure 1.1: The behavior of a European call option with a maturity of one year, $T = 1$, strike price $K = 90$, underlying asset $S = 100$, deterministic interest rate $r = 0.05$, $LGD = 0.05$, and volatility $sigma = 0.4$ is as follows: $Call\_Price\_Default\_Free = 20.9094$;
$Call\_Price\_Defaultable = 20.5963$;
$UCVA = 0.3132$

Before 2007, banks estimated counterparty credit risk using unilateral CVA. If both parties are defaultable, this valuation is asymmetric and it contradicts the accounting principle that an asset for one party is a liability for another.

Bilateral adjustments address this discrepancy by adding a new unit of measure BCVA (*Bilateral Credit Value Adjustment*).

[2]We do not specify here the dynamics of the underlying asset and the Monte Carlo method used. It will be addressed in more detail in the subsequent chapters.

1.2 Bilateral CVA

After 2009, the Basel Committee released Basel III, which required financial institutions to evaluate any party's default risk also when evaluating products having cashflows in both directions.

Bilateral default risk appears in [6, 13] for the first time, who price both the CVA and DVA of a derivative deal.

In [14, 15], a fully rigorous formula is developed, which shows that the bilateral [3] counterparty risk adjustment computed by one of the two parties is obtained as a difference of two terms: an adjustment of the valuation of credit (CVA) due to the counterparty's default and a new measure called *Debt Value Adjustment* (DVA), which was introduced as an accrual of the claim's value due to the investor's default risks. The potential for self-default benefits the contract value, hence it has a negative sign.

Now, we assume that both financial entities might default, and we denote with $\tau = min(\tau_I, \tau_C)$ the event first-to-default between two parties, where τ_C and τ_I denote, respectively, the counterparty's and investor's time default.

- If $\tau > T$, both on investor and the counterparty are not in default during the contract's life.
- If $\tau \leq T$, it means that one of the two parties is in default. As in the unilateral case, we have the discounted cashflow until the default time and the discounted expected value at τ. We have two situations:
  - $i.$ if $\tau = \tau_C$ and $V_\tau \leq 0$, the investor must pay to the counterparty, if $V_\tau > 0$ the investor receives only a recovery fraction R_C of V_τ paying by the counterparty.
  - $ii.$ if $\tau = \tau_I$, the situation is symmetric, namely, $V_\tau > 0$, the counterparty will receive the payment from the investor, otherwise if $V_\tau \leq 0$ the counterparty will pay only a fraction R_I of V_τ to the investor.

The general payoff under bilateral counterparty default risk is

$$\begin{aligned} \widetilde{V}_t = \mathbb{E}_t^{\mathcal{G}}\left[\Pi^D(t,T)\right] =& \mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_{\{\tau>T\}}\Pi(t,T)\right] \\ &+\mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_{\{\tau=\tau_C\}}\mathbf{1}_{\{\tau\leq T\}}\Big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\big(R_C V_\tau^+ + V_\tau^-\big)\Big)\right] \\ &-\mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_{\{\tau=\tau_I\}}\mathbf{1}_{\{\tau\leq T\}}\Big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\big(V_\tau^+ + R_I V_\tau^-\big)\Big)\right] \end{aligned} \tag{1.4}$$

Theorem 1.2.1 (***General bilateral counterparty risk pricing formula***[13]). [4] *For any fixed* $t \in [0,T]$ *on the event* $\{\tau > t\}$ *the price of the financial contract under bilateral counterparty risk is*

[3] Bilateral emphasizes the investor's default in the framework. Thus, the investor's counterparty risk position price is opposite to the counterparty's.

[4] The proof following the same idea of Theorem 1.1.1, for completeness we include it

given by

$$
\begin{aligned}
\mathbb{E}_t^{\mathcal{G}}\Big[\Pi^D(t,T)\Big] =&\mathbb{E}_t^{\mathcal{G}}\Big[\Pi(t,T)\Big]\\
&-\mathbb{E}_t^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau=\tau_C\}}\mathbf{1}_{\{\tau\le T\}}LGD_C \mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+\Big]\\
&+\mathbb{E}_t^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau=\tau_I\}}\mathbf{1}_{\{\tau\le T\}}LGD_I \mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^-\Big],
\end{aligned}
\tag{1.5}
$$

where $LGD_i = 1 - R_i$ is the Loss Given Default, for $i = I, C$ and R_I, R_C denote the recovery fractions of the transaction market value.

Proof. We know that

$$\Pi(t,T) = \mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau_C\le T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau_I\le T\}}\Pi(t,T),$$

hence the fist term in the expectation in (1.5) is

$$
\begin{aligned}
&\mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau_C\le T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau_I\le T\}}\Pi(t,T)\\
&-\mathbf{1}_{\{\tau\le T\}}\Big(\mathbf{1}_{\{\tau=\tau_C\}}LGD_C \mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+ + \mathbf{1}_{\{\tau=\tau_I\}}LGD_I \mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^-\Big)\\
=&\mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau_C\le T\}}\Pi(t,T) + \mathbf{1}_{\{t<\tau_I\le T\}}\Pi(t,T)\\
&-\mathbf{1}_{\{\tau\le T\}}\Big[\mathbf{1}_{\{\tau=\tau_C\}}(1-R_C)\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+ + \mathbf{1}_{\{\tau=\tau_I\}}(1-R_I)\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^-\Big]\\
=&\mathbf{1}_{\{\tau>T\}}\Pi(t,T)\\
&+\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_C\}}\big(\Pi(t,T) - \mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+ + R_C\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+\big)\\
&+\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_I\}}\big(\Pi(t,T) - \mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^- + R_I\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^-\big).
\end{aligned}
$$

Remember that

$$
\begin{aligned}
V_\tau &= \mathbb{E}_\tau^{\mathcal{G}}\Big[\Pi(\tau,T)\Big],\\
\mathbf{1}_{\{t<\tau\le T\}}\Pi(t,T) &= \mathbf{1}_{\{t<\tau\le T\}}\big(\Pi(t,\tau) + \mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T)\big).
\end{aligned}
$$

So, we have

$$
\begin{aligned}
&\mathbf{1}_{\{\tau>T\}}\Pi(t,T)\\
&+\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_C\}}\big(\Pi(t,\tau) + \mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T) - \mathrm{e}^{-\int_t^\tau r_s ds}\mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^+ + R_C\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+\big)\\
&+\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_I\}}\big(\Pi(t,\tau) + \mathrm{e}^{-\int_t^\tau r_s ds}\Pi(\tau,T) - \mathrm{e}^{-\int_t^\tau r_s ds}\mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^- + R_I\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^-\big).
\end{aligned}
$$

Conditionally on the information available at time τ, and by linearity of the expectation, we get

$$
\begin{aligned}
&\mathbb{E}_\tau^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau>T\}}\Pi(t,T)\Big] \\
&+\mathbb{E}_\tau^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_C\}}\Big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\Big(\Pi(\tau,T)-\mathbb{E}_\tau^{\mathcal{G}}[\Pi(\tau,T)]^+\Big)+R_C\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+\Big)\Big] \\
&+\mathbb{E}_\tau^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_I\}}\Big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\Big(\Pi(\tau,T)-\mathbb{E}_\tau^{\mathcal{G}}[\Pi(\tau,T)]^-\Big)+R_I\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^-\Big)\Big].
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\mathbb{E}_\tau^{\mathcal{G}}\Big[\Pi(\tau,T)-\mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^+\Big] &= \mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]-\mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^+ = \mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^- = V_\tau^- \\
\mathbb{E}_\tau^{\mathcal{G}}\Big[\Pi(\tau,T)+\mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^-\Big] &= \mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]+\mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^- = \mathbb{E}_\tau^{\mathcal{G}}\big[\Pi(\tau,T)\big]^+ = V_\tau^+.
\end{aligned}
$$

Finally,

$$
\begin{aligned}
&\mathbf{1}_{\{\tau>T\}}\Pi(t,T) \\
&+\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_C\}}\Big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\Big(R_C V_\tau^+ + V_\tau^-\Big)\Big) \\
&+\mathbf{1}_{\{\tau\le T\}}\mathbf{1}_{\{\tau=\tau_I\}}\Big(\Pi(t,\tau)+\mathrm{e}^{-\int_t^\tau r_s ds}\Big(V_\tau^+ + R_I V_\tau^-\Big)\Big).
\end{aligned}
$$

Conditioning the obtained result on the information available at t, and using the fact that $\mathbb{E}_t^{\mathcal{G}}\big[\mathbb{E}_\tau^{\mathcal{G}}[\cdot]\big] = \mathbb{E}_t^{\mathcal{G}}\big[\,\cdot\,\big]$ due to $t<\tau$, we obtain (1.4). □

The bilateral pricing formula (1.4), unlike the unilateral pricing formula (1.1), is symmetric, meaning both the investor and the counterparty obtain the same value. The value of a defaultable claim is the value of the corresponding default-free claim plus two terms. The first gives a nonzero contribution only if the counterparty defaults first, and the second gives a nonzero contribution only if the investor defaults first. This is known as Bilateral Credit Value Adjustment

$$BCVA = -CVA - DVA \tag{1.6}$$

where

$$
\begin{aligned}
CVA =& \mathbb{E}_t^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau=\tau_C\}}\mathbf{1}_{\{\tau\le T\}}LGD_C\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^+\Big] \\
DVA =& \mathbb{E}_t^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau=\tau_I\}}\mathbf{1}_{\{\tau\le T\}}LGD_I\mathrm{e}^{-\int_t^\tau r_s ds}V_\tau^-\Big].
\end{aligned} \tag{1.7}
$$

In particular, we can write

$$\widetilde{V}_t = V_t + BCVA.$$

1.3 Trading under Collateralization

In order to minimize counterparty credit risk, financial institutions use collateralization. A collateral account is jointly opened by both parties and it is used to prevent losses. The parties estimate the contract's value and compare it daily with the current value. The party that estimates a low value compared to the current one has to post cash or assets equal to the difference on the account, otherwise he/she draws them. We call the Collateral Taker the investor or counterparty withdraws from that the account, the other Collateral Provider. This role is interchangeable according to the sign of the collateral. We say, if the sign of the collateral is positive, the Collateral Taker is the investor, otherwise he/she is the counterparty. Moreover, the Collateral Taker can rehypothecate the collateral's excess, which means he/she can invest the account's surplus to self-finance.

1.3.1 Collateralization

Collateralization has been analyzed in [23] and more recently in [13, 34].

We denote the *collateral account value* $\{C_t\}_t$ which has to be $\mathcal{F}_t$-adapted stochastic process, since we assume that it is a risk-free cash account. We also assume that $C_t = 0$ for all $t \leq 0$ and $t \geq T \wedge \tau$, that means the collateral account is opened for each new contract and closed when the deal defaults or ends. If the account is closed, any collateral must be returned to the original party.

We describe the mechanism of collateral posting. If the Collateral Taker is the investor, this means that by time t the overall collateral account is composed of the counterparty's excess of collateral posted, in this case the collateral account is positive $C_t > 0$, and it can be used by the investor to reduce his exposure. When $C_t < 0$ the investor is Collateral Provider and the counterparty is the Collateral Taker.

1.3.2 Close-Out value

The ISDA specifies how parties must respond to defaults to fulfill their contractual obligations. The surviving party should assess the final transactions, and claim for a reimbursement only after the application of collateralization.

The ISDA Master Agreement defines this term close-out value, which is the the amount of losses or costs the surviving party would incur in replacing or providing for an economic equivalent at time of counterparty default. Notice that the close-out amount is not a symmetric quantity with respect to the exchange of roles between two parties since it is valued by one party after the default of the other one. The close-out value is agreed upon at the beginning of the contract, and it depends only on market information, which means it has to be $\mathbb{F}$-measurable. We denote the close-out value as $\varepsilon_\tau := \mathbf{1}_{\{\tau=\tau_C\}}\varepsilon_{I,\tau} + \mathbf{1}_{\{\tau=\tau_I\}}\varepsilon_{C,\tau}$, where

- $\varepsilon_{I,t}$ is the investor's close-out value based on the counterparty's default at time τ. If it is positive, the investor is the counterparty's creditor;

- $\varepsilon_{C,t}$ is the close-out amount when the investor is defaulting. A negative value for $\varepsilon_{C,t}$, means that the counterparty is a creditor to the investor.

1.3.3 Collateral Rehypothecation

The Collateral Provider expects to receive the remaining collateral from the Collateral Taker at maturity of the contract or in case of default, after it has been used to cover the exposure.

However, if the Collateral Taker has rehypothecated the excess collateral, the Collateral Provider may only receive a fraction of the original collateral, which leaves the Collateral Taker as an unsecured creditor.

We denote as R_i', for $i = C, I$ the recovery fraction of rehypothecation, when the Collateral Taker is respectively the counterparty or the investor and with $LGD_i' = 1 - R_i'$ the respective loss fractions.

1.3.4 Bilateral CVA Formula under Collateralization and Close-out rules

In this subsection we aim at constructing a valuation framework that includes both collateralization and close-out value.

- If the investor measures a positive close-out value on counterparty $\varepsilon_{I,\tau} > 0$, we have two cases

  $i.$ $C_\tau > 0$, the investor is Collateral Taker and he/she uses the collateral to reduce the close-out value $(\epsilon_{I,\tau} - C_\tau)$, and if

    $a.$ $(\varepsilon_{I,\tau} - C_\tau) > 0$, means the collateral is not enough, the investor suffers a loss and he/she recovers only a fraction R_C of exposure;

    $b.$ $(\varepsilon_{I,\tau} - C_\tau) \leq 0$ the remaining collateral (if any) is returned to the counterparty

$$R_C(\varepsilon_{I,\tau} - C_\tau)^+ + (\varepsilon_{I,\tau} - C_\tau)^-.$$

  $ii.$ $C_\tau < 0$, the counterparty is the Collateral Taker, the investor may suffer a loss from the exposure, and he/she may only be able to recover a fraction of the collateral R_C. Collateral is returned to the investor if it is not rehypothecated ($R_C' = 1$), otherwise only a portion of it is returned

$$R_C \varepsilon_{I,\tau} - R_C' C_\tau.$$

- if the close-out value of the contract is negative for the investor $\varepsilon_{I,\tau} < 0$, even if the counterparty defaults, the investor is still obligated to pay it.

  $i.$ If $C_\tau > 0$, the investor must return the entire collateral to the counterparty.

$$\varepsilon_{I,\tau} - C_\tau.$$

$ii.$ If $C_\tau < 0$, the collateral is used by the counterparty to reduce the investor's debt,

a. if $(\varepsilon_{I,\tau} - C_\tau) < 0$, the investor pays the remaining exposed;

b. otherwise, $(\varepsilon_{I,\tau} - C_\tau) > 0$, the collateral covers the investor's debt, and the remainder (if any, and if it is not rehypothecated) gets back to the investor.

$$(\varepsilon_{I,\tau} - C_\tau)^- + R'_C(\varepsilon_{I,\tau} - C_\tau)^+.$$

Similarly, one can compute the situation when the investor defaults before the counterparty.

Now, we may combine both parties' default, non-default, and collateral account cashflows obtaining

$$\begin{aligned}
\Pi^D(t,T;C) =& \mathbf{1}_{\{\tau>T\}}\Pi(t,T) + \mathbf{1}_{\{\tau<T\}}(\Pi(t,\tau) + \mathrm{e}^{-\int_t^\tau r_s ds}C_\tau) \\
&+\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{I,\tau}<0\}}\mathbf{1}_{\{C_\tau>0\}}(\varepsilon_{I,\tau} - C_\tau) \\
&+\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{I,\tau}<0\}}\mathbf{1}_{\{C_\tau<0\}}\big((\varepsilon_{I,\tau} - C_\tau)^- + R'_C(\varepsilon_{I,\tau} - C_\tau)^+\big) \\
&+\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{I,\tau}>0\}}\mathbf{1}_{\{C_\tau>0\}}\big((\varepsilon_{I,\tau} - C_\tau)^- + R_C(\varepsilon_{I,\tau} - C_\tau)^+\big) \\
&+\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{I,\tau}>0\}}\mathbf{1}_{\{C_\tau<0\}}(R_C\varepsilon_{I,\tau} - R'_C C_\tau) \\
&+\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{C,\tau}>0\}}\mathbf{1}_{\{C_\tau<0\}}(\varepsilon_{C,\tau} - C_\tau) \\
&+\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{C,\tau}>0\}}\mathbf{1}_{\{C_\tau>0\}}\big((\varepsilon_{C,\tau} - C_\tau)^+ + R'_I(\varepsilon_{C,\tau} - C_\tau)^-\big) \\
&+\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{C,\tau}<0\}}\mathbf{1}_{\{C_\tau<0\}}\big((\varepsilon_{C,\tau} - C_\tau)^+ + R_I(\varepsilon_{C,\tau} - C_\tau)^-\big) \\
&+\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\mathbf{1}_{\{\varepsilon_{C,\tau}<0\}}\mathbf{1}_{\{C_\tau>0\}}(R_I\varepsilon_{C,\tau} - R'_I C_\tau),
\end{aligned} \tag{1.8}$$

with $\Pi^D(t,T;C)$ denoting the analogous net cashflow inclusive of collateralization.

A more compact form is provided by

$$\begin{aligned}
\Pi^D(t,T;C) =& \Pi(t,T) \\
&-\mathbf{1}_{\{\tau<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(\Pi(\tau,T) - \mathbf{1}_{\{\tau=\tau_C\}}\varepsilon_{I,\tau} - \mathbf{1}_{\{\tau=\tau_I\}}\varepsilon_{C,\tau}\big) \\
&-\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(1 - R_C\big)\big(\varepsilon^+_{I,\tau} - C^+_\tau\big)^+ \\
&-\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(1 - R'_C\big)\big(\varepsilon^-_{I,\tau} - C^-_\tau\big)^+ \\
&-\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(1 - R_I\big)\big(\varepsilon^-_{C,\tau} - C^-_\tau\big)^- \\
&-\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(1 - R'_I\big)\big(\varepsilon^+_{C,\tau} - C^+_\tau\big)^-
\end{aligned} \tag{1.9}$$

Notice that the collateral account entry only as a term reducing the exposure of each party upon the default of the other, keeping in mind which party posted the collateral.

1.3.5 Collateralization BCVA General Formula

The bilateral valuation adjustment in presence of collateralization is them given by

$$\begin{aligned} BCVA(C) :=& \mathbb{E}_t^{\mathcal{G}}\big[\Pi^D(t,T;C)\big] - \mathbb{E}_t^{\mathcal{G}}\big[\Pi(t,T)\big] \\ =& -\mathbb{E}_t^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(\Pi(\tau,T) - \mathbf{1}_{\{\tau=\tau_C\}}\varepsilon_{I,\tau} - \mathbf{1}_{\{\tau=\tau_I\}}\varepsilon_{C,\tau}\big)\Big] \\ & -\underbrace{\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(LGD_C\big(\varepsilon_{I,\tau}^+ - C_\tau^+\big)^+ + LGD_C'\big(\varepsilon_{I,\tau}^- - C_\tau^-\big)^+\big)\big]}_{:=CVA(C)} \\ & -\underbrace{\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(LGD_I\big(\varepsilon_{C,\tau}^- - C_\tau^-\big)^- + LGD_I'\big(\varepsilon_{C,\tau}^+ - C_\tau^+\big)^-\big)\big]}_{:=DVA(C)}, \end{aligned} \tag{1.10}$$

Special cases

We introduce a new term, called mark-to-market exposure [5] ε_u, with $t < u \leq T$, as given by

$$\varepsilon_u = \mathbb{E}_u^{\mathcal{G}}\big[\Pi(u,T)\big],$$

which represents the default-free price of all cashflows remaining after time u up to maturity T. Hence, the (1.10) can rewrite as

$$\begin{aligned} BCVA(C) =& -\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(\varepsilon_\tau - \mathbf{1}_{\{\tau=\tau_C\}}\varepsilon_{I,\tau} - \mathbf{1}_{\{\tau=\tau_I\}}\varepsilon_{C,\tau}\big)\big] \\ & -\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(LGD_C\big(\varepsilon_{I,\tau}^+ - C_\tau^+\big)^+ + LGD_C'\big(\varepsilon_{I,\tau}^- - C_\tau^-\big)^+\big)\big] \\ & -\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(LGD_I\big(\varepsilon_{C,\tau}^- - C_\tau^-\big)^- + LGD_I'\big(\varepsilon_{C,\tau}^+ - C_\tau^+\big)^-\big)\big], \end{aligned}$$

If we assume $\varepsilon_\tau = \varepsilon_{C,\tau} = \varepsilon_{I,\tau}$ the expression for $BCVA$ is

$$\begin{aligned} BCVA(C) =& -\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(LGD_C\big(\varepsilon_{I,\tau}^+ - C_\tau^+\big)^+ + LGD_C'\big(\varepsilon_{I,\tau}^- - C_\tau^-\big)^+\big)\big] \\ & -\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(LGD_I\big(\varepsilon_{C,\tau}^- - C_\tau^-\big)^- + LGD_I'\big(\varepsilon_{C,\tau}^+ - C_\tau^+\big)^-\big)\big], \end{aligned}$$

If collateral rehypothecation is not permitted ($LGD_C' = LGD_I' = 0$), the above formula simplifies to

$$\begin{aligned} BCVA(C) =& -\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_C<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}LGD_C\big(\varepsilon_{I,\tau}^+ - C_\tau^+\big)^+\big] \\ & -\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau=\tau_I<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}LGD_I\big(\varepsilon_{C,\tau}^- - C_\tau^-\big)^-\big]. \end{aligned}$$

[5] Mark-to-market (MTM) is a financial accounting practice that involves valuing an asset or liability at its current market value. In other words, MTM is the process of adjusting the value of an asset or liability to reflect its current market price. When an asset is marked-to-market, its value is adjusted daily to reflect changes in market conditions. For example, if an investor owns a stock that is marked-to-market, the value of that stock will be adjusted each day based on the current market price. If the price of the stock goes up, the value of the investor's position will increase, and if the price goes down, the value will decrease.

