Abstract

In the wake of the 2008 financial crisis, regulatory capital requirements for banks have increased significantly through Basel III. As this raised awareness for the capital burden on the derivative businesses of banks, demand has grown for models that assess the costs of holding capital. Amidst a recent trend of pricing valuation adjustments, known as XVAs, a valuation adjustment has been developed that captures precisely this capital cost: the Capital Valuation Adjustment\(^1\). In this thesis, two approaches to modeling KVA are studied and compared. Although the models have different mathematical fundamentals, the resulting KVA formulae are surprisingly similar. Both allow for Monte Carlo simulation of regulatory capital profiles to calculate KVA numbers. A computer implementation is considered, for both the existing and future regulatory landscape.

\(^1\)Capital Valuation Adjustment is often abbreviated as KVA, where the K stands for the German word Kapital, to avoid ambiguity with the term Credit Valuation Adjustment.
Preface

This thesis is written in partial fulfillment of the requirements for the master’s degree Stochastics and Financial Mathematics at the University of Amsterdam. After 18 months of taught courses, the master program expects a student to conduct individual research on an advanced topic of choice, either at university or in cooperation with a company. As I personally felt very eager to learn about practical implications of the stochastic financial models I had been taught, I chose to do the latter. The forthcoming report is the result of a half-year internship at a bank.

A bank is a very interesting place from a financial mathematics perspective. It forms the heart of the derivatives business, where technical aspects collide with IT infrastructure and business decisions. Being an intern at the bank broadened my perspective on derivatives and made me realise there is much more to them than models. It brought me to the conclusion that I want to pursue a career in the financial markets. Moreover, I greatly enjoyed the mathematical technicalities as well. Hence, I am grateful for the opportunity to write my thesis at such a place.

I would like to thank everyone who helped me conduct this research. First of all, my university supervisor Peter Spreij, for his overall guidance and for helping me work through the more theoretical side of the project. As I spent most of my days at the bank, I moreover owe a great debt of gratitude to Raoul Pietersz and Matteo Michielon. Both supervised me on a daily basis and showed an endless amount of patience and knowledge answering my questions. Apart from the substantive matters, they helped me navigate through the bank and to orient myself on a further career. I also owe a lot of insights to Leo Kits, who brought the project in perspective and taught me how the bank’s business works. Along with Raoul and Matteo, I would like to thank the entire team for being such a fun and inspiring group of people to work with. Moreover, I would like to thank Erik Winands for willing to act as a second reader of my thesis.

Lastly, this thesis project marks the end of a two-year journey through financial mathematics. I would like to thank my friends who shared the enthusiasm about the subject along the way and who kept things lively when I needed it. In particular, I want to thank Bastiaan Frerix for being a supportive companion from day one. Finally, I would not be where I am today without the everlasting support of my parents, both morally and financially. Thank you.

Wessel Martens
September 15, 2018, Amsterdam
# Contents

Introduction ................................................. 7

1 Capital .................................................. 10
   1.1 Economic capital ...................................... 12
       1.1.1 Fundamentals of credit risk modeling .......... 12
       1.1.2 The Asymptotic Single Risk Factor Model .... 16
   1.2 Regulatory capital .................................... 20
       1.2.1 Counterparty Credit Risk capital .............. 23
       1.2.2 CVA capital ................................... 28

2 Capital Valuation Adjustment ......................... 30
   2.1 The Semi-replication model ......................... 31
       2.1.1 Stochastic Differential Equations ............. 31
       2.1.2 The Semi-replication formula .................. 34
   2.2 The BSDE model ..................................... 40
       2.2.1 Backward Stochastic Differential Equations ... 40
       2.2.2 The BSDE formula ................................ 45

3 Implementation .......................................... 51
   3.1 Interest rate modeling ............................... 52
       3.1.1 Fundamentals of Interest Rate Modeling ....... 52
       3.1.2 The Libor Market Model ....................... 55
   3.2 Monte Carlo simulation .............................. 60
       3.2.1 Least Squares Monte Carlo .................... 62
       3.2.2 Exposure and capital profiles .................. 64
   3.3 The KVA ............................................. 68
       3.3.1 Non-capital parameters ......................... 68
       3.3.2 The KVA integral ................................ 68

4 Results .................................................. 70
   4.1 Introduction: a case study ......................... 71
   4.2 Impact of model parameters ......................... 74
       4.2.1 Impact under CEM regulations .................. 74
       4.2.2 Impact under SACCR regulations ............... 77
   4.3 Practical considerations ............................. 79

5 Conclusion .............................................. 83

Populaire Samenvatting .................................. 84

Appendix A: Terms and Acronyms ....................... 85

Appendix B: The ASRF Model ............................. 89
Introduction

Before the financial crisis of 2008, derivatives valuation worked very differently from how it does today. Since the 1973 Black-Scholes-Merton seminal papers the pricing of financial derivatives had been centered in the no-arbitrage framework, based on a spectrum of assumptions that did not reflect reality properly. Amongst others, it considered the existence of a risk-free rate and the endless and immediate availability to trade products. Even then it was clear that this could not hold, and over the years efforts were made to relax many of the simplifying assumptions. Rigorous measure-theoretic fundamentals replaced the original mathematical setup, but the essential framework of risk-neutral pricing remained largely unchanged up until the crisis.

The collapse of Lehman Brothers in 2008 demonstrated that the concept of “too big to fail” was fictitious and that no counterparty would ever be free of default risk. Before, such risk was taken into account via an income deferral based on historic default probabilities. A change in regulations required banks to value their derivative books based on so-called exit prices, that aim to reflect the value at which another market participant would price in counterparty default risks. This lead to the use of credit spreads opposed to historical data and ultimately, an industry-uniform price add-on known as Credit Valuation Adjustment (CVA). It allowed banks to hedge their counterparty risks on a bank-wide scale. On the other hand, the implication that the issuer’s credit risk should then also be taken into account, was reflected in the Debt Valuation Adjustment (DVA). In the aftermath of the Lehman Brothers debacle, the spread between the three-month Libor and the OIS rate blew up\(^2\), indicating that liquidity risk could no longer be ignored either. Banks were much more hesitant to lend interbank, leading to funding costs that are imperative for their derivative businesses. It became clear that alongside CVA, such funding costs should be quantified in the form of a Funding Valuation Adjustment (FVA).

As a consequence, the exit prices of derivative trades post-crisis look very different from those pre-crisis. The pricing of a vanilla interest rate swap has changed from a single yield curve discount model to a discount curve construction and multiple yield curve projections for the baseline valuation, and an extensive Monte Carlo simulation framework to calculate the various valuation adjustments, often abbreviated as XVA. An overview can be found in Table 0.1 below.

<table>
<thead>
<tr>
<th>Pre-crisis</th>
<th>Post-crisis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk-neutral price (Libor discounting)</td>
<td>Risk-neutral price (OIS discounting)</td>
</tr>
<tr>
<td>Operational and hedge costs</td>
<td>Operational and hedge costs</td>
</tr>
<tr>
<td>CVA (historical)</td>
<td>CVA and DVA (credit spreads)</td>
</tr>
<tr>
<td></td>
<td>FVA (KVA)</td>
</tr>
</tbody>
</table>

Table 0.1: A comparison between pre-crisis and post-crisis derivative pricing.

\(^2\)The spread grew to over twenty times its average a year earlier, prompting banks to switch to multiple projection curves for rates of different tenors. As spreads were now no longer negligible, banks started to use the OIS curve for discounting, as the three-month Libor became inappropriate.
Apart from the conclusion that risks were not adequately addressed by the existing framework, the financial crisis lead to the realisation that banks must be subject to much stricter regulation and capital requirements. The collapse of Lehman Brothers did not only demonstrate “too big to fail” to be an obsolete phrase; it moreover exhibited the severe consequences of an actual failure of a financial institution that is vital to the system. As it turned out, the regulations in place at the time allowed for insufficient capital levels, excessive leverage and systematic risk. It is therefore not very surprising that post-crisis, new regulations started to emerge very quickly. The regulators, via the Dodd-Frank Act in the US and Basel III in Europe, imposed much stricter regulations to provide stability in over-the-counter derivatives markets. In order to strengthen the capital bases, Basel III introduced new liquidity and leverage measures. As the largest part of losses in the financial crisis came from changes in credit worthiness, rather than actual defaults, the notable CVA risk capital charge was introduced. Finally, the regulators set up advantages for derivatives traded via centralised clearing counterparties, in order to mitigate risk.

Although the new regulations provide more stability in the financial system, the capital requirements imposed pressure on the banks’ derivatives business, as regulatory capital now became a significant cost. Holding regulatory capital has a cost, because shareholders expect a return on their invested equity. Historically, banks have implicitly charged for capital by setting a limit on the amount of capital a trade is allowed to consume. Conforming to the development of the XVA framework\(^3\) in recent years, however, banks have sought more sophisticated ways to incorporate this cost into exit prices. In response, academia developed a model for what is now known as Capital Valuation Adjustment: a valuation adjustment to account for the cost of capital, KVA\(^4\).

The first model to formalise the notion of KVA was developed in 2014 by Andrew Green and Chris Kenyon, who extended an existing all-round XVA model to incorporate the cost of capital. In the following years, the subject was picked up in the financial mathematics community and various ‘challenger models’ came to life, and some are still in development at the time of writing. As will be elaborated on later, a wide range of model approaches currently exist, due to the lack of consensus on the exact purpose and definition of KVA. Feedback from the industry is rather ambiguous, leading to different approaches. The original model considered a replication approach, whereas later models deploy an expectation derivation or a full balance sheet approach.

The aim of this report is to demonstrate the mathematical foundations and a potential computer implementation of two KVA models: the original model by Green and Kenyon (GK), and one of the latest models by Albanese and Crépey (AC). This report is organised as follows. A background on capital modeling and the mathematical foundation of the existing capital requirements, including their exact specifications, can be found in Chapter 1. The aim of this chapter is to provide the reader broader insight into what capital is and how it is modeled, but is not strictly necessary to understand KVA. Chapter 2 describes two models to price capital into a derivative transaction. The former, due to Green and Kenyon, makes use of Stochastic Differential Equations, which are briefly reviewed before diving into the derivation of KVA. The latter model deploys Backward Stochastic Differential Equations, which are stochastic equations that run backward in time. Although in general an explicit solution to such equations does not exist, we are lucky in the case of KVA. It should be noted that the model presented here is not the exact AC model, but a modification of it tailored to regulatory (and not economic) capital.

Once the theoretical KVA formulae have been established, a computer implementation for

---

\(^3\)XVA refers to the collection of valuation adjustments, where X is replaces the various indication letters.

\(^4\)As can be seen in Table 0.1, the term is generally abbreviated as KVA, to avoid ambiguity with CVA.
in particular interest rate swaps will be considered in Chapter 3. Due to its generic setup, it can easily be extended to any (reasonable) kind of derivative. As the report considers interest rate swaps, the Libor Market Model is briefly introduced first. Subsequently the Monte Carlo methodology, including the Longstaff Schwartz regression algorithm, is described. The expected capital profiles, as required for KVA, are numerically integrated to find the final KVA values. The thesis concludes with results of the computer implementation in Chapter 4, where various practical aspects of KVA are demonstrated. The final chapter concludes with some remarks on this implementation, as well as potential issues and recommendations for further research.
1 Capital

Capital is more important for banks than ever, as it has become a scarce resource after the crisis, when many banks had to reallocate or even reduce their portfolios (McKinsey, 2011 [4]). As a consequence, capital management is enjoying a renaissance in modern banking, where capital has a twofold function: first, capital acts as a risk measure. The credit worthiness of a bank is intrinsically related to its capital levels. Indeed, capital acts as a buffer to absorb losses in turbulent times. In that sense, a bank aims to hold as much capital as possible. On the other hand, banks strive for maximum profit, and naturally will want to deploy their utmost amount of capital for business. Indeed, capital secondly serves the purpose of a performance measure, via return on equity. Any bank naturally wants to have a balance between risk and profit, which is safeguarded by capital\(^1\).

Capital levels for risk are generally generated from economic capital models. A comparison between the expected and unexpected loss often forms the core of such a model. Consider a bank with potential portfolio returns over a given risk horizon distributed as in Figure 1.1. In order to generate a long-term profit, the investment portfolio must have an expected return that is higher than its expected loss. However, as this refers to the profit on average, it may happen that over some risk horizons, the loss is greater than the return, and the bank requires capital to remain solvent. A proper way to set this capital level, often denoted economic capital (EC), is to take the difference between the expected loss and a very low quantile of this return distribution, dubbed unexpected loss, such that with high probability potential ‘temporary’ losses can be covered.

Figure 1.1: A hypothetical portfolio return distribution, to calculate economic capital. (Ruiz, 2015 [38]).

As this is merely an illustration of the economic capital concept, one can imagine the real difficulty arising in obtaining a portfolio return (or loss) distribution, by estimating various stand-alone quantities as well as asset correlations. In practice, simplifications are inevitable, and economic capital modeling is a large field of study in itself (c.f. Lütkebohmert, 2009 [26]). Section 1.1 of this chapter gives a brief overview of how economic capital modeling works. The

\(^1\)As it will turn out, in derivatives pricing, the KVA price add-on is precisely the tool to obtain this balance.
concept of economic capital is formalised by example of the ASRF model (Gordy, 2002, [16]), where simplification is made through the assumption of a single (global) underlying risk factor for all portfolio returns. Under this assumption, the notion of economic capital as described above, turns out to be mathematically well defined.

Capital levels are also key in measurement of performance and risk taking. Banks with a higher risk appetite will have a much wider distribution of returns, cf. Figure 1.1, as they are willing to take bigger swings in their profit and loss. Accordingly, their economic capital levels will be higher. The performance of a portfolio, or bank as a whole, is then measured via the Risk-adjusted Return on Capital\(^2\) (RaRoC), given by the fraction of profit and capital usage as

\[
\text{RaRoC} = \frac{\text{expected profit}}{\text{economic capital}} = \frac{\text{profit} - \text{expected loss} - \text{expenses}}{\text{economic capital}}.
\]

Although banks may have different risk appetites, this number provides a natural balance between return and risk. Moreover, it allows a bank to assess internally the (capital relative) performance of different business lines and reallocate capital accordingly (Ruiz, 2015 [38]).

The supervisory bodies that design the post-crisis framework to calculate economic capital in a standardised way face the very same risk-return balance: capital requirements must be “high enough to contribute to a very low possibility of failure, but not so severe as to unfairly penalise the bank and create adverse consequences for their clients and ultimately the economy as a whole” (Gregory, 2016, [19]). Another challenge that arises is the complexity of the requirements they impose. A simple approach would be transparent and well implementable, but may be too narrow to capture more than a few key risk aspects of a complex system. On the contrary, frameworks that assess risks more properly are often hard to implement, in particular for smaller banks which may not have the appropriate resources readily available. As a consequence, regulators attempt to compromise and provide multiple layers of complexity to accommodate all banks.

The regulations relevant for financial derivatives, as set out by the Basel committee after the recent financial crisis (BCBS, 2011, [32]), will be exhibited in Section 1.2. Accommodating all banks, the capital requirements are separated according to risk type and complexity level, resulting in a grid of calculation methodologies a bank might face. Most of the credit related risk calculations connect to economic capital models from the preceding section. Although the technical origin of these regulatory calculations in Section 1.1 is not strictly relevant for KVA, the calculation methods in Section 1.2 themselves are key to understanding the potential issues for KVA calculations later on.

\(^2\)Or sometimes, simply the Return-on-Equity (RoE), which is a return measure insensitive to risk.
1.1 Economic capital

1.1.1 Fundamentals of credit risk modeling

As seen from the introduction of this section, capital management is essential in modern day banking. A prominent factor in capital management is risk, in particular credit risk. But what actually is credit risk? The European Central Bank (ECB) refers in its glossary to credit risk as “the risk that a counterparty will not settle an obligation in full - neither when it becomes due, nor at any time thereafter”. Traditionally, this applied to loans and bonds. Holders of debt were afraid that their counterparty would default on a payment and as a consequence incur losses. In this section, the basics of credit risk modeling are presented much along the lines of (Lütkebohmert, 2009, [26]).

It is clear from the above definition that credit risk entails uncertainty. The main purpose of credit risk modeling is then to assess, in a probabilistic setting, the likelihood of defaults. The severeness of such events depend on several variables of risk. First and foremost, default events are generally quite rare and occur unexpectedly. The uncertainty whether an obligor will default or not, is measured by its Probability-of-default (PD). The probability is specified over a given risk horizon, typically one year, to allow for comparisons. Although default events occur very seldom, the probability that a certain obligor might default is rarely zero. We have seen recently even very high rated borrowers default on their financial obligations\(^3\). Closely related to the default probability is also the notion of migration risk, i.e. the risk of losses due to changes in credit rating (in fact, default probability) of a counterparty. See Section 1.2.2.

Conditional on an obligor’s default, the resulting loss might be very significant. The Exposure-at-default (EAD) is the total value of the financial obligations to the creditor, for example the bank, at the moment of default. There is a chance the obligor will partly recover, meaning that the creditor might receive a fraction of the notional value of the claim. The recovery risk describes this uncertainty about the severity of the loss and the fraction is denoted by the Recovery-rate. In order to calculate portfolio losses, one generally works with its complement: the Loss-given-default (LGD). The LGD denotes the percentage of the exposure that is lost upon default of the counterparty. Combining the three variables, one can define the portfolio loss in a formal setup.

**Definition 1.1.** Let \( P = \{1,...,n\} \) be a credit portfolio of \( n \) obligors, each with Exposure-at-default \( \delta_i \), Loss-given-default \( \eta_i \) and Default-event \( D_i \). The portfolio loss is the random variable

\[
PL_n = \sum_{i=1}^{n} \delta_i \eta_i 1_{D_i},
\]

and the portfolio percentage loss, given exposure fraction \( w_i = \delta_i / \sum_{j=1}^{n} \delta_j \), is

\[
L_n = \sum_{i=1}^{n} w_i \eta_i 1_{D_i}.
\]

A fundamental assumption underlying many credit risk models, is that for any obligor \( i \) in the portfolio the EAD \( \delta_i \), LGD \( \eta_i \) and PD \( p_i \) are independent. We adapt this assumption.

\(^3\)Consider for example the fall of Lehman Brothers in September 2008, at the time the fourth-largest investment bank in the United States. Five days before the firm filed for bankruptcy, its credit rating according to Moody’s Investor Services was still A2, the second highest possible rating in Moody’s framework.
Remark 1.1. Under the above assumption, the default events of different obligors, say \(i\) and \(j\), can (and will) still be correlated. See for example equation (1.13).

The portfolio percentage loss \((1.1)\) allows for a portfolio invariant measure and is key to any credit risk model. It provides insight in the most important elements of the credit portfolio: the Expected loss (EL) and the Unexpected loss (UL) over the specified risk horizon. Formally the EL is simply the expectation of the loss distribution (as follows from the independence assumption)

\[
\mathbb{E}[L_n] = \sum_{i=1}^{n} w_i \eta_i p_i, \tag{1.3}
\]

where \(p_i\) denotes the PD for obligor \(i\), i.e. \(p_i = P(D_i)\). The expected loss represents a kind of risk premium which a bank can charge for taking the default risk of an obligor.

In order to specify the unexpected loss, we first present the probably most widely used risk measure in financial institutions: the Value-at-Risk (VaR) paradigm. It provides an estimate of the losses occurred on a credit portfolio over a given time horizon with a specific confidence level.

**Definition 1.2.** Credit Value-at-Risk at confidence level \(\alpha \in (0, 1)\) over a given risk measurement horizon, is the largest portfolio percentage loss \(l\) such that the probability of a loss \(L_n\) exceeding \(l\) is at most \((1 - \alpha)\). I.e.,

\[
\text{VaR}_\alpha(L_n) = \inf \{ l \in \mathbb{R} : P(L_n \geq l) \leq 1 - \alpha \}. \tag{1.4}
\]

Remark 1.2. In fact, the credit VaR is simply the \(\alpha\)-quantile of the loss distribution. Remember the quantile of a random variable \(X\) to be defined as \(q_\alpha(X) = \inf \{ x \in \mathbb{R} : P(X \leq x) \geq \alpha \}\).

In general VaR can be derived for different time periods and different confidence levels. The most typical values are one year and 95% or 99% respectively. Since the financial crisis of 2008, even higher values are becoming more common. The confidence level of the second Basel Accord, cf. equation (1.26), is for example 99.9%.

The VaR paradigm has a couple of drawbacks. By definition, VaR provides no information about the severity of losses that occur with a probability less than \(\alpha\), i.e. losses in the tail of the distribution of \(L_n\). If this distribution is heavy-tailed, the framework might not be so effective. Moreover, the Value-at-Risk is not a coherent risk measure, in the sense that it is not sub-additive: for two credit portfolios \(L_n^{(1)}\) and \(L_n^{(2)}\) it does not necessarily hold that

\[
\text{VaR}_\alpha(L_n^{(1)} + L_n^{(2)}) \leq \text{VaR}_\alpha(L_n^{(1)}) + \text{VaR}_\alpha(L_n^{(2)}), \tag{1.5}
\]

meaning that the VaR of the merged portfolio is not necessarily bounded from above by the sum of the individual VaR’s. Intuitively, this contradicts the diversification benefit of merging portfolios. An alternative risk measure, that we will provide but not extensively discuss, is the Expected Shortfall (ES) of a portfolio \(L_n\), denoted \(ES_\alpha\).

**Definition 1.3.** The Expected Shortfall at confidence level \(\alpha \in (0, 1)\) over a risk measurement horizon, is the average portfolio percentage loss taken over all values exceeding the \(\alpha\)-quantile. I.e.,

\[
\text{ES}_\alpha(L_n) = \frac{1}{1 - \alpha} \int_{\alpha}^{1} \text{VaR}_\alpha(L_n) du. \tag{1.6}
\]

Remark 1.3. By definition of the Expected Shortfall we have \(\text{ES}_\alpha \geq \text{VaR}_\alpha\). If \(L_n\) is integrable with a continuous distribution function, \(\text{ES}_\alpha = \mathbb{E}[L_n 1_{\{L_n \geq \text{VaR}_\alpha(L_n)\}}]\) (McNeil et al., 2005, [27]).
It is can be seen from formula (1.6) that Expected Shortfall takes into account the shape of the tail of the loss distribution, and moreover it is sub-additive (McNeil et al., 2005, [27]).

The unexpected loss of the portfolio refers to a quantification of large losses that far exceed the expected loss; where the expected loss reserve might not be appropriate. It is often defined as \( \text{VaR}_\alpha(L_n) \) for some large quantile \( \alpha \). Considering the fact that economic (credit risk) capital is held to cover unexpected default losses on the portfolio exceeding expectations, it makes sense to formalize the notion of economic capital as follows.

**Definition 1.4.** Credit risk capital at confidence level \( \alpha \in (0, 1) \) over a given risk measurement horizon, is the difference between the VaR and the expected loss, i.e.,

\[
K_\alpha(L_n) = \text{VaR}_\alpha(L_n) - \mathbb{E}[L_n].
\] (1.7)

In general, the variance of the loss distribution \( L_n \) causes the yearly default loss on a credit portfolio \( \mathcal{P} \) to often exceed its expectation \( \mathbb{E}[L_n] \). The capital level \( K_\alpha \) is defined exactly such that in \( \alpha \% \) of the cases, larger losses can be covered by economic capital.

Consider now a credit portfolio \( \mathcal{P} = \{1, ..., n\} \) of \( n \) obligors. In order to define a full credit risk model and quantify capital levels, it remains to model the three parameters of equation (1.2): the EAD, LGD and PD of each obligor. As we will work towards the Asymptotic Single Risk Factor (ASRF) model in the next section, which considers deterministic EAD and LGD, we only briefly touch upon those factors. The more relevant parameter for Section 1.1.2 is the PD.

1. The exposure-at-default. The parameter EAD quantifies the exposure of the bank to its borrower, at the moment of default. At this moment, it consists of two parts: the outstandings \( O \) and the commitments \( C \). The first denotes the portion of the exposure that is already drawn by the obligor, whereas the latter refers to commitments, that may potentially be drawn in the future. In case of default, the outstandings and future drawn commitments might be lost. As such, the EAD is given by

\[
\delta = O + CCF \cdot C,
\] (1.8)

where \( O \) denotes the amount of outstandings and \( CCF \) denotes the credit conversion factor, that expresses the percentage of the commitment \( C \) that will be drawn and outstanding at default. In practice, a bank calibrates this parameter w.r.t. the credit worthiness of the borrower and the type of product involved. In over-the-counter derivative portfolios, the exposure-at-default is defined as the sum of the replacement cost (the current net present value of the portfolio of trades) and a potential future exposure term.

2. The loss-given-default. The parameter LGD quantifies the fraction of exposure that is lost upon default of the counterparty. It is the complement of the so-called recovery rate. Although loss-given-default is a key factor of expected loss and capital, there are few successful LGD models. It turns out that LGD modeling is challenging, as recovery rates depend on many driving factors. Examples are the state of the economy (cf. Remark 1.4 below), the quality of collateral and the seniority of the bank’s claim on the assets of the obligor. As a consequence, LGD values are often modeled as a very simple function of counterparty credit rating, business sector and location. Values tend to be in the range of 40 to 80%.

---

Some authors, e.g. (Bluhm et al., 2003, [5]), refer to the unexpected loss as the variance of the loss variable \( L_n \), rather than some large quantile of it, adjusted by the expected loss.
**Remark 1.4.** In practice, one distinguishes between *regular* LGD and *downturn* LGD. The latter refers to the loss-given-default value during a ‘downturn’ in a business cycle, such as a stressed period. In general, losses from default tend to be higher in such situations.

3. The probability of default. The parameter PD quantifies the likelihood of the default of the counterparty, i.e. the event $D_i$ in equation (1.2). Generally speaking, default models can be divided into two fundamental classes: *structural* models and *reduced-form* models. Both revolve around modeling a random default time, but based on different underlying processes.

The earliest and most intuitive models are structural models. Structural models describe a firm’s likelihood of default via economic fundamentals. The model prescribes an underlying asset value process, which causes the firm to default in case it falls below some predefined default threshold. A simple example is an asset value process $V_t^{(i)}$ and default threshold (barrier) $B_i$. The default event $D_i$ is evaluated periodically and given by

$$D_i = \{V_T^{(i)} < B_i\},$$

where $T$ is the predefined risk horizon, for example one year. Assuming that the asset value process follows a geometric Brownian motion, as originally considered in (Merton, 1974, [28]), the default probability turns out to be given by the price of a European put option, i.e.,

$$p_i = P(V_T^{(i)} \leq B_i) = \Phi\left(\frac{\log(B_i/V_0^{(i)}) - (\mu V - \frac{1}{2}\sigma^2 V)T}{\sigma V \sqrt{T}}\right),$$

(1.10)

where $\mu_V$ and $\sigma^2_V$ denote the drift and volatility of the asset value process. Hence it can be seen that structural models require strong assumptions on the dynamics on a firm’s asset. Nonetheless, it is attractive to model default from a fundamental economic perspective, because of its intuitive picture and endogenous explanation for default. Therefore, the ASRF model in the forthcoming section, connects conditional and unconditional default probabilities based on a structural model.

Another approach to model default is the use of reduced-form models. Rather than modeling the economic value of a firm, reduced-form models deal with default by specifying an exogenous jump-to-default process. The default time is defined as the first jump of this process. A model of this kind often depends on a separate hazard rate processes, conditional on which default probabilities are then defined. Taking for example a Poisson process with a deterministic hazard rate, the default probability then becomes

$$p_i = P(\tau_i \leq T) = 1 - e^{-\int_0^T \lambda_{(i)} ds},$$

(1.11)

where $T$ is again the predefined risk horizon and $\lambda_s$ is the hazard rate function. Although such an indirect approach seems less reliable, this is compensated by the ease with which a reduced-form model can be calibrated to credit instrument market data. Therefore, in the derivative market context, defaults are almost always modeled this way. An example of this in the XVA (even KVA) context will be showed in Section 3.3.1.

Having established the fundamentals of credit risk modeling and economic capital, we are able to dive deeper into the model that forms the basis of the Basel regulatory capital formulas.
1.1.2 The Asymptotic Single Risk Factor Model

Underlying the Basel II regulatory formulae is the Asymptotic Single Risk Factor model (ASRF). It is an asymptotic extension of a factor model, where the values of assets are partially determined by a (single) global economic factor. The asymptotic extension relies on the assumption of a well diversified portfolio. As a consequence, the capital requirement for a set of risky loans does not depend on the portfolio decomposition, a principle called portfolio invariance. The model originates in (Gordy, 2002, [16]) and was adopted by the Basel committee in 2005. The current section gives an outline, whereas Appendix B: The ASRF Model elaborates on the details.

Consider a bank’s portfolio of \( n \) borrowers. Assume the default of obligor \( i \in \{1, \ldots, n\} \) is modeled via an asset value process \( V^{(i)}_t \) and a pre-defined default threshold \( B_i \) at the end of a risk horizon \([0, \tau]\). Assuming the asset value process follows a geometric brownian motion, the default event \( D_i \), in terms of the standardized log-returns \( r^{(i)}_t \), becomes

\[
D_i = \{ V^{(i)}_\tau \leq B_i \} = \{ r^{(i)}_\tau \leq q_i \}, \tag{1.12}
\]

where \( q_i = \Phi^{-1}(p_i) \) is the standard normal quantile in terms of default probability \( p_i \) as in (1.10).

Assume we wish to explain the firms’ successes by means of some global underlying influences. We consider the standardized log-returns \( r^{(i)}_t \), and therefore the geometric Brownian motions \( V^{(i)}_t \), as a composition of a systematic and an idiosyncratic factor. In such an approach, one is able to interpret the correlation between single loss variables in terms of global, underlying economic variables. Large losses on the portfolio are then explained via these economic factors.\(^6\)

Hence, borrower \( i \)'s asset value standardized log-returns \( r_i := r^{(i)}_\tau \) are modeled as

\[
r_i = \gamma_i Y + \sqrt{1 - \gamma^2_i} Z_i, \tag{1.13}
\]

where random variables \( Z_1, \ldots, Z_n \) and \( Y \) are standard Gaussian and mutually independent. \( Y \) is the economic composite factor and the \( Z_i \) form the idiosyncratic shocks, one for each obligor. The parameters \( \gamma^2_1, \ldots, \gamma^2_n \in (0, 1) \) are correlation parameters that capture borrower \( i \)'s sensitivity to systematic risk.

Substituting representation (1.13) into default condition (1.12), one can express the default probability of obligor \( i \) conditional on realisation \( y \in \mathbb{R} \) of systematic risk factor \( Y \) as

\[
p_i(y) = \mathbb{P}(D_i | Y = y) = \mathbb{P}(r_i < \Phi^{-1}(p_i)) | Y = y) = \mathbb{P}(\gamma_i y + \sqrt{1 - \gamma^2_i} Z_i < \Phi^{-1}(p_i))
\]

\[
= \mathbb{P}(Z_i < \frac{\Phi^{-1}(p_i) - \gamma_i y}{\sqrt{1 - \gamma^2_i}}) = \Phi\left( \frac{\Phi^{-1}(p_i) - \gamma_i y}{\sqrt{1 - \gamma^2_i}} \right), \tag{1.14}
\]

where \( p_i \) is the (unconditional) default probability of obligor \( i \) as in (1.10). Equation (1.14) is due to Vasicek and transforms \emph{unconditional} default probabilities into default probabilities

---

\(^5\)Standardized should be interpreted in the sense that the log-returns are displaced and re-scaled such that they are standard normally distributed, i.e.

\[
r^{(i)}_t = \frac{\log(V_t/V_0) - (\mu - 1/2\sigma^2)t}{\sigma\sqrt{t}}.
\]

\(^6\)In general, such factor models lead to a reduction of the computational effort, which can also be controlled by the number of factors considered in the model.
conditional on the state of the systematic risk factor $Y$.

Figure 1.2 displays the relationship between conditional and unconditional default probabilities for three different states of the composite factor $Y$. In accordance with intuition, default probabilities conditional on a bad state of the economy are larger than those on a good state.

Consider any random variable regarding the portfolio of obligors $\mathcal{P} = \{1, ..., n\}$. The Vasicek equation allows one to transform this variable into a variable conditional on the state of the systematic risk factor $Y$. For example, consider the portfolio percentage loss defined in (1.2) as

$$L_n = \sum_{i=1}^{n} w_i \eta_i \mathbb{1}_{D_i}.$$ 

The Vasicek equation yields that the expected loss conditional on the factor state $Y = y$ is given by

$$\mathbb{E}[L_n|Y = y] = \mathbb{E} \left[ \sum_{i=1}^{n} w_i \eta_i \mathbb{1}_{\{Z_i < \zeta_i(y)\}} \right] = \sum_{i=1}^{n} w_i \eta_i \Phi \left( \frac{\Phi^{-1}(p_i) - \gamma_i y}{\sqrt{1 - \gamma_i^2}} \right),$$

(1.15)

where $\zeta_i = \Phi^{-1}(p_i(y)) = \frac{\Phi^{-1}(p_i) - \gamma_i y}{\sqrt{1 - \gamma_i^2}}$ as in equation (1.14).

The main idea of the asymptotic single risk factor model, now, is that under certain conditions, the conditional portfolio percentage loss converges to the unconditional loss as more obligors are added to the portfolio. The individual risks of single obligors no longer matter, and all losses depend on the global state of the economy. Capital levels can then be set in terms of (quantiles of) the systematic risk factor, which allows one to generalise across all kinds of portfolios\(^7\).

In 2002 Gordy set out two assumptions that establish exactly this result. First, the credit portfolio must be asymptotically fine-grained, in the sense that no single exposure in the portfolio can account for more than an arbitrarily small share of the total portfolio exposure. Hence, idiosyncratic risk must vanish as more obligors are added to the portfolio: the portfolio is well

\(^7\)Of course, the exact purpose of regulatory capital requirements is to be as generally applicable as possible.
diversified. Second, the exposures to different obligors must be mutually independent conditional on the state of the systematic risk factor. This means that all correlations between exposures must stem from the global economy factor. Under these conditions, mathematically defined in Definitions B.1 and B.2 in the Appendix B: The ASRF Model, the portfolio percentage loss converges almost surely to its conditional expectation, as the portfolio approaches granularity.

**Proposition 1.1.** Assume a conditional independence model for an asymptotic credit portfolio. Then,

\[
\lim_{n \to \infty} \left( L_n - \sum_{i=1}^{n} w_i \eta_i p_i(Y) \right) = 0, \ P \ - \ a.s.
\]  

(1.16)

*Proof.* The proof can be found in Proposition B.1 of Appendix B: The ASRF Model.

As can be seen from Proposition 1.1, the (conditional) distribution of \( L_n \) degenerates to its conditional expectation in the asymptotic limit, even under quite general conditions. In intuitive terms, it states that the obligor-specific risk in the portfolio loss is diversified away as the exposure share of each obligor goes to zero. In the limit, the portfolio percentage loss depends merely on the systematic risk factor \( Y \). Limit wise, it is thus sufficient to know the distribution of \( \mathbb{E}[L_n|Y] \) to answer questions about the unconditional distribution of \( L_n \).

In turn this result leads, subject to additional technical conditions, to the following outcome: quantiles of the distribution of the conditional expectation of portfolio percentage loss, may be substituted for quantiles of the original portfolio loss distribution. A practical consequence is that VaR values of the loss distribution can be derived from quantiles of the distribution of portfolio loss conditional on the economic state variable.

**Proposition 1.2.** Consider a credit portfolio comprising \( n \) obligors, and denote by \( L_n \) the portfolio percentage loss. Let \( Y \) be a random variable with continuous and strictly increasing distribution function \( H \). Denote by \( \psi_n(Y) \) the conditional expectation of portfolio percentage loss \( \mathbb{E}[L_n|Y] \). Assuming that various technical conditions hold, it follows that

\[
\lim_{n \to \infty} \mathbb{P}(L_n \leq \psi_n(\Phi^{-1}(1-\alpha))) = \alpha,
\]  

(1.17)

and moreover

\[
\lim_{n \to \infty} \left| \text{VaR}_\alpha(L_n) - \psi_n(\Phi^{-1}(1-\alpha)) \right| = 0.
\]  

(1.18)

*Remark* 1.5. The various technical conditions and comments on these conditions, can be found under Proposition B.2 in Appendix B: The ASRF Model.

*Proof.* The proof can be found in (Gordy, 2002, [16]).

The Asymptotic Single Risk Factor model, in particular in the form of Proposition 1.2, allows for easy capital calculations under asymptotic conditions. Remember from Definition 1.4 that credit risk capital is defined as

\[
K_\alpha(L_n) = \text{VaR}_\alpha(L_n) - \mathbb{E}[L_n].
\]  

(1.19)

*Remark* 1.6. Notice that it is defined as a percentage of the EAD on a credit portfolio comprising \( n \) obligors, such that for actual portfolio capital calculations it should still be scaled by the EAD. Using the previous results, the following asymptotic credit risk capital formula can be derived.
Proposition 1.3. Assume an asymptotic conditional independence model of a credit portfolio. Then the credit risk capital is of the asymptotic form

$$\lim_{n \to \infty} K_\alpha(L_n) = \lim_{n \to \infty} \sum_{i=1}^{n} w_i \eta_i \Phi_i \left( \frac{\Phi_i^{-1}(p_i) - \gamma_i \Phi_i^{-1}(1-\alpha)}{\sqrt{1 - \gamma_i^2}} \right) - \lim_{n \to \infty} \sum_{i=1}^{n} w_i \eta_i p_i, \quad (1.20)$$

assuming the latter limits exist.

Proof. The proof can be found in Proposition B.3 of Appendix B: The ASRF Model.

The result in Proposition 1.3 above involves an asymptotic, conditional independence credit portfolio model. As any real-world portfolio consists of only a finite number of loans, the capital statement does not directly apply. Empirical studies however (Rutkowski, Tarca, 2016, [39]), suggest that international banks’ portfolios very well approximate the asymptotic limit of equation (1.20). Consequently, an adequate capital requirement is given by the following formula.

Definition 1.5. Assume a bank’s credit portfolio satisfies sufficient asymptotic granularity, then the credit risk capital held against unexpected losses (at confidence level $\alpha$ over a given risk measurement horizon) is defined as

$$\tilde{K}_\alpha(L_n) = \mathbb{E}[L_n|Y = \Phi^{-1}(1-\alpha)] - \mathbb{E}[L_n] \quad (1.21)$$

$$= \sum_{i=1}^{n} w_i \eta_i \Phi_i \left( \frac{\Phi_i^{-1}(p_i) - \gamma_i \Phi_i^{-1}(1-\alpha)}{\sqrt{1 - \gamma_i^2}} \right) - \sum_{i=1}^{n} w_i \eta_i p_i. \quad (1.22)$$

We will see in Section 1.2 how the regulator employs this formula to provide counterparty credit risk capital requirements. Setting the confidence quantile to $\alpha = 0.1\%$.

A fair, final question is to what extent the ASRF model applies to real-world banks. The model relies on two assumptions, namely the diversified portfolio and the single systematic risk factor. At first sight, these seemed not so restrictive. As formulated in (Gordy, 2002, [16]):

“The result is obtained with very minimal restrictions on the make-up of the portfolio and the nature of credit risk. The assets may be of quite varied PD, expected LGD and exposure sizes. There are no restrictions on the behaviour of conditional expected loss functions $\mathbb{E}[L_n(y)]$. These functions may be discontinuous and non-monotonic and can vary in form per obligor. Most importantly, there is no restriction on the vector of risk factors $Y$. It may be a vector of any finite length and with any distribution, continuous or discrete.”

The Basel committee adopted the model in their regulatory framework for banks in 2005. Although various papers have expressed criticism on the VaR framework (e.g., (Jarrow, 2006, [23])) and the granularity assumption (e.g., (Tarashev, Zhu, 2007, [42]), the ASRF model is still in place as its simple, closed form capital rules provide transparency, verifiability and ease of implementation, which are important considerations in the regulatory landscape.

---

8Sufficient asymptotic granularity is here to be understood in the intuitive sense: a well-diversified portfolio.
1.2 Regulatory capital

The Basel committee, officially named the Committee on Banking Regulations and Supervisory Practices (BCBS), was established after the banking crisis that lead to the collapse of the West German bank Bankhaus Herstatt, in 1974. The G10\(^9\) created the committee to improve the quality and consistency of banking supervision worldwide. The principles were laid out in the Basel Concordat in 1975 and have been revised several times since. Most notable are the committee’s landmark publications on capital adequacy commonly known as Basel I, II and III.

The Basel I Accord (BCBS, 1988, [29]) was issued by the Basel committee in 1988 and focused mainly on capital adequacy. Over the years, capital standards for banks had eroded and the document set out minimum capital standards for banks as a counterbalance. The document defined regulatory capital as 8% of Risk Weighted Assets (RWAs). Risk Weighted Assets are a measure - defined by the committee - to determine the amount of assets of a bank, corrected for risk. This approach allowed regulations to better address banks’ risk taking and compare banks across different geographies. Over the years, various amendments were made to recognize netting effects and asset class differences for derivatives. In particular, the 1996 Market Risk Amendment deployed capital requirements for market risks arising from banks’ exposures to derivative securities. As a consequence, the regulations then covered both credit and market risk.

After a long period of consultation starting in 1999, the Basel committee published the Revised Capital Framework (BCBS, 2004, [30]), the accord often referred to as Basel II. The new regulations extended the scope, from strict capital requirements to all round banking discipline. As such, the revised framework comprised three pillars:

1. Pillar 1: Minimum Capital Requirement,
2. Pillar 2: Supervisory Review,

The first pillar related best to the Basel I framework and focused on minimum capital requirements. The aim was to ensure capital requirements properly reflect the underlying risks, in particular credit risk, market risk and operational risk. Pillar 2 encouraged banks to perform their risk and capital assessment in a holistic fashion. This enabled regulatory supervisors to evaluate capital strength on an institution-wide level. The last pillar aimed to “lever disclosure of bank information to strengthen market discipline and encourage sound banking practices”.

Already before the financial crisis of 2007, the Basel committee identified excessive amounts of leverage and small liquidity buffers at a number of banks worldwide. As a response to these risk factors, the Basel II framework was strengthened with two intermediate provisions. A few years later, moreover, a broader capital and liquidity reform package was introduced: Basel III (BCBS, 2011, [32]). The accord further broadened the scope of international regulations in order to reflect the lessons learnt during the financial crisis. The capital requirements under Pillar 1 were strengthened, both in quantity and quality, via increased minima on common equity as well as capital and a conservation buffer. The latter enables regulators to increase capital requirements when systemic risk increases. Also the RWA calculations were revised: the definition of credit RWAs was extended not to only consider counterparty credit risk, but also CVA risk. CVA

\(^9\)The Group of Ten, abbreviated G10, refers to the group of countries that was formed to lend money to the International Monetary Fund, IMF.
risk refers to potential losses due to deteriorating (counterparty) credit ratings\textsuperscript{10}. Apart from updated capital charges, the Basel III accord introduced under Pillar 1 additional risk measures such as the leverage ratio, various liquidity ratios and an incremental risk charge. The leverage ratio covers loss-absorbing capital for a bank’s assets and off-balance sheet exposures, whereas the liquidity requirements ensure sufficient cash levels for banks to cover funding needs over a stressed period. Lastly, several policies under Pillars 2 and 3 were sharpened: valuation and accounting standards were updated, as well as incentives for banks to better manage long term risks.

As of today, the Basel III accord provides a very comprehensive framework that touches upon many aspects of modern banking, from assessment and calculation of various risk types to reporting and best banking practice. The items with derivative pricing impact are the RWA calculations under Pillar 1, as they impose capital requirements (and thus costs) on the derivative business. Summarizing in a simplified\textsuperscript{11} formula, regulatory capital must be calculated as

\[
K_{\text{reg}} = 8\% \cdot \left( CR-RWA + MR-RWA + OR-RWA \right),
\]

where CR, MR and OR refer to credit, market and operational risk (RWAs) respectively.

In the KVA context, we focus on the former: credit risk capital, which can be subdivided into counterparty credit risk (CCR) capital and credit valuation adjustment (CVA) capital. This is driven by the fact that the single trade credit capital cost is both well quantifiable and considerable in size. Although large in size, market risk capital is mostly considered on a bank-wide level, making it difficult to estimate the capital impact of a single trade, which is often small and inconsistent\textsuperscript{12}. Operational risk capital costs are either very simple or too difficult to calculate (under the standard and advanced approaches respectively), and in any case, assumed to be relatively small. Hence, operational risk is uninteresting from a KVA perspective.

The regulatory calculation methods for CCR and CVA risk capital can be found in Table 1.1.

<table>
<thead>
<tr>
<th>Risk type</th>
<th>Calculation method</th>
<th>Calculation type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counterparty credit risk</td>
<td>\textit{EAD calculation}</td>
<td>Function of netting set value</td>
</tr>
<tr>
<td></td>
<td>CEM</td>
<td>Function of netting set value</td>
</tr>
<tr>
<td></td>
<td>Standardised</td>
<td>Function of netting set value</td>
</tr>
<tr>
<td></td>
<td>SA-CCR</td>
<td>Function of netting set value</td>
</tr>
<tr>
<td></td>
<td>Internal method</td>
<td>Exposure profile</td>
</tr>
<tr>
<td>Weight calculation</td>
<td>Standardised</td>
<td>External ratings</td>
</tr>
<tr>
<td></td>
<td>FIRB</td>
<td>Internal/external ratings</td>
</tr>
<tr>
<td></td>
<td>AIRB</td>
<td>Internal/external ratings, internal LGD</td>
</tr>
<tr>
<td>CVA risk</td>
<td>Standardised</td>
<td>Function of EAD</td>
</tr>
<tr>
<td></td>
<td>Internal method</td>
<td>VaR and SVaR</td>
</tr>
</tbody>
</table>

Table 1.1: Available approaches to calculate credit related regulatory capital.

\textsuperscript{10}BCBS observed that during the financial crisis, two-thirds of losses were due to mark-to-market changes as a consequence of credit market volatility, opposed to one-third of losses from actual defaults (BCBS, 2010, [31]).

\textsuperscript{11}In reality, regulatory capital is divided into Common Equity Tier 1, Additional Tier 1 and Tier 2 instruments.

\textsuperscript{12}This follows from the fact that market risk capital is calculated from the bank-wide derivatives portfolio VaR (and stressed VaR), which is hardly impacted by a single trade. Moreover, the impact can be positive or negative depending on the trade’s risk, relative to the accumulated portfolio risk.
As mentioned in the introduction, under Basel III there are standardised methods for smaller banks, and internal model methods (IMM) for banks with appropriate supervisory authorisation. The aim of the latter is to better reflect risks underlying the derivative portfolio in the capital levels. The counterparty credit risk calculation consists of an EAD component and a Risk Weight (RW) component. The EAD component can be calculated under the Current-Exposure-Method (CEM), Standardised method or the Standard Approach Counterparty Credit Risk (SA-CCR), which is due to revise the former two from the 1st of January 2021. The weight component may take weights from a standardised table, or calculate weights under the Foundation Internal Ratings Based (FIRB) or Advanced Internal Ratings Based (AIRB) approach. The terms ‘foundation’ and ‘advanced’ refer to the degree to which a bank can supply their own inputs to this calculation, as can be seen in Table 1.1. The CVA risk calculation again has multiple methods: a standardised and an internal method. The former is a direct calculation across all counterparties, where risk weights (RW) are prescribed by the regulator and EAD numbers are taken from the appropriate CCR methodology. The latter is a simulation based method based on the portfolio VaR and stressed VaR (SVaR). Most calculation methods will be described thoroughly in the following section, and demonstrate the key issues that will rise later, in the KVA context.

The Basel III regulations are being implemented since 2011 and should be fully in force as of January 2019. Meanwhile, the Basel committee has begun even another round of revision of elements of the regulatory framework, and even Basel IV plans have been outlined. Although not all have yet been finalized, forthcoming revisions with impact on derivatives pricing are the Fundamental Review of the Trading Book, (FRTB) (BCBS, 2016, [34]) and the Revised Standard Approach to Counterparty Credit Risk (SA-CCR) (BCBS, 2014 [33]). The latter changes the standard calculation methods for CCR capital and is expected to come into force the 1st of January 2021. The following section will therefore cover both current and future methodologies from Table 1.1.
1.2.1 Counterparty Credit Risk capital

A bank is supposed to hold capital to cover unexpected counterparty default losses. Of course, default risk is priced into products such that on average, default losses can be covered. In some exceptional cases however, losses may far exceed these values and additional capital is required. The level of capital that should be held in order to withstand losses in $\alpha \cdot 100\%$ of economic scenarios, cf. Definition 1.4, is the value-at-risk for this $\alpha$ corrected by the expected loss. In the forthcoming section, we will see how the regulator adapted capital requirements to this concept.

Consider a bank’s portfolio constituted by $n$ different trades with a single counterparty $j$. The portfolio loss variable, according to Definition 1.1, is then given by

$$PL_n^{(j)} = \sum_{i=1}^{n} \delta_i^{(j)} \eta_i^{(j)} D_i^{(j)},$$

(1.24)

where $\delta_i^{(j)}$, $\eta_i^{(j)}$ and $D_i^{(j)}$ denote the exposure-at-default, the loss-given-default and default event respectively. Notice that the latter two are trade-independent, as trades are with the same counterparty. The regulator now specifies counterparty credit risk Risk Weighted Assets, CCR-RWA, as

$$RWA_n^{(j)} = 12.5 \cdot RW^{(j)} \cdot EAD_n,$$

(1.25)

where $RW^{(j)} := RW(\eta^{(j)}, p^{(j)})$ is a function of the counterparty LGD and PD, where $p^{(j)} = P(D^{(j)})$, and $EAD_n := EAD_n(\delta_1, \ldots, \delta_n)$ is a function of the portfolio exposures $\delta_i$. The multiplier $12.5$ reflects the transition from capital to RWA as 8% of RWA is now $12.5 \cdot 8\% \cdot RW \cdot EAD_n = RW \cdot EAD_n$. The calculation methodologies for these two constituents of default risk capital, as presented in Table 1.1, will be described below.

The Risk Weight

All risk weight calculation methods can be found in the Basel III document (BCBS, 2011, [32]).

1. Banks without supervisory approval to use IRB approaches must rely on the standardised method. The method simply assigns a risk-weight to a counterparty based on its external rating and the sector in which it operates. Table 1.2 below provides an overview.

<table>
<thead>
<tr>
<th>Corp. risk weight</th>
<th>S&amp;P</th>
<th>Moody’s</th>
<th>Fitch</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>AAA to AA-</td>
<td>Aaa to Aa3</td>
<td>AAA to AA-</td>
</tr>
<tr>
<td>50%</td>
<td>A+ to A-</td>
<td>A1 to A3</td>
<td>A+ to A-</td>
</tr>
<tr>
<td>50%</td>
<td>BBB+ to BBB-</td>
<td>Baa1 to Baa3</td>
<td>BBB+ to BBB-</td>
</tr>
<tr>
<td>100%</td>
<td>BB+ to BB-</td>
<td>Baa1 to Ba3</td>
<td>BB+ to BB-</td>
</tr>
<tr>
<td>100%</td>
<td>B+ to B-</td>
<td>B1 to B3</td>
<td>B+ to B-</td>
</tr>
<tr>
<td>150%</td>
<td>CCC+ or lower</td>
<td>Caa1 or lower</td>
<td>CCC+</td>
</tr>
</tbody>
</table>

Table 1.2: Standardised risk-weights for Counterparty Credit Risk capital.

The table above only assigns risk-weights to institutions for which a qualifying credit assessment is available. Counterparties without such a rating are assigned a risk-weight of 100%.

2. Large banks with advanced $PD$ and $LGD$ models are allowed to use a more technical framework. This internal ratings based approach is based on the Asymptotic Single Risk Factor model.
and in particular on the capital formula (1.22). The counterparty risk weight is given by

\[ RW = c \left[ \eta \cdot \Phi \left( \frac{1}{\sqrt{1 - \rho}} \Phi^{-1}(p) + \frac{\sqrt{\rho}}{\sqrt{1 - \rho}} \Phi^{-1}(0.999) \right) - \eta \cdot p \right], \]

(1.26)

where \( c \) is a supervisory correction factor and \( \rho \) is the ASRF model correlation.

**Remark 1.7.** Compare the risk-weight formula (1.26) to capital equation (1.22). The EAD part of (1.26) is stripped out via RWA formula (1.25) and a supervisory correction is added.

**Remark 1.8.** Effectively, the internal ratings based approach embraces two complexity levels. Banks with FIRB (foundation) status can use their own PD models, but must use regulatory LGD numbers. Banks enjoying the AIRB (advanced) status also supply their own LGD.

Two parameters in formula (1.26) are provided by the regulator: the asset correlation and the supervisory correction. The asset correlation, originating in the factor model (1.13) as \( \rho = \gamma_i^2 \) for \( i = 1, ..., n \), depends on the counterparty type of the portfolio, i.e. corporate, retail or financial institution, and counterparty size. The general formula\(^\text{13}\) is given by

\[ \rho = L_1 \cdot \frac{1 - e^{-\lambda p}}{1 - e^{-50}} + L_2 \cdot \left[ 1 - \frac{1 - e^{-\lambda p}}{1 - e^{-50}} \right]. \]

(1.27)

Here \( L_1 \) and \( L_2 \) denote the lower and upper limit of the correlations respectively. The correlation decreases as a function of the default probability of the counterparty. The smaller the default probability, the higher the correlation to the systematic risk factor. The exponential function decreases rather fast; its pace is determined by the so-called \( \lambda \)-factor, depending on the counterparty type. For corporate exposures, the rate is set to \( \lambda = 50 \) and the correlation limits are 12% and 24%. A size adjustment factor distinguishes between up to medium sized corporates and large financial sector entities.

The supervisory correction factor is a function of default probability and residual trade maturity. Its aim is to correct risk weights (and hence capital requirements) for instruments of different maturities; in general long-term credits are riskier than short-term credits. Moreover, the maturity adjustment is larger for counterparties with a low default probability. Intuitively, this follows from the fact that downgrades in credit rating are more likely for low-pd counterparties. The precise formulation of the supervisory adjustment is

\[ c := c(b, M) = \frac{1 + (M - 2.5)b}{1 - 1.5b}, \]

(1.28)

\(^{13}\)The supervisory asset correlations of the Basel risk-weight formula for corporate, bank and sovereign exposures have been derived by analysis of data sets from G10 supervisors. Time series of corporate accounting and default data have been used to determine default rates as well as correlation between borrowers.
where

\[ b = \left( 0.11852 - 0.05478 \log(p) \right)^2, \]  
(1.29)

\[ M = \min \left\{ 5.0, \max \left\{ 1.0, \frac{\sum_{i=1}^{n} m_i n_i}{\sum_{i=1}^{n} m_i} \right\} \right\}, \]  
(1.30)

for \( m_i \) and \( n_i \) respectively the residual trade maturity and notional of trade \( i \) in the portfolio.