On the other hand, if rehypothecation is permitted, the surviving party is always forced to face the worst case scenario ($LGD_i = LGD'_i$ for $i = C, I$)

$$\begin{aligned} BCVA(C) = &- \mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_{\{\tau=\tau_C<T\}} \mathrm{e}^{-\int_t^\tau r_s ds} LGD_C\big(\varepsilon_{I,\tau} - C_\tau\big)^+\right] \\ &- \mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_{\{\tau=\tau_I<T\}} \mathrm{e}^{-\int_t^\tau r_s ds} LGD_I\big(\varepsilon_{C,\tau} - C_\tau\big)^-\right]. \end{aligned} \tag{1.11}$$

Finally, if we remove collateralization we get (1.6), and and we consider a default-free investor ($\tau_I \to \infty$) we have (1.3).

1.4 Example of Collateralization Schemes

We examine two distinct collateralization mechanisms. The first one is the more realistic, and it is called *margining procedure*, in which both parties post or withdraw collateral to or from the account on a fixed set of dates based on their current exposure. The second mechanism, known as *perfect collateralization*, where the collateral account covers all exposure risk.

1.4.1 Collateral Management under Margining Procedures

The CSA agreement between the parties establishes the terms and conditions for collateralization, including the specific collateral assets that can be used, the margin threshold, and the frequency of margin calls. In addition, the agreement ensures that the Collateral Taker is required to compensate the collateral account at a predetermined accrual rate, which reflects the cost of borrowing funds in the market. The accrual rate may be set as a fixed percentage or as a variable rate linked to a benchmark, such as the overnight interest rate.

We introduce two adapted processes that represent (forward) collateral accrual rates, $\{c_t^>(T)\}_t$ when the investor takes the collateral assets (he/she is the Collateral Taker), and $\{c_t^<(T)\}_t$ when he/she posts them (the counterparty is the Collateral Taker). Furthermore, we define (collateral) zero-coupon bonds as $\{P_t^{\widetilde{c}}(T)\}_t$ as

$$P_t^{\widetilde{c}}(T) := \frac{1}{1 + (T-t)\widetilde{c}_t(T)},$$

where, $\widetilde{c}$ is the effective collateral accrual rate, defined as

$$\widetilde{c}_t(T) := \mathbf{1}_{\{C_t<0\}} c_t^<(T) + \mathbf{1}_{\{C_t>0\}} c_t^>(T).$$

We can represent the timeline of a deal using a time grid $\{t_1, \ldots, t_n = T\}$, that includes the dates when both parties post or withdraw collateral to or from the collateral account. We can begin by listing all cashflows from the investor to counterparty when a default event does not occur:

1. the investor opens the account at the first margining date t_1 if $C_{t_1} < 0$ (the counterparty is the Collateral Taker);

2. The investor posts to the account at each t_k as long as $C_{t_k} < 0$. As Collateral Taker, the counterparty pays interest on the collateral at the accrual rate $c^{<}_{t_k}(t_{k+1})$ between two subsequent margin dates t_k and t_{k+1};

3. if no default event has occurred, the investor closes the account at the last margining date t_n if $C_{t_n} < 0$.

At each margining date, the counterparty considers the same cashflows for opposite values of the collateral account, that is, when the investor is the Collateral Taker and he/she pays a rate $c^{>}_{t_k}(t_{k+1})$. As a result, we denote γ as the sum of the discounted margining costs over the period $(t, T \wedge \tau]$

$$\begin{aligned}
\gamma(t, T\wedge\tau; C) :=& \sum_{k=1}^{n-1} \mathbf{1}_{\{t\le t_k < T\wedge\tau\}} \mathrm{e}^{-\int_t^{t_k} r_s ds} \left(C_{t_k} - C^{-}_{t_k} \frac{P_{t_k}(t_{k+1})}{P^{c<}_{t_k}(t_{k+1})} - C^{+}_{t_k} \frac{P_{t_k}(t_{k+1})}{P^{c>}_{t_k}(t_{k+1})} \right) \\
=& \sum_{k=1}^{n-1} \mathbf{1}_{\{t\le t_k < T\wedge\tau\}} \mathrm{e}^{-\int_t^{t_k} r_s ds} C_{t_k} \left(1 - \frac{P_{t_k}(t_{k+1})}{P^{\widetilde{c}}_{t_k}(t_{k+1})} \right) \\
=& \sum_{k=1}^{n-1} \mathbf{1}_{\{t\le t_k < T\wedge\tau\}} \mathrm{e}^{-\int_t^{t_k} r_s ds} C_{t_k} \left(1 - \frac{\frac{1}{1+(t_{k+1}-t_k) r_{t_k}(t_{k+1})}}{\frac{1}{1+(t_{k+1}-t_k)\widetilde{c}_{t_k}(t_{k+1})}} \right) \\
=& \sum_{k=1}^{n-1} \mathbf{1}_{\{t\le t_k < T\wedge\tau\}} \mathrm{e}^{-\int_t^{t_k} r_s ds} C_{t_k} \left(1 - \frac{1+(t_{k+1}-t_k)\widetilde{c}_{t_k}(t_{k+1})}{1+(t_{k+1}-t_k) r_{t_k}(t_{k+1})} \right) \\
=& \sum_{k=1}^{n-1} \mathbf{1}_{\{t\le t_k < T\wedge\tau\}} \mathrm{e}^{-\int_t^{t_k} r_s ds} C_{t_k} \left(\frac{(t_{k+1}-t_k)(r_{t_k}(t_{k+1}) - \widetilde{c}_{t_k}(t_{k+1}))}{1+(t_{k+1}-t_k) r_{t_k}(t_{k+1})} \right)
\end{aligned} \tag{1.12}$$

with the risk-free zero coupon bond, related to the risk-free rate r, given by $P_t(T)$. We may approximate (1.12) as the first order expansion, for small $\widetilde{c}$ and r,

$$\gamma(t, T\wedge\tau; C) \approx \sum_{j=1}^{n-1} \mathbf{1}_{\{t\le t_j < T\wedge\tau\}} \mathrm{e}^{-\int_t^{t_j} r_s ds} C_{t_j} \alpha_j \Big(r_{t_j}(t_{j+1}) - \widetilde{c}_{t_j}(t_{j+1}) \Big), \tag{1.13}$$

with α_j is the year fraction between t_j and t_{j+1}.

This last expression clearly shows the cost of carrying a structure for collateral costs. If $C > 0$, the investor is the Collateral Taker and he/she will have to pay (hence the minus sign) an interest $c^{>}$, while receiving the natural growth r for cash. In the opposite case, the investor is the Collateral Provider (the counterparty is the Collateral Taker) and receives (pays) interest $c^{<}$.

We define the Bilateral Credit Valuation Adjusted price $\widetilde{V}_t(C)$, with collateral management under margining procedures, as

$$\begin{aligned}
\widetilde{V}_t(C) :=& \mathbb{E}^{\mathcal{G}}_t \left[\Pi^D(t, T; C) \right] \\
=& \mathbb{E}^{\mathcal{G}}_t \left[\Pi(t, T\wedge\tau) + \gamma(t, T\wedge\tau; C) + \mathbf{1}_{\{\tau<T\}} \mathrm{e}^{-\int_t^{\tau} r_s ds} \theta_\tau(C, \varepsilon) \right]
\end{aligned} \tag{1.14}$$

where $\theta_\tau(C, \varepsilon)$ is the on-default cashflows given by

$$\begin{aligned}\theta_\tau(C;\varepsilon) :=& \mathbf{1}_{\{\tau=\tau_C<\tau_I\}}\Big(\varepsilon_{I,\tau} - LGD_C\big(\varepsilon^+_{I,\tau} - C^+_\tau\big)^+ - LGD'_C\big(\varepsilon^-_{I,\tau} - C^-_\tau\big)^+\Big)\\ &+\mathbf{1}_{\{\tau=\tau_I<\tau_C\}}\Big(\varepsilon_{C,\tau} - LGD_I\big(\varepsilon^-_{C,\tau} - C^-_\tau\big)^- - LGD'_I\big(\varepsilon^+_{I,\tau} - C^+_\tau\big)^-\Big).\end{aligned} \tag{1.15}$$

Namely, to price a deal we have to sum up three components:

i. $\Pi(t, s)$ is the discounted cashflows from the contract's payoff structure over the period $(t, s]$;

ii. $\gamma(t, s; C)$ represents the discounted cashflows of collateral margining costs within the interval $(t, s]$;

iii. $\theta_\tau(C; \varepsilon)$ is the on-default cashflow with close-out amount ε.

1.4.2 Perfect Collateralization

The perfect collateralization scheme is defined by collateralization in continuous time, with continuous mark-to-market of the portfolio at default events, and with collateral account inclusive of margining costs at any time u, i.e

$$C^T_u = \mathbb{E}^{\mathcal{G}}_u\big[\Pi(u, T) + \gamma(u, T; C)\big],$$

with close-out amount chosen to be equal to collateral price, i.e.

$$C_\tau = \varepsilon_{I,\tau} = \varepsilon_{C,\tau}.$$

Then, the price (1.14) becomes

$$\begin{aligned}\widetilde{V}_t(C) =& \mathbb{E}^{\mathcal{G}}_t\Big[\Pi(t, T\wedge\tau) + \gamma(t, T\wedge\tau; C) + \mathbf{1}_{\{\tau<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}C_\tau\Big]\\ =& \mathbb{E}^{\mathcal{G}}_t\Big[\Pi(t, T) + \gamma(t, T; C) - \mathbf{1}_{\{\tau<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\big(\Pi(\tau, T) + \gamma(\tau, T; C)\big) + \mathbf{1}_{\{\tau<T\}}\mathrm{e}^{-\int_t^\tau r_s ds}C_\tau\Big]\\ =& \mathbb{E}^{\mathcal{G}}_t\big[\Pi(t, T) + \gamma(t, T; C)\big]\\ =& C^T_t.\end{aligned}$$

1.5 Funding Value Adjustment

When managing a trading position, liquidity, which is provided by either the Treasury or the market, is an important consideration. The lender must be repaid, and borrowers must pay interest, while those who lend must earn it. It is critical to include these financial expenses in the contract evaluation. For this reason, a new adjustment known as the Funding Value Adjustment (FVA) has to be introduced. More precisely, FVA can be viewed as a cost or benefit for hedging a series of transactions. In recent years, academics and practitioners have contributed to a discussion on FVA (see for instance [22, 25–27]).

Various more or less sophisticated approaches have been suggested for calculating this adjustment. In [17], FVA is examined as a pricing component that can only be calculated recursively due to its dependence on the price itself.

To enhance comprehension, consider the following example of a collateralized swap.

Suppose to have a swap with collateral between two parties. When the market value of the swap changes, the parties send or withdraw cashflows from the collateral account. If the market value of the swap becomes negative, the party that is required to transfer cash or assets to the collateral account may borrow them from the Treasury Department or the market at an unguaranteed rate. Meanwhile, the collateral will continue to accrue interest at a rate set by the CSA.

Alternately, if the value is positive, the party will get collateral and the CSA rate paid to its counterparty.

The swap transaction faces additional expenses due to the asymmetrical nature of the collateral cost, known as FVA. However, it should be noted that the previous example, though understandable, is unrealistic as it oversimplifies the complexity of the actual transaction.

Let's consider a single transaction to explore a general scenario. As we showed, CVA and DVA are additively decomposable within the deal's cashflow, but financing costs are not, because they depend on future financing choices. As a result, valuing the product requires a recursive equation, which is challenging to run due to the product's path-dependent nature and the need for both backward induction and forward simulation.

1.5.1 Trading under funding risk

The strategy for hedging that accurately replicates the arbitrage-free valuation of a derivative is composed of a cash position, typically referred to as a current account denoted by $\{F_t\}_t$, and a portfolio position of hedging instruments denoted by $\{\overline{H}_t\}_t$ (the risky-asset account[6]). The current account $\{F_t\}_t$ is determined by considering both the positions taken by the trader at time t, which could be either borrowing or investing. The account F_t is positive when the trader obtains the required amount to establish the hedging strategy, meaning that the trader borrows cash. Conversely, F_t is negative when the trader invests the surplus cash into the hedging strategy. We get

$$\widetilde{V}_t(F) = F_t + \overline{H}_t$$

where $\widetilde{V}_t(F)$ is the derivative risky price inclusive of funding and investing costs. Therefore the financing account as

$$F_t = \widetilde{V}_t(F) - \overline{H}_t.$$

[6]In the classical Black-Scholes-Merton framework, a hedge's risky aspect would entail a δ-position stance in the underlying equity, with the risk-free component being held in a secure bank account.

If the contract is collateralized and rehypothecation is available, the Collateral Taker can use collateral assets for funding, reducing or eliminating the requirement for cash,

$$F_t = \widetilde{V}_t(C, F) - C_t - \overline{H}_t$$

where $\widetilde{V}_t(C, F)$ is the derivative risky price with collateral management under margining procedures, funding and investing costs. Notice that this will produce a recursive equation because the price of the product at time t depends on the funding strategy $F((t, T])$ which depends on the price of the product. This will be made clear in the sections that come after this one.

1.5.2 Liquidity policies

The trader's position for financing or investing is determined by its liquidity policy.

Let be $\{t_1, \dots, t_n\}_t$ a discrete time-grid, and assume that the trader enters a funding position, between two adjacent funding times t_j and t_{j+1} for $1 \le j \le n-1$ we have that

- at t_j the trader requests from the funder a cash sum that is equivalent to the amount of F_{t_j}.
- at t_{j+1} the trader must repay the funder for cash and finance charges. The latter expenses are constant at the start of each financing cycle and charged at the conclusion.

We can follow the same logic also for investing cash amounts ($F_t < 0$) not directly used by the trader, and to consider investing periods along with funding periods.

The contracts used by the investor to finance the deal can be introduced as suitable pricing processes. Let $\{P_t^{f^>}(T)\}_t$ be the price of the $\mathbb{F}$-adapted borrowed contract, where the trader pays a unit of cash at maturity T, and let $\{P_t^{f^<}(T)\}_t$ be the price of the $\mathbb{F}$-adapted lent contract where the dealer receives a unit of cash at maturity T.

The corresponding financial/investment rate (forward) is

$$f_t^{\gtrless}(T) := \frac{1}{T-t}\left(\frac{1}{P^{f^{\gtrless}}(T)} - 1\right).$$

In other words, if the hedging strategy of the deal requires borrowing cash, this can be done at the funding rate $f^>$, while surplus cash can be invested at the lending rate $f^<$. We define the effective funding rate $\{\widetilde{f}_t\}_t$ faced by the dealer as

$$\widetilde{f}_t(T) := \mathbf{1}_{\{F_t<0\}} f^<(T) + \mathbf{1}_{\{F_t>0\}} f^>(T).$$

The sum of discounted cashflows from funding costs during the life of the deal is equal to

$$\varphi(t, T \wedge \tau; F) := \sum_{j=1}^{n-1} \mathbf{1}_{\{t \le t_j < T \wedge \tau\}} \mathrm{e}^{-\int_t^{t_j} r_s ds} F_{t_j} \left(1 - \frac{P_{t_j}(t_{j+1})}{P_{t_j}^{\widetilde{f}}(t_{j+1})}\right), \tag{1.16}$$

where the zero-coupon bond corresponding to the effective funding rate is defined as

$$P_t^{\widetilde{f}}(T) := \frac{1}{1 + (T-t)\widetilde{f}_t(T)}.$$

In the same way, (1.12) may be rewritten as a first order approximation with continuously compounded rates $\widetilde{f}$ and r associated to the relevant bonds

$$\varphi(t, T \wedge \tau; F) \approx \sum_{j=1}^{n-1} \mathbf{1}_{\{t \leq t_j < T \wedge \tau\}} \mathrm{e}^{-\int_t^{t_j} r_s ds} F_{t_j} \alpha_j \Big(r_{t_j}(t_{j+1}) - \widetilde{f}_{t_j}(t_{j+1}) \Big). \tag{1.17}$$

Thus Bilateral Credit Valuation Adjusted price $\widetilde{V}_t(C, F)$, inclusive of funding and investing costs, can be written in the following form:

$$\widetilde{V}_t(C, F) = \mathbb{E}_t^{\mathcal{G}} \Big[\Pi(t, T \wedge \tau) + \gamma(t, T \wedge \tau; C) + \varphi(t, T \wedge \tau; F) + \mathbf{1}_{\{t \leq \tau \leq T\}} \mathrm{e}^{-\int_t^{\tau} r_s ds} \theta_\tau(C, \varepsilon) \Big]. \tag{1.18}$$

1.5.3 Implementing Hedging Strategies

At any fixed time t, when the investor borrows a risky asset, $\overline{H}_t > 0$, while $\overline{H}_t \leq 0$, when he/she lends it. We want to consider both financing costs and benefits of employing the risky asset in our framework.

Let us introduce the adapted processes $\{h_t^>(T)\}_t$, that represents the effective asset lending rates from t to T, and $\{h_t^<(T)\}_t$, the asset borrowing rate. We define the (hedging) zero-coupon bonds $\{P_t^{h^{\gtrless}}(T)\}_t$ as given by

$$P_t^{h^{\pm}}(T) := \frac{1}{1 + (T-t) h_t^{\pm}(T)}.$$

It is also useful to introduce the effective lending/borrowing rate $\widetilde{h}_t$ defined as

$$\widetilde{h}_t(T) := \mathbf{1}_{\{\overline{H}_t < 0\}} h_t^<(T) + \mathbf{1}_{\{\overline{H}_t > 0\}} h_t^>(T).$$

If we implement the hedging strategy using the same time grid as the funding procedure, we can add the costs of funding and hedging in a single term and re-define φ to take the hedging strategy into account.

$$\begin{aligned} \varphi(t, T \wedge \tau; F, \overline{H}) := & \sum_{j=1}^{n-1} \mathbf{1}_{\{t \leq t_j < T \wedge \tau\}} \mathrm{e}^{-\int_t^{t_j} r_s ds} F_{t_j} \left(1 - \frac{P_{t_j}(t_{j+1})}{P_{t_j}^{\widetilde{f}}(t_{j+1})} \right) \\ & - \sum_{j=1}^{n-1} \mathbf{1}_{\{t \leq t_j < T \wedge \tau\}} \mathrm{e}^{-\int_t^{t_j} r_s ds} \overline{H}_{t_j} \left(\frac{P_{t_j}(t_{j+1})}{P_{t_j}^{\widetilde{f}}(t_{j+1})} - \frac{P_{t_j}(t_{j+1})}{P_{t_j}^{\widetilde{h}}(t_{j+1})} \right), \end{aligned} \tag{1.19}$$

which might be approximated as

$$\begin{aligned}\varphi(t,T\wedge\tau;F,\overline{H}) \approx &\sum_{j=1}^{n-1}\mathbf{1}_{\{t\le t_j<T\wedge\tau\}}\mathrm{e}^{-\int_t^{t_j} r_s ds}F_{t_j}\alpha_j\Big(r_{t_j}(t_{j+1})-\widetilde{f}_{t_j}(t_{j+1})\Big)\\ &-\sum_{j=1}^{n-1}\mathbf{1}_{\{t\le t_j<T\wedge\tau\}}\mathrm{e}^{-\int_t^{t_j} r_s ds}\overline{H}_{t_j}\alpha_j\Big(f_{t_j}(t_{j+1})-\widetilde{h}_{t_j}(t_{j+1})\Big).\end{aligned} \tag{1.20}$$

In conclusion, we can rewrite (1.18) as

$$\widetilde{V}_t(C,F) = \mathbb{E}_t^{\mathcal{G}}\Bigg[\Pi(t,T\wedge\tau)+\gamma(t,T\wedge\tau;C)+\varphi(t,T\wedge\tau;F,\overline{H})+\mathbf{1}_{\{t\le\tau\le T\}}\mathrm{e}^{-\int_t^{\tau} r_s ds}\theta_\tau(C,\varepsilon)\Bigg]. \tag{1.21}$$

1.6 Continuous-time generalizated Pricing Equation

If we assume a continuous-time approximation for the general valuation equation (1.21), this means that collateral margining (1.13), funding, and hedging strategies (1.20) are carried out continuously. As we approach the time limit, we get

$$\begin{aligned}\gamma(T\wedge\tau;C) &= \int_t^{T\wedge\tau}\mathrm{e}^{-\int_t^u r_s ds}(r_u-\widetilde{c}_u)C_u du,\\ \varphi(T\wedge\tau;F,\overline{H}) &= \int_t^{T\wedge\tau}\mathrm{e}^{-\int_t^u r_s ds}\big((r_u-\widetilde{f}_u)F_u+(\widetilde{f}_u-\widetilde{h}_u)\overline{H}_u\big)du,\end{aligned}$$

moreover, we express the discount cashflow $\Pi(t,T\wedge\tau)$ as

$$\Pi(t,T\wedge\tau) = \int_t^{T\wedge\tau}\mathrm{e}^{-\int_t^u r_s ds}\Pi_u du,$$

with $\{\Pi_t\}_t$ is the payoff coupon process of the derivative contract.

Then, putting all the above terms together the recursive pricing equation yields

$$\begin{aligned}\widetilde{V}_t := \mathbb{E}_t^{\mathcal{G}}\Bigg[\int_t^{T\wedge\tau}\mathrm{e}^{-\int_t^u r_s ds}\Big(&\Pi_u-(\widetilde{c}_u-r_u)C_u-(\widetilde{f}_u-r_u)F_u\\ &-(\widetilde{f}_u-\widetilde{h}_u)\overline{H}_u\Big)du+\mathbf{1}_{\{t\le\tau\le T\}}\mathrm{e}^{-\int_t^{\tau} r_s ds}\theta_\tau\Bigg].\end{aligned} \tag{1.22}$$

By recalling that, under rehypothecation, $F_t = \widetilde{V}_t - C_t - \overline{H}_t$, (1.22) became

$$\begin{aligned}\widetilde{V}_t := \mathbb{E}_t^{\mathcal{G}}\Bigg[\int_t^{T\wedge\tau} \mathrm{e}^{-\int_t^u r_s ds}\Big(&\Pi_u - (\widetilde{c}_u - r_u)C_u - (\widetilde{f}_u - r_u)(\widetilde{V} - C_u) \\ &-(r_u - \widetilde{h}_u)\overline{H}_u)\Big)du + \mathbf{1}_{\{t\leq\tau\leq T\}}\mathrm{e}^{-\int_t^\tau r_s ds}\theta_\tau\Bigg]\end{aligned} \tag{1.23}$$

with θ_τ in (1.23) representing the one-default cashflow , it is defined in (1.15).

According to the traditional no-arbitrage theory and in a market with no credit risk, the hedging process $\overline{H}$ would be equivalent to a δ-hedging strategy account. This implies that the portfolio would be adjusted continuously to maintain a delta-neutral position, in order to eliminate the risk of price fluctuations in the underlying asset.