The Exposure-at-Default

The regulatory exposure at default is calculated over all \( n \) trades in the portfolio, allowing for netting effects between different exposures. In this section, we consider the Current Exposure Method and the Standardised Approach for measuring Counterparty Credit Risk (SA-CCR). The latter due to revise the former by the first of January 2021 and hence is essential for Capital Valuation Adjustment, which may require expected capital profiles beyond this date. Details of both methods can be found in (BCBS, 2011, [32]) and (BCBS, 2014, [33]), respectively.

Remark 1.9. In fact, Basel III allows for two more calculation methods. The Standardised Approach is a standard approach on the same regulatory level as CEM, and calculates the EAD as a function of notionals. Secondly, the Internal Model Method (IMM) allows advanced banks to simulate maximum positive exposures over the next year and use those for regulatory EAD.

1. Under the Current Exposure Method, the exposure at default is given by the accounting value of the trade, dubbed replacement cost (RC), and an add-on that aims to capture the exposure of the transaction over its remaining life: the regulatory potential future exposure (PFE).

\[ EAD_n = \sum_{i=1}^{n} \delta_i = \sum_{i=1}^{n} \left( V_i^+ + \psi(m_i, N_i, \text{asset class}_i) \right), \]  
(1.31)

where \( \mu_i \) denotes the mark-to-market. The add-ons are a percentage of the trade notional depending on the asset class and residual maturity of the trade, as can be seen in Table 1.3.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Interest rates</th>
<th>FX and gold</th>
<th>Equities</th>
<th>Precious metals</th>
<th>Other commodities</th>
</tr>
</thead>
<tbody>
<tr>
<td>One year or less</td>
<td>0.00%</td>
<td>1.00%</td>
<td>6.00%</td>
<td>7.00%</td>
<td>10.00%</td>
</tr>
<tr>
<td>Over one year, to five years</td>
<td>0.50%</td>
<td>5.00%</td>
<td>8.00%</td>
<td>7.00%</td>
<td>12.00%</td>
</tr>
<tr>
<td>Over five years</td>
<td>1.50%</td>
<td>7.50%</td>
<td>10.00%</td>
<td>8.00%</td>
<td>15.00%</td>
</tr>
</tbody>
</table>

Table 1.3: Add-ons for the CEM calculation

Netting effects are supported through the net-to-gross ratio (NGR) adjustment, allowing up to 60% netting benefits for add-ons of trades in the same netting set:

\[ \psi_{net} = 0.4 \cdot \psi_{gross} + 0.6 \cdot \nu \cdot \psi_{gross}, \]  
(1.32)

where \( \nu \) is the net-to-gross ratio that is given by the fraction of positive mark-to-markets

\[ \nu = \frac{\left( \sum_{i=1}^{n} V_i \right)^+}{\sum_{i=1}^{n} (V_i)^+}, \]  
(1.33)

if at least one trade has a positive value, and is zero otherwise.
2. Under the *Standardised Approach for measuring Counterparty Credit Risk*, calculations are more involved. This hybrid framework aims to capture strengths and avoid weaknesses of both the existing SA and the CEM method. The regulatory exposure-at-default is calculated via

\[ EAD_n = \alpha \cdot (R_n + \psi_n), \]  

where \( R_n \) denotes the portfolio replacement cost (RC) and \( \psi_n \) the potential future exposure (PFE). The supervisory factor \( \alpha \) is in principle set to 1.4, but can be lowered to 1.2 upon regulatory approval.

First, the replacement cost is given by

\[ R_n = \max \{V_n - C, 0\} \quad \text{(unmargined transaction)} \]  

\[ = \max \{V_n - C, TH + MTA - NICA\} \quad \text{(margined transaction)} \]

where \( V_n \) denotes the net MtM value of the trade portfolio and \( C \) is the value of collateral for this set after haircuts. In the case of a margined transaction, \( TH \) is the collateral threshold, \( MTA \) is the minimum transfer amount and \( NICA \) is the net independent collateral amount, calculated by

\[ NICA = \text{Coll}_{\text{received}} - \text{Coll}_{\text{posted(unsegregated)}}. \]

It is thus the amount of collateral available to the bank to offset losses in case of counterparty default. The effective collateral threshold is given by \( TH + MTA \) and as such the term \( TH + MTA - NICA \) denotes the largest exposure that would not trigger a call for variation margin.

Secondly, consider the potential future exposure calculation. It is a straightforward but tedious calculation, that involves many different levels on which exposures are aggregated. Amongst others, there are maturity buckets, risk factors and asset classes. The PFE is of the form

\[ \psi_n = m(V_n, C_n, A^{agg}) \cdot A^{agg}, \]

where \( m \) is a multiplier that gives benefits for out-of-the-money and overcollateralised trades and \( A^{agg} \) an aggregated add-on similar to the CEM add-on.

The multiplier is given by\(^\text{14}\)

\[ m(V, C, A^{agg}) = \min \left\{ 1, f + (1 - f) \exp \left( \frac{V - C}{2(1 - f)A^{agg}} \right) \right\}, \]

where \( f \) is the regulatory floor set to 5%.

The add-on \( A^{agg} \) is decomposed into five asset classes

\[ A^{agg} = A^{IR} + A^{FX} + A^{Eq} + A^{Cred} + A^{Comm}. \]

Per asset class the add-on is calculated as

\[ A^{(c)} = \sum_{j: HS^{(c)}} A_j^{(c)} = \sum_{j: HS^{(c)}} \gamma_j^{(c)} \cdot EN_j^{(c)}, \]

where \( \gamma_j^{(c)} \) is the supervisory factor and \( EN_j^{(c)} \) the effective notional amount. The sum is taken over all hedging sets (HS) for the given asset class. The hedging sets per asset class are summarised in Table 1.4 below.

\(^{14}\)The multiplier aims to imitate the internal model method expected exposure calculation for a normally distributed asset value. The multiplier is an evaluation of its probability density function, floored at 5%. 

26
<table>
<thead>
<tr>
<th>Asset class</th>
<th>Hedging set</th>
<th>Offsetting</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR</td>
<td>Derivatives in the same currency</td>
<td>Full offsetting within same maturity category, partial offsetting within neighbouring maturity categories</td>
</tr>
<tr>
<td>FX</td>
<td>Derivatives in the same currency pair</td>
<td>Full offsetting between positions per hedging set</td>
</tr>
<tr>
<td>CR &amp; EQ</td>
<td>One hedging set for equity derivatives and one credit derivatives</td>
<td>Partial offsetting between positions</td>
</tr>
<tr>
<td>CO</td>
<td>Energy, metals, agricultural and other</td>
<td>Partial offsetting between positions</td>
</tr>
</tbody>
</table>

Table 1.4: Hedging sets for SA-CCR calculation

The supervisory factors $\gamma^{(c)}_j$ are left to be determined by (supra)national supervisors, but will depend on asset class, hedging set and potentially counterparty. The factor is meant to convert the effective notional amount into effective expected positive exposure based on the volatility the supervisor observed for the asset class.

The effective notional (EN) amount $EN^{(c)}_j$ is calculated over different maturity buckets, via

$$EN^{(c)}_j = \sum_{k:MB^{(c)}_j} |D^{(c)}_{jk}|,$$

where each $D_{jk}$ is calculated as the sum over a product of three terms over trades in the same maturity bucket (MB) $k$ within hedging set $j$:

$$D^{(c)}_{jk} = \sum_{i \in MB_k \subset HS_j} \delta_i \cdot d_i^{(c)} \cdot MF_i.$$

The maturity buckets divide a portfolio into trades with a maturity up to one year, one to five years and trades with a maturity beyond five years. The components of equation (1.43) are:

1. The factor $\delta_i$ is the supervisory delta adjustment depending on a trade’s primary risk driver and position. The delta values, in terms of standardnormal cdf $\Phi$, are shown in Table 1.5.

<table>
<thead>
<tr>
<th>Delta value</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_i = +\Phi(q_i)$</td>
<td>Long call options</td>
</tr>
<tr>
<td>$\delta_i = -\Phi(q_i)$</td>
<td>Short call options</td>
</tr>
<tr>
<td>$\delta_i = +\Phi(-q_i)$</td>
<td>Long put options</td>
</tr>
<tr>
<td>$\delta_i = -\Phi(q_i)$</td>
<td>Short put options</td>
</tr>
<tr>
<td>$\delta_i = +1$</td>
<td>Other instrument, long in primary risk factor</td>
</tr>
<tr>
<td>$\delta_i = -1$</td>
<td>Other instrument, short in primary risk factor</td>
</tr>
</tbody>
</table>

Table 1.5: Supervisory delta adjustments for netting in SA-CCR calculation

Here $q_i$ is defined as $\frac{\log(P_i/K_i) + \frac{1}{2}\sigma^2 T_i}{\sigma \sqrt{T_i}}$, for $P_i$ the price of the underlying, $K_i$ strike price, $T_i$ latest contractual exercise date and $\sigma_i$ the supervisory volatility parameter.
2. The factor \( d^{(c)}_i \) denotes the trade level adjusted notional amount of contract \( i \) and is defined as
\[
d^{(c)}_i = N_i \cdot SD_i = N_i \cdot \frac{\exp(-0.05 \cdot S_i) - \exp(-0.05 \cdot E_i)}{0.05},
\]
where \( N_i \) is the notional and \( S_i \) and \( E_i \) are the start date and end dates of the time period referenced by the derivative contract respectively.

3. The factor \( MF_i \) is the minimum time risk horizon calculated as
\[
MF_i = \sqrt{\min\{M_i, 1y\}} \quad \text{(unmargined transaction)} \quad (1.45)
\]
\[
= \frac{3}{2} \sqrt{\frac{\text{MPOR}_i}{1y}} \quad \text{(margined transaction)} \quad (1.46)
\]
where \( M_i \) is the transaction \( i \) remaining maturity floored by 10 business days and \( \text{MPOR}_i \) is the margin period of risk appropriate for the margin agreement containing transaction \( i \).

The SA-CCR method provides netting benefits on maturity bucket level via the supervisory delta adjustments. It is similar to the CEM method as it gives an approximate PFE for each asset class and hence provides a measure of future exposure to the counterparty. While it is not risk sensitive in the sense of an IMM model, it is more realistic than both the SA and the CEM method. Lastly, it remains formula based, which is valuable in a KVA context15.

1.2.2 CVA capital

As stated in the introduction16, the Basel committee found that during the recent financial crisis the majority of losses was a consequence of credit spread volatility, when mark-to-market values plunged as credit ratings deteriorated. Products lost value because the market perception of credit risk on those products increased. A much smaller portion of losses, on the other hand, came from actual defaults. Adapting to this situation, Basel III introduced alongside CCR capital, the notion of CVA risk capital. CVA capital is held against losses due to credit valuation adjustment.

Opposed to counterparty credit risk capital, CVA risk capital is calculated on a global level, in the sense that it involves all of a bank’s counterparties. At first sight, the calculation of CVA capital impact of a single trade thus seems vastly more complex than CCR capital impact. Both of the available CVA capital approaches, standardised and advanced (cf. Table 1.1), have their own solutions. As our focus is on the standardised method, this will be presented below.

Under this approach, the CVA capital charge (BCBS, 2011, [32]) is given by the formula
\[
K_{CVA} = 2.33 \sqrt{\left(\frac{1}{2} \sum_{i:\text{cpt}} w_i M_i EAD_i\right)^2 + \frac{3}{4} \sum_{i:\text{cpt}} w_i^2 M_i^2 EAD_i^2}, \quad (1.47)
\]
where the sum is taken over all counterparties \( i \) in the bank portfolio. The counterparty parameters appearing in the formula are counterparty weight \( w_i \), counterparty notional weighted

15As we will see later, Capital Valuation Adjustment calculations require estimations of capital profiles via simulation. Formula-based capital calculations thus reduce the need for nested simulations.

16Cf. footnote 10.
Table 1.6: CVA risk weights as function of Standard & Poor’s ratings.

<table>
<thead>
<tr>
<th>S&amp;P Rating</th>
<th>Weight $w_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.7%</td>
</tr>
<tr>
<td>AA</td>
<td>0.7%</td>
</tr>
<tr>
<td>A</td>
<td>0.8%</td>
</tr>
<tr>
<td>BBB</td>
<td>1.0%</td>
</tr>
<tr>
<td>BB</td>
<td>2.0%</td>
</tr>
<tr>
<td>B</td>
<td>3.0%</td>
</tr>
<tr>
<td>CCC</td>
<td>10.0%</td>
</tr>
</tbody>
</table>

maturity $M_i$ and counterparty exposure-at-default $EAD_i$, as calculated in Section 1.2.1. The risk-weights depend on the current counterparty credit rating as in Table 1.6.

The formula is global and spans all counterparties. Consequently, in order to find the capital impact of each single trade, the entire formula (1.47) has to be recalculated, which might not be practicable. Assuming a large number of counterparties, this problem can be circumvented by the following approximation: the first term in (1.47) is a square of sum, which is for a large number of counterparties much greater than the second term, a sum of squares. Hence, it follows (Green, 2016, [18]) that

$$K_{CVA} \approx 2.33 \sqrt{\frac{1}{2} \sum_{i \in cpt} w_i M_i EAD_i}^2 \approx \frac{2.33}{2} \sum_{i \in cpt} w_i M_i EAD_i.$$  (1.48)

It is clear from this formula that the impact of a new trade $j$ is approximately the stand-alone trade capital value $2.33/2 \cdot w_j M_j EAD_j$. In practice, this approximation can be made more accurate if the fraction

$$R = \frac{\sqrt{\frac{1}{2} \sum_{i \in cpt} w_i M_i EAD_i}^2 + \frac{3}{4} \sum_{i \in cpt} w_i^2 M_i^2 EAD_i^2}{\sum_{i \in cpt} w_i^2 M_i^2 EAD_i^2}$$

is fairly constant over time, such that in (1.48) the factor $2.33/2$ can be replaced by $R$. 

29
2 Capital Valuation Adjustment

Any financial institution holding derivatives, as exhibited in the previous chapter, is required to hold capital. Holding capital brings costs to the business, as the capital cannot be used to generate profit via other business lines. The increase of capital requirements over the last few years has incentivized banks not only to quantify these costs, but also to price them into their derivative products (Sherif, 2015, [40]), leading to a next generation valuation adjustment: the Capital Valuation Adjustment. The essence of KVA is to pass the capital cost of manufacturing a derivative to the buyer.

The concept of Capital Valuation Adjustment was first formalized by Green and Kenyon in 2014 (Green, Kenyon, 2014, [17]). The paper initiated a stream of studies towards the price of capital in derivatives, but as of 2018 only a few distinct models can be identified. Most notable approaches are, apart from the initial paper, the expectation setup in (Elouerkhaoui, 2016, [13]), the indifference approach of (Brigo et al., 2017, [7]) and the balance sheet approach in (Albanese, Crepey, 2018, [1]). The thesis at hand focuses on the very first paper and the latter.

Section 2.1 describes the classical KVA model of Green and Kenyon (GK), which is based on semi-replication (Green, Kenyon, 2014, [17]). In this model, capital costs are derived via a (partially) replicating portfolio in a risk-neutral framework. The model is an extension of (Burgard, Kjaer, 2011, [9]), an earlier XVA model (BK), where all valuation adjustments are derived simultaneously. The end result is a KVA formula that is an integral over the expected capital profile over the lifetime of a trade. The mathematical foundations will be established in Section 2.1.1 and the final KVA formula in Section 2.1.2.

Section 2.2 illustrates the approach based on backward stochastic differential equations, which is established from a balance sheet approach (AC) (Albanese, Crepey, 2018, [1]). The underlying assumptions are different from the previous method, in the sense that the KVA itself is assumed to be a risk margin part of the regulatory capital. This induces an implicit relationship between KVA and capital, resulting in a backward stochastic differential equation. Equations of this kind will be introduced in Section 2.2.1. The resulting equation can, fortunately, be solved explicitly and results in a KVA formula that is again the integral over the expected capital profile. The derivation and solution can be found in Section 2.2.2.

Although the two presented approaches are very different in nature, both in terms of assumptions and underlying mathematics, the resulting KVA formulae are very similar. It turns out that for regulatory capital, the difference between the two is only a discount factor. Implementation wise this is very efficient, as the same simulation procedure for capital profiles can be deployed in both methodologies. In Chapter 3, this computational implementation will be described.
2.1 The Semi-replication model

2.1.1 Stochastic Differential Equations

This section lays the mathematical foundations to conduct the semi-replication KVA approach in Section 2.1.2. The material covered are general semimartingales up to the relevant Itô formula and basic results on Stochastic Differential Equations (SDEs), much along the lines of (Spreij, 2018, [41]) and extended via (Jacod, Shiryaev, 2002, [22]). Understanding of this material is more or less assumed and the results are primarily stated for reference purposes.

Consider a filtered probability space \((\Omega, \mathcal{F}, \mathbb{P})\) where \(\mathcal{F}\) satisfies the usual conditions (UC).

1. First, we provide some results on discontinuous (semi)martingales and jump processes. Remember that a continuous semimartingale is a process \(X\) that admits a decomposition

\[
X = X_0 + A + M,
\]

where \(A\) is a continuous process with paths of bounded variation and \(A_0 = 0\) a.s. and \(M\) is a continuous local martingale. The decomposition is unique up to indistinguishability, which follows from the fact that continuous local martingales are a.s. constant in time, given they are of bounded variation. See Lemma 2.15 in (Spreij, 2018, [41]). The stochastic integral for a locally bounded process \(Y\) and a continuous semimartingale \(X\) is well defined by

\[
Y \cdot X = Y \cdot A + Y \cdot M,
\]

where the integrals are interpreted as Stieltjes and stochastic integrals, respectively. This allows one to formalize the well-known Itô formula for semimartingales, which will not be restated here.

In what follows, the theory of semimartingales will be extended to consider also the discontinuous case. Some basic properties, the stochastic integral and the Itô formula will be presented, much along the lines of (Jacod, Shiryaev, 2002, [22]). As the purpose of the text below is merely to illustrate the theory, proofs are often omitted. A complete, formal derivation of the theory can be found in the book above.

**Definition 2.1.** A local martingale is purely discontinuous if \(X_0 = 0\) a.s. and

\[
\langle X, M \rangle = 0 \quad \forall M \in \mathcal{M}^{loc}.
\]

Any local martingale can be decomposed into a continuous and a purely discontinuous part.

**Proposition 2.1.** Any local martingale \(M\) admits a unique (up to indistinguishability) decomposition

\[
M = M_0 + M^c + M^d,
\]

where \(M_0^c = M_0^d = 0\), \(M^c\) is a continuous local martingale and \(M^d\) is a purely discontinuous local martingale.

**Proof.** Theorem 4.18 in (Jacod, Shiryaev, 2002, [22]).

The class of continuous semimartingales can be extended into the general class of semimartingales. Similar to equation (2.1), a semimartingale is a process \(X\) with a decomposition

\[
X = X_0 + A + M,
\]
where $A$ is still of bounded variation, but $A$ and $M$ are no longer necessarily continuous. If $A$ is also predictable, then the decomposition in (2.5) is unique. In any case, all semimartingales are càdlàg and adapted. Lastly, the unique decomposition of Proposition 2.1 yields the result below.

**Proposition 2.2.** Let $X$ be a semimartingale. There exists a unique (up to indistinguishability) continuous local martingale $X^c$ with $X_0 = 0$, such that any decomposition $X = X_0 + M + A$ meets $M^c = X^c$ up to indistinguishability. $X^c$ is called the continuous martingale part of $X$.

**Proof.** This is a nearly direct consequence of the previous proposition.

The stochastic integral as known from equation (2.2), naturally extends to the general class of semimartingales and keeps the core of its properties. Details can be found in Jacod and Shiryaev. Moreover, the definition of quadratic (co-)variation in the general case is as follows.

**Definition 2.2.** The quadratic co-variation of the two semimartingales $X$ and $Y$ is given by

$$[X,Y]_t = XY - X \cdot Y - Y \cdot X - X_0 Y_0.$$  

(2.6)

**Remark 2.1.** The quadratic variation of a semimartingale $X$ is defined as $[X,X]_t$.

The definition relates to the continuous version as follows.

**Proposition 2.3.** If $X$ and $Y$ are semimartingales with continuous parts $X^c$ and $Y^c$, then

$$[X,Y]_t = \langle X^c, Y^c \rangle_t + \sum_{s \leq t} \Delta X_s \Delta Y_s,$$  

(2.7)

where $\Delta X_t$ denotes $X_t - X_{t-}$.

**Proof.** Theorem 4.52 in (Jacod, Shiryaev, 2002, [22]).

Finally, we extend the well-known Itô formula to the general class of semimartingales.

**Theorem 2.1.** Let $X = (X^1, ..., X^d)$ be a $d$-dimensional semimartingale and $f$ a twice differentiable function s.t. all partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ exist and are continuous. Then $f(X)$ is again a semimartingale and it holds that

$$f(X_t) = f(X_0) + \sum_{i=1}^d \frac{\partial}{\partial x_i} f(X_{t-}) \cdot X^i_{t-} + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j} f(X_{t-}) \cdot \langle X^{i,c}, X^{j,c} \rangle_t$$

$$+ \sum_{s \leq t} \left[ f(X_s) - f(X_{s-}) - \sum_{i=1}^d \frac{\partial}{\partial x_i} f(X_{s-}) \Delta X^i_s \right],$$  

(2.8)

where all terms in this formula are well defined.

**Proof.** Theorem 4.57 in (Jacod, Shiryaev, 2002, [22]).

As expected, the Itô formula now also contains discontinuous parts (i.e., jumps) of $X$ via the latter sum. This general case, as will be demonstrated in KVA Theorem 2.4, allows us to calculate the dynamics of a portfolio with defaultable assets, of which default is described by a jump-to-default process.

2. Second, we provide some results on stochastic differential equations. SDE’s for short, these equations form the core of a lot of problems in mathematical finance, many of which are related
to pricing and hedging of financial derivatives. In the Valuation Adjustment framework that has been developed since the recent financial crisis, they again play a central role. Below, the equations are formally defined, solved, and related to partial differential equations.

Consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a $k$-dimensional Brownian motion $W := (W_t)_{0 \leq t \leq T}$. Let $\mathbb{F}$ be the Brownian filtration, that satisfies the usual conditions.

**Definition 2.3.** Let $\xi_0$ be a $\mathbb{R}^d$-valued initial condition, $f$ an $\mathbb{R}^d$-valued function and $g$ an $\mathbb{R}^{d \times k}$-valued function, both of which are $\mathcal{S}_\mathcal{F} \times B(\mathbb{R}^d)$ measurable. A (strong) solution for the SDE associated with parameters $(\xi_0, f, g)$ is a process $X \in S^0$ such that for all $T \geq 0$

$$\int_0^T |f(s, X_s)| ds + \int_0^T |g(s, X_s)|^2 ds < \infty \ a.s., \tag{2.10}$$

and

$$X_t = \xi_0 + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dW_s, \ t \geq 0 \ a.s. \tag{2.11}$$

Under certain conditions on the functions $f$ and $g$, existence and uniqueness of a strong solution can be established. The classical and most common are Lipschitz conditions.

**Theorem 2.2.** Consider stochastic differential equation (2.11) for some continuous $f$ and $g$ that satisfy condition (2.10). Assume that for some $l_f, l_g$ the Lipschitz conditions

$$|f(t, x) - f(t, y)| \leq l_f |x - y| \ \forall x, y \in \mathbb{R}^d, \tag{2.12}$$

$$|g(t, x) - g(t, y)| \leq l_g |x - y| \ \forall x, y \in \mathbb{R}^d, \tag{2.13}$$

hold, then there exists a unique solution $X \in S^0$ to the stochastic differential equation (2.11).

**Proof.** See Theorem 3.17 in (Pardoux, Rascanu, 2010, [36]). \qed

Consider now a simple example, the one-dimensional SDE for coefficients $(x, f, g)$ where the initial condition $x$ is deterministic. Assume the triple satisfies the conditions of Theorem 2.2. In relation to this equation, define a family of generators $\{\mathcal{L}_t : t \geq 0\}$ acting on functions $h \in C^{1,2}((0, \infty) \times \mathbb{R})$ as

$$\mathcal{L}_t h(t, x) := f(t, x) h_x(t, x) + \frac{1}{2} g^2(t, x) h_{xx}(t, x), \tag{2.14}$$

where $f$ and $g$ are the coefficients of the SDE. The definition below establishes a connection between the stochastic differential equation we are considering and partial differential equations.

**Definition 2.4.** Let $T > 0$ and let functions $h : \mathbb{R} \to \mathbb{R}$ and $c, k : [0, T] \times \mathbb{R} \to \mathbb{R}$ be given. The Cauchy problem is defined as the problem to find a solution $v : [0, T] \times \mathbb{R}$ in $C^{1,2}([0, T] \times \mathbb{R})$ such that

$$v_t + \mathcal{L}_t v = kv - c, \ v(T, \cdot) = h. \tag{2.15}$$

If moreover, the conditions are imposed that $k$ is continuous and nonnegative and $h$ and $c$ are continuous and satisfy for some constants $L > 0$ and $\lambda \geq 1$ the growth condition

$$|h(x)| + \sup_{0 \leq t \leq T} |c(t, x)| \leq L(1 + |x|^\lambda), \tag{2.16}$$

we denote the problem Constrained Cauchy problem.
The theorem below establishes a link between the Cauchy problem and solutions of stochastic differential equations. It offers a method to solve partial differential equations by simulating random paths of a stochastic process and to compute expectations by taking sample averages. See Chapter 3. It is generally referred to as the Feynman-Kac (FK) formula.

**Theorem 2.3.** Assume that a solution $v$ to the Constrained Cauchy problem of Definition 2.4 exists which also satisfies the growth condition

$$
\sup_{0 \leq t \leq T} |v(t,x)| \leq M(1 + |x|^{2\mu}),
$$

(2.17)

for some $M > 0$ and $\mu \geq 1$. Let $\{X_{t,x}, s \geq t\}$ be the solution to equation (2.11), unique in law. Then $v$ admits the unique stochastic representation on $[0,T] \times \mathbb{R}$

$$
v(t,x) = \mathbb{E}
\left[h(X^{t,x}_T) \exp\left(-\int_t^T k(u,X^{u,x}_u)du\right)\right] + \mathbb{E}
\left[\int_t^T c(r,X^{r,x}_r) \exp\left(-\int_t^r k(u,X^{u,x}_u)du\right)dr\right].
$$

(2.18)

**Proof.** The proof can be found in Theorem 11.2 in (Spreij, 2018, [41]).

---

2.1.2 The Semi-replication formula

The semi-replication approach to calculating the Capital Valuation Adjustment for a trade (portfolio) is based on the Burgard Kjaer (BK) semi-replication model (Burgard, Kjaer, 2011, [9]). It is a PDE approach that provides transparency of cashflows and simultaneous derivation of credit (debt), funding and capital valuation adjustments. The original BK model was extended to include capital costs in (Kenyon, Green, 2014, [17]), which will be presented in this section.

In order to deal with capital, a parameter $\phi$ is introduced to represent the fraction of capital $K$ used for funding. Capital used for funding represents the use of funds from issued equity capital. Clearly we have

$$
\phi \in [0, 1].
$$

(2.19)

Obviously there are practical issues regarding this parameter, as to whether capital can be used to fund the derivative. The base case $\phi = 0$ may best reflect market practice, where there is no explicit use of capital to fund derivatives. Although capital may not be used explicitly to fund derivatives, its existence changes the funding requirements of a bank. Hence, the edge case $\phi = 1$ could be considered the most realistic. The general case will be derived below, although the implementation of Chapter 3 considers the base case $\phi = 0$ to reflect market practice.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$. The base processes are a Brownian motion $(W_t)_{t \geq 0}$ and two jump processes $(J^B_t)_{t \geq 0}$ and $(J^C_t)_{t \geq 0}$, the latter describing a jump to default. The processes are assumed to be adapted to the filtration $\mathbb{F}$, which is assumed to be large enough to carry these two types of processes and to satisfy the usual conditions. A few relevant asset processes are defined in the economy below.

**Definition 2.5.** A semi-replication economy is given by four tradeable instruments: a spot asset $S$ with no default risk; a default risk-free zero-coupon bond $P_R$; a default risky zero-recovery zero-coupon bond $P_C$ of party $C$; and two default-risky, partial recovery zero-coupon bonds $P_1$ and $P_2$. Thus whilst a derivative can be funded by explicitly issuing and-buying-back bonds, specific derivatives, or strategies, cannot be funded by issuing-and-buying-back capital “(Kenyon, Green, 2014, [17]).”
\( P_2 \) of party \( B \) with recovery rates \( R_1 < R_2 \) respectively. Here, a zero coupon bond is a contract that guarantees its holder the payment of a single unit of currency at maturity. Essentially it is a loan\(^2\), with current value \( P_t = P(t,T) \) and terminal value \( P_T = P(T,T) = 1 \). The different assets in the economy have dynamics governed by the stochastic differential equations

\[
\begin{align*}
    &dS_t = \mu_t S_t dt + \sigma_t S_t dW_t \\
    &dP_t^B = rP_t^B dt \\
    &dP_t^C = rC P_t^C dt + P_t^C dJ_t^C \\
    &dP_t^i = ri P_t^i dt - (1 - R_i)P_t^i dJ_t^B 	ext{ for } i \in \{0,1\},
\end{align*}
\]

where parameters satisfy conditions of Theorem 2.2. All the bonds pay 1 at some maturity time \( T \) if the issuing party has not defaulted; otherwise \( P_i \) return \( R_i \) and \( P_C \) returns zero.

The bonds are assumed to have a zero basis between different maturities, such that

\[
\begin{align*}
    &r_i - r = (1 - R_i)\lambda_B, \quad i \in \{1,2\}, \\
    &r_C - q_C = \lambda_C.
\end{align*}
\]

holds. A proof of this statement can be found in (Burgard, Kjaer, 2013, [10]).

We will often refer to \( B \) as the bank (seller) and \( C \) as the counterparty (buyer of a derivative). Consider in this economy a derivative portfolio with economic value \( \hat{V}_t = \hat{V}(t,S_t,J^B_t,J^C_t) \). The economic value incorporates default and funding risk and can hence be decomposed as \( \hat{V} = V + U \), where \( V \) denotes the classic risk-neutral value and \( U \) what turns out to be, the valuation adjustment. This \( VA \) is the part of the price that accounts for default, funding and capital risk.

Consider the value of the portfolio on default of the issuer. It is assumed to take the values

\[
\begin{align*}
    &\hat{V}(t,S,1,0) = g_B(V,X) \\
    &\hat{V}(t,S,0,1) = g_C(V,X),
\end{align*}
\]

where \( g \) are two functions that allow a degree of flexibility in the model around the value of the portfolio after default. \( X \) denotes the collateral posted. Case (2.26) refers to when the seller defaults first and (2.27) refers to the case where the counterparty defaults. It makes sense to define specifically the default conditions

\[
\begin{align*}
    &\hat{V}(t,S,1,0) = (V(t,S) - X)^+ - R_B(V(t,S) - X)^- + X \\
    &\hat{V}(t,S,0,1) = R_C(V(t,S) - X)^+ - (V(t,S) - X)^- + X,
\end{align*}
\]

reflecting the full payment in case of a positive mark-to-market (from the view of the surviving party) and a partial recovery for a negative mark-to-market.\(^3\)

**Definition 2.6.** The bank’s bonds are used to fund or invest any excess cash not funded by collateral. Hence, define the following funding constraint, to be satisfied by the portfolio.

\[
\hat{V} - X + \alpha_1 P_1 + \alpha_2 P_2 = 0,
\]

where \( \alpha_1 \) and \( \alpha_2 \) denote the positions in the portfolio of the bank.

---

\(^2\)The full, formal definition of a zero-coupon bond is given in Section 3.1.1.

\(^3\)In the case of a semi-replication economy of definition (2.5), \( R_C = 0 \) as there is no (partial) recovery.
Denote by $K$ the regulatory capital requirement on the portfolio. The value is generated not only by the capital requirements on the derivative itself, but also on potential hedging transactions, reflecting that regulatory capital applies at the level of the whole portfolio, not individual trades or counterparties. In fact, we have

$$K_t = K(t, V_t, X_t, S_t, P^B_t, P^C_t).$$  \hspace{1cm} (2.31)

The trading desk has to pay for the costs associated with their portfolio. The desk borrows the money $K_t$ from shareholders at a remuneration cost of $\gamma^K_t$ per unit of capital. The cost of capital is, in fact, the cost of the return expected by shareholders for putting their capital at risk. The theorem below formalizes the (portfolio) lifetime cost KVA of holding the regulatory capital $K$ in the semi-replication setup described above.

**Theorem 2.4.** Assume a semi-replication economy, which satisfies funding equation (2.30). In this economy the replication cost of capital, the KVA, is given by

$$KVA_t = -\int_t^T \gamma^K_u e^{-\int_u^t r + \lambda^B_u + \lambda^C_u} ds \mathbb{E}[K_u|\mathcal{F}_t] du.$$  \hspace{1cm} (2.32)

Here $\gamma^K_t$ denotes the cost of capital and $\lambda^B_t$ and $\lambda^C_t$ the default intensities of the seller and buyer of the derivative respectively.

**Remark 2.2.** Indeed, the Capital Valuation Adjustment of a derivative can be represented as the integral over the expected capital profile, taking into account interest rate, the default intensities of both the seller and the counterparty, and the cost of capital.

**Proof.** As in the Black-Scholes framework (Black, Scholes, 1976), the derivative $\hat{V}$ is hedged via a replicating portfolio $\Pi$. The market risk (i.e., the risk-neutral part $V$) can be hedged via the underlying $S_t$, and the default risk of the counterparty is hedged via the defaultable bond $P_C$. The seller’s own default remains unhedged. The portfolio that the seller of the derivative sets up consists of $\delta_t$ units of $S_t$, $\alpha^1_t$ units of $P^B_t$ and $\alpha^2_t$ units of $P^C_t$, $\alpha^C_t$ units of $P_C$ and $\beta_t$ units of cash. The purpose of the portfolio is to hedge out at time $t$ (except at $t = \tau_B$), the value of the derivative, i.e.

$$\hat{V}_t + \Pi_t = 0.$$  \hspace{1cm} (2.33)

The portfolio build-up yields the dynamics

$$d\Pi_t = \delta dS_t + \alpha^1_t dP^B_t + \alpha^2_t dP^C_t + \alpha^C_t dP_C + d\beta_t,$$  \hspace{1cm} (2.34)

where the growth in the cash account $d\beta_t$ can be decomposed as $d\beta_t = d\beta^S_t + d\beta^C_t + d\beta^K_t + d\beta^K_t$.

Here the components represent

1. $\beta^S_t$: the cash change due to underlying (share) dividends. The share provides a dividend income at rate $\delta(t)\gamma^K_t S_t$ and has financing cost at rate $\delta(t)q^K_t S_t$, where $q^K_t$ depends on the risk-free rate $r_t$ and the repo-rate of $S_t$. It follows that $d\beta^S_t = \delta_t (\gamma^K_t - q^K_t) S_t dt$.

2. $\beta^C_t$: the counterparty bond financing cost. The seller will short sell the counterparty bond through a repurchase agreement, and incur financing costs at rate $\beta^C_t = -\alpha^C_t q^K_t P^C_C$. Here $q^K_t$ denotes the bond’s repo rate.

---

4Hence the name semi-replication model; the hedge is perfect except for the case where the seller defaults.

5A repurchase agreement is an agreement to borrow an instrument for a period of time, in exchange for a repo-rate.
3. $\beta_t^K$: the collateral yield. The collateral yield has to be returned to the buyer, resulting in the negative cash flow $d\beta_t^K = -r_X X dt$ for the seller.

4. $\beta_t^K$: the cost of the capital position held by the seller, given by $\beta_t^K = -\gamma_t^K K_t dt$. The equation reflects the cost at which capital is borrowed from shareholders to support trading activities, in order to fulfill regulatory requirements.

As a consequence, the portfolio dynamics (2.34) are given by

$$d\Pi_t = \delta_t dS_t + \alpha_t^1 dP_t^1 + \alpha_t^2 dP_t^2 + \alpha_t^C dP_t^C + \left[ \delta_t (\gamma_t^S - q_t^S) S_t - \alpha_t^C q_t^C P_t^C - r_X X - \gamma_t^K K_t \right] dt,$$

where in fact $\alpha_t^idP_t^i = \alpha_t^1 r_t^i P_t^i dt - \alpha_t^i (1 - R_i) P_t^i dt$ for the bank’s bonds $i \in \{1, 2\}$ and moreover $\alpha_t^C dP_t^C = \alpha_t^C r_t^C P_t^C dt - \alpha_t^C P_t^C dJ_t^C$ for the (zero recovery) counterparty bond. On the other hand, the derivative’s economic value dynamics follow from generalized Itô’s lemma 2.1 as

$$d\hat{V}_t = \frac{\partial \hat{V}_t}{\partial t} + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 \hat{V}_t}{\partial S^2} + \frac{\partial \hat{V}_t}{\partial S} dS_t + \Delta \hat{V}_t^B dJ_t^B + \Delta \hat{V}_t^C dJ_t^C,$$

where $\Delta \hat{V}_t^B = \hat{V}(t, S, 1, 0) - \hat{V}(t, S, 0, 0) = g_B - \hat{V}_t$ and $\Delta \hat{V}_t^C = \hat{V}(t, S, 0, 1) - \hat{V}(t, S, 0, 0) = g_C - \hat{V}_t$. Combining equations (2.35) and (2.36) and condition (2.33) it follows that

$$d\hat{V}_t + d\Pi_t = \left[ \frac{\partial \hat{V}_t}{\partial t} + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 \hat{V}_t}{\partial S^2} + \alpha_t^1 r_t^1 P_t^1 + \alpha_t^2 r_t^2 P_t^2 + \alpha_t^C (r_t^C - q_t^C) P_t^C + \delta_t (\gamma_t^S - q_t^S) S_t - r_t^X X_t - \gamma_t^K K_t \right] dt + \left[ g_B - X_t + P_D - \phi K_t \right] dJ_t^B + \left[ g_C - \hat{V}_t - \alpha_t^C P_t^C \right] dJ_t^C,$$

where $P_t^D = \alpha_t^1 (1 - R_t) P_t^1 + \alpha_t^2 (1 - R_t) P_t^2$ and $(\hat{V} + \alpha_t^1 P_t^1 + \alpha_t^2 P_t^2)$ was replaced by $(\phi K_t + X_t)$ via the funding equation (2.30). The hedging error on default of the issuer $B$ is then $\epsilon_B = g_B - X_t + P_D - \phi K_t$. The remaining sources of risk in (2.37) at $t \neq \tau_B$ can be eliminated by choosing hedging coefficients

$$\delta_t = -\frac{\partial \hat{V}_t}{\partial S} \text{ and } \alpha_t^C = \frac{g_C - \hat{V}_t}{P_t^C}, \quad (2.38)$$

resulting in

$$d\hat{V}_t + d\Pi_t = \left[ \frac{\partial \hat{V}_t}{\partial t} - \frac{\partial \hat{V}_t}{\partial S} (\gamma_t^S - q_t^S) S_t + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 \hat{V}_t}{\partial S^2} + \alpha_t^1 r_t^1 P_t^1 + \alpha_t^2 r_t^2 P_t^2 + (g_C - \hat{V}_t) (r_t^C - q_t^C) 
-r_t^X X_t - \gamma_t^K K_t \right] dt + \epsilon_B dJ_t^B.$$ \hspace{1cm} (2.39)

Use the bond funding equation (2.30) and the issuer bond yields $r_t = r_t^i = (1 - R_t) \lambda^B_t$ to find

$$\alpha_t^1 r_t^1 P_t^1 + \alpha_t^2 r_t^2 P_t^2 = r_t X_t - (r + \lambda^B_t) V_t - \lambda^B_t (g_B - X_B) + r_t \phi K_t.$$ \hspace{1cm} (2.40)

Also applying the counterparty bond yield $\lambda^C_t = r_t^C - q_t^C$ turns equation (2.39) into

$$d\hat{V}_t + d\Pi_t = \left[ \frac{\partial \hat{V}_t}{\partial t} - \frac{\partial \hat{V}_t}{\partial S} (\gamma_t^S - q_t^S) S_t + \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2 \hat{V}_t}{\partial S^2} - (r + \lambda^B_t + \lambda^C_t) V_t + \lambda^B_t g_B + \lambda^C_t g_C - \epsilon_B \lambda^B_t 
+(r_t - r_t^X) X_t + (r_t \phi - \gamma_t^K K_t) \right] dt + \epsilon_B dJ_t^B.$$ \hspace{1cm} (2.41)
At issuer default \( t = \tau_B \), the jump term \( \epsilon_h dJ^B_t \) gives rise to a hedge error of size \( \epsilon_h \). The hedge error can be a windfall or shortfall, where sign and size depend on the post-default value of the own bond portfolio \( P_B \) and the funding strategy employed. On the other hand, until the issuer \( B \) defaults, it will accrue a corresponding cost or gain of size \( -\epsilon_h \lambda_B \) per unit of time. See this as the running spread to pay for potential windfall or shortfall upon default.

Assuming the portfolio \( \Pi_t \) of the seller to be evolving in a self-financing fashion while \( t < \tau_B \), the drift term in (2.41) is required to be zero. Hence, the following partial differential equation must hold for the economic value \( \hat{V} \) of the portfolio:

\[
\frac{\partial \hat{V}}{\partial t} + \mathcal{L}_t \hat{V} - (r_t + \lambda_t B + \lambda_t C) \hat{V} + \lambda_t g_C + \lambda_t B g_B - \epsilon_h \lambda_B - (r_t^X - r_t) X - \gamma_t^K K + r_t \phi K_t = 0,
\]

with terminal condition \( \hat{V}(T, S_T) = H(S_T) \), where \( H \) is the payout of the derivative at maturity. Here \( \mathcal{L}_t \) denotes the operator \( \mathcal{L}_t := \frac{1}{2} \sigma_t^2 S_t^2 \frac{\partial^2}{\partial S^2} - (\gamma_t^C - \delta_t^C) S_t \frac{\partial}{\partial S} \).

Remember that \( \hat{V} = V + U \) and \( V \) satisfies the Black-Scholes partial differential equation

\[
\frac{\partial V}{\partial t} + \mathcal{L}_t V - r_t V_t = 0,
\]

with terminal condition \( V(T, S_T) = H(S_T) \). The decomposition \( \hat{V} = V + U \) allows us to subtract (2.43) from (2.42) to obtain a partial differential equation for \( U \), the aggregated valuation adjustment:

\[
\frac{\partial U}{\partial t} + \mathcal{L}_t U = (r_t + \lambda_t B + \lambda_t C) U - h, \quad U(T, S_T) = 0,
\]

where \( h = (V - g_C) \lambda_C + (V - g_B) \lambda_B + \epsilon_h \lambda_B + s^X X + (\gamma^K - r \phi) K \), denoting \( s^X = r^X - r \). An application of the Feynman-Kac formula in Theorem 2.3 yields

\[
U(t, S_t) = \mathbb{E} \left[ \int_t^T h(u, S_u) \exp \left( -\int_t^u r_s + \lambda_s^B + \lambda_s^C ds \right) du | \mathcal{F}_t \right].
\]

Slicing \( h(u, S_u) \) into the terms \( (V - g_C) \lambda_C, (V - g_B) \lambda_B, \epsilon_h \lambda_B - r \phi K, s^X X \) and \( \gamma^K K \) yields the individual components of the total valuation adjustment \( \hat{U} \) as

\[
U = \text{CVA} + \text{DVA} + \text{FCA} + \text{COLVA} + \text{KVA},
\]

where

\[
\text{CVA}_t = -\int_t^T \lambda_u^C e^{-\int_t^u r+\lambda_s^B+\lambda_s^C ds} \mathbb{E} \left[ V(u) - g_C(V_u, X_u) | \mathcal{F}_t \right] du
\]

(2.47)

\[
\text{DVA}_t = -\int_t^T \lambda_u^B e^{-\int_t^u r+\lambda_s^B+\lambda_s^C ds} \mathbb{E} \left[ V(u) - g_B(V_u, X_u) | \mathcal{F}_t \right] du
\]

(2.48)

\[
\text{FCA}_t = -\int_t^T \lambda_u^B e^{-\int_t^u r+\lambda_s^B+\lambda_s^C ds} \mathbb{E} \left[ \lambda_u^B \epsilon_h - r_u \phi K_u | \mathcal{F}_t \right] du
\]

(2.49)

\[
\text{COLVA}_t = -\int_t^T s_u^X e^{-\int_t^u r+\lambda_s^B+\lambda_s^C ds} \mathbb{E} \left[ X_u | \mathcal{F}_t \right] du
\]

(2.50)

\[
\text{KVA}_t = -\int_t^T \gamma_u^K e^{-\int_t^u r+\lambda_s^B+\lambda_s^C ds} \mathbb{E} \left[ K_u | \mathcal{F}_t \right] du.
\]

(2.51)

\[\text{Some authors collect the third and fourth terms as } FVA = \text{FCA} + \text{COLVA}, \text{ the funding valuation adjustment.}\]
In the case $\phi = 0$ all capital related terms are in equation (2.51). Hence we conclude the Capital Valuation Adjustment is given by (2.51), which matches formula (2.32).

\[ \square \]

Remark 2.3. In the FCA term (and hence, the FVA) there is a dependence on the capital term $K$ both via the hedging error $\epsilon_h$ and the $r \phi K$ term. This reflects the use of capital as funding. In the case $\phi = 0$ this capital term will disappear, such that the KVA is the only capital related valuation adjustment.
2.2 The BSDE model

2.2.1 Backward Stochastic Differential Equations

In this section, backward stochastic differential equations will be introduced. Abbreviated as BSDE’s, the equations are a popular tool in mathematical finance to describe various problems. As we will see later, problems in which they naturally appear relate to pricing and hedging of contingent claims. The backward stochastic differential equations of particular interest for KVA are of the linear type, for which an explicit solution exists. The solution can often be written as a conditional expectation of an integral, as we would (by now) expect in this context. As the classical results on linear BSDE’s of Pardoux and Peng (Pardoux, Peng, 1990, [35]) are sufficient, they will be presented, though along the lines of El Karoui’s BSDE survey (El Karoui, 2008, [12]).

Consider a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) on which is defined a \(d\)-dimensional Brownian motion \(W := (W_t)_{0 \leq t \leq T}\). Define the augmented Brownian filtration \(\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}\) by

\[
\mathcal{F}_t := \mathcal{F}_t^W \vee \mathcal{N},
\]

where \(\mathcal{N}\) denotes the collection of sets with zero probability contained in sets of \(\mathcal{F}_T^W\). It follows from (Spreij, 2018, [41]) Proposition 8.2 that \(\mathcal{F}_t^W\) satisfies the usual conditions and \(W_t\) is still a Brownian motion w.r.t. \(\mathbb{F}\). An \(\mathbb{F}\)-martingale is called a Brownian martingale.

Define in this setting the following spaces.

**Definition 2.7.** Let the spaces \(L^2_{\mathbb{F}}(\mathcal{F}_t), \mathcal{P}_n, S^2_n(0, T), H^2_n(0, T)\) and \(H^4_n(0, T)\) be

1. \(L^2_{\mathbb{F}}(\mathcal{F}_t)\) the set of \(\mathcal{F}_t\)-measurable \(\mathbb{R}^n\)-valued variables \(Y\) with \(E[|Y|^2] < \infty\).
2. \(\mathcal{P}_n\) the set of \(\mathcal{F}_t\)-progressively measurable, \(\mathbb{R}^n\)-valued processes \(X\) on \(\Omega \times [0, T]\).
3. \(S^2_n(0, T)\) the space of processes \(X\) in \(\mathcal{P}_n\) s.t. \(E[\sup_{t \leq T} |X_t|^2] < \infty\).
4. \(H^2_n(0, T)\) the space of processes \(X\) in \(\mathcal{P}_n\) s.t. \(E[\int_0^T |X_s|^2 ds] < \infty\).
5. \(H^4_n(0, T)\) the space of processes \(X\) in \(\mathcal{P}_n\) s.t. \(E[(\int_0^T |X_s|^2 ds)^{1/2}] < \infty\).

Let us now introduce backward stochastic differential equations.

**Definition 2.8.** Let \(\xi_T \in L^2_{\mathbb{F}}(\mathcal{F}_T)\) be a \(\mathbb{R}^m\)-valued terminal condition and \(g\) be a \(\mathbb{R}^m\)-valued coefficient, \(S_T \times \mathcal{B}(\mathbb{R}^n \times \mathbb{R}^{m \times d})\)-measurable, where \(S_T\) denotes the \(\sigma\)-algebra of \(\mathcal{F}_t\)-progressively measurable subsets of \(\Omega \times [0, T]\). A solution for the \(m\)-dimensional BSDE associated with parameters \((g, \xi_T)\) is a pair of progressively measurable processes \((Y, Z) := (Y_t, Z_t)_{0 \leq t \leq T}\) with values in \(\mathbb{R}^m \times \mathbb{R}^{m \times d}\) such that \(Y \in S^2_{m, n}, Z \in H^2_{m \times d}\) and

\[
Y_t = \xi_T + \int_t^T g(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T.
\]

A Backward Stochastic Differential Equation is thus given by a coefficient \(g\) and terminal condition \(\xi_T\).

**Remark 2.4.** The differential form of (2.53) is given by \(-dY_t = g(t, Y_t, Z_t) dt - Z_t dW_t, \quad Y_T = \xi_T\).
A context in which an equation like (2.53) could appear, is the following. Suppose one wants to find the price of a contingent claim \( \xi \geq 0 \) with maturity \( T \). In a complete market, it is possible to construct a replicating portfolio for the claim. The value of this portfolio \( Y \) would be given by the BSDE with linear generator \( g \) and hedging portfolio \( Z \). The price of the claim at time \( t \) is then naturally associated with the value of the hedging portfolio. Under reasonable conditions, BSDE theory can show the uniqueness of the hedging portfolio.

Under specific assumptions on the coefficient \( g \), the BSDE (2.53) has a unique solution. The following (classical) result is due to Pardoux and Peng (Pardoux, Peng, 1990, [35]).

**Theorem 2.5.** If for the coefficient \( g \) it holds that

1. \((g(t,0,0))_{0 \leq t \leq T} \in H^2_m\),
2. \( g \) is uniformly Lipschitz with respect to \((y,z)\), \( dt \otimes d\mathbb{P} \text{ a.s.} \),

then there exists a unique \(^7\) solution \((Y,Z)\) of the BSDE (2.53) with parameters \((g,\xi_T)\).

**Remark 2.5.** If instead of assumption 2 it is only required that \( g \) is uniformly Lipschitz continuous w.r.t. \( z \) and moreover satisfies the following monotonicity condition w.r.t. \( y \): for any \( y,y' \in \mathbb{R}^m \)

\[
g_t(y,z) - g_t(y',z)(y - y') \leq M(y - y')^2,
\]

then the BSDE associated with \((g,\xi_T)\) has a unique solution.

**Proof.** We prove the statement for increasing complexity of the function \( g \).

1. First assume \( g = 0 \). A solution \((Y_t,Z_t)\) of equation (2.53) satisfies

\[
Y_t = \xi_T - \int_t^T Z_s dW_s, \quad Y_T = \xi_T.
\]

(2.55)

It follows that \( Y_t \) is necessarily the continuous version of the \( S^2 \) martingale \( \mathbb{E}[\xi_T | \mathcal{F}_t] \).

The martingale representation theorem (Revuz, Yor, [37]) then yields a unique process \( Z \in H^{2m \times d}_m \) such that

\[
Y_t = \mathbb{E}[\xi_T | \mathcal{F}_t] = Y_0 + \int_0^t Z_s dW_s.
\]

(2.56)

Applying this equation twice, for \( Y \) at times \( t \) and \( T \), yields

\[
Y_t = Y_0 + \int_0^t Z_s dW_s = [\xi_T - \int_0^T Z_s dW_s] + \int_0^t Z_s dW_s = \xi_T - \int_t^T Z_s dW_s,
\]

(2.57)

and hence \((Y,Z)\) is a unique solution for the BSDE with \( g = 0 \).

2. When \( g \) does not depend on \((y,z)\), i.e. \( g(t,y,z) = g(t) \in H^2_m \), it is easy to reduce to case 1. Consider solution \((Y_t,Z_t)\) of equation (2.55) with terminal condition \( \xi_T = \xi_T + \int_0^T g(s) ds \).

Then \((Y_t + \int_0^t g(s) ds, Z_t)\) solves

\[
Y_t = \xi_T - \int_t^T g(s) ds - \int_t^T Z_s dW_s.
\]

(2.58)

As follows from part 1, the solution \((Y_t + \int_0^t g(s) ds, Z_t)\) is unique.