Chapter 2

Backward Stochastic Differential Equation

In this chapter, we briefly present the theory of Backward Stochastic Differential Equation, in order to apply it to our specific case of Value Adjustments. This theory was born in 1973 with the seminal paper by Bismut [7] and since then it has enormously developed in several directions and we refer the interested reader to [30, 31, 42, 44, 49] for a detailed account of both theory and application. The Chapter mainly follows the Zhang's book [49].

2.1 Preliminary notions

In this section, we introduce the notations we are going to use in the rest of chapter.

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space on which is defined a dimensional standard Brownian motion $W = (W_t)_{0 \le t \le T}$, such that $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ is the natural filtration of W, and T is a fixed finite horizon, augmented by all the $\mathbb{P}$-nulls sets.

Let $p, q > 0$, we define

- $L^0(\mathcal{F})$ the space of $\mathcal{F}$-measurable random variables;
- $L^p(\mathcal{F})$ the space of $X \in L^0(\mathcal{F})$ such that $\|X\|_p^p := \mathbb{E}\big[|X|^p\big] < \infty$;
- $L^0(\mathbb{F})$ the space of $\mathbb{F}$-adapted stochastic processes $\{X\}_{t \in [0,T]}$;
- $L^{p,q}(\mathbb{F})$ the space of $\{X\}_{t \in [0,T]} \in L^0(\mathbb{F})$ such that $\|X\|_{p,q}^p := \mathbb{E}\big[\big(\int_0^T |X_t|^p dt\big)^{\frac{q}{p}}\big] < \infty$. For $p = q$ we abbreviate it by $L^p(\mathbb{F}) := L^{p,p}(\mathbb{F})$ and $\|X\|_p := \|X\|_{p,p}$ for $\{X\}_{t \in [0,T]} \in L^p(\mathbb{F})$;
- $\mathbb{S}^p(\mathbb{F})$ the set of $\{X\}_{t \in [0,T]} \in L^0(\mathbb{F})$ continuous a.s. such that $\|X\|_{\infty,p}^p := \mathbb{E}[X^{*p}] < \infty$, and we denote by $X^* = \sup_{0 \le t \le T} |X_t|$;
- $L^\infty(\mathbb{F})$ the space of bounded processes in $L^0(\mathbb{F})$, with L^∞-norm denoted by $\| \cdot \|_\infty$;
- $L^p_{loc}(\mathbb{F})$ the space of processes in $L^0(\mathbb{F})$ such that $\int_0^T |X_t|^p dt < \infty, \mathbb{P}$-a.s.

Let us recall some important inequalities:

Propotion 2.1.1. *Let p, q be conjugates.*

i) (Young's Inequality) Assume $p, q > 1$. For any $x, y \in \mathbb{R}$, it holds:

$$|xy| \leq \frac{|x|^p}{p} + \frac{|y|^q}{q}.$$

In particular, for $p = q = 2$ and for any $\alpha > 0$, we have

$$2xy \leq \alpha x^2 + \alpha^{-1} y^2 \quad (x+y)^2 \leq (1+\alpha)x^2 + (1+\alpha^{-1})y^2.$$

ii) (Hölder's Inequality) Let $\{X\}_{t\in[0,T]} \in L^p(\mathbb{F}), \{Y\}_{t\in[0,T]} \in L^q(\mathbb{F})$. Then

$$\|XY\|_1 \leq \|X\|_p \|Y\|_q.$$

iii) (Gronwall Inequality) Assume the function $a : [0,T] \to [0,\infty)$ satisfies

$$a_t \leq C_0 + C_1 \int_0^t a_s ds,$$

for some $C_0, C_1 \geq 0$. Then $a_t \leq C_0 \mathrm{e}^{C_1 t}, 0 \leq t \leq T$.

Theorem 2.1.1 (*Burkholder-Davis-Gundy*). [1] *For any $p > 0$ and $\sigma \in L^{2,p}(\mathbb{F}) \subset L^2_{loc}(\mathbb{F})$, define*

$$M_t = \int_0^t \sigma_s dW_s \ \text{ and } \ M^* = \sup_{0\leq s\leq t} |M_s|.$$

There exist universal constants $0 < c_p < C_p$, depending only on p, such that

$$c_p \mathbb{E}\left[\left(\int_0^T |\sigma_t|^2 dt\right)^{\frac{p}{2}}\right] \leq \mathbb{E}\Big[|M_T^*|^p\Big] \leq C_p \mathbb{E}\left[\int_0^T \left(|\sigma_t|^2 dt\right)^{\frac{p}{2}}\right] \tag{2.2}$$

All vectors are considered to be column vectors, namely we take the convention that $\mathbb{R}^n = \mathbb{R}^{n\times 1}$ for some dimension$n$. A n-dimensional random vector is a mapping $X = (X_1, \dots, X_n)^\top : \Omega \to \mathbb{R}^n$ such that $X_i \in L^0(\mathcal{F}), i = 1, \cdots, n$, where $\top$ stands for transpose. We say $X = (X^1, \cdots, X^n)^\top$ is a n-dimensional process if $X^1, \cdots, X^n$ are processes such that $X^i \in L^0(\mathbb{F})$. We note that all of the preceding notations can be extended to multidimensional settings; to emphasise the dimension we write, for example $L^0(\mathbb{F}, \mathbb{R}^n)$.

[1] We observe that for $p = 2$ is exactly the *Doob's maximum inequality*:

$$\mathbb{E}\big[|M_T|^2\big] \leq \mathbb{E}\big[|M_T^*|^2\big] \leq 4\mathbb{E}\big[|M_T|^2\big]. \tag{2.1}$$

2.2 Motivation

Classical Stochastic Differential Equations (SDEs) usually describe the stochastic time evolution of some phenomenon starting from a known initial condition. The evolution may have a finite horizon, say T, and the final value of the solution process is a random variable that needs to be determined. In nature and finance, it often happens to have evolutionary phenomena that depend on a given final condition rather than an initial one. These phenomena are not simply time-reversed evolutions, since randomness requires the measurability properties of the solution with respect to the underlying filtration that represents the progressive accumulation of information. This opens up a series of mathematical issues that have been addressed by the theory of Backward Stochastic Differential Equations.

To understand better the peculiarities of the topic, let us look at the following example.

We want to construct a process $\{Y_t\}_{t\in[0,T]}$ that verifies a final condition $Y_T = \xi \in L^2(\mathcal{F}_T)$.

Certainly the constant process

$$Y_t = \xi, \quad t \in [0,T] \tag{2.3}$$

verifies our requirement, but it is $\mathcal{F}_t$-adapted only if ξ is a constant. If ξ is a random variable, then $\{Y_t\}_{t\in[0,T]}$ is certainly adapted, if it is defined as the martingale

$$Y_t := \mathbb{E}[\xi|\mathcal{F}_t],$$

which verifies of course $Y_T = \xi$. If $\mathcal{F}_t$ is the natural filtration generated by the Brownian motion W, then the martingale representation theorem gives the existence of a progressively measurable stochastic process $\{Z_t\}_{t\in[0,T]} \in L^2(\mathbb{F})$ such that

$$dY_t = Z_t dW_t \quad \forall t \in [0,T], \quad Y_0 = \mathbb{E}[\xi].$$

By integrating on $[0,t]$ and $[0,T]$

$$\begin{aligned} Y_t - Y_0 &= \int_0^t Z_s dW_s \\ Y_T - Y_0 &= \int_0^T Z_s dW_s, \end{aligned}$$

by subtracting the two quantities, we have

$$Y_t = Y_T - \int_t^T Z_s dW_s.$$

Remember that $Y_T = \xi$, hence

$$Y_t = \xi - \int_t^T Z_s dW_s \quad \forall t \in [0,T] \quad \mathbb{P} - a.s. \tag{2.4}$$

We remark that the representation theorem gives the existence and uniqueness of Z, but not an explicit expression.

Definition 2.1. *Given an $\mathcal{F}_t$-measurable random variable ξ, and $f : [0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}^n$* [2] *is $\mathbb{F}$-progressively measurable function in all variables. A solution to the BSDE is a pair (Y,Z) satisfying*

$$Y_t = \xi + \int_t^T f(s,Y_s,Z_s)ds - \int_t^T Z_s dW_s, \quad Y_T=\xi, \quad 0\le t\le T, \quad \mathbb{P}-a.s. \tag{2.5}$$

2.3 Existence and uniqueness

We want to study the following BSDE

$$-dY_t = f(t,Y_t,Z_t)dt - Z_t dW_t, \quad Y_T=\xi, \quad 0\le t\le T, \quad \mathbb{P}-a.s \tag{2.6}$$

where $Y\in\mathbb{S}^2(\mathbb{F},\mathbb{R}^n), Z\in L^2(\mathbb{F},\mathbb{R}^{n\times d})$, for some dimension n,d and the pair (ξ,f).

We prove an existence and uniqueness result for the above BSDE. Wihout loss generality we shall assume $n=d=1$ in most proofs.

2.3.1 Linear BSDE

We see the following first result.

Theorem 2.3.1. *Let $\xi\in L^2(\mathcal{F}_T,\mathbb{R}^n)$ and $f:[0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^{n\times d}\to\mathbb{R}^n$ such that $f(\cdot,0,0)\in L^{1,2}(\mathbb{F},\mathbb{R}^n)$. Then the following linear BSDE has a unique solution $(Y,Z)\in L^2(\mathbb{F},\mathbb{R}^n)\times L^2(\mathbb{F},\mathbb{R}^{n\times d})$:*

$$Y_t = \xi + \int_t^T f(s,0,0)ds - \int_t^T Z_s dW_s \tag{2.7}$$

Proof. Let us set

$$Y_t := \mathbb{E}\Big[\xi + \int_t^T f(s,0,0)ds|\mathcal{F}_t\Big]. \tag{2.8}$$

By the integrability assumptions we know that

$$M_t = \mathbb{E}\Big[\xi + \int_0^T f(s,0,0)ds|\mathcal{F}_t\Big]$$

is a square integrable thus, by the martingale representation theorem, there exists unique progressively stochastic process $Z\in L^2(\mathbb{F},\mathbb{R}^{n\times d})$, such that

$$M_t = M_0 + \int_0^t Z_s dW_s.$$

[2] For simplicity we will omit the dependence on $\omega\in\Omega$

Hence

$$Y_t = M_0 - \int_0^t f(s,0,0)ds + \int_0^t Z_s dW_s$$
$$Y_T = M_0 - \int_0^T f(s,0,0)ds + \int_0^T Z_s dW_s.$$

By calculating the difference $Y_t - Y_T$, and remember that $Y_T = \xi$, we get immediately (2.7). □

Let $\alpha \in L^\infty(\mathbb{F}, \mathbb{R}^n), \beta \in L^\infty(\mathbb{F}, \mathbb{R}^{n\times d})$, we consider the linear BSDE with $n = d = 1$

$$Y_t = \xi + \int_t^T [\alpha_s Y_s + Z_s\beta_s + f(s,0,0)]ds - \int_t^T Z_s dW_s, \tag{2.9}$$

or

$$dY_t = -[\alpha_t Y_t + Z_t\beta_t + f(t,0,0)]dt + Z_t dW_t, \tag{2.10}$$

Here we provide a representation formula for its solution (2.5).

Theorem 2.3.2. *Let $\xi \in L^2(\mathcal{F}_T), \alpha, \beta \in L^\infty(\mathbb{F})$, and $f : [0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^{n\times d} \to \mathbb{R}^n$ such that $f(\cdot,0,0) \in L^{1,2}(\mathbb{F})$, and f is uniformly Lipschitz continuous in (y,z), i.e. there exists $L > 0$ such that*

$$\|f(t,y_1,z_1) - f(t,y_2,z_2)\| \leq L\big(\|y_1 - y_2\| + \|z_1 - z_2\|\big), \quad y_1, y_2 \in \mathbb{R}^n, z_1, z_2 \in \mathbb{R}^{n\times d}, \forall t \in [0,T].$$

If $(Y,Z) \in \mathbb{S}^2(\mathbb{F})\times L^2(\mathbb{F})$ satisfies the linear BSDE (2.10)*, then*

$$Y_t = \Gamma_t^{-1}\mathbb{E}\Big[\Gamma_T\xi + \int_t^T \Gamma_s f(s,0,0)ds|\mathcal{F}_t\Big], \tag{2.11}$$

where

$$\begin{aligned} d\Gamma_t =& \Gamma_t\alpha_t dt + \Gamma_t\beta_t dW_t \quad \textit{or} \\ \Gamma_t =& exp\Big\{\int_0^t \beta_s dW_s + \int_0^t \big[\alpha_s - \frac{1}{2}|\beta_s|^2 ds\Big\}. \end{aligned} \tag{2.12}$$

Proof. By applying Itô's formula

$$\begin{aligned} d(\Gamma_t Y_t) &= \Gamma_t dY_t + Y_t d\Gamma_t + d\langle Y,\Gamma\rangle_t \\ &= -\Gamma_t\big(\alpha_t Y_t + Z_t\beta_t + f(t,0,0)\big)dt + \Gamma_t Z_t dW_t + Y_t\Gamma_t\alpha_t dt + Y_t\Gamma_t\beta_t dW_t + \Gamma_t\beta_t Z_t dt \\ &= -\Gamma_t\alpha_t Y_t - \Gamma_t Z_t\beta_t - \Gamma_t f(t,0,0)dt + \Gamma_t Z_t dW_t + Y_t\Gamma_t\alpha_t dt + Y_t\Gamma_t\beta_t dW_t + \Gamma_t Z_t\beta_t dt \\ &= \big[-\Gamma_t\alpha_t Y_t + \Gamma_t\alpha_t Y_t - \Gamma_t Z_t\beta_t + \Gamma_t Z_t\beta_t - \Gamma_t f(t,0,0)\big]dt + \big[Y_t\Gamma_t\beta_t + \Gamma_t Z_t\big]dW_t \\ &= -\Gamma_t f(t,0,0)dt + \Gamma_t\big[Y_t\beta_t + Z_t\big]dW_t \end{aligned}$$

Denote

$$\widehat{Y}_t := \Gamma_t Y_t, \quad \widehat{Z}_t := \Gamma_t\big[Y_t\beta_t + Z_t\big], \quad \widehat{\xi} := \Gamma_T\xi, \quad \widehat{f}(t,0,0) := \Gamma_t f(t,0,0). \tag{2.13}$$

Then

$$d\widehat{Y}_t = -\widehat{f}(t,0,0)dt + \widehat{Z}_t dW_t. \tag{2.14}$$

We integrate over the intervals $[0,t]$ and $[0,T]$

$$\widehat{Y}_t + \int_0^t \widehat{f}(s,0,0)ds = \widehat{Y}_0 + \int_0^t \widehat{Z}_s dW_s,$$
$$\widehat{Y}_T + \int_0^T \widehat{f}(s,0,0)ds = \widehat{Y}_0 - \int_0^T \widehat{Z}_s dW_s,$$

therefore

$$\widehat{Y}_t - \widehat{Y}_T = \int_t^T \widehat{f}(s,0,0)ds + \int_t^T \widehat{Z}_s dW_s. \tag{2.15}$$

Since α and β are bounded, we see that $\mathbb{E}[\sup_{0\le t\le T}|\Gamma_t|^2] < \infty$, and by denoting by $\widehat{\beta}$ the upper-bound of β, we have

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_0^T |\widehat{Z}_s|^2 ds\Big)^{1/2}\Big] &\le \mathbb{E}\Big[\Big(\int_0^T |\Gamma_s|^2 |Y_s\beta_s + Z_s|^2 ds\Big)^{1/2}\Big] \le \mathbb{E}\Big[\Big(\int_0^T \sup_{0\le t\le T}|\Gamma_t|^2 (Y_s\beta_s + Z_s)^2 ds\Big)^{\frac{1}{2}}\Big] \\
&\le \mathbb{E}\Big[\underbrace{\sup_{0\le t\le T}|\Gamma_t|}_{A}\underbrace{\Big(\int_0^T (Y_s\beta_s + Z_s)^2 ds\Big)^{\frac{1}{2}}}_{B}\Big] \text{ (Hölder's Inequality)} \\
&\le \Big(\mathbb{E}[A^2]\Big)^{\frac{1}{2}}\Big(\mathbb{E}[B^2]\Big)^{\frac{1}{2}} = \underbrace{\Big(\mathbb{E}\Big[\sup_{0\le t\le T}|\Gamma_t|^2\Big]\Big)^{\frac{1}{2}}}_{a}\underbrace{\Big(\mathbb{E}\Big[\int_0^T (Y_s\beta_s + Z_s)^2 ds\Big]\Big)^{\frac{1}{2}}}_{b} \\
&\le \frac{1}{2}(a^2+b^2) = \frac{1}{2}\mathbb{E}\Big[\sup_{0\le t\le T}|\Gamma_t|^2 + \int_0^T (Y_s\beta_s + Z_s)^2 ds\Big] \\
&\le \frac{1}{2}\mathbb{E}\Big[\sup_t |\Gamma_t|^2 + \int_0^T |Y_t\beta_s|^2 dt + \int_0^T |Z_t|^2 dt + \int_0^T \underbrace{2(Y_s\beta_s)Z_s}_{2xy\le x^2+y^2} ds\Big] \\
&\le \frac{1}{2}\mathbb{E}\Big[\sup_t |\Gamma_t|^2 + 2\widehat{\beta}^2\int_0^T |Y_t|^2 dt + 2\int_0^T |Z_t|^2 dt\Big] < \infty.
\end{aligned}$$

This shows that the local martingale in (2.15) is a uniformly integrable martingale. By taking the its expectation, we obtain

$$\widehat{Y}_t = \mathbb{E}\Big[\widehat{\xi} + \int_t^T \widehat{f}(s,0,0)ds|\mathcal{F}_t\Big]. \tag{2.16}$$

by (2.13), and observing that Γ is invertible, this it implies (2.11) immediately. □

2.3.2 A Priori Estimates for BSDEs

Now, we look into the nonlinear BSDE (2.5), and let's start by introducing two important results that will be useful in proving the uniqueness of its solution.

Theorem 2.3.3. *Let $\xi \in L^2(\mathcal{F}_T)$, and $f : [0,T]\times\Omega\times\mathbb{R}^n\times\mathbb{R}^{n\times d} \to \mathbb{R}^n$ such that $f(\cdot,0,0) \in L^{1,2}(\mathbb{F})$, and f is uniformly Lipschitz continuous in (y,z), if $(Y,Z) \in L^2(\mathbb{F},\mathbb{R}^n)\times L^2(\mathbb{F},\mathbb{R}^{n\times d})$ is* *a solution of the BSDE* (2.5)*. Then* $Y \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}^n)$ *and there exists a constant* C*, depending only on* T, L, n, d*, such that*

$$\|(Y,Z)\|^2 := \mathbb{E}\left[|Y_T^*|^2 + \int_0^T |Z_t|^2 dt\right] \le C\mathbb{E}\left[|\xi|^2 + \left(\int_0^T |f(t,0,0)|dt\right)^2\right]. \tag{2.17}$$

In this proof and in the sequel, we shall denote by C a generic constant, which depends only on T, L, n, d, and may vary from line to line.

Proof. For simplicity, we assume $n = d = 1$. The proof is divided into several step.

Step 1. We show that

$$\mathbb{E}\big[|Y_T^*|^2\big] \le C\mathbb{E}\left[\int_0^T \big[|Y_t|^2 + |Z_t|^2\big]dt\right] + C\mathbb{E}\left[|\xi|^2 + \left(\int_0^T |f(t,0,0)|dt\right)^2\right] < \infty, \tag{2.18}$$

Since (Y, Z) is a solution of BSDE (2.5), from the triangular inequality, and the Lipshitz property of f

$$\begin{aligned} |Y_t| &\le |\xi| + \int_t^T |f(s, Y_s, Z_s)|ds + \left|\int_t^T Z_s dW_s\right| \\ &\le |\xi| + \int_t^T |f(s,0,0)| + L\big(|Y_s| + |Z_s|\big)ds + \left|\int_t^T Z_s dW_s\right|. \end{aligned}$$

Moreover,

$$Y_T^* \le C\left[|\xi| + \int_0^T \big(|f(s,0,0)| + |Y_s| + |Z_s|\big)ds + \sup_{0\le t\le T}\left|\int_0^t Z_s dW_s\right|\right].$$

Square both sides, take the expectation, and apply the Burkholder-Davis-Gundy Inequality, we have

$$\mathbb{E}\big[|Y_T^*|^2\big] \le C\mathbb{E}\left[|\xi|^2 + \Big(\int_0^T |f(t,0,0)|dt\Big)^2 + \int_0^T \big[|Y_t|^2 + |Z_t^2|\big]dt\right]$$

with (2.18) resulting immediately.

Step 2. For any $\varepsilon > 0$, we show that

$$\sup_{0\le t\le T} \mathbb{E}\big[|Y_t|^2\big] + \mathbb{E}\left[\int_0^T |Z_t|^2 dt\right] \le \varepsilon\mathbb{E}[|Y_T^*|^2] + C\varepsilon^{-1}\mathbb{E}\left[|\xi|^2 + \left(\int_0^T |f(t,0,0)|dt\right)^2\right]. \tag{2.19}$$

By applying It$\hat{o}$ formula to $|Y_t|^2$ a

$$d|Y_t|^2 = 2Y_t dY_t + d\langle Y, Y\rangle_t = -2Y_t f(t, Y_t, Z_t)dt + 2Y_t Z_t dW_t + |Z_t|^2 dt$$

By integrating on $[0, t]$ and $[0, T]$, and then taking $|Y_t|^2 - |Y_T|^2$ we have

$$|Y_t|^2 - |Y_T|^2 = \int_t^T 2Y_s f(s, Y_s, Z_s)ds - \int_t^T 2Y_s Z_s dW_s - \int_t^T |Z_s|^2 ds$$

Thus,

$$|Y_t|^2 + \int_t^T |Z_s|^2 ds = |\xi|^2 + \int_t^T 2Y_s f(s, Y_s, Z_s) ds - \int_t^T 2Y_s Z_s dW_s. \tag{2.20}$$

Note that

$$\int_0^T |Y_s Z_s|^2 ds \le 2 \int_0^T \big(|Y_s|^2 + |Z_s|^2\big) ds < \infty, \tag{2.21}$$

then $\int_0^T Y_s Z_s dW_s$ is $\mathbb{F}$-martingale.