\(^7\)The notion of uniqueness here has to be understood \( dt \otimes d\mathbb{P} \text{ a.s.} \)
3. Finally, consider a general Lipschitz coefficient \( g = g(t, y, z) \). Define on the space of possible solutions for \((Y, Z)\), \(\mathcal{H}_m^2 \times \mathcal{H}_m^{2,d}\), the following norm for some constant \( \alpha > 0 \):

\[
||| (Y, Z)|||_\alpha = \left( \mathbb{E} \left[ \int_0^T e^{\alpha s}(|Y_s|^2 + |Z_s|^2)\,ds \right] \right)^{\frac{1}{2}} .
\] (2.59)

The solution of the Lipschitz BSDE (2.53) will be obtained as a fixed point of the operator \( \Phi : \mathcal{H}_m^2 \times \mathcal{H}_m^{2,d} \to \mathcal{H}_m^2 \times \mathcal{H}_m^{2,d} \) (with respect to norms \(|| . ||_\alpha\)) defined by

\[
(u, v) := (u, v)_{t \leq T} \mapsto \Phi(u, v) := (Y_u^{u,v}, Z_u^{u,v})_{t \leq T},
\] (2.60)

where \((Y^{u,v}, Z^{u,v})\) is the \(\mathcal{H}_m^2 \times \mathcal{H}_m^{2,d}\)-valued solution of the BSDE with coefficient \( g^{u,v}(t) := g(t, u_t, v_t) \), which exists in the scope of case 2. The existence of the fixed point of \( \Phi \) relies on its contraction property, obtained for a norm \(|| . ||_\alpha\) with a large enough \( \alpha \) factor. So take \((u, v), (u', v') \in \mathcal{H}_m^2 \times \mathcal{H}_m^{2,d}\). An application of Itô’s lemma to \( e^{\alpha T}(Y_t^{u,v} - Y_t^{u',v'}) \) yields

\[
e^{\alpha T}(Y_t^{u,v} - Y_t^{u',v'}) - e^{\alpha T}(Y_t^{u,v} - Y_t^{u',v'}) = \alpha \int_t^T e^{\alpha s}|Y_s^{u,v} - Y_s^{u',v'}|^2\,ds + 2 \int_t^T e^{\alpha s}(Y_s^{u,v} - Y_s^{u',v'})(\dot{Y}_s^{u,v} - \dot{Y}_s^{u',v'})\,ds
\]

\[+ \frac{1}{2} \int_t^T e^{\alpha s}(Y_{ss}^{u,v} - Y_{ss}^{u',v'})\,ds,\] (2.61)

where we used that \( Y_s^{u,v} \) and \( Y_s^{u',v'} \) are solutions of the equation (2.53). Taking into account also that both processes have the same terminal value \( \xi_T \), one finds that

\[
e^{\alpha T}(Y_t^{u,v} - Y_t^{u',v'}) + \int_t^T e^{\alpha s}|Z_s^{u,v} - Z_s^{u',v'}|^2\,ds + \alpha \int_t^T e^{\alpha s}|Y_s^{u,v} - Y_s^{u',v'}|^2\,ds
\]

\[= (M_T - M_t) + 2 \int_t^T e^{\alpha s}(Y_s^{u,v} - Y_s^{u',v'})(g(s, u_s, v_s) - g(s, u'_s, v'_s))\,ds,\] (2.62)

where \( M_t = 2 \int_0^t e^{\alpha s}(Y_s^{u,v} - Y_s^{u',v'})(Z_s^{u,v} - Z_s^{u',v'})\,dW_s \), which will turn out to be a uniformly integrable martingale. First, note that for large enough \( K \in \mathbb{R} \)

\[
\mathbb{E} \left[ \left( \int_0^T e^{2\alpha s}[Y_s^{u,v} - Y_s^{u',v'}]^2 \cdot |Z_s^{u,v} - Z_s^{u',v'}|^2\,ds \right)^{\frac{1}{2}} \right]
\]

\[\leq K \mathbb{E} \left[ \sup_{s \leq T} |Y_s^{u,v} - Y_s^{u',v'}| \left( \int_0^T |Z_s^{u,v} - Z_s^{u',v'}|^2\,ds \right)^{\frac{1}{2}} \right],\] (2.63)

\[\leq K \mathbb{E} \left[ \sup_{s \leq T} |Y_s^{u,v} - Y_s^{u',v'}|^2 \right] + K \mathbb{E} \left[ \int_0^T |Z_s^{u,v} - Z_s^{u',v'}|^2\,ds \right],\] (2.64)

where we used \( 2ab \leq a^2 + b^2 \) and in the last step fact that \( Y^{u,v}, Y^{u',v'} \in \mathcal{S}_m^2 \) and \( Z^{u,v}, Z^{u',v'} \in \mathcal{H}_m^{2,d} \) by definition of a BSDE solution. Application of the Burkholder-Davis-Gundy inequalities [Spreij, 2018, [41]] yields for some \( K_1 \in \mathbb{R} \)

\[
\mathbb{E} \left[ \sup_{t \leq T} |M_t| \right] \leq K_1 \mathbb{E} \left[ (M_T)^{\frac{1}{2}} \right] < \infty
\] (2.65)
Consider a backward stochastic differential equation (Definition 2.9. approximation method, for instance stochastic Picard iteration. Y see corollary 2.1 below, the process differential equation given by the pair \((g, \xi)\) recursively as \(Y\) Proposition 2.4. In the third step we use the polarization \(-\alpha x^2 + 2Cxy = -\alpha(x - \frac{C}{\alpha}y)^2 + \frac{C^2}{\alpha^2}y^2 \leq \frac{C^2}{\alpha^2}y^2 \) for all \(x, y \in \mathbb{R}\).
Proof. We apply estimate (2.75) to the Picard iteration sequence. It is clear from the definition of $\Phi$ and equation (2.76) that $(u^k, v^k) = (Y_{t-1}^k, Z_{t-1}^k)$ maps to $(Y_t^k, v^k, Z_{t-1}^k)^k = (Y_t^k, Z_t^k)$. But then the contraction property (2.75) yields
\[
||Y_{t+1}^k - Y_t^k, Z_{t+1}^k - Z_t^k||^2 \leq \frac{K}{\alpha} ||(Y_{t+1}^k - Y_t^k, Z_{t+1}^k - Z_t^k)||^2 \leq (\frac{K}{\alpha})^k ||Y_1^k, Z_1^k||^2, \quad (2.77)
\]
It holds that $\frac{K}{\alpha} < 1$, hence the norm series $||(Y_{t+1}^k - Y_t^k, Z_{t+1}^k - Z_t^k)||^2$ is geometric in $k$. Therefore, both series $(Y_{t+1}^k - Y_t^k)$ and $(Z_{t+1}^k - Z_t^k)$ converge in the Hilbert space $(H^2_{m \times d}, ||\cdot||_2)$. In particular, the sequence of partial sums of the series converges in this space, and $dt \otimes dP$-a.e.

Remark 2.6. In fact, the sequence $(Y_t^k)_{k \in \mathbb{N}}$ of the Picard iteration above converges uniformly almost surely. A proof can be found in (El Karoui, 2008, [12]).

In particular, one can find an explicit expression for $Y$: when $m = 1$ and $g$ is linear.

Corollary 2.1. Let $(\beta, \mu)$ be a bounded $(\mathbb{R}, \mathbb{R}^d)$-valued progressively measurable process. Let $\phi \in \mathcal{H}^2_{\mathbb{R}}(0, T)$ and $\xi_T \in L^2_\mathbb{R}(0, T)$. The linear BSDE given by
\[
dY_t = (\phi_t + \beta_t Y_t + \mu_t Z_t)dt - Z_t dW_t, \quad Y_T = \xi_T, \quad (2.78)
\]
has a unique solution $(Y, Z) \in S^2_T(0, T) \times \mathcal{H}^2_{\mathbb{R}}(0, T)$. Further, $Y$ is given explicitly by
\[
Y_t = \mathbb{E}[\xi_t \Gamma_{t,T} + \int_t^T \phi_s \Gamma_{t,s} ds | \mathcal{F}_t], \quad (2.79)
\]
where $(\Gamma_{t,s})_{s \geq t}$ is the adjoint process defined by $d\Gamma_{t,s} = \Gamma_{t,s}(\beta ds + \mu_s dW_s)$, where $\Gamma_{t,t} = 1$.

Proof. It follows directly from Theorem 2.5 that equation (2.78) has a unique solution $(Y, Z)$. In order to find the representation for $Y$, apply Itô's lemma to find
\[
d(\Gamma_s Y_s) = \Gamma_s dY_s + Y_s d\Gamma_s + (Y, \Gamma)_s = -\phi_s \Gamma_s ds + \Gamma_s(Z_s + \mu_s Y_s) dW_s. \quad (2.80)
\]
Hence $\Gamma_s Y_s + \int_t^s \phi_u \Gamma_u du$ is a local martingale on $[t, T]$ as
\[
\Gamma_s Y_s + \int_t^s \phi_u \Gamma_u du = Y_t + \int_t^s \Gamma_u Z_u + \mu_u Y_u dW_u, \quad (2.81)
\]
where we used $\Gamma_t = 1$. As $\beta$ and $\mu$ are bounded, it follows that $\mathbb{E}[\sup_{s \leq t \leq T} |\Gamma_u|^2] < \infty$. Hence for $K_\mu$ the upper bound of $\mu$ it holds that
\[
\mathbb{E}[\left(\int_t^T \Gamma_u^2 |Z_u + \mu_u Y_u|^2 du\right)^{\frac{1}{2}}] \leq 2^{\frac{1}{2}} \mathbb{E}[\left(\sup_{s \leq t \leq T} |\Gamma_u|^2\right)^{\frac{1}{2}}] + 2 \mathbb{E}[\int_t^T |Z_u|^2 du + 2K_\mu^2 \int_t^T |Y_u|^2 du] < \infty. \quad (2.82)
\]
As $\int_t^s \Gamma_u^2 |Z_u + \mu_u Y_u|^2 du$ is the quadratic variation of $\Gamma_s Y_s + \int_t^s \phi_u \Gamma_u du$ in (2.81), it follows from the classic Burkholder-Davis-Gundy inequality that for some $K_1 \in \mathbb{R}$
\[
\mathbb{E}\left[\sup_{t \leq s \leq T} |\Gamma_s Y_s + \int_t^s \phi_u \Gamma_u du|\right] \leq K_1 \mathbb{E}[\left(\int_t^T \Gamma_u^2 |Z_u + \mu_u Y_u|^2 du\right)^{\frac{1}{2}}] < \infty. \quad (2.83)
\]
But then the process $\Gamma_s Y_s + \int_t^s \phi_u \Gamma_u du$ is a martingale, as it is uniformly bounded. In fact, the martingale property then yields
\[
Y_t = \mathbb{E}[\Gamma_T \xi + \int_t^T \phi_u \Gamma_u du | \mathcal{F}_t]. \quad (2.84)
\]
2.2.2 The BSDE formula

The backward stochastic differential equation approach to calculating the Capital Valuation Adjustment is a modified version of Albanese and Crépey (Albanese, Crépey, 2018, [1]). According to these authors the semi-replication is flawed, by the assumption that KVA is a liability; that it contributes to the trading profit-and-loss (PnL) of the bank. Albanese et al. rather define the KVA as a risk premium stored in a risk margin account, that is used to absorb unexpected losses. The KVA charge must then be gradually released to shareholders as dividend, opposed to directly as profit and loss. The result is an implicit relationship between KVA and capital, described by a backward stochastic differential equation.

As it turns out, this approach is not easy to formalize. The above reasoning requires an extensive accounting overview of the bank, that reviews the purpose of all cashflows. Moreover Albanese and Crépey focus on economic capital, rather than regulatory, defined via expected shortfall\(^9\). As a consequence, their derivations touch upon many accounting aspects. In what follows, a simpler, modified approach is presented for regulatory capital.

\textit{Remark 2.7.} It should be remarked that a modified presentation (for regulatory capital) of the work of Albanese and Crépey can sometimes heavily compress the story and derivations. In case of a desire for more context, the reader is referred to the original paper.

The financial setup for the KVA model via the BSDE approach is based on two distinct but intertwined sources of market incompleteness. First, the instruments to hedge counterparty credit risk, such as Credit Default Swaps (CDSs), are often illiquid. Although efforts are being made to make the market more centralized, most trades are done over the counter, resulting in a lack of data available. Moreover, small counterparties require credit data estimation as a pure CDS is not available. Second, and even harder to replicate, is the risk of losses due to a banks’ own default. Therefore, the sources of market incompleteness are summarized as

i A bank cannot perfectly hedge counterparty default losses,

ii A bank cannot hedge its own jump-to-default exposure.

The bank setup

The bank model assumes two main stakeholders of a bank, shareholders and creditors. Shareholders are in control of the bank before default, and make or influence investment decisions during that time. At potential default of the bank, their investments are lost. The cashflows that arise as part of a default resolution, affect only creditors, i.e., the holders of debt.

The derivatives trading activities of the bank are organised across several desks.

i The CA desk: counterparty default losses and funding expenditures are called contra-assets and are handled at portfolio level of the bank. The CA desk is the (mathematical) entity merged from the classic CVA desk and FVA desk of a bank. The CA desk values the contra-assets and charges the corresponding capital and funding valuation adjustments to the clients, storing the charges as reserve capital. The desk is exposed to the corresponding payoffs: as time proceeds, eventual counterparty default losses and funding expenditures are covered by the desk from the reserve capital account.

\(^9\)Remember Definitions 1.2 and 1.3 of value-at-risk and economic shortfall respectively. The regulatory capital measures are based on value-at-risk and are analytically tractable. Expected shortfall however is more complex and requires additional computation effort.
ii The clean (CL) desks: the clean desks deal with the portfolio after counterparty default risks and (risky) funding expenditures are “stripped off”. They are left with the management of market risk. The clean desks fund their trading from a collateral account, called the clean margin account, with the treasury of the bank\textsuperscript{10}. Hence the desks do not need any cash account.

In fact the bank now has two portfolios. There is the client portfolio, managed by the CA desk, involving all counterparty cashflows. Moreover there is the cleaned portfolio, between the CA desk and the clean desks, that handles only contractually promised cashflows.

Observe that the CA desk is not in charge of the KVA. The capital valuation adjustment is of a different nature and treated separately by the management of the bank: on top of the reserve capital, the bank maintains a separate risk margin account. This account contains the risk margin (i.e., KVA) sourced from clients at each deal, to cover the cost of holding capital for unexpected losses. Moreover, the risk margin account can be used to cover counterparty losses beyond the reserve capital of the CA desk. Hence, the risk margin is also part of the economic capital\textsuperscript{11}. The KVA charge is then only gradually released to shareholders over the lifetime of the portfolio, as the economic capital requirement decreases. This idea will be formalized in the forthcoming sections.

We denote by CM, RC and RM the respective amounts on the clean margin, reserve capital and risk margin accounts for the bank. Moreover each desk, CA or clean, may setup a related hedge. The hedges generate a cumulative hedging loss process intended to control the fluctuations of their respective trading loss-and-profit.

An important assumption, in line with the idea that shareholders lose control over the bank at default, is the following: any residual value on the reserve capital and the risk margin account is transferred from the shareholders to the creditors of the bank, upon default of the bank. This notion will be made explicit when dealing with KVA for a defaultable bank.

A formal approach

The probabilistic setup is as follows. Consider a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) carrying a Brownian motion \((W_t)_{t \geq 0}\) with the augmented Brownian filtration \(\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}\) of equation (2.52).

It is known that \(\mathbb{F}\) satisfies the usual conditions. Let \(\tau\) be a stopping time w.r.t. \(\mathbb{F}\) reflecting the default of the bank\textsuperscript{12} and \([0, T] \subset \mathbb{R}\) denote the portfolio lifetime. Moreover, define \(\bar{\tau} := \tau \wedge T\).

The clean desks and the CA desk both have trading activities and hence generate P&L.

Definition 2.10. The loss-and-profit processes of the respective desks are denoted by \(L_{ca}^t\) and \(L_{cl}^t\), where \(L_{ca}^t = (L_{ca}^t)\tau\) is stopped at bank default. Both processes are assumed to be right-continuous and adapted to \(\mathbb{F}\). The total loss process of the bank, bearing the same conditions, is then

\[
L_t := (L_{ca}^t + L_{cl}^t)\tau = L_{ca}^t + (L_{cl}^t)\tau, \tag{2.85}
\]

which is again stopped at bank default, in line with shareholder interest.

\textsuperscript{10}The exact funding dynamics of the clean desks can be found in (Crepey, 2018).

\textsuperscript{11}This dual role of the risk margin account, and thus the KVA charge, is the key assumption that distinguishes the BSDE approach to KVA from the ‘traditional’ Green-Kenyon approach.

\textsuperscript{12}Consider for example, a structural default model depending on geometric Brownian motion.
Remark 2.8. In the original setup (Albanese, Crépey, 2018, [1]), both loss processes are defined through underlying wealth processes, involving the cashflows $CM$ and $RC$. This decomposition is only relevant for CVA and FVA calculations and will hence not be considered here.

The trading activities of the desks impact the wealth of the bank (shareholders and creditors together). In view of the distinction between shareholders and creditors in the previous section, trading only affects the shareholders’ wealth up to default of the bank. The clean desks ignore default in their pricing, which is then corrected by the CA desk. The CA desk, by nature, has a modeling view at the level of the bank as a whole. Hence, in Definition 2.10, the $P&L$ generated by the CA desk is stopped at $\tau$, while the $P&L$ generated by the clean desks is not.

In practice, losses-and-earnings are typically realized on a quarterly basis. Profits are released as dividends or a recapitalization process might be initiated when losses have occurred. However, for the purpose of modeling, this setup considers the instantaneous reset assumption, according to which losses (and earnings) are marked to model and realized in real time. As a consequence, the trading gains $-L_t = -L^\tau_t$ continuously flow into the shareholder dividend stream. Also taking into account the KVA payments from the risk margin account, until default of the bank, the shareholder cumulative discounted dividends net to

$$-(RM^\tau_t + \bar{L}), \quad (2.86)$$

where the bar denotes (and subsequently will denote) a risk-free discounted process.

Now, we establish an assumption that ensures consistency of valuation across the different desks of the bank. Although markets are assumed to be incomplete, it follows in the spirit of Andersen et al. (Andersen, Duffie, Song, 2016, [2]) that to guarantee consistency of valuation across different desks, one must have martingale pricing. Hence, we assume that $L^{cl}_t$ is an $\mathbb{F}$-local martingale on $[0,T]$ and $L^{ca}_t$ is a $\mathbb{F}$-local martingale on $[0,\bar{\tau}]$ without jump at $\tau$. “The respective martingale assumptions should be viewed as broad shareholder no-arbitrage conditions, as seen from the respective perspectives of the clean and CA desks”, (Crepey, 2018). As a consequence:

**Proposition 2.5.** The trading loss $L_t$ of the bank is a $\mathbb{F}$-local martingale on $[0,\bar{\tau}]$ without jump at $\tau$.

**Proof.** It follows directly from Definition 2.10 and the local martingale assumptions on $L^{ca}_t$ and $L^{cl}_t$. \qed

The bank cannot perfectly replicate its counterparty default losses due to the (second) market incompleteness assumption. Hence, as it maintains its derivative trading activities, the loss-and-profit process $L_t$ is in fact non-vanishing. As a consequence, the bank needs shareholders to put capital at risk to cope with exceptional losses beyond the size of the reserve capital account of the CA desk. The size of this shareholders’ capital at risk SCR is given by

$$SCR_t = C_t - RM_t, \quad (2.87)$$

where $C_t$ is the economic capital level and $RM_t$ the risk margin account\(^{13}\). The value of the risk margin account is subtracted as - by assumption - the risk margin account itself contributes to economic capital. It is clear that the shareholders should be remunerated for putting their capital at risk. Defining $\gamma_K$ as the cost of capital, the shareholder instantaneous average return

\[^{13}\text{Technical conditions on these processes depend on the default-ability of the bank and will be made subsequently.}\]
should be

\[
(\text{Shareholder instantaneous average return})_t = \gamma_K \cdot \text{SCR}_t,
\]

\(= \gamma_K \cdot (C - RM)_t. \)  

(2.88)

(2.89)

Remark 2.9. The risk margin account held by the CA desk preserves capital for unexpected losses in the form of risk margin. The risk margin is sourced from clients at each deal as a capital valuation adjustment, and then gradually released from the RM account into the shareholders dividend stream. Moreover, the bank sends the KVA risk-free accrual into the account. Hence, as the account only carries KVA-related cashflows, from now on we will denote the RM process as \( K_t = K(C_t) \), the size of the capital valuation adjustment account given the capital level \( C_t \).

**KVA in the case of a default-free bank**

Consider the case where the bank cannot default, i.e. \( \tau := \infty \). The capital process for the bank \( C_t \) is assumed to be in \( H^2(0,T) \), such that \( \mathbb{E}[\int_0^T |C_s|^2 ds] < \infty \). Combining equations (2.86) and (2.89), for the KVA charge \( K_t = K_t(C) \) it must hold that the ensuing dividend stream to shareholders\(^{14}\)

\[-(dL_t + dK_t - r_t K_t dt) \]

(2.90)

corresponds to an average expected remuneration to shareholders of

\[\gamma_K(C_t - K_t) dt. \]

(2.91)

The local martingale property of the loss process \( L_t \) then yields the conditions

\[K_T = 0 \text{ and } dK_t + (\gamma_K(C_t - K_t) - r_t K_t) dt \text{ is a local martingale on } [0,T]. \]

(2.92)

Expanding the local martingale property using that \( K_T = 0 \) one finds

\[K_t = \mathbb{E}\left[ \int_t^T \gamma_K C_s - (r_s + \gamma_K) K_s \right] ds | F_t, \ t \in [0,T]. \]

(2.93)

In fact, this is a Backward Stochastic Differential Equation for the KVA process \( K_t \). This can be seen via the martingale representation theorem as follows. For \( t \in [0,T] \) define the martingale

\[M_t := \mathbb{E}\left[ \xi + \int_0^T (\gamma_K C_s - (r_s + \gamma_K) K_s) ds | F_t \right], \]

(2.94)

where \( \xi \) denotes \( K_T \). Assuming that \( K_s \) is square integrable and adapted to \( \mathbb{F} \), the process (2.94) is a square integrable Brownian martingale. As follows from the martingale representation theorem (Revuz, Yor, 1990, [37]), there exists a progressive process \( Z_t \) integrable w.r.t. the Brownian motion \( W_t \) such that for \( t \in [0,T] \), \( M_t \) can be represented as

\[M_t = M_0 + \int_0^t Z_s dW_s. \]

(2.95)

Applying this identity twice yields

\[\xi + \int_0^T (\gamma_K C_s - (r_s + \gamma_K) K_s) ds = M_T = M_0 + \int_0^T Z_s dW_s = M_t + \int_t^T Z_s dW_s, \]

(2.96)

\(^{14}\)Here \( r_t \) denotes the (deterministic) risk-free rate at which the KVA account accrues.
and the defining equation (2.94) for $M_t$ yields the familiar BSDE form

$$K_t = M_t - \int_0^t (\gamma K_s - (r_s + \gamma K_s)K_s)ds$$

$$= \xi + \int_t^T (\gamma K_s - (r_s + \gamma K_s)K_s)ds - \int_t^T Z_s dW_s. \tag{2.97}$$

In fact, this BSDE is of the linear type (2.78). We now solve it for the KVA process $K_t$.

**Proposition 2.6.** Let $C_t$ be a capital process in $H^2_t$. Assume that $r_t$ is bounded from below and in $H^2_t(0,T)$. The linear BSDE (2.97), stated in differential form as

$$-dK_t = (\gamma K_t - (r_t + \gamma K_t)K_t)dt - Z_t dW_t, \quad K_T = 0, \tag{2.98}$$

has a unique, well-posed solution $(K_t, Z_t) \in S^2_t \times H^2_t(0,T)$, where $K_t$ is explicitly given by

$$K_t = \gamma K_t \mathbb{E}\left[\int_t^T e^{-\int_t^s (r_u + \gamma K_u)du} C_s ds \mid \mathcal{F}_t\right]. \tag{2.99}$$

**Proof.** We apply the result of Corollary 2.1. Consider in formula (2.78) the processes $\phi_t = \gamma K_t C_t$, $\beta_t = -(r_t + \gamma K_t)$ and $\mu_t = 0$. It holds by assumption that $\phi_t \in H^2_t(0,T)$ and $\beta_t$ is bounded in $\mathbb{R}$. Hence the equation (2.98) has a unique solution in $S^2_t \times H^2_t(0,T)$. It follows, moreover, that the adjoint process $\Gamma_{t,s}$ satisfies

$$d\Gamma_{t,s} = -(r_s + \gamma K_s)\Gamma_{t,s}ds, \tag{2.100}$$

which is solved by $\Gamma_{t,s} = e^{-\int_t^s (r_u + \gamma K_u)du}$, where indeed $\Gamma_{t,t} = 1$. It follows from $\xi = 0$ that

$$K_t = \gamma K_t \mathbb{E}\left[\int_t^T e^{-\int_t^s (r_u + \gamma K_u)du} C_s ds \mid \mathcal{F}_t\right]. \tag{2.101}$$

Hence the proof is complete. \qed

Hence, the KVA charge in the BSDE setup can be represented as in (2.99). It is again an integral over the expected capital profile. However, we see the cost of capital appear in the discount factor, reflecting the fact that there is less shareholder capital at risk to be remunerated.

Let us now reconsider the capital process $C_t$. It should be clear that the capital process, whether referring to economic or regulatory capital, should be greater than zero. Moreover, shareholder capital at risk cannot be negative. This self-consistency condition leads to the stronger constraint: $C_t - K_t(C)$ should be nonnegative. It reflects the natural idea that the KVA charge should not be larger than the capital of which it reflects the cost. Consider the following definition.

**Definition 2.11.** Define the space $A$ of admissible capital processes as

$$A = \{C_t \in H^2_t(0,T) : C_t \geq \max(C^\text{reg}_t, K_t(C))\}, \tag{2.102}$$

where $C^\text{reg}_t$ denotes the regulatory capital requirement process.

**Remark 2.10.** In view of the definition above, the natural definition for a capital process is $C_t = C^\text{reg}_t \lor K_t(C^\text{reg}_t)$). This is the minimal (hence cheapest) admissible capital process that satisfies both the regulatory acceptability condition of $C_t \geq C^\text{reg}_t$ and the self-consistency condition of $C_t \geq K_t$.
In the context of this minimal admissible capital process, the appropriate reconsideration of equation (2.91) would be \( \gamma_K (C_{t_{reg}}^r \lor K_t) - K_t \). The BSDE (2.98) for KVA would then become of the nonlinear form

\[
-dK_t = (\gamma_K (C_{t_{reg}}^r \lor K_t) - (r_t + \gamma_K)K_t)dt - Z_t dW_t, \quad K_T = 0.
\] (2.103)

As the equation is nonlinear, there is no explicit expression for the process \( K_t \). It is however of the Lipschitz type and existence of a solution is guaranteed by Theorem 2.5.

**Proposition 2.7.** Let \( C_t \) be a capital process in \( H_2^1(0,T) \). Assume that \( r_t \) is bounded from below and in \( H_2^1(0,T) \). The solution of the linear BSDE (2.103) for KVA is well-posed in \( S_2^1(0,T) \).

**Proof.** The equation (2.103) has terminal condition \( \xi_T = 0 \) and coefficient

\[
g_t(y, z) = g_t(y) := \gamma_K (C_{t_{reg}}^r \lor y) - (r_t + \gamma_K)y, \quad y \in \mathbb{R}.
\] (2.104)

It holds for \( t \in [0,T] \) and \( y,y' \in \mathbb{R} \) that

\[
(g_t(y) - g_t(y'))(y - y') = \gamma_K (y - y')[((C_{t_{reg}}^r \lor y) - (C_{t_{reg}}^r \lor y')) - (r_t + \gamma_K)(y - y')^2
\]
\[
= -r_t(y - y')^2 + \gamma_K (y - y')[((C_{t_{reg}}^r \lor y) - y) - ((C_{t_{reg}}^r \lor y') - y')]
\]
\[
\leq -r_t(y - y')^2 \leq M(y - y')^2,
\] (2.105)

as \( r_t \) was assumed to be bounded from below, say by \( M \). The BSDE thus satisfies the monotonicity condition of Remark 2.5. We also have \( g_t(0,0) = \gamma_K C_{t_{reg}}^r \in H_2^1(0,T) \) by assumption. Hence, the BSDE associated with \( (g, \xi_T) \) has a unique solution with \( K_t \) in \( S_2^1(0,T) \). \( \square \)

**KVA in the case of a defaultable bank**

The case of a defaultable bank is more involved. As it no longer holds that \( \bar{\tau} = T \), a random portfolio terminal time enters into the KVA equations. According to Albanese and Crépey, moreover, the clean desks and the CA desk can no longer work in the same mathematical framework, as they view default of the bank in different ways. The clean desks are not concerned with bank default and work with a smaller filtration \( F_{CL} \), to which \( \tau \) is not a stopping time, opposed to the filtration \( \mathbb{F} = F_{CA} \) that contains all model information. A connection between the two is made via \( F_{CL} \) “reductions” of \( F_{CA} \) local martingales, which are linked via an isometry in terms of the default time \( \tau \), as in (Crépey, Song, 2017, [11]). The isometry allows Albanese and Crépey to prove the existence of a KVA-BSDE solution even when the bank may default, at time \( \tau \).

As the primary reason for existence of KVA is the default of a bank’s clients, and the extension to defaultable bank KVA is rather vague in the BSDE context, this thesis considers only the unilateral case of KVA, where the bank cannot default. If interested in the defaultable bank case, the reader is referred to (the rather technical) Section 7.3 of (Albanese, Crepey, 2018, [1]).
3 Implementation

As the principles of Capital Valuation Adjustment have been established in the previous chapter, it is time to convert the theory into a computer implementation. The formulae for unilateral KVA in the SDE and BSDE approaches are respectively

\[
KVA_{0}^{SDE} = \gamma_K \mathbb{E} \left[ \int_{0}^{T} e^{-\int_{0}^{t}(r_u + \lambda^C_u)du} K_t dt | \mathcal{F}_0 \right], \tag{3.1}
\]

\[
KVA_{0}^{BSDE} = \gamma_K \mathbb{E} \left[ \int_{0}^{T} e^{-\int_{0}^{t}(r_u + \gamma_K + \lambda^C_u)du} K_t dt | \mathcal{F}_0 \right], \tag{3.2}
\]

where \(\gamma_K\) denotes a constant cost of capital. It can be seen from the formulae that a practical implementation of both methods will be very similar, as the difference is only a discount factor. The computer implementation will involve simulation of the various parameters over time in a risk-neutral framework. In the end, we expect the KVA under the BSDE approach to be smaller, as there is additional discounting with the cost of capital \(\gamma_K\), opposed to the SDE case.