By taking expectation on both sides (2.20)

$$\begin{aligned}
\mathbb{E}\left[|Y_t|^2 + \int_t^T |Z_s|^2 ds\right] &= \mathbb{E}\left[|\xi|^2 + 2\int_t^T Y_s f(s, Y_s, Z_s) ds\right] \\
&\le \mathbb{E}\left[|\xi|^2 + 2\int_t^T |Y_s||f(s, Y_s, Z_s)| ds\right] \\
&\le \mathbb{E}\left[|\xi|^2 + 2\int_t^T |Y_s|\Big[|f(s,0,0)| + C\big(|Y_s| + |Z_s|\big)\Big] ds\right] \\
&\le \mathbb{E}\left[|\xi|^2 + 2Y_T^* \int_0^T |f(s,0,0)| ds + C\int_t^T |Y_s|^2 ds + \int_t^T \underbrace{|Y_s||Z_s|}_{2xy \le x^2 + y^2} ds\right] \\
&\le \mathbb{E}\left[|\xi|^2 + 2Y_T^* \int_0^T |f(s,0,0)| ds + C\bigg(\int_t^T \frac{3}{2}|Y_s|^2 ds + \frac{1}{2}\int_t^T |Z_s|^2 ds\bigg)\right].
\end{aligned}$$

This leads to

$$\mathbb{E}\left[|Y_t|^2 + \frac{1}{2}\int_t^T |Z_s|^2 ds\right] \le \mathbb{E}\left[|\xi|^2 + CY_T^* \int_0^T |f(s,0,0)| ds + C\int_t^T |Y_s|^2 ds\right], \tag{2.22}$$

by splitting the expectation and applying Fubini's theorem,

$$\underbrace{\mathbb{E}\Big[|Y_t|^2\Big]}_{a_t} \le \underbrace{\mathbb{E}\left[|\xi|^2 + CY_T^* \int_0^T |f(s,0,0)| ds\right]}_{C_0} + \underbrace{C\int_t^T \mathbb{E}\Big[|Y_s|^2\Big] ds}_{C_1 \int_t^T a_s ds}$$

applying backward Gronwall inequality $\big(a_t \le C_0 \mathrm{e}^{C_1(T-t)}\big)$, we get

$$\mathbb{E}\big[|Y_t|^2\big] \le C\mathbb{E}\left[|\xi|^2 + Y_T^* \int_0^T |f(s,0,0)| ds\right], \quad \forall t \in [0,T]. \tag{2.23}$$

Then, by letting $t = 0$ and plugging (2.23) into (2.22) we have

$$\mathbb{E}\left[\int_0^T |Z_s|^2 ds\right] \le C\mathbb{E}\left[|\xi|^2 + \underbrace{Y_T^* \int_0^T |f(s,0,0)| ds}_{ab}\right]. \tag{2.24}$$

By (2.23) and (2.24) and noting that $2ab \le \varepsilon a^2 + \varepsilon^{-1} b^2$, we obtain (2.19) immediately.

Step 3. Plugging (2.19) into (2.18), we get

$$\mathbb{E}\big[|Y_T^*|^2\big] \le C\varepsilon\mathbb{E}\big[|Y_T^*|^2\big] + C\varepsilon^{-1}\mathbb{E}\left[|\xi|^2 + \Big(\int_0^T |f(t,0,0)|dt\Big)^2\right] \tag{2.25}$$

By choosing $\varepsilon = \dfrac{1}{2C}$ for the constant C above, we obtain

$$\mathbb{E}\big[|Y_T^*|^2\big] \le C\mathbb{E}\left[|\xi|^2 + \Big(\int_0^T |f(t,0,0)|dt\Big)^2\right]. \tag{2.26}$$

This, together with (2.19) proves (2.17). □

Theorem 2.3.4. *For $i = 1, 2$, assume (ξ^i, f^i) satisfy the assumptions of Theorem 2.3.3 and $(Y^i, Z^i) \in L^2(\mathbb{F}, \mathbb{R}^n) \times L^2(\mathbb{F}, \mathbb{R}^{n\times d})$ are solutions to BSDE* (2.5) *respectively with coefficients (ξ^i, f^i). Then*

$$\|(\Delta Y, \Delta Z)\|^2 \le C\mathbb{E}\left[|\Delta\xi|^2 + \Big(\int_0^T |\Delta f(t, Y_t^1, Z_t^1)|dt\Big)^2\right], \tag{2.27}$$

where

$$\Delta Y := Y^1 - Y^2, \quad \Delta Z := Z^1 - Z^2, \quad \Delta\xi := \xi^1 - \xi^2, \quad \Delta f := f^1 - f^2.$$

Proof. Without loss generality we assume $n = d = 1$. Note that

$$\begin{aligned}\Delta Y_t &= Y_t^1 - Y_t^2 \\ &= \Delta\xi + \int_t^T \big[f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)\big]ds - \int_t^T \Delta Z_s dW_s.\end{aligned}$$

Adding and subtracting $f^2(s, Y_s^1, Z_s^1)$ and $f^2(s, Y_s^2, Z_s^1)$

$$\begin{aligned}= \Delta\xi + \int_t^T \big[&\underbrace{f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^1, Z_s^1)}_{\Delta f(s, Y_s^1, Z_s^1)} + f^2(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^1) \\ &+ f^2(s, Y_s^2, Z_s^1) - f^2(s, Y_s^2, Z_s^2)\big]ds - \int_t^T \Delta Z_s dW_s.\end{aligned}$$

Hence

$$\Delta Y_t = \underbrace{\Delta\xi}_{:=\xi} + \int_t^T \underbrace{\big[\Delta f(s, Y_s^1, Z_s^1) + \alpha_s \Delta Y_s + \beta_s \Delta Z_s\big]}_{:=f(s, Y_s, Z_s)} ds - \int_t^T \underbrace{\Delta Z_s}_{:=Z_s} dW_s, \tag{2.28}$$

where

$$\alpha_t := \frac{f^2(t, Y_t^1, Z_t^1) - f^2(t, Y_t^2, Z_t^1)}{\Delta Y_t} 1_{\{\Delta Y_t \neq 0\}}$$
$$\beta_t := \frac{f^2(t, Y_t^2, Z_t^1) - f^2(t, Y_t^2, Z_t^2)}{\Delta Z_t} 1_{\{\Delta Z_t \neq 0\}}$$

are bounded by liptschiz property of f^i, for $i = 1, 2$. Then, we may view $(\Delta Y, \Delta Z) \in L^2(\mathbb{F}) \times L^2(\mathbb{F})$ as the solution to the BSDE (2.28), given that the coefficients satisfy the assumption required to apply Theorem 2.3.3, and we obtain the result immediately. $\square$

2.3.3 Well-Posedness of BSDEs

We establish the well-posedness of BSDE (2.5).

Theorem 2.3.5. *Under the assumptions of Theorem 2.3.3, BSDE* (2.5) *has a unique solution* $(Y, Z) \in L^2(\mathbb{F}, \mathbb{R}^n) \times L^2(\mathbb{F}, \mathbb{R}^{n\times d})$.

Proof. Uniqueness follows directly from theorem 2.3.4. We now prove the existence by using the Picard iteration. For simplicity we assume $n = d = 1$, since the proof is analogous in the multidimensional case.

Step 1.

Let $\delta > 0$ and $0 < T \leq \delta$ be constants.

Let us set $Y_t^0 := 0, \ \ Z_t^0 := 0$. For $k = 1, 2, \ldots$, let us define (Y^k, Z^k) as the unique solution of the following equation

$$Y_t^k = \xi + \int_t^T f(s, Y_s^{k-1}, Z_s^{k-1})ds - \int_t^T Z_s^k dW_s \tag{2.29}$$

with $(Y^{k-1}, Z^{k-1}) \in L^2(\mathbb{F}) \times L^2(\mathbb{F})$.

Indeed, by the linear growth condition it follows that $f(t, Y_t^{k-1}, Z_t^{k-1}) \in L^{1,2}(\mathbb{F})$, and Theorem 2.3.1 and 2.3.3 imply that (2.29) has a unique solution $(Y^k, Z^k) \in \mathbb{S}^2(\mathbb{F}) \times L^2(\mathbb{F})$. By induction we have $(Y^k, Z^k) \in \mathbb{S}^2(\mathbb{F}) \times L^2(\mathbb{F})$ for all $k \geq 0$.

Denote by $\Delta Y_t^k := Y_t^k - Y_t^{k-1}, \quad \Delta Z_t^k := Z_t^k - Z_t^{k-1}$. Then,

$$\Delta Y_t^k = \int_t^T \left[\alpha_s^{k-1}\Delta Y_s^{k-1} + \beta_s^{k-1}\Delta Z_s^{k-1}\right] - \int_t^T \Delta Z_s^k dW_s$$

where

$$\alpha_t^{k-1} = \frac{f(s, Y_s^{k-1}, Z_s^{k-1}) - f(s, Y_s^{k-2}, Z_s^{k-1})}{\Delta Y_t^{k-1}} 1_{\{\Delta Y_t^{k-1} \neq 0\}},$$
$$\beta_t^{k-1} = \frac{f(s, Y_s^{k-2}, Z_s^{k-1}) - f(s, Y_s^{k-2}, Z_s^{k-2})}{\Delta Z_t^{k-1}} 1_{\{\Delta Z_t^{k-1} \neq 0\}}$$

are bounded due to the Lipschitz property of f. Applying Itô's formula, we have

$$d(|\Delta Y_t^k|^2) = -2\Delta Y_t^k\left[\alpha_t^{k-1}\Delta Y_t^{k-1} + \beta_t^{k-1}\Delta Z_t^{k-1}\right]dt + 2\Delta Y_t^k \Delta Z_t^k dW_t + |\Delta Z_t^k|^2 dt.$$

Note that, as shown in (2.21), $\int_0^t \Delta Y_s^{k-1}\Delta Z_s^{k-1} dW_s$ is a $\mathbb{F}$-martingale. Nothing that $\Delta Y_T^k = 0$, and by integrating over the intervals $[0, t]$ and $[0, T]$, subtracting the two quantities from each other, we get

$$|\Delta Y_t^k|^2 + \int_t^T |\Delta Z_s^k|^2 ds = 2\int_t^T \Delta Y_s^k \big(\alpha_s^{k-1}\Delta Y_s^{k-1} + \beta_s^{k-1}\Delta Z_s^{k-1}\big) ds - 2\int_t^T \Delta Y_s^k \Delta Z_s^k dW_s$$

Taking expectations, we have

$$\begin{aligned} \mathbb{E}\bigg[|\Delta Y_t^k|^2 + \int_t^T |\Delta Z_s^k|^2 ds\bigg] &= \mathbb{E}\bigg[2\int_t^T \Delta Y_s^k \big(\alpha_s^{k-1}\Delta Y_s^{k-1} + \beta_s^{k-1}\Delta Z_s^{k-1}\big) ds\bigg] \\ &\leq C\mathbb{E}\bigg[\int_0^T |\Delta Y_s^k|\big(|\Delta Y_s^{k-1}| + |\Delta Z_s^{k-1}|\big) ds\bigg]. \end{aligned} \tag{2.30}$$

In particular

$$\mathbb{E}\bigg[|\Delta Y_t^k|^2\bigg] \leq C\mathbb{E}\bigg[\int_0^T |\Delta Y_s^k|\big(|\Delta Y_s^{k-1}| + |\Delta Z_s^{k-1}|\big) ds\bigg].$$

By integrating over $[0,T]$ and applying the Fubini's theorem, we have

$$\mathbb{E}\bigg[\int_0^T |\Delta Y_s^k|^2 ds\bigg] \leq \int_0^T \bigg(C\mathbb{E}\Big[\int_0^T |\Delta Y_s^k|\big(|\Delta Y_s^{k-1}| + |\Delta Z_s^{k-1}|\big) ds\Big]\bigg) dt,$$

remembering that $T \leq \delta$

$$\begin{aligned} \mathbb{E}\bigg[\int_0^T |\Delta Y_t^k|^2 ds\bigg] &\leq C\delta\mathbb{E}\bigg[\int_0^T \underbrace{|\Delta Y_s^k|\big(|\Delta Y_s^{k-1}| + |\Delta Z_s^{k-1}|\big)}_{2xy \leq x^2+y^2} ds\bigg] \\ &\leq C\delta\mathbb{E}\bigg[\int_0^T |\Delta Y_s^k|^2 + (|\Delta Y_s^{k-1}| + |\Delta Z_s^{k-1}|)^2 ds\bigg] \\ &= C\delta\mathbb{E}\bigg[\int_0^T \big(|\Delta Y_s^k|^2 + |\Delta Y_s^{k-1}|^2 + |\Delta Z_s^{k-1}|^2 + \underbrace{2|\Delta Y_s^{k-1}||\Delta Z_s^{k-1}|}_{2xy \leq x^2+y^2}\big) ds\bigg] \\ &\leq C\delta\mathbb{E}\bigg[\int_0^T \big(|\Delta Y_s^k|^2 + |\Delta Y_s^{k-1}|^2 + |\Delta Z_s^{k-1}|^2\big) ds\bigg]. \end{aligned}$$

Assume $\delta < \dfrac{1}{2C}$, thus $1 - C\delta \leq 1/2$. Then,

$$\mathbb{E}\bigg[\int_0^T |\Delta Y_t^k|^2 ds\bigg] \leq C\delta\mathbb{E}\bigg[\int_0^T \big(|\Delta Y_s^{k-1}|^2 + |\Delta Z_s^{k-1}|^2\big) ds\bigg].$$

Moreover, setting $t = 0$ in (2.30) we have

$$\mathbb{E}\Big[\int_0^T |\Delta Z_t^k|^2 dt\Big] \leq C\mathbb{E}\Big[\int_0^T |\Delta Y_t^k|^2 dt\Big] + \frac{1}{8}\mathbb{E}\Big[\int_0^T \big[|\Delta Y_t^{k-1}|^2 + |\Delta Z_t^{k-1}|^2\big] dt\Big]$$
$$\leq \Big[C\delta + \frac{1}{8}\Big]\mathbb{E}\Big[\int_0^T \big[|\Delta Y_t^{k-1}|^2 + |\Delta Z_t^{k-1}|^2\big] dt\Big].$$

Set $\delta := \frac{1}{8C}$ for the above C. Then

$$\mathbb{E}\Big[\int_0^T |\Delta Y_t^k|^2 + |\Delta Z_t^k|^2 dt\Big] \leq \frac{1}{4}\mathbb{E}\Big[\int_0^T |\Delta Y_t^{k-1}|^2 + |\Delta Z_t^{k-1}|^2 dt\Big].$$

By induction we have

$$\mathbb{E}\Big[\int_0^T |\Delta Y_t^k|^2 + |\Delta Z_t^k|^2 dt\Big] \leq \frac{C}{4^k}, \quad \forall r \geq 1.$$

This implies that the pair (Y^k, Z^k) is a Cauchy sequence, so it converges in $\mathbb{S}^2(\mathbb{F}) \times L^2(\mathbb{F})$. Therefore, we can conclude that there exists $(Y, Z) \in \mathbb{S}^2(\mathbb{F}) \times L^2(\mathbb{F})$ such that

$$\lim_{k\to\infty} \|(Y_t^k - Y_t, Z_t^k - Z_t)\| = 0.$$

By letting $k \to \infty$ in BSDE (2.29) we know that (Y, Z) satisfies BSDE (2.5).

Step 2.

We now prove the existence for arbitrary T. Let $\delta > 0$ be the constant in *Step 1*. Consider a partition $0 = t_0 < \cdots < t_m = T$ such that $t_{i+1} - t_i \leq \delta$, $i = 0, 1, \ldots, m-1$. Define $Y_{t_m} := \xi$, and for $i = m-1, \ldots, 0$ and $t \in [t_i, t_{i+1})$, let (Y_t, Z_t) be the solution to the following BSDE on $[t_i, t_{i+1}]$:

$$Y_t = Y_{t_{i+1}} + \int_t^{t_{i+1}} f(s, Y_s, Z_s) ds - \int_t^{t_{i+1}} Z_s dW_s, \quad t \in [t_i, t_{i+1}].$$

Since $t_{i+1} - t_i \leq \delta$, by *Step 1* the above BSDE is well posed. Moreover, we see that $(Y, Z) \in L^2(\mathbb{F}) \times L^2(\mathbb{F})$, and thus they are a global solution on the whole interval $[0, T]$. □

Theorem 2.3.6 (***Comparison theorem***). *Assume, for $i = 1, 2$, (ξ^i, f^i) satisfies assumption of Theorem 2.3.3 and $(Y^i, Z^i) \in \mathbb{S}^2(\mathbb{F}) \times L^2(\mathbb{F})$ are the unique solution to the following BSDE:*

$$Y_t^i = \xi^i + \int_t^T f^i(s, Y_s^i, Z_s^i) ds - \int_t^T Z_s^i dW_s \quad i = 1, 2. \tag{2.31}$$

Assume further that $\xi^1 \leq \xi^2$, $\mathbb{P}-$a.s., and $f^1(t, y, z) \leq f^2(t, y, z), dt \times d\mathbb{P}$-a.s. Then,

$$Y_t^1 \leq Y_t^2, \quad 0 \leq t \leq T, \quad \mathbb{P} - a.s. \tag{2.32}$$

Proof. Let us denote by

$$\Delta Y_t := Y_t^1 - Y_t^2; \quad \Delta Z_t := Z_t^1 - Z_t^2; \quad \Delta\xi := \xi^1 - \xi^2; \quad \Delta f := f^1 - f^2.$$

Then,

$$\begin{aligned}\Delta Y_t =& \Delta\xi + \int_t^T \left[f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)\right] ds - \int_t^T \Delta Z_s dW_s \\ =& \Delta\xi + \int_t^T \left[\alpha_s \Delta Y_s + \Delta Z_s \beta_s + \Delta f(s, Y_s^2, Z_s^2)\right] ds - \int_t^T \Delta Z_s dW_s,\end{aligned}$$

where α and β are defined in a similar way as (2.28), and they are bounded due to the Lipschitz property of f. Let us set Γ_s as (2.12), and by (2.11) we have

$$\Delta Y_t = \Gamma_t^{-1} \mathbb{E}\left[\Gamma_T \Delta\xi + \int_t^T \Gamma_s \Delta f(s, Y_s^2, Z_s^2) ds \middle| \mathcal{F}_t\right], \tag{2.33}$$

moreover

$$f^1(t, y, z) \leq f^2(t, y, z), \quad \forall (y, z), \quad dt \times d\mathbb{P} - a.s.,$$

which implies that $\Delta f(, Y^2, Z^2) \leq 0, dt \times d\mathbb{P}$-a.s. Since $\Gamma \geq 0$ and $\Delta\xi \leq 0$, then (2.32) follows from (2.33) immediately. □

2.4 Markov BSDEs and PDEs

We introduce the flow notation by indexing the family of the solutions of a given SDE on the basis of the initial time and of the starting point $\{X_s^{t,x}, t \leq s \leq T, X_t^t = x\}$.

Definition 2.2. *Let $b : [0, T] \times \mathbb{R}^m \to \mathbb{R}^m$ be a continuous and monotone in x, uniformly with respect to t, and $\sigma : [0, T] \times \mathbb{R}^m \to \mathbb{R}^{m \times d}$ be continuous and globally Lipschitz in x uniformly with respect to t. Let $\{X_s^{t,x}, t \leq s \leq T\}$ be the solution of the Forward SDE*

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) dr + \int_t^s \sigma(r, X_r^{t,x}) dW_r$$

and, let $(Y^{t,x}, Z^{t,x})$ be the unique solution

$$Y_s^{t,x} = g(X_T^{t,x}) - \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr + \int_s^T Z_r^{t,x} dW_r \tag{2.34}$$

where $g : \mathbb{R}^m \to \mathbb{R}^n$ and $f : [0, T] \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n \times d} \to \mathbb{R}^n$ are uniformly Lipschitz continuous in (x, y, z) with constant L. We define the following decoupled Forward-Backward SDE (FBSDE) with deterministic coefficients on $[t, T]$:

$$\begin{cases} dX_t^{t,x} &= b(t, X_t^{t,x}) dt + \sigma(t, X_t^{t,x}) dW_t \\ dY_t^{t,x} &= -f(t, X_t^{t,x}, Y_t^{t,x}, Z_t^{t,x}) dt + Z_t^{t,x} dW_t. \end{cases} \tag{2.35}$$

Assume, also that $b(\cdot,0), \sigma(\cdot,0), f(\cdot,0,0,0)$ and $g(0)$ are bounded, and b,σ,f,g are uniformly Hölder-$\frac{1}{2}$ [3] continuous in t with constant H.

Let us set

$$u(t,x) := Y_t^{t,x}.$$

Then $u(t,x)$ is bth $\mathcal{F}_t$-measurable and independent of $\mathcal{F}_t$, and thus is deterministic. Moreover, we have

$$Y_t^{t,X_t^{t,x}} = u(t,X_t^{t,x}), \quad 0 \le t \le T.$$

We shall denote by

$$\mathcal{L} := \sum_{i=1}^m b_i(t,x)\partial_{x_i} + \frac{1}{2}\sum_{i,j=1}^m \sigma_i(t,x)\sigma_j(t,x)\partial^2_{x_i x_j},$$

the infinitesimal generator of the Markov process $\{X_s^{t,x} : t \le s \le T\}$.

2.4.1 Nonlinear Feynman-Kac Formula

We now derive the PDE which the above function u should satisfy.

We consider the following system of backward semilinear parabolic PDEs

$$\begin{cases} \partial_t u(t.x) + \mathcal{L}u(t,x) + f(t,x,u(t,x),\nabla u(t,x)\sigma(t,x)) = 0, & (t,x) \in [0,T]\times\mathbb{R}^m, \\ u(T,x) = g(x), & x \in \mathbb{R}^m, \end{cases} \tag{2.36}$$

where the ∇ is gradient operator.

Theorem 2.4.1. ***(Nonlinear Feyman-Kac Formula)*** *Let $u \in C^{1,2}([0,T]\times\mathbb{R}^m,\mathbb{R}^n)$ be a classical solution of* (2.36)*. Then, for each $(t,x) \in [0,T]\times\mathbb{R}^m$, $\{(u(s,X_s^{t,x}), \nabla u(s,X_s^{t,x})\sigma(s,X_s^{t,x}) : t \le s \le T\}$ is the solution of the FBSDE* (2.35)*. In particular,*

$$u(t,x) = Y_t^{t,x}, \textit{and} \quad \nabla u(s,x)\sigma(s,x) = Z_t^{t,x} \tag{2.37}$$

Proof. For the sake of exposition, we consider $d = m = n = 1$.

Applying Itô's formula to $u(t,X_t^{t,x})$, we have

$$\begin{aligned} du(t,X_t^{t,x}) &= \partial_t u(t,X_t^{t,x})dt + \partial_x u(t,X_t^{t,x})dX_t^{t,x} + \frac{1}{2}\partial^2_{xx}u(t,X_t^{t,x})d\langle X_t^{t,x},X_t^{t,x}\rangle \\ &= \partial_t u(t,X_t^{t,x})dt + \partial_x u(t,X_t^{t,x})\big[b(t,X_t^{t,x})dt + \sigma(t,X_t^{t,x})dW_t\big] \\ &+ \frac{1}{2}\partial_x u(t,X_t^{t,x})^2\sigma(t,X_t^{t,x})^2dt \\ &= \big[\partial_t u(t,X_t^{t,x}) + \partial_x u(t,X_t^{t,x})b(t,X_t^{t,x}) + \frac{1}{2}\partial_x u(t,X_t^{t,x})^2\sigma(t,X_t^{t,x})^2\big]dt \\ &+ \partial_x u(t,X_t^{t,x})\sigma(t,X_t^{t,x})dW_t. \end{aligned} \tag{2.38}$$

[3] The Holder continuity is mainly for the regularity, not for the well-posedness.

Compare thus with

$$dY_t^{t,x} = -f(t, X_t^{t,x}, Y_t^{t,x}, Z_t^{t,x})dt + Z_t^{t,x}dW_t$$

we obtain

$$\begin{aligned}
&\big[\partial_t u(t, X_t^{t,x}) + \partial_x u(t, X_t^{t,x}) b(t, X_t^{t,x}) + \frac{1}{2}\partial_x u(t, X_t^{t,x})^2 \sigma(t, X_t^{t,x})^2\big]dt + \partial_x u(t, X_t^{t,x})\sigma(t, X_t^{t,x})dW_t \\
&= - f(t, X_t^{t,x}, Y_t^{t,x}, Z_t^{t,x})dt + Z_t^{t,x}dW_t
\end{aligned} \tag{2.39}$$

$$\begin{aligned}
&\Rightarrow \partial_t u(t, X_t^{t,x}) + \partial_x u(t, X_t^{t,x}) b(t, X_t^{t,x}) + \frac{1}{2}\partial_x u(t, X_t^{t,x})^2 \sigma(t, X_t^{t,x})^2\big] = -f(t, X_t^{t,x}, Y_t^{t,x}, Z_t^{t,x}) \\
&\quad \partial_x u(t, X_t^{t,x})\sigma(t, X_t^{t,x}) = Z_t^{t,x}
\end{aligned} \tag{2.40}$$

This concludes the proof. □

Chapter 3

Non-linear approximated value adjustments for derivatives under multiple risk factor

The following chapters are summarized in our two papers [1, 32].

3.1 Evaluation of European claims under the intensity approach

We will be discussing the intensity approach for evaluating a contingent claim based on multiple risk factors, which it was previously introduced in Chapter 1. Firstly, we present our market model.

Let $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ be the same complete probability space and $[0, T]$ the finite time interval of Chapter 1, and let us denote by $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ the market filtration generated by the adapted process S_t, representing the asset price, called *underlying*, with the following dynamics under the risk-neutral measure $\mathbb{P}$:

$$dS_t = r_t S_t dt + \sigma_t S_t dW_t, \tag{3.1}$$

where W_t is a Brownian motion, and σ_t represents the underlying volatility and r_t the risk-free interest rate process. We assume that both are deterministic, bounded functions of time.

We assume the so called H-hypothesis (see for details [5, 33, 39, 43])

(H) every $\mathcal{F}_t$-martingale remains a $\mathcal{G}_t$-martingale.

We consider a European claim with maturity T and payoff[1] $\Phi(S_T)$, where Φ is a function as regular as needed, not necessarily non-negative.

Remember that $\tau = min(\tau_C, \tau_I)$ is a stopping time with respect to the enlarged filtration $\mathcal{G}$, but in general, not necessarily with respect to the market filtration $\mathcal{F}$. We assume that τ_I, τ_C are the first jump times of two Cox processes with stochastic $\mathcal{F}$-predictable positive intensities λ^I, λ^C (see [5, 16, 39]). More precisely, let us take two $\mathcal{F}-$adapted, right-continuous, increasing processes Γ^i

[1] Payoff refers to the total amount of money received or paid out as a result of an investment or financial transaction. The term can refer to the profits or losses that an investor incurs from holding or selling a particular asset, such as a stock or option.

defined on our space, for $i = C, I$, called $\mathcal{F}-$hazard processes, such that $\Gamma_0^i = 0$, and

$$\Gamma_t^i = \int_0^t \lambda_u^i du, \quad i = C, I \quad \forall t > 0.$$

We assume that the probability space is sufficiently rich to support two random variables ξ_i, which are uniformly distributed on the interval $[0, 1]$ and independent of the filtration $\mathcal{F}$ under $\mathbb{P}$. Then

$$\tau_i = \inf\{t \geq 0 : \mathrm{e}^{-\Gamma_t^i} \leq \xi_i\} = \inf\{t \geq 0 : \Gamma_t^i \geq -ln(\xi_i)\}.$$

Consequently, we have that the $\mathcal{F}$-adapted increasing processes $F_t^i = \mathbb{P}(\tau_i \leq t|\mathcal{F}_t)$ have the representation

$$F_t^i = 1 - \mathrm{e}^{-\Gamma_t^i} = 1 - \mathrm{e}^{-\int_0^t \lambda_u^i du}.$$

The default times, as defined, are conditionally independent[2] with respect to $\mathcal{F}$, that is

$$\mathbb{P}[\tau_C > t_1, \tau_I > t_2|\mathcal{F}_t] = \mathbb{P}[\tau_C > t_1|\mathcal{F}_t]\mathbb{P}[\tau_I > t_2|\mathcal{F}_t], \quad \forall t_1, t_2 \in [0, t],$$

so that the probability of simultaneous default is 0.

As a consequence, the conditional distribution of the first-to-default time τ has the representation

$$\mathbb{P}[\tau > t|\mathcal{F}_t] = \mathrm{e}^{-\int_0^t \lambda_u du}, \quad \lambda = \lambda^C + \lambda^I,$$

and we denote $\overline{\tau} = min(\tau, T)$.

By following Chapter 1, the $\mathcal{G}_t$-adapted value process of a defaultable derivative $\widetilde{V}_t$ is given by the sum of the discounted default-free price and the adjustments due to default, funding, and collateralization risks, and it is characterized as the solution of the following BSDE

$$\begin{aligned}
\widetilde{V}_t = {} & \mathbb{E}_t^{\mathcal{G}}\Bigg[\underbrace{\mathbf{1}_{\{\tau>T\}}\mathrm{e}^{-\int_t^T r_s ds}\Phi(S_T) + \int_t^{\overline{\tau}} \mathrm{e}^{-\int_t^u r_s ds}\pi_u \, du}_{\text{Contractual cashflows}}\Bigg] - \mathbb{E}_t^{\mathcal{G}}\Bigg[\underbrace{\int_t^{\overline{\tau}} \mathrm{e}^{-\int_t^u r_s ds}\big(c_u - r_u\big)C_u du}_{\text{Cost of carry of collateral account}}\Bigg] \\
& - \mathbb{E}_t^{\mathcal{G}}\Bigg[\underbrace{\int_t^{\overline{\tau}} \mathrm{e}^{-\int_t^u r_s ds}(\overline{f_u} - r_u)(\widetilde{V}_u - C_u)du}_{\text{Costs due to funding account}}\Bigg] - \mathbb{E}_t^{\mathcal{G}}\Bigg[\underbrace{\int_t^{\overline{\tau}} \mathrm{e}^{-\int_t^u r_s ds}(r_u - \overline{h}_u)\overline{H}_u du}_{\text{Costs due to hedging}}\Bigg] \\
& + \mathbb{E}_t^{\mathcal{G}}\Bigg[\mathrm{e}^{-\int_t^{\tau} r_s ds}\mathbf{1}_{\{t\leq\tau\leq T\}}\underbrace{\Big(\varepsilon_\tau - \mathbf{1}_{\{\tau_C<\tau_I\}}LGD_C(\varepsilon_\tau - C_\tau)^+ + \mathbf{1}_{\{\tau_I<\tau_C\}}LGD_I(\varepsilon_\tau - C_\tau)^-\Big)}_{\text{On-default cashflows due to contract}}\Bigg].
\end{aligned} \tag{3.2}$$

[2] It is worth noting that the independence assumption certainly simplifies computations, but it does not consider default contagion effects. Within the intensity framework, more realistic models allowing default dependence were recently proposed (see [8, 9] and the references there in), and we remark that we could extend the method to the correlated case, provided we introduce an additional parameter.

We note the (3.2) corresponds to the theoretical equation (1.22) which we recall below

$$\widetilde{V}_t := \mathbb{E}_t^{\mathcal{G}}\Bigg[\int_t^{T\wedge\tau} \mathrm{e}^{-\int_t^u r_s ds}\Big(\Pi_u - (\widetilde{c}_u - r_u)C_u - (\widetilde{f}_u - r_u)(\widetilde{V} - C_u) - (r_u - \widetilde{h}_u)\overline{H}_u)\Big)du + \mathbf{1}_{\{t\le\tau\le T\}}\mathrm{e}^{-\int_t^{\tau} r_s ds}\theta_\tau\Bigg]$$

where we assumed that Π_t in (3.2) was due to a payoff and, possibly, to a dividend[3] process, $c := \widetilde{c}$, $\overline{f} := \widetilde{f}$, $\overline{h} := \widetilde{h}$ and $\varepsilon := \varepsilon_{C,\tau} = \varepsilon_{I,\tau}$. In Table 3.1, we highlight the measurability properties of the factors involved, which include those predefined by the contract agreement as well as those that depend on the price evolution. We recall that the close-out value ε_u is usually taken as the default-free price or as the adjusted price of the defaultable claim: the first choice gives a solvable linear BSDE, while the second ($\varepsilon_u = \widetilde{V_u}$) determines a non-linear BSDE, not explicitly solvable. We examine this last case.

It is to be noted that the default times are not market observable, thus the theoretical price represented by (3.2) must be projected on to the market filtration $\mathcal{F}_t$.

To do so, we employ the Key Lemma and its extensions, (see Section 3.1 of [5]).

Lemma 3.1.1 (***Key***)**.** *For any $\mathcal{G}$-measurable random variable X and $t > 0$, we have*

$$\mathbb{E}_t^{\mathcal{G}}\big[\mathbf{1}_{\{\tau>t\}}X\big] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{G}}\big[X\big] = \mathbf{1}_{\{\tau>t\}}\frac{\mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}X\big]}{\mathbb{P}(\tau > t\}|\mathcal{F}_t)}. \tag{3.3}$$

In particular, for any $t \le s$

$$\mathbb{P}(t < \tau \le s|\mathcal{G}_t) = \mathbf{1}_{\{\tau>t\}}\frac{\mathbb{P}(t < \tau \le s|\mathcal{F}_t)}{\mathbb{P}(\tau > t|\mathcal{F}_t)},$$

Symbol	Role	Assumption
$\Phi()$	Payoff at maturity	Lipschitz function of S_T
π	Contract dividends	$\mathcal{F}$-predictable
C	Collateral process	$\mathcal{F}$-predictable
$\overline{H}$	Hedging process	$\mathcal{G}$-predictable
ε	Close-out value	$\mathcal{F}$-predictable
c	Collateral rate	$\mathcal{F}$-predictable
$\overline{f}$	Funding rate	$\mathcal{G}$-predictable
$\overline{h}$	Hedging rate	$\mathcal{G}$-predictable
$LGD_i, i = C, I$	Loss Given Default	Constant

Table 3.1: Summary of cashflows and their measurability properties

[3]A dividend is a distribution of a portion of a company's earnings to its shareholders. Typically, dividends are paid out in cash, but they can also be distributed in the form of additional shares of stock or other assets. Dividends are usually paid out on a regular basis, such as quarterly or annually, and are often seen as a way for companies to share their profits with their investors.

and we have that for any $\mathcal{G}_t$-measurable random variable Y there exists an $\mathcal{F}_t$-measurable random variable Z such that $\mathbf{1}_{\{\tau>t\}}Y = \mathbf{1}_{\{\tau>t\}}Z$.

Proof. Multiplying both sides of (3.3) by $\mathbb{P}(\tau > t|\mathcal{F}_t)$, we need to verify that for any $A \in \mathcal{G}_t$ we have

$$\int_A \mathbf{1}_{\{\tau>t\}} X\mathbb{P}(\tau > t|\mathcal{F}_t)d\mathbb{P} = \int_A \mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}X\big]d\mathbb{P}.$$

By Lemma 3.1.1[4] of [5], for any $A \in \mathcal{G}_t$ we have $A \cap \{\tau > t\} = B \cap \{\tau > t\}$ for some $B \in \mathcal{F}_t$, and so

$$\begin{aligned}\int_A \mathbf{1}_{\{\tau>t\}} X\mathbb{P}(\tau > t|\mathcal{F}_t)d\mathbb{P} &= \int_{A\cap\{\tau>t\}} X\mathbb{P}(\tau > t|\mathcal{F}_t)d\mathbb{P} = \int_{B\cap\{\tau>t\}} X\mathbb{P}(\tau > t|\mathcal{F}_t)d\mathbb{P}\\ &= \int_B \mathbf{1}_{\{\tau>t\}} X\mathbb{P}(\tau > t|\mathcal{F}_t)d\mathbb{P} = \int_B \mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}X\big]\mathbb{P}(\tau > t|\mathcal{F}_t)d\mathbb{P}\\ &= \int_B \mathbb{E}_t^{\mathcal{F}}\Big[\mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}X\big]\Big]d\mathbb{P} = \int_{B\cap\{\tau>t\}} \mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}X\big]d\mathbb{P}\\ &= \int_{A\cap\{\tau>t\}} \mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}X\big]d\mathbb{P} = \int_A \mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}X\big]d\mathbb{P}\end{aligned}$$

This ends the proof. □

Now, we illustrate how we can apply the previous Lemma 3.1.1 with a stochastic process. Following the proofs in Appendix B of [16] and Lemma 3.8.1 in [39], we have two particular case. In Lemma 3.1.2, we deal with the case of a change of filtration for a stopped integral, and in Lemma 3.1.3, we deal with the case of a process valued at a default time.

Lemma 3.1.2. *If φ_u is an integrable $\mathcal{G}$-adapted process, then*

$$\mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{G}}\left[\int_t^{\overline{\tau}} \varphi_u du\right] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{F}}\left[\int_t^T \mathrm{e}^{-\int_t^u \lambda_s ds}\overline{\varphi}_u du\right], \quad .$$

where $\overline{\varphi}_u$ is an $\mathcal{F}_u$-measurable variable such that $\mathbf{1}_{\{\tau>u\}}\overline{\varphi} = \mathbf{1}_{\{\tau>u\}}\varphi$.

Proof. We note that

$$\begin{aligned}\mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{G}}\left[\int_t^{\overline{\tau}} \varphi_u du\right] =& \mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{G}}\left[\int_t^{\tau\wedge T} \varphi_u du\right] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{G}}\left[\int_t^T \mathbf{1}_{\{\tau>u\}}\varphi_u du\right]\\ =& \mathbb{E}_t^{\mathcal{G}}\left[\int_t^T \mathbf{1}_{\{\tau>t\}}\mathbf{1}_{\{\tau>u\}}\varphi_u du\right] = \int_t^T \mathbb{E}_t^{\mathcal{G}}\Big[\mathbf{1}_{\{\tau>t\}}\mathbf{1}_{\{\tau>u\}}\varphi_u\Big]du.\end{aligned}$$

By applying the Lemma 3.1.1 we have

$$\int_t^T \mathbf{1}_{\{\tau>t\}}\frac{\mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}\mathbf{1}_{\{\tau>u\}}\varphi_u\big]}{\mathbb{E}_t^{\mathcal{F}}\big[\mathbf{1}_{\{\tau>t\}}\big]}du.$$

[4] Assume that the filtration $\mathbb{G} = (\mathcal{G}_t)_t$ satisfies $\mathcal{G}_t = \mathcal{H}_t \wedge \mathcal{F}_t$, with $\mathcal{H}_t = \sigma(\mathbf{1}_{\{\tau\leq u\}} : u \leq t)$. Then $\mathbb{G}^* \subseteq \mathbb{G}$, where $\mathbb{G}^* = (\mathcal{G}_t^*)_{t\geq 0}$, with $\mathcal{G}_t^* := \{A \in \mathcal{G} : \exists B \in \mathcal{F}_t, A \cap \{\tau > t\} = B \cap \{\tau > t\}\}$.

Remember that $\mathbb{E}_t^{\mathcal{F}}\left[\mathbf{1}_{\{\tau>t\}}\right] = \mathbb{P}(\tau > t|\mathcal{F}_t) = \mathrm{e}^{-\int_0^t \lambda_u du}$, we get

$$\mathbf{1}_{\{\tau>t\}} \int_t^T \mathbb{E}_t^{\mathcal{F}}\left[\mathbf{1}_{\{\tau>u\}}\varphi_u\right]\mathrm{e}^{\int_0^t \lambda_s ds}du.$$

We choose an $\mathcal{F}_u$-measurable variable such that $\mathbf{1}_{\{\tau>u\}}\overline{\varphi}_u = \mathbf{1}_{\{\tau>u\}}\varphi_u$ and obtain

$$\begin{aligned}
&\mathbf{1}_{\{\tau>t\}} \int_t^T \mathbb{E}_t^{\mathcal{F}}\left[\mathbb{E}_u^{\mathcal{F}}[\mathbf{1}_{\{\tau>u\}}]\overline{\varphi}_u\right]\mathrm{e}^{\int_0^t \lambda_s ds}du\\
=&\mathbf{1}_{\{\tau>t\}} \int_t^T \mathbb{E}_t^{\mathcal{F}}\left[\mathrm{e}^{-\int_0^u \lambda_s ds}\overline{\varphi}_u\right]\mathrm{e}^{\int_0^t \lambda_s ds}du\\
=&\mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{F}}\left[\int_t^T \mathrm{e}^{-\int_t^u \lambda_s ds}\overline{\varphi}_u du\right].
\end{aligned}$$

□

Lemma 3.1.3. *If φ_u is an $\mathcal{F}$-predictable process, we have:*

$$\mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_{\{t<\tau<T\}}\mathbf{1}_{\{\tau_C<\tau_I\}}\varphi_\tau\right] = \mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{F}}\left[\int_t^T \mathrm{e}^{-\int_t^u(\lambda_s^C+\lambda_s^I)ds}\lambda_u^C\varphi_u du\right].$$

Proof. Consider $\varphi_u = \mathbf{1}_A\mathbf{1}_{\{s<u\leq v\}}$ for $t \leq s < v \leq T$ and some $A \in \mathcal{F}_s$. We note that

$$\mathbf{1}_{\{s<\tau=\tau_C\leq v\}} = \mathbf{1}_{\{v\wedge\tau\geq\tau_C\}} - \mathbf{1}_{\{s\wedge\tau\geq\tau_C\}}.$$

By adding and subtracting $\int_0^{v\wedge\tau}\lambda_u^C du$ and $\int_0^{s\wedge\tau}\lambda_u^C du$, we have

$$\begin{aligned}
&\left(\mathbf{1}_{\{v\wedge\tau\geq\tau_C\}} - \int_0^{v\wedge\tau}\lambda_u^C du\right) + \int_0^{v\wedge\tau}\lambda_u^C du\\
-&\left(\mathbf{1}_{\{s\wedge\tau\geq\tau_C\}} - \int_0^{s\wedge\tau}\lambda_u^C du\right) - \int_0^{s\wedge\tau}\lambda_u^C du,
\end{aligned}$$

denote as $M_a^C = \mathbf{1}_{\{a\wedge\tau\geq\tau_C\}} - \int_0^{a\wedge\tau}\lambda_u^C du$, it's a $\mathcal{G}$-martingale (refer to Section 3.7.1 of [5] for further details.). Hence

$$\begin{aligned}
&\mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_{\{t<\tau=\tau_C\leq T\}}\mathbf{1}_{\{\tau_C\leq\tau_I\}}\varphi_\tau\right] = \mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_A\mathbf{1}_{\{s<\tau=\tau_C\leq v\}}\right]\\
=&\mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_A\left(M_v^C - M_s^C + \int_{s\wedge\tau}^{v\wedge\tau}\lambda_u^C du\right)\right] = \mathbb{E}_t^{\mathcal{G}}\left[\mathbf{1}_A\mathbb{E}_s^{\mathcal{G}}\left[M_v^C - M_s^C + \int_{s\wedge\tau}^{v\wedge\tau}\lambda_u^C du\right]\right]\\
=&\mathbb{E}_t^{\mathcal{G}}\left[\int_{t\wedge\tau}^{T\wedge\tau}\lambda_u^C\varphi_u du\right] = \mathbf{1}_{\{\tau>t\}}\frac{1}{\mathbb{P}(\tau>t|\mathcal{F}_t)}\mathbb{E}_t^{\mathcal{F}}\left[\int_t^T \lambda_u^C\varphi_u\mathbb{P}(\tau>u|\mathcal{F}_u)du\right]\\
=&\mathbf{1}_{\{\tau>t\}}\mathbb{E}_t^{\mathcal{F}}\left[\int_t^T \mathrm{e}^{-\int_t^u(\lambda_s^C+\lambda_s^I)ds}\lambda_u^C\varphi_u du\right],
\end{aligned}$$

where the last equality follows from the formula

$$\mathbb{E}_t^{\mathcal{G}}\left[\int_{t\wedge\tau}^{T\wedge\tau}\phi_u du\right] = \mathbf{1}_{\{\tau>t\}}\frac{1}{\mathbb{P}(\tau>t|\mathcal{F}_t)}\mathbb{E}_t^{\mathcal{F}}\left[\int_t^T \phi_u \mathbb{P}(\tau>u|\mathcal{F}_u)du\right]$$

which holds for any $\mathbb{F}$-predictable process ϕ such that the right-hand side is well defined. □

Projecting (3.2) on $\mathcal{F}_t$, and employing the previous lemma we may conclude that the $\mathcal{F}_t$-adapted adjusted price V_t, such that $\mathbf{1}_{\{\tau>t\}}V_t = \mathbf{1}_{\{\tau>t\}}\widetilde{V}_t$ verifies the following $\mathcal{F}$-BSDE

$$\begin{aligned} V_t = \mathbb{E}_t^{\mathcal{F}}\Bigg[&\mathrm{e}^{-\int_t^T(r_s+\lambda_s)ds}\Phi(S_T) + \int_t^T \mathrm{e}^{-\int_t^u(r_s+\lambda_s)ds}\Big(\pi_u - (c_u - r_u)C_u - (f_u - r_u)(V_u - C_u) \\ &- (r_u - h_u)H_u + V_u\lambda_u - LGD_C\lambda_u^C(V_u - C_u)^+ + LGD_I\lambda_u^I(V_u - C_u)^-\Big)du\Bigg], \end{aligned} \tag{3.4}$$

where f_u, h_u and H_u are $\mathcal{F}$-adapted processes such that $\mathbf{1}_{\{\tau>t\}}\xi_u = \mathbf{1}_{\{\tau>t\}}\overline{\xi}_u$ for $\overline{\xi} = \overline{f}, \overline{h}, \overline{H}$ ad define in (3.2).