First and foremost, a model must be specified for the underlying interest rate dynamics. Such a model will not only generate the risk-neutral price of a derivative, but also the required exposure profiles to calculate the capital valuation adjustment. An overview of the fundamentals of interest rate modeling will be given in Section 3.1.1. The interest rates that will be generated are so called forward rates, simulated via the Libor Market Model (LMM), (e.g.; Andersen, Piterbarg, 2010, [3]). The model’s construction and basic properties will be described Section 3.1.2.

Once the theoretical basis of rates, products and the Libor Market Model are established, we construct the simulation engine in Section 3.2. The implementation will involve Monte Carlo simulation, a method where (conditional) expectations are estimated using many sample paths of the underlying interest rate dynamics. Once the basics of Monte Carlo are established, the more advanced Least Squares Monte Carlo algorithm (Longstaff, Schwartz, 2001, [25]) is introduced in Section 3.2.1. The algorithm is then deployed to calculate the required capital profiles, \(K_t\) for \(t \in [0, T]\), in Section 3.2.2, similar to the CVA simulation approach in (Joshi, Kwon, 2016, [24]).

The remainder of the chapter will be devoted to translating the capital profiles into capital valuation adjustments via equations (3.1) and (3.2). In Section 3.3.1 there will be a brief description of the estimation of other parameters prevailing in these equations, such as the cost of capital and the survival probabilities. Although these parameters are important for the KVA calculation, the prime focus will be on the capital profile. Lastly, in Section 3.3.2, the integrals of the KVA equations will be approximated using an appropriate discretization method.

Remark 3.1. As the focus is primarily on the methodology rather than the mathematical derivations, and the fact that implementation calculations can be tedious, the level of rigour in the forthcoming section will be lower than before. Formal proofs will sometimes be omitted; for more details the reader is referred to (Andersen, Piterbarg, 2010, [3]) for the Libor Market Model and, moreover, to (Longstaff, Schwartz, 2001, [25]) and (Glasserman, 2004, [15]) for the LSMC methodology.
3.1 Interest rate modeling

3.1.1 Fundamentals of Interest Rate Modeling

The time value of money is an everyday concept. Money today is worth more than the same amount of money tomorrow\(^1\), as money today has potential earning capacity. Such capacity is expressed in terms of interest rate. A primary financial instrument to trade this interest rate is a bond, which is a securitized form of a loan. In the coming section, the mathematical mechanics underlying such instruments and their derivatives will be outlined much along the lines of (Filipovic, 2009, [14]). A reader familiar with the subject, is recommended to skip to Section 3.1.2, as the current section is merely an introduction before moving on to the Libor Market Model.

The most simple bond is a zero coupon bond.

**Definition 3.1.** A maturity \(T\) zero-coupon bond is a contract that guarantees its holder the payment of 1 unit of currency at time \(T\). Its value at time \(t < T\) is denoted \(P(t, T)\).

In our mathematical framework, we will assume a frictionless market for such bonds for every maturity \(T\). Moreover, it is assumed that \(P(t, T)\) is differentiable in \(T\). Clearly \(P(T, T) = 1\), given that we do also not consider credit issues. Lastly, we assume the market to be arbitrage free.

As the introduction implies, interest rates and bonds both describe the time value of money. As such, bond prices and interest rates are heavily intertwined. The definition of a zero-coupon bond leads to different forward rates, that describe interest rates at various times.

**Definition 3.2.** The simple forward rate for \([T, S]\) prevailing at time \(t\) is given by

\[
F(t, T, S) := \frac{1}{S - T} \left( \frac{P(t, T)}{P(t, S)} - 1 \right). \tag{3.3}
\]

**Remark 3.2.** The simple spot rate \(F(t, T)\) for \([t, T]\) is then

\[
F(t, T) := F(t, t, T) = \frac{1}{T - t} \left( \frac{1}{P(t, T)} - 1 \right). \tag{3.4}
\]

A market example of such simple forward rates are the LIBOR rates: London Interbank Offered Rates. The LIBOR rates are rates at which deposits between banks are exchanged. They are governed and published daily by the British Bankers’ Association, and serve as a reference for many fixed income contracts. The maturities range from one day to twelve months. For example, the three-months forward for the period \([T, T + 1/4]\) is given by

\[
L(t, T) = F(t, T, T + 1/4). \tag{3.5}
\]

**Remark 3.3.** Although one expects fractions as above to have forms as 1/52 for a week or 1/4 for three months, in reality such fractions are governed by fairly complicated market conventions\(^2\).

We will later model these rates in the Libor Market Model.

The simple forward (spot) rate in Definition 3.2 refers to growth of money in a linear fashion: it is assumed to be invested at time \(T\) and returned at time \(S\). The simple forward rate thus ignores the idea of earning interest on interest. In order to capture continuous accrual of money, consider the continuously compounded forward rate in the following definition.

\(^1\)Assuming interest rates are positive and there is no inflation.

\(^2\)For more information on day-count conventions, see for example Filipovic(2009)
Definition 3.3. The continuously compounded forward rate for $[T,S]$ prevailing at time $t$ is given by
\[
R(t,T,S) := -\frac{\log P(t,S) - \log P(t,T)}{S - T}.
\] (3.6)

Remark 3.4. The continuously compounded spot rate is then
\[
R(t,T) := R(t,t,T) = -\frac{\log P(t,T)}{T - t}.
\] (3.7)

Note that the rates in Definition 3.3 are defined in such a way that
\[
e^{R(t,T,S)(S-T)} = \frac{P(t,T)}{P(t,S)}.
\]

Although continuous compounding is practically impossible, the concept of continuously compounded rates forms a good theoretical basis for stochastic interest rate modeling. Rates that are often modeled are either the forward rate or the short rate, which are both derived from the continuously compounded rate $R(t,T,S)$, by taking limits to zero.

Definition 3.4. The (instantaneous) forward rate for time $T$ prevailing at time $t$ is
\[
f(t,T) := \lim_{S \downarrow T} R(t,T,S) = -\frac{\partial \log P(t,T)}{\partial T},
\] (3.8)

and the (instantaneous) short rate prevailing at time $t$ is
\[
r(t) := f(t,t) = \lim_{T \downarrow t} R(t,T).
\] (3.9)

There is a long history to interest rate modeling, starting with a growing range of short rate models. As such models proved to not always be flexible enough to cover the observed initial term structure, forward rate models were developed through the famous HJM framework. Forward rate models aim to capture the entire forward curve directly. As the interest rate modeling for KVA will be done via market models (in particular, the Libor Market Model), short and forward rate models are beyond the scope of this thesis. However, the definitions of short and forward rates are required to define what is called the money market account.

Definition 3.5. The money market account $B$ refers to an asset with $B(0) = 1$ that grows at the instantaneous short rate $r(t)$, i.e.,
\[
B(t) = e^{\int_0^t r(s)\,ds}.
\] (3.10)

The money market account allows one to relate amounts of money at different moments in time. Whereas a bond price $P(t,T)$ is a deterministic price, the money market account defines a stochastic discount factor, in the sense that one unit of currency at time $T$ is currently worth
\[
\frac{B(t)}{B(T)} = e^{-\int_T^t r(s)\,ds}.
\]

The bond price $P(t,T)$ will turn out to be the expectation of the discount factor $B(t)/B(T)$ under a certain risk-neutral measure (see Definition 3.6 below).

In order to continue with the definition and pricing of various interest rate products, it is necessary to define a mathematical market in which we operate. The market must reasonably reflect the real world and be mathematically consistent.

---

3HJM refers to Heath, Jarrow and Morton who proposed the framework in the late eighties (Heath, Jarrow, Morton, 1991).
Definition 3.6. Consider a filtered probability space \((\Omega, F, \mathbb{P})\), where \(F\) satisfies the usual conditions. The bond market consists of bonds for every maturity \(T\), such that their price processes \(P(t, T)\) satisfy

1. \(P(t, T)\) is adapted to \(F\) for all \(T\),
2. \(P(t, T)\) is everywhere differentiable in \(T\),
3. \(P(T, T) = 1\) for all \(T\).

Further, assume there exists an equivalent martingale measure \(Q \sim \mathbb{P}\) such that the discounted bond price processes \(P(t, T)/B(t)\) are martingales under \(Q\) for \(t \leq T\), for all maturities \(T\).

Remark 3.5. The existence of an equivalent martingale measure in the bond market induces a no-arbitrage condition via the First Fundamental Theorem of Asset Pricing; see Filipovic.

Remark 3.6. The condition \(P(T, T) = 1\) together with the martingale property of \(P(t, T)/B(t)\) yields the relationship

\[
P(t, T) = E\left[e^{-\int_t^T r(u) \, du} \mid F_t\right].
\]

As zero coupon bonds and the money market account are defined in a formal setup above, one can study more complex fixed income\(^4\) products in the bond market. As the final goal is an interest rate swap, the following section will consider the most basic products. Consider the definition of fixed coupon bonds, floating rate notes and ultimately interest rate swaps below.

Definition 3.7. A fixed coupon bond is a contract specified over a number of future tenors \(T_1 < \ldots < T_n\), a sequence of deterministic coupons \(c_1, \ldots, c_n\) and a nominal value, such that the owner of the contract receives \(c_i\) at time \(T_i\) for \(1 \leq i \leq n\) and an additional notional value \(N\) at terminal time \(T_n\).

The price \(p_{FCB}(t)\) at time \(t \leq T_1\) is given by the sum of discounted cashflows, i.e.,

\[
p_{FCB}(t) = \sum_{i=1}^{n} P(t, T_i)c_i + P(t, T_n)N. \tag{3.11}
\]

We see that the value of the coupon is fixed at the time the bond is issued. However, one might also be interested in coupon bonds whose coupon values are reset for every period. The reset is determined by a rate prevailing in the market, e.g. LIBOR.

Definition 3.8. A floating rate note is a contract specified by a number of future dates \(T_0 < T_1 < \ldots < T_n\) and a nominal value \(N\), such that the owner of the contract receives \(c_i = (T_i - T_{i-1})F(T_{i-1}, T_i)N\) at time \(T_i\) for \(1 \leq i \leq n\). Here \(F(T_{i-1}, T_i)\) is the prevailing (simple) interest rate at time \(T_{i-1}\).

It can be seen from the definition of \(F(T_{i-1}, T_i)\) that \(c_i = N \cdot \left(1/P(T_{i-1}, T_i) - 1\right)\), which by a no-arbitrage argument has a present value of \(N \cdot \left(P(t, T_{i-1}) - P(t, T_i)\right)\). Hence the price of a floating rate note at time \(t \leq T_0\) is given by

\[
p_{FRN}(t) = N \cdot \left(P(t, T_n) + \sum_{i=1}^{n} P(t, T_{i-1}) - P(t, T_i)\right) = N \cdot P(t, T_0). \tag{3.12}
\]

\(^4\)Fixed income is a term that refers to products under which the borrower or issuer is obliged to make payments of a fixed amount on a fixed schedule.
The price formula is surprisingly simple. Intuitively, though, it makes sense: at each time $T_i$ the coupon $c_i$ is reset with respect to the prevailing interest rate, resulting in exactly the then-present value of the cashflow at time $T_i$.

The main product of interest for KVA in this project, is an interest rate swap. An interest rate swap (IRS) is an agreement to periodically exchange interest rates on a given notional value.

**Definition 3.9.** An interest rate swap is a contract specified by a number of future dates $T_0 < T_1 < \ldots < T_n$, a fixed rate $K$ and a nominal value $N$, such that the owner of the contract at time $T_i$ pays fixed $K(T_i - T_{i-1})N$ and receives floating $F(T_{i-1}, T_i)(T_i - T_{i-1})N$, resulting in a net cash flow at time $T_i$ of

$$N \cdot \left( F(T_{i-1}, T_i) - K \right) \cdot \tau_i,$$

where $\tau_i = T_i - T_{i-1}$ denotes the tenor interval length.

The price of an interest rate swap $P_{IRS}(t)$ at time $t \leq T_0$ is

$$P_{IRS}(t) = N \cdot \left( P(t, T_0) - P(t, T_n) - K \sum_{i=1}^n \tau_i P(t, T_i) \right),$$

as can be seen from discounting the coupon cashflows in similar fashion as before.

In practice, one distinguishes between a **payer** and **receiver** interest rate swap, which describes whether the owner of the contract pays or receives the fixed leg of the contract. Definition 3.9 describes a payer swap, and the price of a receiver swap is given by

$$P_{IRS(R)} = -P_{IRS(P)}.$$

Another common practice is to trade interest rate swaps at *par*, which means that the swap rate $SR$ is set such that the value of the traded contract is zero. Inverting equation (3.13) reveals that this par swap rate is given by

$$SR_{par}(t) = \frac{P(t, T_0) - P(t, T_n)}{\sum_{i=1}^n \tau_i P(t, T_i)}.$$  

Banks tend to trade interest rate swaps for various reasons. Clients of a bank often want to hedge interest rate exposure on their portfolio of loans. Such companies swap their floating rate for a fixed rate. On the other hand, banks have to manage their interest exposure as well and hedge risk in the interbank market. In general, interest rate swaps have interbank offered rates as a reference index. Consequently in the late nineties, interest rate models started to focus on modeling the market observable xIBOR\(^5\) rates. The next section presents the LIBOR market model, according to which we will model the price and exposure of interest rate swaps.

**3.1.2 The Libor Market Model**

Over the course of recent history, a wide range of interest rate models has been developed. Until the late nineties, models generally focused on two types of rates: the short rates $r(t)$ and the forward rates $f(t, T)$, as defined in equations (3.9) and (3.8) respectively\(^6\). Although such models allow for successful pricing and hedging of a variety of securities, some of the more exotic

\(^5\)The x here refers to many possible prefixes describing different interbank rates, such as “L” or “EUR”.

\(^6\)A study of short rate and forward rate models can be found in (Filipovic, 2009, [14]).
derivatives require richer interest rate dynamics to be priced properly. In particular, the models were hard to calibrate, as the short rates are not directly observable in the market.

A solution to these and other problems with the existing models was proposed in 1997, by Brace et al. (Brace, Gatarek, Musiela, 1997, [6]), through the Libor Market Model framework. They insisted on modeling a set of (simply compounded) Libor forward rates, of which the current value is directly observable in the market. The advantage of such an approach is not only (relatively) easy calibration, but also the circumvention of more technical issues that we will omit here. In the forthcoming section, the model and various properties will be described, along the lines of (Andersen, Piterbarg, 2010, [3]) and (Brigo, Mercurio, 2006, [8]).

Consider a bond market as in Definition 3.6, that is assumed to be carrying a \(d\)-dimensional Brownian motion \(W(t)\) under the martingale measure \(Q\). The aim of the LMM is to model the Libor forward rates \(L(t, T_j, T_{j+1})\) on a tenor structure \(0 = T_0 < T_1 < ... < T_n\), where \(T_i = T_{i+1} - T_i\) denote the tenor intervals for \(0 \leq i \leq n-1\). At any given time point \(t\) in this structure we focus on the prevailing zero-coupon bonds \(P(t, T_j)\) for \(i\) such that \(t < T_i \leq T_n\).

To this extent, consider the index function \(q(t) \in \{1, ..., n\}\) that denotes the index of the next tenor, i.e. for any \(t\) up to \(T_n\) it holds that \(T_{q(t)-1} \leq t < T_{q(t)}\). In accordance with Definition 3.2 the Libor forward rates are defined as

\[
L_j(t) := L(t, T_j, T_{j+1}) = \frac{1}{\tau_j} \left( \frac{P(t, T_j)}{P(t, T_{j+1})} - 1 \right),
\]

(3.15)

for \(q(t) \leq j \leq n-1\). In the forthcoming story, for any forward Libor rate \(L_j(t)\) it is implicitly assumed that \(q(t) \leq j\), such that the rate at time \(t\) is still stochastic.

At time \(t\), the prevailing rates \(L_{q(t)}(t), L_{q(t)+1}(t), ..., L_{n-1}(t)\) remain to be modeled. Let the dynamics under the risk-neutral measure \(Q\) be given by

\[
dL_j(t) = \sigma_j(t) \mu_j(t) dt + dW(t),
\]

(3.16)

where \(\mu_j\) and \(\sigma_j\) are \(d\)-dimensional adapted processes, such that a strong solution exists (cf. Theorem 2.2).

The modus operandi is to define a smart set of probability measures such that each forward rate \(L(t, T_i, T_{i+1})\) is a martingale under exactly one of those measures, providing driftless dynamics, as opposed to equation (3.16). The dynamics of all the rates \(L(t, T_0, T_1), ..., L(t, T_{n-1}, T_n)\) can then be aligned under the last (terminal) measure, allowing for proper simulation of rates, to calculate price and exposure of interest rate swaps.

The new measures are defined through their Radon-Nikodym derivative, using zero coupon bonds for various maturities. Define for each \(i \in \{1, ..., n\}\), the \(T_i\) forward measure \(Q^{T_i} \sim Q\) by

\[
\frac{dQ^{T_i}}{dQ} = \frac{1}{P(0, T_i)B(T_i)}.
\]

(3.17)

Without further specification, let \(W^{T_i}\) be the Brownian motion \(W\) under the new measure. The most important property of these measures is given by the proposition below.
Proposition 3.1. Let $X$ be a price process under the risk-neutral measure $\mathbb{Q}$. Under the forward measure $\mathbb{Q}^{T_i}$, the $P(t,T_i)$-discounted price process $X_t$ is a martingale. As a consequence, the following pricing formula holds:

$$X_t = P(t,T_i)\mathbb{E}_{\mathbb{Q}^{T_i}}[X_{T_i} | F_t].$$ \hfill (3.18)

Proof. Adaptedness is trivial and integrability follows from $\mathbb{E}_{\mathbb{Q}^{T_i}}[|X|] = \mathbb{E}_{\mathbb{Q}}\left[\frac{|X|}{P(0,T_i)B(T_i)}\right] < \infty$.

It follows from (3.17) that

$$\frac{d\mathbb{Q}^{T_i}}{d\mathbb{Q}} | F_t = \mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{Q}^{T_i}}{d\mathbb{Q}} | F_t\right] = \frac{P(t,T_i)}{P(0,T_i)B(t)}.$$

To prove the martingale property, apply Bayes’ formula to find

$$\mathbb{E}_{\mathbb{Q}^{T_i}}\left[\frac{X_{T_i}}{P(t,T_i)} | F_t\right] = \frac{\mathbb{E}_{\mathbb{Q}}\left[\frac{X_{T_i}}{P(t,T_i)/B(t)} | F_t\right]}{\mathbb{E}_{\mathbb{Q}}\left[\frac{P(t,T_i)/B(t)}{P(0,T_i)/B(0)} | F_t\right]} = \frac{\mathbb{E}_{\mathbb{Q}}\left[\frac{X_{T_i}}{B(T_i)} | F_t\right]}{\mathbb{E}_{\mathbb{Q}}\left[\frac{P(t,T_i)}{B(T_i)} | F_t\right]} = \frac{X_t}{P(t,T_i)},$$

where we used that the fact that the $B(t)$-discounted bond price processes are martingales under the risk-neutral measure $\mathbb{Q}$. The pricing formula then follows from $P(T_i,T_i) = 1$. \hfill $\square$

The proposition outputs the pricing formula (3.18) of claims under the various forward measures in the Libor Market Model, that will be used later. But moreover, it provides a key property of the bond price fractions $P(t,T_i)/P(t,T_{i+1})$ that will make the model work.

Corollary 3.1. Under measure $\mathbb{Q}^{T_{i+1}}$, the $T_{i+1}$-bond discounted bond price $P(t,T_i)$ is a martingale.

Proof. Take forward measure $\mathbb{Q}^{T_i}$ and (bond) price process $P(t,T_i)$ in Proposition 3.1. \hfill $\square$

It can be seen from the defining equation (3.15) of the Libor forward rate $L_j(t)$ why this property is useful. The process $L_j(t)$ is a martingale under a particular measure, if the fraction $P(t,T_j)/P(t,T_{j+1})$ is a martingale under that same measure. Hence, we conclude:

$L_j(t)$ is a martingale under the forward measure $\mathbb{Q}^{T_{j+1}}$.

The $\mathbb{Q}$ dynamics given by (3.16) thus become under the forward measure $\mathbb{Q}^{T_{j+1}}$

$$dL_j(t) = \sigma_j(t)^T dW^{j+1}(t).$$ \hfill (3.19)

At this stage, it is important to emphasize that the $[T_j,T_{j+1}]$ Libor forward rate is only a martingale under the $T_j$ forward measure. The (drift containing) dynamics of $L_j(t)$ in various other forward measures is derived in the following proposition.

Proposition 3.2. Let $L_j(t)$ satisfy the $\mathbb{Q}^{T_{j+1}}$ dynamics of (3.19). Under the previous forward measure $\mathbb{Q}^{T_j}$, the process $L_j(t)$ is given by

$$dL_j(t) = \sigma_j(t)^T \left( \frac{\tau_j \sigma_j(t)}{1 + \tau_j L_j(t)} \right) dW^j(t),$$ \hfill (3.20)

where $W^j(t)$ is a $d$-dimensional Brownian motion under $\mathbb{Q}^{T_j}$.
Proof. Consider the Radon-Nikodym derivative $Z(t)$ that relates $Q_{T_j}$ and $Q_{T_{j+1}}$. It is given by

$$Z(t) = \mathbb{E}_{Q_{T_{j+1}}} \left[ \frac{dQ_{T_j}}{dQ_T} | \mathcal{F}_t \right] = \mathbb{E}_{Q_{T_{j+1}}} \left[ \frac{dQ_T}{dQ_{T_{j+1}}} \frac{dQ_{T_{j+1}}}{dQ} | \mathcal{F}_t \right] = \frac{P(t, T_j)/P(0, T_j)}{P(t, T_{j+1})/P(0, T_{j+1})} = (1 + \tau_j L_j(t)) \frac{P(0, T_{j+1})}{P(0, T_j)}.$$ 

Hence it follows that

$$dZ(t) = \frac{P(0, T_{j+1})}{P(0, T_j)} \tau_j \sigma_j(t)^T dW^{j+1}(t),$$

and thus

$$\frac{dZ(t)}{Z(t)} = \frac{\tau_j \sigma_j(t)^T}{1 + \tau_j L_j(t)} dW^{j+1}(t).$$

As a consequence of the Girsanov theorem, it follows that

$$dW^j(t) := dW^{j+1}(t) - \frac{\tau_j \sigma_j(t)^T}{1 + \tau_j L_j(t)} dt$$

is a Brownian motion under $Q_{T_j}$. Hence, the $L_j(t)$ dynamics under this measure are given by

$$dL_j(t) = \sigma_j(t)^T \left( \frac{\tau_j \sigma_j(t)}{1 + \tau_j L_j(t)} \right) dW^j(t).$$

The proposition can be iterated to find dynamics of $L_j$ under any of the forward measures.

Corollary 3.2. Let $L_j(t)$ satisfy the $Q_{T_{j+1}}$ dynamics of (3.19). Under the terminal measure $Q_{T_n}$ the process $L_j(t)$ is given by

$$dL_j(t) = \sigma_j(t)^T \left( - \sum_{i=j+1}^{n-1} \frac{\tau_i \sigma_i(t)}{1 + \tau_i L_i(t)} dt + dW^n(t) \right), \tag{3.21}$$

where $W_n(t)$ is a $d$-dimensional Brownian motion under $Q_{T_n}$.

Proof. This follows by iteratively applying Proposition 3.2.

We now have dynamics for all the $n$ Libor forward rates $L_j(t)$ under the same probability measure $Q_{T_n}$. The dynamics depend solely on up to $n$ volatility functions. It should be stressed that in the initial $Q$ dynamics of (3.16), a different volatility function $\sigma_j(t)$ can be picked for each Libor forward rate $L_j(t)$. The Libor Market Model thus provides great freedom in the specification of the forward curve dynamics. In practice, the volatility functions are often specified as affine transformations of the Libor rates themselves. The most common choice is a simple scalar transformation, i.e.

$$\sigma_j(t) = \lambda_j(t)L_j(t), \tag{3.22}$$

for some bounded vector-valued deterministic functions $\lambda_j(t)$, $j \in \{1, ..., n\}$. Although terminology can be ambiguous, the functions $\lambda_j(t)$ will further be referred to as instantaneous volatilities.

Summarizing, the Libor Market Model is defined by
1. A tenor structure $T_0 < T_1 < \ldots < T_n$,

2. A set of initially prevailing Libor forward rates $L_0(0), \ldots, L_{n-1}(0)$,

3. A set of volatility functions $\lambda_1(t), \ldots, \lambda_n(t)$.

The forward rate dynamics are then specified as

$$\frac{dL_j(t)}{L_j(t)} = \lambda_j(t)^\top \left( - \sum_{i=j+1}^{n-1} \frac{\lambda_i(t) \tau_i L_i(t)}{1 + \tau_i L_i(t)} dt + dW^n(t) \right).$$

In order to price interest rate derivatives in this model, one needs to specify the instantaneous volatility functions $\lambda_j(t)$ for all Libor rates $L_0(t), \ldots, L_{n-1}(t)$. The functions are in practice calibrated to prevailing market prices\footnote{A key observation here is that a change of measure as conducted in (3.19), does not change the volatility of the underlying process. This is sometimes known as the diffusion invariance principle, see for example Piterbarg(2010). Hence the calibration can be performed regardless of the model measure.}. Although a wide panel of calibration methods exist and the preferred approach depends on the instrument to price, the general calibration procedure can be outlined as follows. One takes actively traded interest rate derivatives, which have an analytical pricing formula that depends on the (norm of the) instantaneous volatility functions $\lambda_j$, such as optionality products like caplets or swaptions. Using these formulae, the prevailing market prices can be inverted to implied volatility norms. Assuming a particular (parametric) form of the volatility functions, the functions can be estimated from the implied volatilities via norm optimization, taking into account correlation information. Once the volatility functions are available, exotic derivative prices can calculated via LMM simulation.

Although the calibration procedure sounds straightforward, it involves many details and as such can be performed in many different ways. Calibration is out of scope for this thesis, but more on calibration approaches can be found, among others, in (Andersen, Piterbarg, 2010, [3]).
3.2 Monte Carlo simulation

Once the Libor Market Model has been calibrated to market data, the parameterized model is ready to calculate price and exposure of interest rate swaps. The model is implemented on a computer via one of the most common numerical methods in mathematical finance: Monte Carlo simulation (MC). The method revolves around computing expectations of a random variable by sampling a large number of realizations. The price of a derivative claim \( X \) at \( T \) in the bond market of Definition 3.6 is given by the risk-neutral expectation:

\[
V(t) = B(t)\mathbb{E}_Q\left[ \frac{X}{B(T)} | \mathcal{F}_t \right].
\]  

(3.23)

The conditional expectation, and hence the price of the derivative, is estimated via Monte Carlo simulation in a simple three step process.

1. Simulate independent paths of the collection of Libor rates through time,

2. Calculate the derivative value on each path by summation of all discounted future cashflows,

3. Take the average value over all paths to approximate the expectation in (3.23).

The approximation works due to the law of large numbers (Glasserman, 2004, [15]).