If $\mathcal{F}_t$ is generated by a (possibly multidimensional) Brownian motion driving the market assets prices, by the martingale representation theorem, taking for granted the necessary integrability conditions, we have the following theorem

Propotion 3.1.1. *The value process* (3.4) *satisfies the following BSDE:*

$$\begin{aligned} V_t =&\Phi(S_T) + \int_t^T \Big(\pi_u + (f_u - c_u)C_u - f_uV_u - (r_u - h_u)H_u - LGD_C\lambda_u^C(V_u - C_u)^+ \\ &+LGD_I\lambda_u^I(V_u - C_u)^-\Big)du - \int_t^T Z_u \cdot dW_u + \mathcal{M}_t, \end{aligned} \tag{3.5}$$

where W_t is a (vector) Brownian motion, Z_t an $\mathcal{F}$-adapted, possibly square integrable, (vector) process, and $\mathcal{M}_t$ is a martingale orthogonal to $\int_t^T Z_u \cdot dW_u$, possibly depending on further stochastic factors. For the sake of simplicity, we assume $\mathcal{M} = 0$.

Proof. Denote as

$$\begin{aligned} \mathcal{Q}(t, V_t, \lambda_t^C, \lambda_t^I) := &\ \pi_t - (c_t - r_t)C_t - (f_t - r_t)(V_t - C_t) - (r_t - h_t)H_t + V_t\lambda_t \\ &- LGD_C\lambda_t^C(V_t - C_t)^+ + LGD_I\lambda_t^I(V_t - C_t)^-. \end{aligned}$$

We rewrite the value process as

$$V_t = \mathbb{E}_t^{\mathcal{F}}\left[\mathrm{e}^{-\int_t^T(r_s+\lambda_s)ds}\Phi(S_T) + \int_t^T \mathrm{e}^{-\int_t^u(r_s+\lambda_s)ds}\mathcal{Q}(u, V_u, \lambda_u^C, \lambda_u^I)du\right]. \tag{3.6}$$

We multiply (3.6) by $e^{-\int_0^t (r_s+\lambda_s)ds}$, and we split the integral

$$e^{-\int_0^t (r_s+\lambda_s)ds} V_t = \mathbb{E}_t^{\mathcal{F}} \Bigg[e^{-\int_0^T (r_s+\lambda_s)ds} \Phi(S_T) - \int_0^t e^{-\int_0^u (r_s+\lambda_s)ds} \mathcal{Q}(u, V_u, \lambda_u^C, \lambda_u^I) du + \int_0^T e^{-\int_0^u (r_s+\lambda_s)ds} \mathcal{Q}(u, V_u, \lambda_u^C, \lambda_u^I) du \Bigg].$$

The right hand side is made up of a $\mathcal{F}_t-$predictable term and a local $\mathcal{F}$-martingale. Then, since it is adapted to the Brownian generating the filtration $\mathcal{F}$, by the martingale representation theorem we have

$$\mathbb{E}_t^{\mathcal{F}} \left[e^{-\int_0^T (r_s+\lambda_s)ds} \Phi(S_T) + \int_0^T e^{-\int_0^u (r_s+\lambda_s)ds} \mathcal{Q}(u, V_u, \lambda_u^C, \lambda_u^I) du \right] = \int_0^t \overline{Z}_u \cdot dW_u$$

for some $\mathcal{F}$-predictable (vector) process $\overline{Z}_u$

$$e^{-\int_0^t (r_s+\lambda_s)ds} V_t + \int_0^t e^{-\int_0^u (r_s+\lambda_s)ds} \mathcal{Q}(u, V_u, \lambda_u^C, \lambda_u^I) du = \int_0^t \overline{Z_u} dW_u.$$

By applying integration parts formula, we have

$$-(r_t + \lambda_t) e^{-\int_0^t (r_s+\lambda_s)ds} V_t dt + e^{-\int_0^t (r_s+\lambda_s)ds} dV_t + e^{-\int_0^t (r_s+\lambda_s)ds} \mathcal{Q}(t, V_t, \lambda_t^C, \lambda_t^I) dt = \overline{Z_t} dW_t.$$

whence, we way rewrite the above as

$$dV_t = \big[(r_t + \lambda_t) V_t - \mathcal{Q}(t, V_t, \lambda_t^C, \lambda_t^I) \big] dt + Z_t \cdot dW_t,$$

with $Z_t = e^{\int_0^t (r_s+\lambda_s)ds} \overline{Z}_t$.

By integrating on interval $[t, T]$, we have

$$V_t = V_T - \int_t^T \big[(r_u + \lambda_u) V_u - \mathcal{Q}(u, V_u, \lambda_u^C, \lambda_u^I) \big] dt - \int_t^T Z_u \cdot dW_u.$$

Recalling $\mathcal{Q}$ and that $V_T = \Phi(S_T)$ we get immediately (3.5). □

Missing a closed form solution for (3.5), one may try to construct an appropriate approximation procedure. In the literature, the most widespread method is Monte Carlo simulations.

Monte Carlo simulations are widely used in the literature. Here, we briefly present the Longstaff-Schwartz method, which provides a tool to discretize conditional expectations. That the method we employed was introduced in a paper written in 2001, we would like to emphasize that our decision to use this method stems from its widespread use in the articles referenced throughout in this thesis. The chosen method serves as a benchmark value, allowing us to conduct our analyses effectively. By doing so, we ensure consistency and comparability with existing research in the field.

It is worth noting that more advanced Monte Carlo methods, such as variance reduction or multilevel techniques, could indeed offer increased competitiveness. However, it is important to clarify that the primary objective of our research is not to find the best Monte Carlo method. Instead, our focus lies in establishing a benchmark value that enables us to carry out our analyses accurately and draw meaningful conclusions.

3.1.1 Monte Carlo techniques

In [41], Longstaff and Schwartz developed a Monte Carlo simulation method for pricing American options[5]. Obviously, the method can be employed to value also European options.

The key feature of this method is the discretization of the conditional expectation at all times and we are going to adapt and employ it to (3.4).

Let us consider our evaluation on a finite partition $\{t_0, t_1, \cdots, t_m = T\}$ of $[0, T]$ and with step size $\Delta t = t_{j+1} - t_j$.

According to the theory of no-arbitrage valuation, the value of an option is determined by taking the expectation of the remaining discounted cashflows with respect to the risk-neutral pricing measure. At time $t_m = T$, the investor knows the option's value, which is given by the payoff $\Phi(S_T)$. However, the value of the option is unknown at earlier times t_k for $k = m-1, m-2, \ldots, 0$, and we want to estimate it.

Discretizing the time integral we may approximate the value of the derivative defined by (3.5), at time t_k by

$$V_{t_k} = \mathbb{E}^{\mathcal{F}}_{t_k}\left[\mathrm{e}^{-\int_{t_k}^{T}(r_s+\lambda_s)ds}\Phi(S_T) + \sum_{j=k+1}^{m} \mathrm{e}^{-\int_{t_k}^{t_j}(r_s+\lambda_s)ds}\mathcal{Q}(t_j, V_{t_j}, \lambda^C_{t_j}, \lambda^I_{t_j})\right]. \tag{3.7}$$

Once this done, it remains to approximation the conditional expectation. The most popular approach to do so (3.7) is based on regression methods.

The method consists in approximating the solution of (3.7) by a linear combination of known functions, called basis functions, of the current state of asset price S. This is possible, because the conditional expectation is a square integrable functions in a Hilbert space, hence, it has a countable orthonormal basis and the conditional expectation can be represented as a linear function of the elements of the basis. We can express (3.7) as

$$V_{t_k} = \sum_{i=0}^{\infty} \beta^i_{t_k}\psi_i(S_{t_k}) \tag{3.8}$$

Here, $\psi_i(S_{t_k})$ represents the basis function, and $\beta^i_{t_k}$ represents the unknown coefficients that de-

[5] An American option is a financial contract that gives the holder the right, but not the obligation, to buy or sell an underlying asset at a predetermined price (strike price) at any time before or on the option's expiration date, unlike European options which can only be exercised on the expiration date.

pend on time t_k.

In our simulation, we use the Laguerre polynomials as our orthonormal basis

$$\psi_i(x) = \exp(x/2)\frac{\mathrm{e}^{\mathrm{x}}}{i!}\frac{d^i}{dx^i}(x^n \mathrm{e}^{-\mathrm{x}}), \quad i = 0, 1, 2, \ldots$$

but Chebyshev, Legendre, Jacobi or Hermite polynomials can also be used [41].

A regression method is then used to estimate the optimal coefficients for the approximation. For this, we use the first $M < \infty$ basis functions and denote this further approximation as

$$V_{t_k}^M = \sum_{i=0}^{M-1} \beta_{t_k}^i \psi_i(S_{t_k}) = \beta_{t_k}^\top \psi(S_{t_k}),$$

with $\beta_{t_k} = (\beta_{t_k}^0, \beta_{t_k}^1, \cdots, \beta_{t_k}^{M-1})^\top$, and $\psi(x) = (\psi_0(x), \psi_1(x), \cdots, \psi_{M-1}(x))$. The regression coefficients can be determined using the following theorem.

Theorem 3.1.1 ([48])**.**

$$\beta_{t_k} = \big(\psi(S_{t_k})\psi(S_{t_k})^\top\big)^{-1}\big(\psi(S_{t_k})V_{t_{k+1}}\big). \tag{3.9}$$

where $\psi(S_{t_k})\psi(S_{t_k})^\top$ *is a* $M \times M$ *matrix and* $\psi(S_{t_k})V_{t_{j+1}}$ *is a vector of length* M.

Proof. By least-square regression we want to find

$$\min_{\beta_{t_k}} \Big(\psi(S_{t_k})^\top \beta_{t_k} - V_{t_{k+1}}\Big)^2.$$

By taking the derivative of the above with respect to β_{t_k} and setting it equal 0, we get

$$\psi(S_{t_k})\Big(\psi(S_{t_k})^\top \beta_{t_k} - V_{t_{k+1}}\Big) = 0.$$

By solving for β_{t_k}, we get

$$\begin{aligned} \psi(S_{t_k})\psi(S_{t_k})^\top \beta_{t_k} &= \psi(S_{t_k})V_{t_{k+1}} \\ \Rightarrow \beta_{t_k} &= \big(\psi(S_{t_k})\psi(S_{t_k})^\top\big)^{-1}\psi(S_{t_k})V_{t_{k+1}}. \end{aligned}$$

□

We remark that (3.4) describes a backward regression.

So starting from $t_m = T$ we may estimate β_{t_k}, step by step by employing Monte Carlo simulations at each time t_k.

More in detail, we simulate N independent sample paths for asset price $\mathcal{S}_{t_k} = (S_{t_1}^{(n)}, S_{t_2}^{(n)}, \ldots, S_{t_m}^{(n)})$ for $n = 1, \ldots, N$. The least-squares estimation of the regression coefficients β_{t_k} is then calculated as follows:

$$\widehat{\beta}_{t_k} = \left(\frac{1}{N}\sum_{n=1}^N \psi(S_{t_k}^{(n)})\psi(S_{t_k}^{(n)})^\top\right)\left(\frac{1}{N}\sum_{n=1}^N \psi(S_{t_k}^{(n)})V_{t_{k+1}}^{(n)}\right),$$

where $V^{(n)}_{t_{k+1}}$ is known, because we are working backwards in time. Then, the approximation of the option's value is given by:

$$\widehat{V}^M_{t_k} = \widehat{\beta}^\top_{t_k}\psi(\mathcal{S}_{t_k}) \quad k = m-1, m-2, \ldots, 0,$$

where $\widehat{V}^M_{t_k} = \big(\widehat{V}^{M(1)}_{t_k}, \widehat{V}^{M(2)}_{t_k}, \ldots, \widehat{V}^{M(N)}_{t_k}\big)^\top$.

In [24] it is showed that

$$\lim_{M\to\infty} \widehat{V}^M_{t_k} = V_{t_k}.$$

In particular, we are interested in the approximation value at the initial time t_0, which can be obtained by:

$$\overline{V}_{t_0} = \frac{1}{N}\sum_{n=1}^{N} \widehat{V}^{M(n)}_{t_0}.$$

When considering a triple of stochastic processes, the Monte Carlo simulations needed to approximate V_t in (3.4) are bound to become extremely costly in terms of machine time. As a result, using alternative numerical methods that have lower computational costs becomes a crucial matter.

3.2 Contract's value expression by PDE

In this section, we derive the PDE associated with (3.5). To do so, we require the processes $(S, \lambda^C, \lambda^I)$ to be in a Markovian context. Using the flow notation, for the underlying we write

$$S^{t,x}_s = x + \int_t^s r_u S^{t,x}_u du + \int_t^s \sigma S^{t,x}_u dW_u, \quad \sigma > 0, \quad t \le s \le T,$$

and we assume σ constant from now on.

As for the default intensities, we propose modeling them using Cox Ingersoll Ross processes

$$\begin{aligned} \lambda^{C,t,y}_s &= y + \int_t^s \gamma_C(\psi_C - \lambda^{C,t,y}_u)du + \int_t^s \eta_C\sqrt{\lambda^{C,t,y}_u}dB^C_u \\ \lambda^{I,t,z}_s &= z + \int_t^s \gamma_I(\psi_I - \lambda^{I,t,z}_u)du + \int_t^s \eta_I\sqrt{\lambda^{I,t,z}_u}dB^I_u, \end{aligned} \tag{3.10}$$

with $\gamma_i, \psi_i, \eta_i \ge 0, \quad i = C, I$, verify the *Feller condition*, $2\gamma_i\psi_i \ge \eta_i^2$, to ensure the processes' positivity, and

$$W_t = \rho_C B^C_t + \rho_I B^I_t + \sqrt{1 - \rho_C^2 - \rho_I^2}B_t, \quad \rho_C^2 + \rho_I^2 \le 1, \quad -1 \le \rho_i \le 1$$

where (B^C_t, B^I_t, B_t) is 3-dimensional standard Brownian motion.

To simplify our discussion we also assume that

- the claim pays no dividends, hence $\pi = 0$;
- the rates f, c, h are deterministic, bounded functions of time;

- the collateral process is a fraction of the process V_u, namely $C_u = \alpha_u V_u$, where $0 \leq \alpha_u \leq 1$ is a function of time.
- the process $H_t = H(t, S_t, V_t, Z_t)$, where $H(u, x, v, z)$ is a deterministic, Lipschitz-continuous function in v, z, uniformly in u. Besides we take $H(u, x, 0, 0)$ continuous in x. This means that we have an explicit representation for the hedging process H_t (see [2, 12, 16, 20]);

Here, we choose the two default intensities independent of each other to simplify calculations, but this assumption may be easily removed by adding a further correlation parameter in the discussion that follows.

We remark that by taking $\gamma_i = \psi_i = \eta_i = 0$, $i = C, I$, we can restrict to the case of deterministic intensities (as in [16]).

Using this representation, (3.5) becomes

$$\begin{aligned} dV_s^{t,x,y,z} =& \Bigg[(1-\alpha_t)\Big[f_s V_s^{t,x,y,z} + LGD_C \lambda_s^{C,t,y} V_s^{t,x,y,z,+} - LGD_I \lambda_s^{I,t,y} V_s^{t,x,y,z,-}\Big] \\ &+ \alpha_s c_s V_s^{t,x,y,z} + (r_s - h_s) H(t, S_s^{t,x}, V_s^{t,x,y,z}, Z_s^{t,x,y,z})\Bigg] ds + Z_s^{t,x,y,z} dW_s + d\mathcal{M}_s \\ V_T^{t,x} =& \Phi(S_T^{t,x}), \end{aligned} \tag{3.11}$$

and the previous equation satisfies the assumptions of Theorem 2.3.5, here reported: $\Phi(x) \in L^2(\mathcal{F}_T)$, and $\mathcal{V} : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}^n$ such that $\mathcal{V}(\cdot, 0, 0) \in L^{1,2}(\mathbb{F})$, and f is uniformly Lipschitz continuous in (y, z), with

$$\begin{aligned} d\mathcal{V}(s,y,z) =& \big[(1-\alpha_t)\Big[f_s y_s + LGD_C \lambda_s^C y_s^+ - LGD_I \lambda_s^I y_s^-\Big] + \alpha_s c_s y_s \\ &+ (r_s - h_s) H(t, x_s, y_s, z_s)\big] ds \end{aligned}$$

As shown in [2], assuming a δ-hedging for this product, an appropriate change of probability may be applied to include the hedging function H in the dynamics [6], so that (3.11) may be rewritten as

$$\begin{aligned} dV_s^{t,x,y,z} =& \Bigg[(1-\alpha_s)\Big[f_s V_s^{t,x,y,z} + LGD_C \lambda_s^{C,t,y} V_s^{t,x,y,z,+} - LGD_I \lambda_s^{I,t,y} V_s^{t,x,y,z,-}\Big] \\ &+ \alpha_s c_s V_s^{t,x,y,z} + (r_s - h_s) \frac{\partial V_s^{t,x,y,z}}{\partial S} S_s^{t,x}\Bigg] ds + Z_s^{t,x,y,z} dW_s + d\mathcal{M}_s \\ V_T^{t,x} =& \Phi(S_T^{t,x}). \end{aligned} \tag{3.12}$$

Since the triple $(S^{t,x}, \lambda^{C,t,y}, \lambda^{I,t,z})$ is Markovian, $V_s^{t,x,y,z}$ is a deterministic function of the state variables, $u(s, S_s^{t,x}, \lambda_s^{C,t,y}, \lambda_s^{I,t,z})$. Assuming $u(t,x,y,z) \in C^{1,2}([0,T] \times \mathbb{R}_+^3)$, by applying Ito's

[6] We do not address the hedging problem any further.

formula

$$
\begin{aligned}
du(t,x,y,z) =&\partial_t u(t,x,y,z)dt + \partial_x u(t,x,y,z)dS + \partial_y u(t,x,y,z)d\lambda^C + \partial_z u(t,x,y,z)d\lambda^I \\
&+\frac{1}{2}\Big(\partial^2_{xx}u(t,x,y,z)d\langle S,S\rangle + \partial^2_{yy}u(t,x,y,z)d\langle \lambda^C,\lambda^C\rangle + \partial^2_{zz}u(t,x,y,z)d\langle \lambda^I,\lambda^I\rangle\Big) \\
&+2\partial^2_{xy}u(t,x,y,z)d\langle S,\lambda^C\rangle + 2\partial^2_{xz}u(t,x,y,z)d\langle S,\lambda^I\rangle \\
=&\partial_t u(t,x,y,z) + \partial_x u(t,x,y,z)(rxdt + \sigma x dW_t) + \partial_y(t,x,y,z)\big(\gamma_C(\psi_C - y)dt + \eta_C\sqrt{y}dB^C_t\big) \\
&+\partial_z u(t,x,y,z)\big(\gamma_I(\psi_I - z)dt + \eta_I\sqrt{z}dB^I_t\big) + \frac{1}{2}\Big(\partial^2_{xx}u(t,x,y,z)\sigma^2x^2dt + \partial^2_{yy}u(t,x,y,z)\eta^2_C y dt \\
&+\partial^2_{zz}u(t,x,y,z)\eta^2_I z dt + 2\partial^2_{xy}u(t,x,y,z)\rho_C\sigma x\eta_C\sqrt{y}dt + 2\partial^2_{xz}u(t,x,y,z)\sigma x\eta_I\sqrt{z}dt\Big) \\
=&\big[\partial_t u(t,x,y,z) + rx\partial_x u(t,x,y,z) + \big(\gamma_C(\psi_C - y)\big)\partial_y u(t,x,y,z) + \big(\gamma_I(\psi_I - z)\big)\partial_z u(t,x,y,z) \\
&+\frac{1}{2}\sigma^2x^2\partial^2_{xx}u(t,x,y,z) + \frac{1}{2}\eta^2_C y\partial^2_{yy}u(t,x,y,z) + \frac{1}{2}\eta^2_I z\partial^2_{zz}u(t,x,y,z) \\
&+\rho_C\sigma x\eta_C\sqrt{y}\partial^2_{xy}u(t,x,y,z) + \rho_I\sigma x\eta_I\sqrt{z}\partial^2_{xz}u(t,x,y,z)\big]dt \\
&+\sigma x\partial_x u(t,x,y,z)dW_t + \eta_C\sqrt{y}\partial_y u(t,x,y,z)dB^C_t + \eta_I\sqrt{z}\partial_z u(t,x,y,z)dB^I_t,
\end{aligned}
\tag{3.13}
$$

and comparing the two expressions (3.12) and (3.13), it can be shown that $u(t,x,y,z)$ verifies the non-linear PDE independent of the interest rate r. Indeed,

$$
\begin{aligned}
&\big((1-\alpha)\big(fu(t,x,y,z) + LGD_C yu(t,x,y,z)^+ - LGD_I zu(t,x,y,z)^-\big) \\
&\quad + \alpha cu(t,x,y,z) + (r-h)x\partial_x u(t,x,y,z)\big]dt \\
=&\big[\partial_t u(t,x,y,z) + rx\partial_x u(t,x,y,z) + \big(\gamma_C(\psi_C - y)\big)\partial_y u(t,x,y,z) + \big(\gamma_I(\psi_I - z)\big)\partial_z u(t,x,y,z) \\
&+\frac{1}{2}\sigma^2x^2\partial^2_{xx}u(t,x,y,z) + \frac{1}{2}\eta^2_C y\partial^2_{yy}u(t,x,y,z) + \frac{1}{2}\eta^2_I z\partial^2_{zz}u(t,x,y,z) \\
&+\rho_C\sigma x\eta_C\sqrt{y}\partial^2_{xy}u(t,x,y,z) + \rho_I\sigma x\eta_I\sqrt{z}\partial^2_{xz}u(t,x,y,z)\big]dt
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow&\big[\partial_t u(t,x,y,z) + hx\partial_x u(t,x,y,z) + \big(\gamma_C(\psi_C - y)\big)\partial_y u(t,x,y,z) + \big(\gamma_I(\psi_I - z)\big)\partial_z u(t,x,y,z) \\
&+\frac{1}{2}\sigma^2x^2\partial^2_{xx}u(t,x,y,z) + \frac{1}{2}\eta^2_C y\partial^2_{yy}u(t,x,y,z) + \frac{1}{2}\eta^2_I z\partial^2_{zz}u(t,x,y,z) \\
&+\rho_C\sigma x\eta_C\sqrt{y}\partial^2_{xy}u(t,x,y,z) + \rho_I\sigma x\eta_I\sqrt{z}\partial^2_{xz}u(t,x,y,z) - \alpha cu(t,x,y,z) \\
&-\big((1-\alpha)\big(fu(t,x,y,z) + LGD_C yu(t,x,y,z)^+ - LGD_I zu(t,x,y,z)^-\big)\big]dt = 0.
\end{aligned}
$$

Now, we denote

$$
\begin{aligned}
\mathcal{L}(u)(t,x,y,z) =& \partial_t u(t,x,y,z) + \gamma_C(\psi_C - y)\partial_y u(t,x,y,z) + \gamma_I(\psi_I - z)\partial_z u(t,x,y,z) \\
&+ hx\partial_x u(t,x,y,z) + \frac{1}{2}\eta^2_C y\partial^2_{yy}u(t,x,y,z) + \frac{1}{2}\eta^2_I z\partial^2_{zz}u(t,x,y,z) \\
&+ \frac{1}{2}\sigma^2x^2\partial^2_{xx}u(t,x,y,z) + \rho_C\sigma x\eta_C\sqrt{y}\partial^2_{xy}u(t,x,y,z) \\
&+ \rho_I\sigma x\eta_I\sqrt{z}\partial^2_{xz}u(t,x,y,z) - \alpha cu(t,x,y,z) - (1-\alpha)fu(t,x,y,z),
\end{aligned}
$$

and we rewrite the non-linear PDE

$$
\begin{cases}
\mathcal{L}(u)(t,x,y,z) - (1-\alpha)\Big[LGD_C yu(t,x,y,z)^+ - LGD_I zu(t,x,y,z)^-\Big] = 0 \\
u(T,x,y,z) = \Phi(x).
\end{cases}
\tag{3.14}
$$

In the next chapter we are going to suggest a discretization of (3.14) that seems to work efficiently in terms of computational times and accuracy.