An application of this procedure to a Libor Market Model parameterization works as follows. Fix a time grid \( 0 = t_0 < t_1 < \ldots < t_m \) over which to simulate. It is practical to include tenor dates \( T_1, \ldots, T_n \) among the simulation dates, or even set the simulation dates equal to tenor dates. In the following, we assume the latter; \( m = n \) and \( t_i = T_i \) for all \( 0 \leq i \leq m \). The initially available forward rates are given by \( L_0(0), \ldots, L_{n-1}(0) \). Furthermore, remember from Corollary 3.2 and equation (3.22) the forward rate dynamics under the terminal measure to be given by the SDE’s

\[
dL_j(t) = -L_j(t) \sum_{i=j+1}^{n-1} \frac{\lambda_i(t) \tau_i L_i(t)}{1 + \tau_i L_i(t)} dt + L_j(t) \lambda_j(t) dW^n(t).
\]  

(3.24)

The dynamics can be rewritten by applying Itô’s lemma with \( f(x) = \log(x) \), to the exponential form

\[
dL_j(t) = L_j(t) \exp \left( -\lambda_j(t) T \sum_{i=j+1}^{n-1} \frac{\lambda_i(t) \tau_i L_i(t)}{1 + \tau_i L_i(t)} - \frac{1}{2} \lambda_j^2(t) \right) dt + \lambda_j(t) dW^n(t).
\]  

(3.25)

These equations can be simulated over the simulation grid \( t_0, \ldots, t_m \) via Euler approximation. Given the value \( L_j(t_k) \) of a forward rate at a simulation point \( t_k \), the value at the next simulation point \( t_{k+1} \) is approximated by one of the equations

\[
L_j(t_{k+1}) = L_j(t_k) \exp \left( -\lambda_j(t_k) T \sum_{i=j+1}^{n-1} \frac{\lambda_i(t_k) \tau_i L_i(t_k)}{1 + \tau_i L_i(t_k)} \right) \tau_k + \lambda_j(t_k) L_j(t_k) \sqrt{\tau_k} Z_{k+1},
\]  

(3.26)

\[
L_j(t_{k+1}) = L_j(t_k) \exp \left( -\lambda_j(t_k) T \sum_{i=j+1}^{n-1} \frac{\lambda_i(t_k) \tau_i L_i(t_k)}{1 + \tau_i L_i(t_k)} - \frac{1}{2} \lambda_j^2(t_k) \right) \tau_k + \lambda_j(t_k) \sqrt{\tau_k} Z_{k+1},
\]  

(3.27)

*This follows directly from the martingale property in Definition 3.6.
where \( Z_{k+1} \) is a standard normal random variable. The equations (3.26) and (3.27) are often referred to as the Euler scheme and Log-Euler scheme respectively. The latter method has the advantage of Gaussian increments, which improves the convergence of the regular Euler scheme. However, without displacement, the second method does not allow for negative interest rates\(^9\).

An example simulation is demonstrated below.

![Figure 3.1: LMM forward rate sample paths with full Brownian motion correlation (left) and no Brownian motion correlation (right). The rate \( L_i(t) \) is only stochastic up to \( T_i \). Image from (Filipovic, 2009, [14]).](image)

**Remark 3.7.** The Euler schemes presented here are not very sophisticated, but are appealing in terms of presentability and computational efficiency. Advanced schemes can be found in, for example, (Glasserman, 2010, [15]).

Once \( M \) paths of the relevant forward rates are simulated over the time grid \( t_1, ..., t_n \), one is able to price the derivative in question. A path \( m \) for \( 1 \leq m \leq M \) is denoted by the time sequence of the vector \( (L_{0}^{(m)}(t), ..., L_{n-1}^{(m)}(t)) \), where it should be noted that \( L_{k}^{(m)}(t) = L_{k}^{(m)}(t_k) \) for \( t > t_k \) as the rate is no longer stochastic on this interval. See Figure 3.1.

As the simulation is done under the terminal measure \( Q^T_n \), it follows from Proposition 3.1 that the price of a \( T_j \)-claim \( X \) is given by

\[
V_j(t) = \mathbb{E}_{Q^T_n} \left[ \frac{P(t,T_j)}{P(T_j,T_n)} X | \mathcal{F}_t \right].
\]

Hence, it can be seen from Definition 3.9 that simulation is not strictly necessary to price an interest rate swap. However, the simulation methodology was illustrated here as it is crucial to generate exposure profiles, in the coming section.

\[^9\text{Note that the regular Euler-scheme does allow for negative interest rates, even when considering a non-negative interest rate model, by construction of the approximation (3.26).}\]
3.2.1 Least Squares Monte Carlo

As follows from Section 1.2, calculating the regulatory capital requirement today is rather simple. In a Capital Valuation Adjustment context, one is interested in the future capital profile of a trade. The future capital profile $K_t$ depends on the exposure of a trade: how positive (or negative) is the value of the trade given the market parameters prevailing at time $t$? As calculating the value of a product requires an idea of future payoffs, the capital $K_t$ calculation is thus a future outlook in a future outlook. The brute force approach would be to do a nested Monte Carlo simulation, where at each future time point, in each simulation path, a new simulation is initiated. As such an approach results in combinatorial explosion, it is not a viable solution.

As it turns out, fortunately, the concept of outlook in outlook has already been studied extensively in the context of American options. The valuation of American options, that can be exercised at any given time before maturity, entails finding the optimal exercise rule: at any time point $t$, one needs future information to determine whether or not to exercise the option. A wide range of solutions has been established over time, the most prominent being the Least Squares Monte Carlo (LSMC) simulation developed by Longstaff and Schwartz (Longstaff, Schwartz, 2001, [25]). The outline of this regression-based approach will be given here for American options, whereas in the next section the method is tailored to the KVA problem, for exposure and capital profiles.

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the augmented filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$. On this space, consider an asset $S_t$ that constitutes an arbitrage-free economy with a risk-neutral measure $\mathbb{Q}$. The goal is to price an American option on this asset, that can be exercised on a discrete set of times

$$0 = t_0 < t_1 < ... < t_n = T,$$

by finding the optimal stopping strategy. In order to do so, a Monte Carlo simulation for the underlying value $S_t$ is performed on (at least) this grid, according to some risk neutral dynamics.

Now for each simulation path $m$, define the value $C^{(m)}_t$ as

$$C^{(m)}_t = \sum_{s \geq t} c^{(m)}_s ,$$

the sum of the option cashflows in path $m$ between time $t$ and $T$, conditional on two premises:

1. the option has not yet been exercised at time $t$
2. the option holder follows the optimal stopping strategy between $t$ and $T$.

The aim of the LSMC procedure that follows is to approximate the optimal stopping strategy in a pathwise fashion. The approximation is performed by backwards induction as follows: at time $t_n = T$, it is clear that the option holder exercises the option if it is in the money. At any time $t_k$ for $k < n$, the option should be exercised if the immediate exercise value is greater than or equal to the (expected) value of continuation. The continuation value depends on the cashflows after time $t_k$ and is hence unknown at this point. The theory of risk-neutral pricing yields that the continuation value is given by the expectation of discounted optimal future cashflows, i.e.

$$V^{(m)}_{t_k} = \mathbb{E}_Q \left[ \sum_{j=k+1}^n e^{-\int_{t_j}^{t_{j+1}} r(m) ds} C^{(m)}_{t_j} | \mathcal{F}_{t_k} \right],$$

As remarked in the introduction of this chapter, technical details will be omitted. Optimal in the sense that the exercise policy for the option yields maximum profit.
where $r^{(m)}_t$ denotes the discount rate on path $m$. The optimal strategy is thus found by comparing the continuation value (3.30) at $t_k$ with the direct exercise value of the option at $t_k$.

One (naive) solution would be to calculate in each simulated path $m$ at each exercise date $t_k$, the conditional expectation (3.30) via a new Monte Carlo simulation using information $F_{t_k}$, i.e. an asset value simulation initiated from the value $S^{(m)}_{t_k}$. A nested Monte Carlo simulation like this is however computationally unfeasible. The LSMC method provides a solution by approximation of the continuation value, not by “horizontal” simulation, but via a “vertical” regression.

Assume the conditional expectation (3.30) is an element of the $L^2$ space of square integrable functions with respect to $Q$. As this is a separable Hilbert space, it has a countable, orthonormal basis. Hence, the continuation value at time $t_k$ can be expressed as the infinite sum

$$V^{(m)}_{t_k} = \sum_{i=0}^{\infty} \alpha^{t_k}_i F_i(S^{(m)}_{t_k})$$

for coefficients $\alpha^{t_k}_i \in \mathbb{R}$ and basis functions $F_i(x)$ for all $i$, e.g., Laguerre or Hermite polynomials. The key observation here is that the continuation value at time $t_k$, can thus be expressed as a linear combination of basis functions of the underlying state variable at time $t_k$, which is in $F_{t_k}$.

The LSMC method applies this observation in the following way. According to equation (3.31), there exists a single function that describes for every path $m$, the continuation value $V^{(m)}_{t_k}$ in terms of the random variable $S_{t_k}$. But a set of state values of this variable is readily available through the initial Monte Carlo simulation; namely the paths $S^{(j)}_{t_k}$ for $1 \leq j \leq M$. Hence, one can use the information of all simulated paths at time $t_k$ to estimate the continuation value function coefficients that works for all continuation values $V^{(m)}_{t_k}$, by minimizing the $L^2$ distance

$$\sum_{m=1}^{M} \left[ \sum_{i=0}^{I} \alpha^{t_k}_i F_i(S^{(m)}_{t_k}) - \sum_{j=k+1}^{n} e^{-\int_{t_j}^{t_k} r^{(m)}_s \, ds} C^{(m)}_{t_j} \right]^2$$

between the regression value and the realized future optimal cashflows per path. Note that the sum is taken over a finite number $I$ of regression coefficients. The regression works backwards starting at $t_{n-1}$ and is re-performed at each exercise date $t_{n-1}, t_{n-2}, ..., t_0$.

As the regression step is now defined, the algorithm procedure can be summarized in the following way. At time $t_n$, determine the optimal cashflows by evaluating the exercise value. Proceed then, recursively for time $t_k$ in $t_{n-1}, t_{n-2}, ..., t_0$, as follows:

1. At time $t_k$, estimate the conditional expectation (3.30) via regression,

2. Identify the optimal stopping strategy at time $t_k$ by comparing the (estimated) continuation value and the value of immediate exercise,

3. Determine, if $t_k > t_0$, the option cash flow paths $C^{(m)}_{t_k}$ at the previous time step for all $m$.

---

12 A lengthy but very illuminating example of calculating such a pathwise optimal stopping procedure for an American put option is provided in (Longstaff, Schwartz, 2001, [25]).

13 In equation (3.31) it is implicitly assumed that the asset value process is Markov, such that only current state variables are necessary. For non-Markovian processes, one would include both current and past state variable realizations in the basis functions $F_i(\cdot)$ (Longstaff, Schwartz, 2001, [25]).

14 Hence the name, “vertical” regression.
Having determined the optimal exercise strategy for each path, the option is valued as follows: average over all paths the value of the discounted cashflow over that path, at the optimal stopping time for that path. Hence, denoting the optimal time to exercise on path \( m \) by \( \tau_m \in \{ t_0, \ldots, t_n \} \), the price is given by

\[
V_{t_0} = \frac{1}{M} \sum_{m=1}^{M} e^{-\int_{t_0}^{\tau_m} C^{(m)}}.
\] (3.33)

The value of the option is equal to the maximized value of the discounted cash flows of the option, where the maximum is taken over all stopping times with respect to the relevant filtration. Under reasonable assumptions, Longstaff and Schwartz prove that the estimate (3.33) converges to the true option value in probability for an underlying process \( S_t \) that is Markov.

Subsequently, this methodology will be tailored to the KVA context, where the regression procedure is used to estimate moneyness of a product, rather than optimal exercise.

### 3.2.2 Exposure and capital profiles

An application of this regression-based approach has already been deployed in the XVA context. Here, a modified version of the CVA approach in (Joshi, Kwon, 2016, [24]) is presented for KVA.

The final goal of the implementation are the integrals (3.1) and (3.2), as a result from the SDE and BSDE approaches respectively. Key element of both formulae are the future capital profiles \( K_t \) for \( t \in [0, T] \). The capital levels according to equations (1.25) and (1.48) are given by

\[
K^{CCR}_t = RW^{CCR}_t \cdot EAD_t,
\] (3.34)

\[
K^{CVA}_t = 2 \sqrt{h} \cdot RW^{CVA}_t \cdot M_t \cdot EAD_t,
\] (3.35)

for counterparty credit risk and CVA capital respectively. The variables risk weight \( RW_t \) and maturity \( M_t \) are deterministic, hence for expected capital profiles it suffices to consider the stochastic exposure-at-default \( EAD_t \). The \( EAD_t \) is provided through two regulatory measures, the current-exposure method and the \( SA-CCR \) method, defined by

\[
EAD^{CEM}_t = V_t 1_{\{V_t > 0\}} + \psi^{CEM}_t \] (3.36)

\[
EAD^{SA-CCR}_t = 1.4 \cdot (V_t 1_{\{V_t > 0\}} + \psi^{SA-CCR}_t), \] (3.37)

where \( V_t \) denotes the value of the product and the \( \psi_t \) denote the potential future exposure. It should be noted that in the latter formula, the \( (SA-CCR) \) potential future exposure is stochastic as well as the product value. See equation (1.38). The value is rewritten as (cf. equation (3.28))

\[
V_t = \mathbb{E}_{Q^{T_n}} \left[ P(t, T_n) \sum_{j=q(t)}^{n} \frac{c_j}{P(T_j, T_n)} | \mathcal{F}_t \right],
\] (3.38)

where \( c_j = N \cdot \tau_j \cdot (L_{j-1}(T_{j-1}) - K) \) denotes the cashflow of the swap at time \( T_j \).\(^{16}\)

---

\(^{15}\)In stochastics and mathematical finance, this pricing principle relates to the Snell envelope, which is the smallest supermartingale dominating a stochastic process.

\(^{16}\)Note here that although the procedure is demonstrated for an interest rate swap, it can easily be extended to any derivative with future cashflows \( c_j \).
The exposure-at-default processes are treated separately per regulatory regime.

1. Consider first the current-exposure-method version \(EAD_{t}^{CEM}\). The main concern of this process is the time-zero expectation of \(V_t \mathbb{1}_{\{V_t > 0\}}\), for which the Least Squares Monte Carlo methodology will be employed. The time-zero expectation of \(EAD_{t}^{CEM}\) is calculated as

\[
E_{Q} \mathbb{E}_{t}^{CEM} EAD_t^{CEM} | F_0 = E_{Q} \mathbb{E}_{t}^{CEM} V_t \mathbb{1}_{\{V_t > 0\}} + \psi_t | F_0 
\]

\[
= E_{Q} \mathbb{E}_{t}^{CEM} [P(t, T_n) \sum_{j=q(t)}^{n} \frac{c_j}{P(T_j, T_n)} \mathbb{1}_{\{V_t > 0\}} | F_0] + E_{Q} \mathbb{E}_{t}^{CEM} \psi_t | F_0 
\]

\[
= E_{Q} \mathbb{E}_{t}^{CEM} [P(t, T_n) \sum_{j=q(t)}^{n} \frac{c_j}{P(T_j, T_n)} \mathbb{1}_{\{V_t > 0\}} | F_0] + E_{Q} \mathbb{E}_{t}^{CEM} \psi_t | F_0 
\]

\[
= E_{Q} \mathbb{E}_{t}^{CEM} [P(t, T_n) \sum_{j=q(t)}^{n} \frac{c_j}{P(T_j, T_n)} \mathbb{1}_{\{V_t > 0\}} | F_0] + E_{Q} \mathbb{E}_{t}^{CEM} \psi_t | F_0 . \tag{3.39}
\]

A Monte Carlo simulation will be performed in order to calculate the value of the expected exposure-at-default. A crucial consequence of the derivation above, is that the condition on information \(F_t\) appears only in the indicator \(\mathbb{1}_{\{V_t > 0\}}\) and no longer in the actual product value, as a consequence of the tower property. The time \(t\) swap values can hence be calculated from the paths simulated at time zero, whereas the “moneyness” indicator \(\mathbb{1}_{\{V_t > 0\}}\) must be estimated given the information at time \(t\). The Least Squares Monte Carlo methodology will be deployed.

Consider again the simulation grid \(0 = t_0 < t_1 < ... < T_n\) which contains the swap coupon dates. Abbreviating the notation of the numeraire and coupon values, one derives similarly to equations (3.28) and (3.29) the numerical approximation\(^{17}\) at time \(t \in \{t_1, ..., t_n\}\)

\[
E_{Q} \mathbb{E}_{t}^{CEM} EAD_t^{CEM} | F_0 = E_{Q} \mathbb{E}_{t}^{CEM} [P(t, T_n) \sum_{j=q(t)}^{n} \frac{c_j}{P(T_j, T_n)} \mathbb{1}_{\{V_t > 0\}} | F_0] + E_{Q} \mathbb{E}_{t}^{CEM} \psi_t | F_0 \tag{3.40}
\]

\[
\approx \frac{1}{M} \sum_{m=1}^{M} \sum_{j=q(t)}^{n} \frac{P(m)(t, T_n)}{P(m)(T_j, T_n)} c_j \mathbb{1}_{\{y_t^{(m)} > 0\}} + \psi_t, \tag{3.41}
\]

where \(y_t^{(m)}\) is a regression value to estimate the time \(t\) conditional expectation \(V_t^{(m)}\).

\(\text{Remark 3.9.}\) Compare at this point the following: in the American option situation, the continuation value (3.30) is estimated by regression to find the indicator \(\mathbb{1}_{\{t_n = t_k\}}\) \(t_k\). Subsequently only this indicator is used to price the option and not the underlying estimate (3.31) itself. Similarly, in the case of an exposure profile, the moneyness indicator \(\mathbb{1}_{\{V_t > 0\}}\) is estimated via a regression procedure, while the size of the exposure comes from the paths \(m\) as in (3.42).

The regression performed is now based on a simple polynomial. At every time \(t \in \{t_{n-1}, ..., t_1\}\) a regression is performed across all simulated paths at time \(t\) to find coefficients \(\alpha_0^t, \alpha_1^t, \alpha_2^t\) such that

\[
\sum_{m=1}^{M} |\alpha_0^t + \alpha_1^t x_t^{(m)} + \alpha_2^t (x_t^{(m)})^2 - \sum_{j=q(t)}^{n} \frac{P(m)(t, T_n)}{P(m)(T_j, T_n)} c_j |^2 \tag{3.42}
\]

\(^{17}\)Note that under the current exposure method \(\psi_t\) is deterministic and hence is not averaged over all paths.
Note that the multiplier \( m \) where \( \psi \) time-zero expectation of the potential future exposure smaller than 1 if \( V \) \( E \) the-money. Hence, simulating the Libor forward rates under the terminal measure \( 18 \) ity is a consequence of Jensen’s inequality. Setting the multiplier equal to the upper bound

\[
y_t^{(m)} := \alpha_0 + \alpha_1 x_t^{(m)} + \alpha_2 (x_t^{(m)})^2.
\]

(3.44)

The question remaining is to find a regression variable \( x_t \) which is able to describe the “mon-
eyness” behaviour of \( V_t \) in a proper manner. In the American option case, the regression was performed on the underlying \( S_t \). In the swap case, generally \( x_t \) is a \( t \)-present value of a simpler product, priced on the same simulated Libor rates as the original swap. In practice many different products can be used, even simultaneously, to obtain good regression accuracy.

2. Consider the second exposure-at-default version \( EAD_i^{SACCR} \). The methodology for this regulatory regime is very similar to the first, however slightly more complicated due to the stochastic nature of the term \( \psi \). Remember the potential future exposure to be given by, see equation (1.38),

\[
\psi_t = m_t \cdot A_t,
\]

(3.45)

where \( A_t \) denotes the deterministic add-on and \( m_t \) the multiplier depending, again, on the moneyness of the swap. The multiplier, where \( f = 0.05 \), is given by

\[
m_t = m_t(V_t, A_t) = \min \left\{ 1, f + (1 - f) \exp \left( \frac{V_t}{2(1 - f) A_t} \right) \right\}.
\]

(3.46)

Note that the multiplier \( m_t \) at time \( t \) is given by 1 if the mark-to-market value \( V_t > 0 \) and smaller than 1 if \( V_t < 0 \), reflecting the reduction of capital requirements if the swap is out-of-the-money. Hence, simulating the Libor forward rates under the terminal measure \( E_{Q^T_n} \), the time-zero expectation of the potential future exposure \( \psi_t \) in (3.45) is given by

\[
E_{Q^T_n} \left[ \psi_t | F_0 \right] = E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t > 0\}} + A_t \mathbb{1}_{\{V_t < 0\}} \left( f + (1 - f) \exp \left( \frac{V_t}{2(1 - f) A_t} \right) \right) \right] | F_0
\]

(3.47)

\[
= E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t > 0\}} \right] | F_0
\]

\[
+ E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t < 0\}} \left( f + (1 - f) \exp \left( \frac{E_{Q^T_n} \left[ Z | F_t \right]}{2(1 - f) A_t} \right) \right) \right] | F_0
\]

\[
\leq E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t > 0\}} \right] | F_0
\]

\[
+ E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t < 0\}} \left( f + (1 - f) \exp \left( \frac{Z}{2(1 - f) A_t} \right) \right) \right] | F_0
\]

\[
= E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t > 0\}} \right] | F_0
\]

\[
+ E_{Q^T_n} \left[ E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t < 0\}} \left( f + (1 - f) \exp \left( \frac{Z}{2(1 - f) A_t} \right) \right) \right] | F_t \right] | F_0
\]

\[
= E_{Q^T_n} \left[ A_t \mathbb{1}_{\{V_t > 0\}} + A_t \mathbb{1}_{\{V_t < 0\}} \left( f + (1 - f) \exp \left( \frac{Z}{2(1 - f) A_t} \right) \right) \right] | F_0
\]

(3.48)

where \( Z = P(t, T_n) \sum_{j=0(t)}^{n} e^{c_j} P(t_j, T_n) \) denotes the time \( t \) value of future cashflows, and the inequality is a consequence of Jensen’s inequality. Setting the multiplier equal to the upper bound\(^{18} \)

\(^{18}\)It follows from the Taylor approximation \( e^x \approx 1 + x + \ldots \) for \( |x| < 1 \) that for trades with \(-1 < \frac{V_t}{2(1 - f) A_t} < 0\),

this approximation works sufficiently well. If a trade has a very negative mark-to-market \( V_t \) with respect to the add-on value \( A_t \), the potential future exposure value might be well overestimated.
in (3.48), one obtains for the exposure-at-default under the SA-CCR method

$$E_{Q^*} \left[ EAD_t^{SACCR} | \mathcal{F}_0 \right] = 1.4 \cdot E_{Q^*} \left[ V_t \mathbb{1}_{\{V_t > 0\}} + \psi_t | \mathcal{F}_0 \right]$$

$$= 1.4 \cdot E_{Q^*} \left[ V_t \mathbb{1}_{\{V_t > 0\}} | \mathcal{F}_0 \right] + 1.4 \cdot E_{Q^*} \left[ \psi_t | \mathcal{F}_0 \right]$$

$$= 1.4 \cdot E_{Q^*} \left[ Z \mathbb{1}_{\{V_t > 0\}} | \mathcal{F}_0 \right] + 1.4 \cdot E_{Q^*} \left[ A_t \mathbb{1}_{\{V_t > 0\}} + A_t \mathbb{1}_{\{V_t < 0\}} \left( f + (1 - f) \exp \left( \frac{Z}{2(1 - f)A_t} \right) \right) | \mathcal{F}_0 \right]$$

(3.49)

where $Z$ is given by $Z = P(t, T_n) \sum_{j=q(t)}^{n} \frac{e_j}{T_j - T_n}$. Again, cf. equation (3.40), it follows from the derivation above that the condition on $\mathcal{F}_t$ only appears in the indicator $\mathbb{1}_{\{V_t > 0\}}$ and no longer in the actual product value. The time $t$ swap values can be calculated from the paths simulated at time zero, whereas the moneyness indicator must be estimated given the information at time $t$. Applying the LSMC methodology as before, it follows that

$$E_{Q^*} \left[ EAD_t^{SACCR} | \mathcal{F}_0 \right]$$

$$= 1.4 \cdot E_{Q^*} \left[ Z \mathbb{1}_{\{V_t > 0\}} | \mathcal{F}_0 \right]$$

$$+ 1.4 \cdot E_{Q^*} \left[ A_t \mathbb{1}_{\{V_t > 0\}} + A_t \mathbb{1}_{\{V_t < 0\}} \left( f + (1 - f) \exp \left( \frac{Z}{2(1 - f)A_t} \right) \right) | \mathcal{F}_0 \right]$$

(3.50)

$$= \frac{1}{M} \sum_{m=1}^{M} Y_t^{(m)} \mathbb{1}_{\{y_t^{(m)} > 0\}} + A_t \mathbb{1}_{\{y_t^{(m)} > 0\}} + A_t \mathbb{1}_{\{y_t^{(m)} > 0\}} \left( f + (1 - f) \exp \left( \frac{Z_t^{(m)}}{2(1 - f)A_t} \right) \right)$$

(3.51)

$$= \frac{1}{M} \sum_{m=1}^{M} Y_t^{(m)} \mathbb{1}_{\{y_t^{(m)} > 0\}} + A_t \mathbb{1}_{\{y_t^{(m)} > 0\}} + A_t \mathbb{1}_{\{y_t^{(m)} > 0\}} \left( f + (1 - f) \exp \left( \frac{Z_t^{(m)}}{2(1 - f)A_t} \right) \right)$$

(3.52)

where $Z_t^{(m)} = \sum_{j=q(t)}^{n} \frac{e_j^{(m)}}{T_j^{(m)} - T_n^{(m)}}$, similar to (3.42), and $y_t^{(m)}$ represents the time $t$ regressed value

$$y_t^{(m)} = \alpha_0^{(m)} + \alpha_1^{(m)} x_t^{(m)} + \alpha_2^{(m)} (x_t^{(m)})^2$$

Summarizing, the calculation of expected capital profiles come down to calculation of expected exposure-at-default, either with respect to the CEM regime or the SACCR regime. In both cases, the EAD profile relates to the positive exposure, and can be estimated using the Least Squares Monte Carlo method. In the SACCR case, moreover, the potential future exposure is estimated using the same regression variables. In both estimation procedures, the regression values are only employed to define an indicator and not an actual value.

**Remark 3.10.** The section above describes the LSMC methodology in the scope of interest rate swaps. It should be noted that the regression procedure is not strictly necessary for such swaps, as the value $V_t$ can be analytically computed via (3.13) at time $t$ using the simulated rates. The same holds for a variety of products with analytical pricing formulas. In general, however, one wants to calculate KVA for fixed income products without an analytical pricing formula, like caps, floors or swaptions. In that case, the above setup is particularly useful.
3.3 The KVA

3.3.1 Non-capital parameters

The final step before calculating the KVA integral is to define the few non-capital parameters. The parameters in scope are the survival probability and the cost of capital. In this section, a brief overview is given of how these parameters are obtained.

1. The survival probabilities at the various simulation points are taken from a survival curve. Similar to the discount curve, a credit (survival) curve can be bootstrapped from market prices of liquidly traded assets; in this case, for products for which the value can be expressed in terms of the survival probability of the underlying - e.g. credit default swaps. As an intermediate step, the hazard rate function is calibrated to the market prices, yielding survival probabilities via a reduced form model as in equation (1.11). An example model is in (Hull, White, 2000, [21]). The construction of the survival curve is not simple, due to lack of market data availability. Relatively small counterparties may not have single name CDSs written on them, and for larger counterparties the longer dated maturities are often illiquid. Another issue is the parametrisation of the hazard rate function, which is a modeler’s choice. The most conventional option is a piecewise constant function, as there is no information on CDSs between maturities. However, because this construction process is not directly relevant for KVA, and already in place at most banks, the survival probabilities for the implementation are considered to be readily available.