We remark that the solution to (3.14) might be intended in viscosity sense, and we postpone the analysis of the regularity of the solution to further future work.

Chapter 4

Problem discretization

The method of lines (see, for instance, [28, 29, 38, 40, 45–47]) is a numerical technique commonly used to solve partial differential equations (PDEs) by discretizing them in one or more spatial dimensions and solving the resulting system of ordinary differential equations (ODEs). Typically, the method of lines involves dividing the spatial domain into a set of grid points and approximating the spatial derivatives in the PDE with finite difference approximations. This discretization transforms the PDE into a system of ODEs, which can be solved by a suitable time integration method, such as the Euler one.

4.1 Finite difference approximations

In this section we brefly review the finite difference method often employed to approximate the solution of a PDE, that we will later use for (3.14). However, for a more detailed discussion we refer to the reader [40, 46, 47].

The finite difference method is a numerical technique used to solve differential equations. It involves approximating the derivatives in the equation with difference quotients and discretizing the spatial domain by dividing it into a grid of points. The method then iteratively computes approximate solutions of the equation at each point on the grid. The accuracy of the solution is determined by the size of the grid and the order of the used difference approximation. The finite difference method is widely used in various fields, including physics, engineering, and finance. For simplicity, we consider the one-dimensional case. Let's suppose that a function ν is sufficiently smooth to be expanded in a Taylor series in the neighborhood of x with an increment $h > 0$. Truncating the expansion it at the first order, we have:

$$\nu(x+h) = \nu(x) + h\nu'(x) + O(h^2) \Rightarrow \nu'(x) \approx \frac{\nu(x+h) - \nu(x)}{h} \tag{4.1}$$

where the term $O(h^2)$ indicates that the error is proportional to h^2. From (4.1), we deduce that there exists a constant $K > 0$, such that for $h > 0$ sufficiently small we have:

$$\left| \frac{\nu(x+h) - \nu(x)}{h} - \nu'(x) \right| \leq Kh, \quad K = \sup_{y \in [x, x+h_0]} \frac{|\nu''(y)|}{2},$$

for $h_0 > 0$ and $h \leq h_0$. The error committed by replacing the derivative $\nu'(x)$ by the differential quotient is then of order h and the approximation of ν' at point x is said to be consistent at the first order. This approximation is known as the *forward difference approximation* of ν'. Likewise, we can define the first-order *backward difference approximation* of ν' at point x as:

$$\nu(x-h) = \nu(x) - h\nu'(x) + O(h^2) \Rightarrow \nu'(x) \approx \frac{\nu(x-h) - \nu(x)}{h}. \tag{4.2}$$

In order to improve the accuracy of the approximation, we define a consistent approximation, called the *central difference approximation*, by taking the points $x - h$ and $x + h$ into account. Expanding the function ν at the points $x + h$ and $x - h$, truncating it at the third order, we have

$$\begin{aligned} \nu(x+h) &= \nu(x) + h\nu'(x) + \frac{h^2}{2}\nu''(x) + O(h^3) \\ \nu(x-h) &= u(x) - h\nu'(x) + \frac{h^2}{2}\nu''(x) - O(h^3). \end{aligned}$$

By subtracting these two expressions we obtain

$$\frac{\nu(x+h) - \nu(x-h)}{2h} = \nu'(x) + 2O(h^3) \Rightarrow \nu'(x) \approx \frac{\nu(x+h) - \nu(x-h)}{2h} \tag{4.3}$$

Hence, for every $h \in (0, h_0)$, we have the following bound on the approximation error

$$\left| \frac{\nu(x+h) - \nu(x-h)}{2h} - \nu'(x) \right| \leq Ch^2, \quad K = \sup_{y \in (x-h_0, x+h_0)} \frac{|\nu'''(y)|}{6}.$$

This defines a second-order consistent approximation to ν'.

In a similar way, an approximation can be found for the second derivative, using the Taylor expansions up to the fourth order to achieve the result:

$$\begin{aligned} \nu(x+h) &= \nu(x) + h\nu'(x) + \frac{h^2}{2}u''(x) + \frac{h^3}{6}\nu'''(x) + O(h^4) \\ \nu(x-h) &= \nu(x) - h\nu'(x) + \frac{h^2}{2}\nu''(x) - \frac{h^3}{6}\nu'''(x) + O(h^4). \end{aligned}$$

Like before, we can write

$$\frac{\nu(x+h) - 2\nu(x) + \nu(x-h)}{h^2} = \nu''(x) + 2O(h^4) \Rightarrow \nu''(x) \approx \frac{\nu(x+h) - 2\nu(x) + \nu(x-h)}{h^2} \tag{4.4}$$

As far as numerical solutions are concerned, we will only deal with discrete solutions, i.e., solutions denoted only by a discrete set of x-values. Let us assume a grid of points $\{x_0 < x_1 < \cdots < x_m\}$ with step $\Delta x = x_{i+1} - x_i$. In general, a method provides an approximation $\nu_i \approx \nu(x_i)$ on each grid point. So, now we can approximate (4.1) (4.2) (4.3) (4.4) as follows

$$\nu_i' = \frac{\nu_{i+1} - \nu_i}{\Delta x} \quad \text{forward finite difference for first derivative}$$
$$\nu_i' = \frac{\nu_{i-1} - \nu_i}{\Delta x} \quad \text{backward finite difference for first derivative}$$
$$\nu_i' = \frac{\nu_{i+1} - \nu_{i-1}}{2\Delta x} \quad \text{central finite difference for first derivative}$$
$$\nu_i'' = \frac{\nu_{i+1} - 2\nu_i + \nu_{i-1}}{(\Delta x)^2} \quad \text{central finite difference for second derivative}$$

4.2 Example: discretizing a parabolic problem

We consider the diffusion equation with boundary and initial condition

$$\begin{cases} \partial_t u(t,x) - \partial^2_{xx} u(t,x) = f(t,x), & t > 0, \quad x \in (0,L) \\ u(t,0) = u(t,L) = 0, & t > 0 \\ u(0,x) = u_0(x), & x \in (0,L). \end{cases} \tag{4.5}$$

To solve the heat equation numerically we have to discretize the spatial domain following the finite difference. We sub-divide the space interval into m uniform sub-interval with step $\Delta x = x_{i+1} - x_i$, and denote by $u_i(t)$ an approximation of $u(t, x_i)$, $i = 1, \cdots, m-1$, and approximate the problem (4.5) by the scheme

$$u_i'(t) = \frac{1}{\Delta x^2}\big(u_{i-1}(t) - 2u_i(t) + u_{i+1}(t)\big) = f_i(t), \quad i = 1, \cdots, m-1, \forall t > 0$$
$$u_0(t) = u_m(t) = 0, \quad \forall t > 0,$$
$$u_i(0) = u_0(x_i), \quad i = 1, \cdots, m-1,$$

where $f_i(t) = f(t, x_i)$, and it is a system of ordinary differential equations of the following form

$$\begin{cases} \mathbf{u}'(t) & = -A\mathbf{u}(t) + \mathbf{f}(t), \quad \forall t > 0 \\ \mathbf{u}(0) & = \mathbf{u}_0 \end{cases} \tag{4.6}$$

where $\mathbf{u}(t) = (u_1(t), \cdots, u_{m-1}(t))^\top$ is the vector of unknowns, $\mathbf{f}(t) = (f_1(t), \cdots, f_{m-1}(t))^\top$, $\mathbf{u}_0 = (u_1(x_1), \cdots, u_0(x_{m-1}))^\top$ and A is the $(m-1) \times (m-1)$ tridiagonal matrix of the form

$$A = \frac{1}{\Delta x^2} \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 1 & -2 \end{pmatrix}.$$

A scheme for the integration of (4.6) with respect to time is the explicit Euler scheme. To construct the scheme, we consider a temporal grid $\{t_0, t_1, \cdots, t_n\}$, with step $\Delta t = t_{j+1} - t_j$ for $j = 0, \cdots, n-1$ then, the explicit Euler scheme for the time-integration of (4.6) is

$$\begin{cases} \dfrac{\mathbf{u}(t_{j+1}) - \mathbf{u}(t_j)}{\Delta t} = -A\mathbf{u}(t_j) + \mathbf{f}(t_j) \\ \mathbf{u}(t_0) = \mathbf{u}_0, \end{cases} \tag{4.7}$$

and we can get $\mathbf{u}(t_{j+1})$ explicitly. Regarding stability [38, 40, 45, 47], the explicit Euler method is conditionally stable[1], and the time-step Δt should decay as the square of the grid spacing Δx. In other words, the method is stable if $\Delta t < \frac{\Delta x^2}{2}$. This condition is called *Courant–Friedrichs–Lewy condition (CFL)*. An unconditionally stable method is the implicit Euler scheme that, applied to Equation (4.6) leads to

$$\begin{cases} \dfrac{\mathbf{u}(t_{j+1}) - \mathbf{u}(t_j)}{\Delta t} = -A\mathbf{u}(t_{j+1}) + \mathbf{f}(t_{j+1}) \\ \mathbf{u}(t_0) = \mathbf{u}_0. \end{cases} \tag{4.8}$$

4.3 Discretization of the contract's value

Returning to our PDE (3.14), we note that the spatial domain $\mathbb{R}^3_+$ is unbounded, so we need to restrict it to an appropriate bounded rectangle $[a_x, b_x] \times [a_y, b_y] \times [a_z, b_z] \subset \mathbb{R}^3_+$. This truncation requires defining appropriate boundary conditions, which can be done by identifying, when possible, the asymptotic behaviour of the solution. Here, we decided to exploit the knowledge of the Black & Scholes formula, with adjusted rates to include the default intensities

$$\begin{aligned} u(t, a_x, y, z) &= 0, & u(t, b_x, y, z) &= \phi(t, b_x; r+\lambda, \sigma), \\ u(t, x, a_y, z) &= \phi(t, x; r+\lambda, \sigma), & u(t, x, b_y, z) &= \phi(t, x; r+\lambda, \sigma), \\ u(t, x, y, a_z) &= \phi(t, x; r+\lambda, \sigma), & u(t, x, y, b_z) &= \phi(t, x; r+\lambda, \sigma), \end{aligned} \tag{4.9}$$

[1] A method is conditionally stable if its stability depends on certain conditions being satisfied. In other words, the method may be stable for some ranges of input parameters or initial conditions, but unstable for others.In numerical analysis, a method is considered stable if small perturbations in the input or the algorithm do not cause large errors in the output. Conditionally stable methods may have regions in their parameter space where they become unstable, which can lead to numerical errors or incorrect results.

where $\phi(t,x;w,\sigma)$ is the Black and Scholes's pricing function. The choice of the Black and Scholes's pricing function to set the Dirichlet boundary conditions is somewhat arbitrary. The rationale behind such choice is that it is exactly what we would have, when considering only the CVA without any other feature. We sub-divide the three space intervals into m uniform [2] sub-intervals by taking, $x_i = a_x + i\Delta x$, $y_i = a_y + i\Delta y$, $z_i = a_z + i\Delta z$ with $\Delta x = \frac{(b_x - a_x)}{m}$, $\Delta y = \frac{(b_y - a_y)}{m}$, $\Delta z = \frac{(b_z - a_z)}{m}$ for $i = 0, \ldots, m$, and we apply the finite difference method to approximate the space partial derivatives,

$$
\begin{aligned}
\partial_x u(t,x_k,y_i,z_j) &\approx \frac{u(t,x_{k+1},y_i,z_j) - u(t,x_k,y_i,z_j)}{\Delta x} \quad k=0,\ldots,m-1, \quad i,j=0,\ldots,m \\
\partial^2_{xx} u(t,x_k,y_i,z_j) &\approx \frac{u(t,x_{k+1},y_i,z_j) - 2u(t,x_k,y_i,z_j) + u(t,x_{k-1},y_i,z_j)}{\Delta x^2} \\
& k=1,\ldots,m-1, \quad i,j=0,\ldots,m \\
\partial^2_{xy} u(t,x_k,y_i,z_j) &\approx \frac{u(t,x_{k+1},y_{i+1},z_j) - u(t,x_k,y_{i+1},z_j) - u(t,x_{k+1},y_i,z_j) + u(t,x_k,y_i,z_j)}{\Delta x \Delta y} \\
& i,k=0,\ldots,m-1, \quad j=0,\ldots,m.
\end{aligned}
$$

We write the equation at each point x_k, y_i, z_j, and we denote the piecewise approximation of $u(t,x,y,z)$ by $u_{k,i,j}(t) = u(t,x_k,y_i,z_j)$ for $x \in [x_k, x_{k+1})$,$y \in [y_i, y_{i+1})$, $z \in [z_j, z_{j+1})$ with $i,j,k = 0,\ldots,m-1$. For fixed x_k, y_i, z_j we get the following non-linear ODE

$$
u_{k,i,j}(t)' = \mathcal{DL}(u_{k,i,j})(t) - (1-\alpha)\big[LGD_C y_i u_{k,i,j}(t)^+ - LGD_I z_j u_{k,i,j}(t)^-)\big], k,i,j=0,\ldots,m, \tag{4.10}
$$

$\mathcal{DL}(u_{k,i,j})$ is the discretized operator of $\mathcal{L}(u)$.

We can write the non-linear ODE system in matrix form

$$
\overline{u}(t)' = \mathbf{A}(\overline{x},\overline{y},\overline{z})\overline{u}(t) - (1-\alpha)\big[LGD_C \overline{yu}(t)^+ - LGD_I \overline{zu}(t)^-\big], \tag{4.11}
$$

where $\overline{u}(t)'$, $\overline{u}(t)$, and $\mathbf{A}(\overline{x},\overline{y},\overline{z})$ is a 3-dimensional tensor respectively, and $\overline{x},\overline{y},\overline{z}$ are the vectors in $\mathbb{R}^{m+1}$ given by

$$
\overline{x} = (a_x, x_1, \ldots, x_{m-1}, b_x)^\top, \quad \overline{y} = (a_y, y_1, \ldots, y_{m-1}, b_y)^\top, \quad \overline{z} = (a_z, z_1, \ldots, z_{m-1}, b_z)^\top
$$

with final condition $u(T,\overline{x},\overline{y},\overline{z}) = \overline{\Phi}(x)$ holds, where $\overline{\Phi}(x) = \big(\Phi_0(x), \Phi_1(x), \ldots, \Phi_m(x)\big)^\top$.

Accordingly with the choice described before, we pose the boundary conditions

$$
\begin{aligned}
u(t,x_0,y_i,z_j) =& 0 & i,j=0,\ldots,m, \\
u(t,x_m,y_i,z_j) =& \phi(t,x_m;r+\lambda,\sigma) & i,j=0,\ldots,m, \\
u(t,x_k,y_0,z_j) =& u(t,x_k,y_m,z_j) = \phi(t,x_k;r+\lambda,\sigma) & k,j=1,\ldots,m-1, \\
u(t,x_k,y_i,z_0) =& u(t,x_k,y_i,z_m) = \phi(t,x_k;r+\lambda,\sigma) & k,i=1,\ldots,m-1.
\end{aligned}
$$

[2] There may be sub-regions of the spatial sub-intervals that may be more probable than others, so it would be worthwhile to perform non-uniform discretization, for details see [37, 47].

To solve system (4.11), we use the explicit Euler scheme with N ($0 = t_0 < t_1 < \cdots < t_{N-1} < t_N = T$) time sub-intervals of uniform length $\Delta t = t_{i+1} - t_i$ for $i = 0, \ldots, N-1$, so we get

$$\overline{u}(t_i) = \overline{u}(t_{i+1}) - \Delta t\big[\mathbf{A}(\overline{x}, \overline{y}, \overline{z})\overline{u}(t_{i+1}) - (1-\alpha)\big(LGD_C\overline{yu}(t_{i+1})^+ - LGD_I\overline{zu}(t_{i+1})^-\big)\big] \quad . \quad (4.12)$$

We are aware that the explicit Euler method could produce serious numerical instabilities, and therefore favouring an implicit scheme would be a better choice, since it has no limitations on the time integration step, even through implying quite lengthy computations. In the next section, we compare some numerical results from the explicit, implicit, and semi-explicit Euler schemes. In our setting, we are indeed able to achieve a good and competitive accuracy by the explicit scheme in highly shorter computational times.

For $t = 0$, we are interested in computing the value $u(0, x, y, z)$ for given x, y, z. To do so, we simply choose the closest points of the grid such that $x_k \approx x$,$y_i \approx y$, $z_j \approx z$ for some $k, i, j = 0, \ldots, m$ and we approximate the solution value by $u(0, x_k, y_i, z_j)$, or, as suggested in [47], the specific option value is determined by spline interpolation.

We remark that for the solution of (4.12) to remain stable, the $min\big(\frac{\Delta t}{\Delta x^2}, \frac{\Delta t}{\Delta y^2}, \frac{\Delta t}{\Delta z^2}\big)$ must remain below a critical value. Hence, if one wishes to increase the accuracy of (4.12) by using smaller Δx or $\Delta y, \Delta z$, also a smaller value of Δt is required to keep the CFL number below its critical value. Thus, there is a conflicting requirement between improving accuracy and maintaining stability (for more detail on the stability theory, see Chapter 9 of [40], or Chapter 9 of [38]), which may imply an increase in computational time. For completeness, we also give the expression using the implicit Euler method

$$\overline{u}(t_i) = \overline{u}(t_{i+1}) - \Delta t\big[\mathbf{A}(\overline{x}, \overline{y}, \overline{z})\overline{u}(t_i) - (1-\alpha)\big(LGD_C\overline{yu}(t_i)^+ - LGD_I\overline{zu}(t_i)^-\big)\big] \quad . \quad (4.13)$$

and the semi-implicit Euler method [4, 10]

$$\overline{u}(t_i) = \overline{u}(t_{i+1}) - \Delta t\big[\mathbf{A}(\overline{x}, \overline{y}, \overline{z})\overline{u}(t_{i+1}) - (1-\alpha)\big(LGD_C\overline{yu}(t_i)^+ - LGD_I\overline{zu}(t_i)^-\big)\big] \quad . \quad (4.14)$$

4.4 Numerical Results

In this section, we present some numerical results of our method for the European call price. First, we looked at the case with constant intensities to test the method's accuracy, comparing with the results obtained in [16] by Monte Carlo simulations, with the same set of parameters. In this case, only one state variable, represented by the underlying price, is present.

All the algorithms were implemented in MatLab(R2021a) on a Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz 2.11 GHz computer.

We consider a European call option with six months maturity, strike price $K = 90$, and we set (as in [16]) $r = 0.005$, $\sigma = 0.4$, $LGD_C = 0.6$, $LGD_I = 0.6$, $c = 0.002$, $f = r$, $\alpha = 0.5$ $\lambda_C = 0.04$ and $\lambda_I = 0.02$.

As the computational time was not reported in [16], we replicated their simulations, moreover a 95% confidence interval has been built, with $M = 10^6$ sample independent paths and with $N_t = 1000$ temporal nodes, obtaining the value 16.4494, in about 7 minutes of machine time (fairly close to 16.4534 in [16]). In Table 4.1, we report the results of our method with the relative computational times and we compare them with the results in [16]. We remark that with only

Monte Carlo simulations				
N_t	Seconds	confidence interval	Price	by Brigo
1000	416	(16.4405;16.4583)	16.4494	**16.4534**

Method of lines			
N_t	N_x	Seconds	Price
100	30	0.31	16.4545771
500	50	0.28	16.4272255
1000	90	0.57	16.4643242
5000	150	1.93	16.4555334
5000	200	1.52	16.4568071
10000	300	3.7	16.4574087
50000	500	18.48	16.4574889

Table 4.1: Prices of a European call with maturity 6 months and deterministic intensities with explicit scheme.

30 spatial nodes, we get about the same value as by Monte Carlo simulations, with almost nihil computational time. From Table 4.1, we achieve better performance and comparable accuracy also with respect to [3] [3], where the computational time is about 25 seconds.

Moreover, increasing the number of spatial nodes and of temporal nodes, the second and third digits stabilize, showing the convergence of the method. The first two decimal digits coincides with those obtained by [16], and thanks to the convergence, we probably achieve a better accuracy. Indeed, the digits seem to stabilize progressively.

In Table 4.2, we compare the explicit, semi-implicit and implicit [4] methods as the strike price varies ($K = 90; 100; 110$), with the benchmark values from [16] and from our Monte Carlo simulations. Finally in the Table 4.3, we run the same analysis for varying volatility ($\sigma = 0.3; 0.4; 0.6$). Given a mesh dense enough, all Euler schemes produce faster results than Monte Carlo simulations, as shown in Table 4.2, and they approximate the benchmark very well. Furthermore, we observe that the explicit method achieves the same results as the implicit one, but in remarkably shorter times. This indicates that, in our particular setting, the explicit method might be preferable even through unstable. Again, the explicit method results considerably quicker than the semi-implicit one, marginally faster, yet less accurate, than the implicit one. Since the explicit method, unlike the implicit one, imposes stability constraints on the time step, as the underlying's volatility increases, we expect the CFL constraint to become stricter. Actually, we observe instability also with the implicit scheme (see lines $1-5$ Table 4.3), due to a space step problem, while in the explicit scheme the problem is due to the time step. These considerations are worth investigation

[3] All tests have been performed by using Matlab on an Intel(R) Xeon(R) CPU E3-1241 3.50 GHz computer.
[4] We use the Matlab function "fsolve" to implement the implicit and semi-implicit approaches.

K=90							
P.Brigo	**16.4534**						
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(16.4405;16.4583)			**16.4494**		416	
N_t	N_x	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
100	30	0.579	**16.454577**	0.723	**16.454279**	1.209	**16.454577**
1000	90	0.544	**16.464324**	11.56	**16.464295**	12.64	**16.464324**
3000	150	1.16	**16.455686**	81.5	**16.455677**	81.88	**16.455686**
K=100							
P.Brigo	**11.2858**						
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(11.3064;11.3160)			**11.3112**		437.36	
N_t	N_x	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
100	30	0.137	**11.207271**	0.71	**11.206931**	0.90	**11.207271**
1000	90	0.67	**11.290589**	12.3	**11.290555**	15.5	**11.290589**
3000	150	1.954	**11.296083**	87.6	**11.296071**	89.1	**11.296083**
K=110							
P.Brigo	**7.4999**						
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(7.5416;7.5463)			**7.5439**		448.22	
N_t	N_x	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
100	30	0.250	**7.498880**	1.12	**7.498544**	1.23	**7.498880**
1000	90	0.460	**7.519851**	25.18	**7.519817**	23.02	**7.519851**
3000	150	1.648	**7.513870**	84.5	**7.513859**	83.45	**7.513870**

Table 4.2: We compare explicit, semi-implicit and implicit methods with Monte Carlo simulations and with [16] results in the case of deterministic intensities.

and they might bring up new lines of research for future work. Nevertheless, in Table 4.3, we show this has no impact on our issue as long as an appropriately dense mesh is selected.

In the case of stochastic intensities, we additionally set the following values for the parameters of the CIR processes

$\gamma_i = 0.02, \quad \psi_i = 0.161, \quad \eta_i = 0.08, i = C, I.$

In Table 4.4, we report the results of our method with the corresponding computational times. To the best of our knowledge, in the literature, we could not find numerical methods covering this general case, so we had to resort again to Monte Carlo simulations to provide a benchmark.

As shown in Table 4.4, Monte Carlo simulations give 16.4416 in about 9 minutes, while with 30 nodes for each space interval, we get a result close to the benchmark in less than a second. To stabilize the first two decimal digits, we increased the spatial nodes to 100, still with a very reasonable computational time. To achieve better accuracy, we increased the number of spatial and time nodes even further, inevitably paying a cost in terms of time machine. Certainly Monte Carlo simulations may be optimized, nevertheless a our approach provides consistent improvement in machine.

The Table 4.5 compare the results of Table 4.1 with a symmetric finite difference for the ap-

$\sigma = 60\%$							
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(21.22770;21.26145)			**21.2445788**		459.6	
N_t	N_x	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
100	30	0.166	**8.46E+24**	0.89	**0.000000**	0.9	**0.000000**
1000	90	0.5	**20.854963**	20.74	**10.104217**	26.16	**20.854963**
2000	90	0.8	**21.328809**	19.93	**21.328788**	21.16	**21.328809**
3000	150	1.15	**20.854866**	117.09	**10.003868**	105.93	**10.003868**
6000	150	2.10	**21.323061**	155.5	**21.323054**	140.8	**21.323061**
$\sigma = 40\%$							
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(16.4387;16.4567)			**16.44777**		433.06	
N_t	N_x	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
100	30	0.150	**16.454577**	0.59	**16.454279**	0.79	**16.454577**
1000	90	0.420	**16.464324**	10.69	**16.464295**	11.49	**16.464324**
3000	150	1.160	**16.455686**	78.5	**16.455677**	76.57	**16.455686**
$\sigma = 30\%$							
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(14.0337; 14.0457)			**14.03972**		476.8	
N_t	N_x	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
100	30	0.158	**14.054061**	0.57	**14.053846**	0.83	**14.054061**
1000	90	0.420	**14.071745**	11.69	**14.071724**	12.15	**14.071745**
3000	150	1.18	**14.060971**	75.87	**14.060971**	75.21	**14.060978**

Table 4.3: Comparison between explicit, semi-implicit and implicit methods for various volatilities in the case of deterministic intensities.

Monte Carlo simulation			
N_t	Seconds	Price	confidence interval
1000	527	16.4416729	(16.433;16.4506)
Method of lines			
N_t	N_{xyz}	Seconds	Price
100	30;30;30	0.66	16.42897944
500	50;50;50	4.16	16.42209551
1000	90;90;90	17.68	16.46349986
1500	100;100;100	46.32	16.45064167
2000	120;120;120	173.43	16.45825752
5000	150;150;150	487.6	16.45475883

Table 4.4: Prices of a European call with maturity 6 months and stochastic intensities, $\rho_1 = \rho_2 = 0$ with explicit scheme.

proximation of the first spatial derivative. This new discretization may certainly give benefits in terms faster, but there is no significant gain in accuracy.

		Forward finite difference		Symmetric finite difference	
N_t	N_{xyz}	Seconds	Price	Seconds	Price
50	10	0.226	16.16472574	0.150	16.41309
100	30	0.66	16.42897944	0.320	16.48178
500	50	4.16	16.42209551	1.630	16.46791
1000	90	17.68	16.46349986	12.800	16.51000
1500	100	46.32	16.45064167	32.090	16.49799

Table 4.5: Prices of a European call with maturity 6 months and stochastic intensities, $\rho_C = \rho_I = 0$ with classic and symmetric finite difference for the approximation of the first spatial derivatives.

4.5 Sensitivity analysis

In this section, we run a short sensitivity analysis for our method in the case of stochastic intensities. This is done employing the explicit Euler scheme. Indeed, again a similar accuracy is achieved by both the explicit and the implicit scheme, but with much larger computational times for the second. Indeed a better accuracy can be obtained by increasing the number of spatial and temporal nodes, but it becomes prohibitive timewise when applying the implicit scheme. We further remark that when using the explicit scheme, the increase of computational times is due solely to the thickening of the spatial nodes, while they remain stable (about 1 second) as the number of temporal nodes increases.

In Table 4.6, we compare the explicit, semi-implicit and implicit Euler schemes. Especially, when using an semi-implicit and implicit schemes, computational times grow considerably when increasing the number of spatial nodes. Hence we were forced to keep the number of spatial nodes equal to 15 with consequently far less accuracy.

The Table 4.6 emphasizes that, despite the explicit technique's potential instability, it is precise and extraordinarily fast in solving this particular problem.

From Table 4.4 one might conclude that the introduction of randomness for the intensities did really affect the price. To understand whether this was due to the particular choice of parameters or it was a general feature, fixing 100 spatial nodes, we performed a short sensitivity analysis, with respect to the intensity parameters, maturity, and strike price. In Table 4.7 we consider a European call option with different maturity (six months, nine months, and one year) and different strike prices and we compared the results with the constant intensities case (taking the initial value of the CIR processes), to underline the effect of introducing randomness for the intensities. We used the explitic scheme for this comparison.

As expected, the price appears to be decreasing with respect to the strike price, and increasing with respect to maturity. Table 4.7 shows also that the randomness of the intensities affects the price up to the first decimal digit when maturity increases, confirming it might be significant to consider stochastic intensities models for longer maturities.

Fixing $S = 100, K = 90, T = 0.5, \alpha = 0.5, \rho_C = \rho_I = 0$, we also explored the sensitivity of the model varying the intensities parameters of λ^C and λ^I. Being the derivative a call, the most relevant effect comes, as it is to be expected, by the parameters (regression speed and long term

K=90							
$\sigma = 60\%$							
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(21.1139; 21.1475)			**21.130690**		598.69	
N_t	N_{xyz}	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
50	10	0.15	**20.8416**	19.80	**20.9650**	37.76	**20.7396**
100	15	0.24	**21.4150**	320.23	**21.5437**	759.70	**21.3704**
500	15	0.74	**21.3971**	1154.24	**21.3977**	2757.91	**21.3882**
$\sigma = 40\%$							
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(16.433;16.4506)			**16.4416729**		527	
N_t	N_{xyz}	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
50	10	0.20	**16.1647**	21.20	**16.2614**	42.23	**16.1024**
100	15	0.21	**16.5132**	330.05	**16.6135**	599.97	**16.4902**
500	15	0.67	**16.5039**	1165.80	**16.6044**	1872.27	**16.4993**
$\sigma = 30\%$							
N_t	Confidence Interval			**Price MC**		Seconds	
1000	(14.0511;14.0631)			**14.05715**		587.5	
N_t	N_{xyz}	Seconds	**Explicit**	Seconds	**Semi-Implicit**	Seconds	**Implicit**
50	10	0.14	**13.9204**	20.08	**14.0037**	44.38	**13.8786**
100	15	0.21	**14.2040**	322.50	**14.2903**	354.83	**14.1904**
500	15	0.93	**14.1985**	1174.17	**14.2849**	5186.40	**14.1957**

Table 4.6: Prices of a European call with different volatility, with explicit, semi-implicit and implicit schemes in the stochastic case.

6 months		*9 months*		*1 year*	
K=90					
16.4518463	16.450641	18.6724165	18.6835031	20.4575975	20.554191
K=100					
11.292111115	11.291427	13.7820539	13.7948775	15.76652711	15.868368
K=110					
7.51052959	7.5102467	10.0105913	10.0248779	12.0356766	12.142071

Table 4.7: Prices of a European call with explicit scheme with different maturities (6 months, 9 months, and 1 year) and strike prices $(90, 100, 110)$, in the deterministic and stochastic case.

average) of the counterparty's default intensity, while the investor's intensity parameters influence the price almost irrelevantly (Table 4.8). Finally, in Table 4.9, we show how the volatility affects the explicit method's convergence.

4.6 Conclusion

In this work, we developed a simple approximation procedure for the adjusted value of a derivative contract subject to counterparty risk, collateralization and founding costs, assuming a diffusion model for the default intensities and close-out values as a portion of the adjusted price itself. This

$\eta_C = 0.08$					
$\psi_C = 0.161$		$\psi_C = 0.25$		$\psi_C = 0.4$	
γ_C	Price	γ_C	Price	γ_C	Price
0.02	16.4510	0.02	16.4499	0.02	16.4480
0.03	16.4502	0.03	16.4486	0.03	16.4458
0.05	16.4488	0.05	16.4459	0.05	16.4415
0.1	16.4451	0.1	16.4397	0.1	16.4307
0.2	16.4380	0.2	16.4274	0.2	16.4096
$\eta_I = 0.08$					
$\psi_I = 0.161$		$\psi_I = 0.25$		$\psi_2 = 0.4$	
γ_I	Price	γ_I	Price	γ_I	Price
0.02	16.45102548	0.02	16.451025498	0.02	16.4510255148
0.03	16.45102549	0.03	16.451025512	0.03	16.4510255277
0.05	16.45102551	0.05	16.451025528	0.05	16.4510255387
0.1	16.45102553	0.1	16.451025541	0.1	16.4510255434
0.2	16.45102554	0.2	16.451025544	0.2	16.4510255436

Table 4.8: Sensitivity analysis for different regression speeds and fixed long term averages with explicit scheme.

$\gamma_C = \gamma_I = 0.1$ $\psi_C = \psi_I = 0.05$			
K=90			
$\sigma = 40\%$	$N_t = 1500$	$N_t = 2500$	$N_t = 3000$
η_C/η_I	0.08	0.1	0.2
0.08	16.42465705	16.42431168	16.34466761
0.1	16.42466191	16.42431655	16.34467243
0.2	NaN	16.42160683	16.34200124
$\sigma = 60\%$	$N_t = 2500$	$N_t = 3000$	$N_t = 4500$
η_C/η_I	0.08	0.1	0.2
0.08	21.39176544	21.39161946	21.28965302
0.1	21.39177163	21.39162566	21.28964689
0.2	NaN	21.38742986	21.28551264

Table 4.9: Sensitivity analysis for different volatilities with explicit scheme.

generate a non-linear BSDE, with an associated non-linear PDE characterizing the price.

By the simple method of lines applied to this PDE, we showed that accurate approximations could be achieved in very manageable computational times, differently from what happens when employing Monte Carlo simulations. We ran a short sensitivity analysis to estimate the effects of the introduction of stochastic intensities.

Bibliography

[1] F. Antonelli, R. D'Ambrosio, and I. Gallo. Analysis of non linear approximated value equation under multiple risk factors and stochastic intensities. *Computers and Mathematics with Applications, https://doi.org/10.1016/j.camwa.2023.03.014*, 2023.

[2] F. Antonelli, A. Ramponi, and S. Scarlatti. Approximate value adjustments for european claims. *European Journal of Operational Research*, 300(3):1149–1161, 2022.

[3] I. Arregui, B. Salvador, and C. Vázquez. Pde models and numerical methods for total value adjustment in european and american options with counterparty risk. *Applied Mathematics and Computation*, 308:31–53, 2017.

[4] U. M. Ascher, S. J. Ruuth, and R. J. Spiteri. Implicit-explicit runge-kutta methods for time-dependent partial differential equations. *Applied Numerical Mathematics*, 25(2-3):151–167, 1997.

[5] T. R. Bielecki, M. Jeanblanc, and M. Rutkowski. Credit risk modeling. osaka university csfi lecture notes series 2, 2009.

[6] T. R. Bielecki and M. Rutkowski. *Credit risk: modeling, valuation and hedging*. Springer Science & Business Media, 2013.

[7] J.-M. Bismut. Théorie probabiliste du contrôle des diffusions. *American Mathematical Soc.*, 167, 1976.

[8] L. Bo, A. Capponi, and P.-C. Chen. Credit portfolio selection with decaying contagion intensities. *Mathematical Finance, Forthcoming*, 2017.

[9] L. Bo and C. Ceci. Locally risk-minimizing hedging of counterparty risk for portfolio of credit derivatives. *Applied Mathematics & Optimization*, 82(2):799–850, 2020.

[10] M. Briani, R. Natalini, and G. Russo. Implicit–explicit numerical schemes for jump–diffusion processes. *Calcolo*, 44(1):33–57, 2007.

[11] D. Brigo. Counterparty risk faq: credit var, pfe, cva, dva, closeout, netting, collateral, re-hypothecation, wwr, basel, funding, ccds and margin lending. *http://dx.doi.org/10.2139/ssrn.1955204*, 2011.

[12] D. Brigo, C. Buescu, M. Francischello, A. Pallavicini, and M. Rutkowski. Nonlinear valuation with xvas: two converging approaches. *Mathematics*, 10(5):791, 2022.

[13] D. Brigo, A. Capponi, and A. Pallavicini. Arbitrage-free bilateral counterparty risk valuation under collateralization and application to credit default swaps. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 24(1):125–146, 2014.

[14] D. Brigo, A. Capponi, A. Pallavicini, and V. Papatheodorou. Pricing counterparty risk including collateralization, netting rules, re-hypothecation and wrong-way risk. *International Journal of Theoretical and Applied Finance (IJTAF)*, 16(02):1–16, 2013.

[15] D. Brigo and K. Chourdakis. Counterparty risk for credit default swaps: Impact of spread volatility and default correlation. *International Journal of Theoretical and Applied Finance*, 12(07):1007–1026, 2009.

[16] D. Brigo, M. Francischello, and A. Pallavicini. Nonlinear valuation under credit, funding, and margins: Existence, uniqueness, invariance, and disentanglement. *European Journal of Operational Research*, 274(2):788–805, 2019.

[17] D. Brigo, Q. Liu, A. Pallavicini, and D. Sloth. Nonlinear valuation under collateral, credit risk and funding costs: a numerical case study extending black-scholes. *Handbook in Fixed-Income Securities, Wiley*, 2014.

[18] D. Brigo, Q. D. Liu, A. Pallavicini, and D. Sloth. Nonlinearity valuation adjustment: Nonlinear valuation under collateralization, credit risk, and funding costs. *Innovations in Derivatives Markets: Fixed Income Modeling, Valuation Adjustments, Risk Management, and Regulation*, 3–35, 2016.

[19] D. Brigo and M. Masetti. Risk neutral pricing of counterparty risk. *Counterparty Credit Risk Modeling: Risk Management, Pricing and Regulation. Risk Books*, 2005.

[20] D. Brigo, M. Morini, and A. Pallavicini. *Counterparty credit risk, collateral and funding: with pricing cases for all asset classes*, volume 478. John Wiley & Sons, 2013.

[21] D. Brigo and A. Pallavicini. Counterparty risk pricing under correlation between default and interest rates. *Numerical methods for finance*, 63:79–98, 2007.

[22] C. Burgard and M. Kjaer. Partial differential equation representations of derivatives with bilateral counterparty risk and funding costs. *The Journal of Credit Risk*, 7(3):1–19, 2011.

[23] U. Cherubini. Counterparty risk in derivatives and collateral policies: the replicating portfolio approach. *ALM of Financial Institutions. Institutional Investor Books*, 2005.

[24] E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression method for american option pricing. *Finance and Stochastics*, 6:449–471, 2002.

[25] S. Crépey. Bilateral counterparty risk under funding constraints—part i: Pricing. *Mathematical Finance*, 25(1):1–22, 2015.

[26] S. Crépey. Bilateral counterparty risk under funding constraints—part ii: Cva. *Mathematical Finance*, 25(1):23–50, 2015.

[27] S. Crépey, T. R. Bielecki, and D. Brigo. *Counterparty risk and funding: A tale of two puzzles*. Chapman and Hall/CRC, 2014.

[28] R. D'Ambrosio, S. D. Giovacchino, and D. Pera. Parallel numerical solution of a 2d chemotaxis-stokes system on gpus technology. *International Conference on Computational Science*, 59–72, 2020.

[29] R. D'Ambrosio, M. Moccaldi, and B. Paternoster. Adapted numerical methods for advection–reaction–diffusion problems generating periodic wavefronts. *Computers & Mathematics with Applications*, 74(5):1029–1042, 2017.

[30] N. El Karoui, E. Pardoux, and M. Quenez. American options. *Numerical methods in finance*, 13:215, 1997.

[31] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[32] I. Gallo. Non-linear approximated value adjustments for derivatives under multiple risk factors. *International Conference on Computational Science and Its Applications*, 13376:217–227, 2022.

[33] P. V. Gapeev, M. Jeanblanc, L. Li, and M. Rutkowski. Constructing random times with given survival processes and applications to valuation of credit derivatives. *Contemporary quantitative finance*, 255–280, 2010.

[34] K. Glau, Z. Grbac, M. Scherer, and R. Zagst. *Innovations in derivatives markets: fixed income modeling, valuation adjustments, risk management, and regulation*. Springer Nature, 2016.

[35] A. Green. *XVA: credit, funding and capital valuation adjustments*. John Wiley & Sons, 2015.

[36] J. Gregory. *Counterparty credit risk and credit value adjustment: A continuing challenge for global financial markets*. John Wiley & Sons, 2012.

[37] T. Haentjens and K. J. In't Hout. Alternating direction implicit finite difference schemes for the heston-hull-white partial differential equation. *The Journal of Computational Finance*, 16(1):83, 2012.

[38] E. Isaacson and H. B. Keller. *Analysis of numerical methods*. Courier Corporation, 2012.

[39] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical methods for financial markets*. Springer Science & Business Media, 2009.

[40] R. J. LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*. SIAM, 2007.

[41] F. A. Longstaff and E. S. Schwartz. Valuing american options by simulation: a simple least-squares approach. *The review of financial studies*, 14(1):113–147, 2001.

[42] J. Ma and J. Yong. *Forward-backward stochastic differential equations and their applications*. Number 1702. Springer Science & Business Media, 1999.

[43] A. Nikeghbali. An essay on the general theory of stochastic processes. *Probability Surveys*, 3:345–412, 2006.

[44] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

[45] W. E. Schiesser. *The numerical method of lines: integration of partial differential equations*. Elsevier, 2012.

[46] W. E. Schiesser and G. W. Griffiths. *A compendium of partial differential equation models: method of lines analysis with Matlab*. Cambridge University Press, 2009.

[47] G. D. Smith, G. D. Smith, and G. D. S. Smith. *Numerical solution of partial differential equations: finite difference methods*. Oxford university press, 1985.

[48] Z. Wu. *Pricing American options using Monte Carlo method [Master's thesis]*. University of Oxford, 2012.

[49] J. Zhang. *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*, volume 86. Springer, 2017.