2. The cost of capital is a major driver of KVA, as it linearly scales the value. It reflects the return a bank must make in order to remunerate its shareholders for providing capital to the business. The shareholders of a bank appoint a board of directors, who allocate capital to the various business lines, as illustrated in the Chapter 1 introduction. As such, the cost of capital depends on the choices made by the management of the bank, and may vary from institution to institution, depending each’s history and culture. Generally speaking the intended return on equity, usually around 10%, is a good proxy. Although the cost of capital is thus idiosyncratic on a bank-to-bank level, one may argue that it must even be evaluated on client or trade level. A bank might take a low return on a particular derivative transaction if it generates other business flow with the same client, resulting in the desired return on equity for the client as a whole. As we will not further enter the discussion on cost of capital, the value in the forthcoming implementation is taken to be 10%.

3.3.2 The KVA integral

As well known by now, the Capital Valuation Adjustments are integrals over the future capital profiles, taking into account also the cost of capital, discounting and in the unilateral case the survival probability of the counterparty. The previous sections have described how the values are generated on a grid

\[ 0 = t_0 < t_1 < ... < t_n = T, \]
where the payment dates of the derivative (i.e. swap) are among the grid\textsuperscript{19}. The integral formulas (3.1) and (3.2) can be approximated on this grid via the trapezoidal rule, as follows.

\[
KVA_0^x = \gamma_K \mathbb{E} \left[ \int_0^T d_u^x \cdot e^{-\int_0^u \lambda_u^c \cdot d_u} \cdot K_t dt | \mathcal{F}_0 \right]
\]

\[
= \frac{\gamma_k}{2} \sum_{i=0}^{n-1} \left( d_{t_i}^x \cdot e^{\lambda_{t_i}^c} \cdot K_{t_i} + d_{t_{i+1}}^x \cdot e^{\lambda_{t_{i+1}}^c} \cdot K_{t_{i+1}} \right) \cdot \left( t_{i+1} - t_i \right)
\]

where \( x \in \{ \text{SDE, BSDE} \} \) and \( d_{t_k}^x \) denotes the relevant discount factor. Here the capital profile is given as in equations (3.34) and (3.35), i.e.

\[
K_{t_i} = K_{t_i}^{CCR} + K_{t_i}^{CVA}
\]

\[
= R W_{t_i}^{CCR} \cdot EAD_{t_i} + \frac{2.33}{2} \cdot \sqrt{1} \cdot RW_{t_i}^{CVA} \cdot M_{t_i} \cdot EAD_{t_i},
\]

where the exposure-at-defaults are calculated from the regression procedure. These exposure values come from either of the CEM or SACCRR methodologies.

\textsuperscript{19}One can apply interpolation methods for the parameters for which this is not the case; this will not be covered here.
4 Results

In the final part of this thesis, we present some KVA calculation results obtained following the approaches described in the previous sections. In particular, this section is organised as follows.

The results of the methodology are presented in a simple, step-by-step case study of an interest rate swap in Section 4.1. The section demonstrates respectively the initial trade and market properties, as well as the simulated exposure-at-default and capital profiles and the resulting KVA values, using real market data. The case study should provide the reader sufficient insight in the methodology to understand the more detailed results later in the next two sections.

Section 4.2 studies the impact of various model input parameters to the resulting KVA numbers. The parameters in scope are the maturity and moneyness of the swap and the credit rating of the counterparty. The results are separated per regulatory framework considered, i.e. CEM and SACCR, as the impact of the parameters above can vary per regulatory calculation.

The last part of Section 4 considers a few practical matters. In practice, KVA calculations are performed at netting set level; simultaneously for many trades in a portfolio. The section demonstrates how netting effects can drastically influence the KVA numbers. Moreover, it shows how to handle incremental KVA, whereby one calculates the KVA of a trade that is added to an existing portfolio. Next, the regulatory switch is explained, which allows us to cover both regulatory regimes for trades with a maturity beyond the SACCR implementation date. Finally, the impact of collateral on KVA numbers is considered.
4.1 Introduction: a case study

An example of a Capital Valuation Adjustment calculation is presented here. We consider an interest rate swap with a 5-year maturity, from the 1st of January 2018 until the 2nd of January 2023. The swap is fixed-vs-floating (EURIBOR) with a coupon tenor of 3 months and a notional value of 10,000,000. The swap is traded at par, such that the initial value is zero, cf. (3.14). The trade counterparty is fictional and is assumed to have a Standard & Poor’s credit rating of A.

The (market) setting as of the 1st of January 2018 is as follows. The yield curve for the lower tenors is negative, as market observable interest rates are negative. The counterparty CDS quotes are converted into hazard rates as per Section 3.3.1, from which in turn a survival curve can be constructed. The curves are illustrated in Figure 4.1 below.

The Libor Market Model calibrated to these market conditions is used to simulate the forward rates. The simulation points match the swap coupon dates, such that the simulation grid

$$0 =: t_0 < t_1 < ... < t_n := T$$

is in fact given by the 1st of January plus 3M, 6M, 9M, etc. In this case, no interpolation of simulated rates is required as the swap can be directly valued on simulation points.

As soon as forward rates have been simulated, the swap can be valued at future dates. At each simulation timepoint $t_i$, the least squares regression is performed to estimate the exposure-at-default for the capital profiles\(^1\). The expected EAD values are calculated under both the CEM and SACCR methodologies. In both cases, the exposure-at-default consists of a replacement cost and a potential future exposure part. The first is the same in both cases, while the latter is different, although under the SACCR method both are multiplied by the alpha factor 1.4. The potential future exposure is a deterministic step function in the CEM case, while in the SACCR case it is a smooth, deterministic function boosted by the stochastic multiplier. The resulting EAD profiles are displayed in Figure 4.2.

\(^1\)Remember the regressed values are used for the replacement cost, as well as the SACCR potential future exposure.
As at each time step the risk weights are also (deterministically) available, such that the conversions to counterparty credit risk capital $K_{\text{CCR}}$ (3.34) and CVA capital $K_{\text{CVA}}$ (3.35) can be performed. The risk weight for counterparty credit risk depends on the remaining maturity while the CVA risk weight is constant. The capital profiles can be found in figure 4.3 below.

Once the expected capital profiles are simulated on the grid, the actual Capital Valuation Adjustment can be calculated. The discrete integral (3.53) is given by

$$KVA_x^0 = \frac{7k}{2} \sum_{j=0}^{n-1} \left( d_{t_j}^x \cdot e^{\lambda_{t_j}^C} \cdot K_{t_j} + d_{t_{j+1}}^x \cdot e^{\lambda_{t_{j+1}}^C} \cdot K_{t_{j+1}} \right) \cdot (t_{j+1} - t_j),$$

where the discount factors $d_{t_j}^x$ depend on the KVA method $x \in \{SDE, BSDE\}$ and the factors $\exp(\lambda_{t_j}^C)$ are taken from the survival curve. The result is a Capital Valuation Adjustment, that should be charged upon inception of the deal, in addition to its risk-neutral price. The latter is zero in this case study, as the swap was traded at par. Table 4.1 displays the results for all combinations of KVA methodology, KVA type and regulatory framework.
Table 4.1: The KVA numbers for the swap in the case study under the respective regulatory regimes. Results are moreover separated per capital type and discount method.

<table>
<thead>
<tr>
<th></th>
<th>SDE</th>
<th>BSDE</th>
<th></th>
<th>SDE</th>
<th>BSDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCR capital charge</td>
<td>708.53</td>
<td>613.31</td>
<td>CCR capital charge</td>
<td>1,785.18</td>
<td>1,590.32</td>
</tr>
<tr>
<td>CVA capital charge</td>
<td>911.71</td>
<td>795.86</td>
<td>CVA capital charge</td>
<td>1,674.72</td>
<td>1,502.17</td>
</tr>
</tbody>
</table>

As can be seen in the table above, the KVA charge is much smaller under the CEM regulation. The result follows from the multiplier of 1.4 in the SACCR EAD as well as the higher potential future exposure; cf. figure 4.2. The SDE method results in a lower value than the BSDE method, which is a simple consequence of the larger discount factor $d_{t_i}^{SDE} > d_{t_i}^{BSDE}$ in the first case. A last, but most interesting observation that can be made is a comparison between counterparty credit risk capital and CVA risk capital numbers. Under the CEM regime, the CVA capital charge is larger than the CCR capital charge. Under the SACCR regime however, the CCR capital charge is larger. The reason is as follows: the SACCR exposure-at-default is generally 1.4 times higher than the CEM exposure-at-default, and enters into both the $CCR$ and $CVA$ charge via equations (3.34) and (3.35). In the latter equation for CVA capital, however, the alpha factor of 1.4 is divided out of the EAD, to convert the EAD back to EEPE, effective expected positive exposure$^2$. Hence the CCR capital is larger, as the actual (non-divided) EAD is used.

$^2$The effective expected positive exposure is defined as the cumulative maximum of the positive exposure, which is used for CVA capital calculations under the internal model method. For more details, consult Green(2015).
4.2 Impact of model parameters

A long list of parameters goes into the KVA model. The primary set are market parameters, such as the prevailing interest rates and interest rate product data, as well as credit product data. These parameters influence primarily the risk-neutral price of a trade, but also the various Valuation Adjustments, such as KVA. As we have seen, the KVA depends on the exposure of a counterparty, which via simulation depends on these market data. The next set of input parameters are trade properties: the maturity, notional and collateral agreements of a trade all influence the KVA value. Capital wise some trades perform much better than others, based on these parameters. A third set of parameters revolve around the counterparty with which the trade is conducted. Consider the credit rating of a counterparty (which relates to prevailing market credit data), but also the regulatory status of the counterparty and the existing portfolio and risks with it. The regulatory status brings us to the last parameter: regulation. As mentioned in the introduction, KVA numbers are very sensitive to regulatory changes. It is important for banks to be well up to date with regulations coming ahead.

In this section, the impact of some of these parameters is studied. The main market parameter that will be considered, is the moneyness of a trade. At inception, a trade can be settled at par, or at a price that is far in- or out-of-the-money. It turns out that this influences the KVA number, as the replacement cost is directly related to moneyness of a trade. In the category of trade properties, maturity is most imperative. For instance, consider the situation where two trades have similar market-risk properties, but different duration. In this case, it is important to quantify the capital impact (or cost) of both trades; values that are hard to estimate without full KVA calculation. Among the set of counterparty parameters, two are particularly relevant. As we will see, the credit rating of a counterparty has a significant impact on the counterparty credit risk and CVA capital values. In the end, the capital requirements aim to reflect the risk of default (CCR capital) and risk of a smaller change in credit worthiness (CVA capital). Another critical counterparty parameter is the existing trade portfolio with that party: the result of netting effects will be considered under Section 4.3, as well as the notion of incremental KVA. As these parameters moneyness of a trade, trade maturity and counterparty credit rating may influence the Capital Valuation Adjustment numbers differently under two regulatory regimes, the analysis will be split over two sections covering them respectively: Section 4.2.1 and 4.2.2.

Here, the same interest rate swap scenario as per Section 4.1 is considered. All KVA numbers will be stated under the SDE method, with smaller discount factors, unless stated otherwise.

4.2.1 Impact under CEM regulations

Reconsider the situation of a Capital Valuation Adjustment calculation for an interest rate swap, as before. We will change the moneyness of the trade, maturity of the trade and the counterparty credit rating one by one and study the impact on the original numbers in the left Table 4.1.

1. The trade moneyness. An interest rate swap can be traded at par, where the fixed rate is set such that the initial value of the product is zero. In some sense, the fixed rate then reflects the market’s perception of the interest rates in the future. One may, however, also do a trade for some other fixed rate. The trade in that case has an initial value which can be positive, when the fixed rate is higher than the par swap rate (in the receiver swap case), or negative when it is lower.

Reconsider the receiver swap case outlined in Section 4.1, which had a par swap rate of about
0.23%. We will bump and revalue the Capital Valuation Adjustment on this rate with ±0.05%, or 5 basis points. As the swap is receiver, the positive bump yields an in-the-money swap and the negative bump yields an out-of-the-money swap. The capital profiles for the respective OTM and ITM swaps are plotted below. See Figure 4.4.

![Figure 4.4: Capital profiles (CEM) for a 5-year OTM(left) and ITM(right) interest rate swaps.](image)

As can be seen in the plots, in-the-money swaps have a higher capital profile and indeed a higher KVA value. This is a direct consequence of the replacement cost being higher for in-the-money products.

<table>
<thead>
<tr>
<th>Swap rate</th>
<th>Value</th>
<th>KVA</th>
<th>CCR</th>
<th>KVA</th>
<th>CVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTM swap</td>
<td>0.18%</td>
<td>-25,038.28</td>
<td>657.12</td>
<td>844.02</td>
<td></td>
</tr>
<tr>
<td>ATM swap</td>
<td>0.23%</td>
<td>0</td>
<td>708.53</td>
<td>911.71</td>
<td></td>
</tr>
<tr>
<td>ITM swap</td>
<td>0.28%</td>
<td>25,038.28</td>
<td>775.67</td>
<td>1,000.92</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: The KVA numbers for the swap for in-the-money, at-the-money and out-of-the-money scenarios under CEM regulation.

2. The trade maturity. The duration of a trade is of significant importance in relation to capital. Not only simply due to the fact that a longer trade means a longer time to hold capital, but also because of the particular construction of the CEM potential future exposure. The pfe level depends directly on the remaining maturity as in Table 1.3, such that a long maturity means a ‘double’ capital boost. A situation in which such a consideration is important, is when two trades have similar hedging or risk properties but different duration.

We will consider three swaps with similar properties as in the case study, but with different maturities: three years, five years and ten years. All swaps will be traded at par. The results can be found in Figure 4.5. The left figure shows a maturity smaller than the case study, while the right a larger maturity. The capital profiles demonstrate in both cases a form similar to the well-known ‘parabola’ exposure profile of a swap. At the time points of 1 and 5 year remaining maturity, the capital profiles show a drop, reflecting the potential future exposure relief of Table 1.3. As a consequence, cf. Table 4.3, the capital numbers are large for a high maturity swap: a maturity greater than five years carries a lot of potential future exposure.

Remark 4.1. It should be emphasized that the impact of potential future exposure is amplified for...
out-of-the-money swaps. In that case, the potential future exposure component of the exposure-at-default is relatively large compared to the replacement cost.

![Figure 4.5: Capital profiles (CEM) for a 3-year (left) and 10-year (right) interest rate swaps.](image)

<table>
<thead>
<tr>
<th>Results CEM</th>
<th>Maturity</th>
<th>Swap rate</th>
<th>KVA CCR</th>
<th>KVA CVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low maturity swap</td>
<td>3</td>
<td>-0.07%</td>
<td>189.24</td>
<td>222.40</td>
</tr>
<tr>
<td>Medium maturity swap</td>
<td>5</td>
<td>0.23%</td>
<td>708.53</td>
<td>911.71</td>
</tr>
<tr>
<td>High maturity swap</td>
<td>10</td>
<td>0.82%</td>
<td>5,364.81</td>
<td>9,617.46</td>
</tr>
</tbody>
</table>

Table 4.3: The KVA numbers for three swaps with different maturity under CEM regulation.

3. The counterparty credit rating. Regulatory capital levels should reflect the risks of losing money on a trade: the riskier a trade, the more capital should be held. One of such risks is counterparty default risk, of which the probability is described by the counterparty credit rating. The trade considered in the case study had a relatively trustable counterparty: an S&P rating of A. Below, we will consider exactly the same trade for different (dummy) counterparties with ratings BBB and B. Of course, one expects their KVA charge to be larger.

![Figure 4.6: Capital profiles (CEM) for an interest rate swap against a rating BBB (left) and rating B (right) counterparty.](image)

The capital profiles are actually similar, apart from a scaling factor. In the case of CCR
capital, the risk weight in formula (3.34) given by (1.26) increases with the regulatory survival probability. The CVA capital formula (3.35) has a risk weight from Table 1.6, which also increases with the credit rating. The results are summarized in Table 4.4 below.

<table>
<thead>
<tr>
<th>Results CEM</th>
<th>Credit rating</th>
<th>RW CCR</th>
<th>RW CVA</th>
<th>KVA CCR</th>
<th>KVA CVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bad credit rating swap</td>
<td>B</td>
<td>1.739</td>
<td>0.030</td>
<td>3,941.84</td>
<td>2,999.88</td>
</tr>
<tr>
<td>Medium credit rating swap BBB</td>
<td>0.757</td>
<td>0.010</td>
<td>1,721.34</td>
<td>1,062.24</td>
<td></td>
</tr>
<tr>
<td>High credit rating swap A</td>
<td>0.318</td>
<td>0.008</td>
<td>708.53</td>
<td>911.71</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: The KVA numbers for three swaps with different credit quality counterparties under CEM regulation.

As expected, the Capital Valuation Adjustment grows up when the counterparty credit worthiness goes down. Interestingly, the CCR risk weight grows faster than the CVA risk weight. As a consequence the CCR risk KVA is lower than the CVA risk KVA for the counterparties on the high end of the credit rating spectrum, while it is higher for others.

4.2.2 Impact under SACCRR regulations

In this section we repeat the analysis conducted in Section 4.2.1 but under the SACCRR regulatory regime. The exposure-at-default as calculated under this methodology is significantly smoother in terms of potential future exposure and moreover reflects netting effects a lot better. The exposure-at-default is generally a lot higher than under the CEM regime, as a consequence of the alpha multiplier of 1.4 as well as the different potential future exposure calculation. The trade scenarios studied below are the same as the scenarios under CEM in Section 4.2.1.

1. The trade moneyness. Consider the scenario of the illustratory example in Section 4.1 with three swaps: an out-of-the-money, an at-the-money and an in-the-money swap, with swap rates equal to before. The expected SACCRR capital profiles are displayed below. In-the-money swaps thus have a higher capital profile. In the case of SACCRR, this is not only due to higher replacement cost, but also because the pfe-multiplier (1.39) is at its maximum value for in-the-money swaps. The results in Table 4.5 confirm this.

3The counterparty credit risk weight shown here is the initial $t = t_0$ risk weight.
Table 4.5: The KVA numbers for the swap for in-the-money, at-the-money and out-of-the-money scenarios under SACCR regulation.

2. The trade maturity. We consider again a swap traded at par for three different maturities: 3 years, 5 years, and 10 years. The KVA numbers increase with the maturity as before. The capital profile however decays much smoother than in Figure 4.8 as a consequence of the pfe multiplier. As a consequence, there is no big, step-wise impact of going from a 5 to a 6-year maturity trade, opposed to the CEM case.

Table 4.6: The KVA numbers for three swaps with different maturity under CEM regulation.

3. The counterparty credit rating. As the credit worthiness of the counterparty is mainly a parameter in the risk weights and not in the exposure-at-default, the impact on the KVA is similar under both regulatory regimes. For completeness, the impacts under SACCR are displayed below.

Table 4.7: The KVA numbers for three swaps with different counterparties under SACCR regulation.

78
4.3 Practical considerations

Up to now, Capital Valuation Adjustments were considered in a relatively simple environment. The illustratory case study demonstrated the workings of the KVA model for a single trade added to an empty portfolio. The simulated exposures were the true exposures, without any form of collateral, and were calculated in a single regulatory regime. Although such a situation is sufficient to demonstrate the structure and properties of the KVA model, it is clear that a practical implementation would require far more twists and turns. The forthcoming section will briefly touch upon the issues and solutions that might appear in practice.

1. Netting effects and incremental KVA. The regulatory capital calculations, as described in Section 1.2, are performed on a netting set basis, which is generally a counterparty. The exposure and capital profiles should thus be simulated simultaneously for all trades, acknowledging the benefit of trades with offsetting risk. To illustrate how one would have KVA for multiple trades, considering the following situation: we trade two interest rate swaps, one payer and one receiver, with maturities of 3 and 5 years respectively. See the trade details in 4.8.

<table>
<thead>
<tr>
<th>Trade</th>
<th>Pay/Rec.</th>
<th>Maturity</th>
<th>Notional</th>
<th>Swap rate</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trade 1</td>
<td>Payer</td>
<td>3 yr.</td>
<td>8,000,000</td>
<td>-0.07%</td>
<td>0.00</td>
</tr>
<tr>
<td>Trade 2</td>
<td>Receiver</td>
<td>5 yr.</td>
<td>10,000,000</td>
<td>0.23%</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 4.8: The KVA numbers for three swaps with different maturity under CEM regulation.

The individual capital profiles for these trades will be simulated and subsequently the profiles for the trades as part of the same netting set. Clearly, the trades have offsetting risk and, as expected, the capital profile of the netting set is smaller than the sum of the individual profiles. In some cases, the netting set capital level is even smaller than the largest of the two individual ones. All (total) capital profiles are displayed in Figure 4.9 below. Under the CEM regime, there is only partial netting: the replacement costs are fully netted, while the potential future exposure is netted up to only 60%; cf. equation (1.31). The SACCR facilitates full netting effects, and indeed the right capital profiles in Figure 4.9 show more netting effects. Note moreover that,

\[ t = t_0 \]

The counterparty credit risk weight shown here is the initial \( t = t_0 \) risk weight.

Figure 4.9: Netting effects of two trades under CEM (left) and SACCR (right) regulatory regimes.
after three years, Trade 1 has matured and the capital profile of the netting set thus matches the capital profile of single Trade 2.

In practice, deals are conducted sequentially, as ever changing market circumstances or company structure lead counterparties to asking for new derivative trades. As follows from the netting principles above, the Capital Valuation Adjustment for a new trade cannot be individually calculated and added to the existing portfolio KVA level. Instead, the full portfolio of trades with the given counterparty should be re-simulated in order to find the appropriate KVA charge. It may in fact happen that the new trade reduces capital requirements on the portfolio, such that the KVA charge would theoretically be negative! The same could happen when an existing trade is unwound, and its capital relieved.

Consider the two-trade example of Table 4.8. We will calculate the KVA changes and charges as if the trades 1 and 2 are conducted one after each other. Examine Table 4.9 below. As follows

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>KVA CCR</th>
<th>KVA CVA</th>
<th>Total KVA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trade 1</td>
<td>459.29</td>
<td>371.64</td>
<td>830.93</td>
</tr>
<tr>
<td>Trade 2</td>
<td>1,735.64</td>
<td>1,628.46</td>
<td>3,364.10</td>
</tr>
<tr>
<td>Trade 1+2</td>
<td>987.75</td>
<td>865.21</td>
<td>1,852.96</td>
</tr>
</tbody>
</table>

Table 4.9: The KVA values (under SACCR) of individual and netted trades.

from the table above, the counterparty would be charged a KVA of 830.93 for the 3-year payer swap. If the counterparty would subsequently conduct the 5-year receiver swap, it would only get an incremental KVA charge of

$$KVA(1 + 2) - KVA(1) = 1,852.96 - 830.93 = 1,022.03,$$

rather than the much higher individual Trade 2 charge of 3,364.10. Hence, the counterparty earned itself a cheap deal by providing capital relief for the bank. In practice, the situation where a counterparty enters a trade with a directly opposite risk sign w.r.t. an existing trade will not often occur. However, it may happen that such capital charge reduction occurs as a result of a risk-related trade or unwinding of a trade. In that sense, it is a very realistic situation.

Remark 4.2. At this point, it should be remarked that this incremental, non-linear behaviour of KVA (and other XVA’s in general) makes XVA one of the most complex challenges in trading. The price of a derivative is now counterparty dependent and no longer unique. Moreover, the non-linearity properties give birth to complex (portfolio) optimization problems for banks.

2. Collateral. Another practical issue for KVA calculations is the notion of collateral. Banks often require clients to post collateral to cover large exposures on their derivative products, in order to reduce credit risk and, hence, capital requirements. The exposure to a client given a collateral level $C \geq 0$ is given by

$$E_t = C \mathbb{1}_{\{V_t > C\}} + V_t \mathbb{1}_{\{C > V_t > 0\}}.$$

Such a threshold thus drastically decrease the replacement cost numbers.

The case study in Section 4.1 had no collateral threshold, i.e. $C = \infty$. A bank may also have a fully collateralized trade where $C = 0$, such that the replacement cost is zero and the capital profile thus consists merely of the potential future exposure part. In a partially collateralized
situation, where $0 < C < \infty$, the situation is more interesting as it requires capping the simulated values. Below, the case study example is reconsidered with a collateral threshold of $C = 10,000$. See Figure 4.10 below.

![Figure 4.10: Collateralized EAD decomposition (left) and capital comparison (right).](image)

As expected, the replacement cost is now below the collateral threshold. The potential future exposure makes up a large part of the EAD and hence the capital profile. In the last year before maturity, the exposure hardly ever crosses the collateral threshold, so the replacement cost is near-zero. As the CEM add-on is actually zero (cf. Table 1.3), the capital requirement vanishes.

3. Regulatory switching. So far we have considered two different regulatory regimes for calculating exposure-at-default: the current-exposure-method and the standard-approach counterparty credit risk. The latter is due to revise the former at some future date, of which the current best estimate is the 1st of January 2021. As the KVA depends on the estimated capital profile over the lifetime of a trade, the longer maturity trades will have capital levels calculated under both measures. Therefore, the simulation engine will have to take into account a regulatory switch. Reconsider the case study example, which has a maturity date of 1 January 2023. The switch date of 1 January 2021 would yield a switch at $3/5$-th of the maturity. The capital profile and KVA numbers look as follows.

![Figure 4.11: Regulatory switch capital decomposition (left) and capital comparison (right).](image)
The left image of Figure 4.11 shows the counterparty credit risk and CVA risk capital profiles under the switching regime. Notice how the CCR capital bumps on the switch date as the SACCR exposure-at-default is larger. The CVA capital does not increase significantly as the alpha factor 1.4 for SACCR is divided out. On right hand side, it is shown how the switch methodology capital profile goes from CEM to following SACCR at switch date.
5 Conclusion

In this thesis, we investigated and compared different models for Capital Valuation Adjustment: the valuation adjustment that assesses the cost of holding regulatory capital for derivatives. The aim was to come up with a model that is both theoretically sound and practically implementable. Among the few models that have been developed over the short life of KVA, two stood out as conforming the requirements: the semi-replication model of Green and Kenyon and the balance sheet approach of Albanese and Crépey, which was to be modified to consider regulatory capital. Although different both in terms of assumptions and mathematical context, the models’ results were very similar: a KVA formula in terms of the risk-neutral expectation over the future capital profile, the difference being additional discounting with the cost of capital in the AC model.

The future capital levels that enter into the risk-neutral expectation for KVA stem from an aggregate of regulatory Credit Risk, Market Risk and Operational Risk capital. We considered only the credit risk component, as this is well quantifiable and significant on a single trade basis. Credit risk capital, consisting of a CCR and a CVA risk component, is calculated on different levels of granularity, but was reduced to a counterparty level calculation via an approximation. The difficult item in this calculation turned out to be the EAD, which is currently calculated under the CEM regulatory method and soon to be calculated under the SA-CCR method. Both methods have their own stochastic components, most notably the replacement cost, but also the net-to-gross ratio in the former and the potential-future-exposure multiplier in the latter method.

A potential implementation of KVA was demonstrated for interest rate swaps. The KVA expectation was calculated via Monte Carlo simulation, in which the least-squares algorithm of Longstaff and Schwartz was employed to obtain regulatory capital profiles. In particular, the replacement cost was estimated using a moneyness indicator from the LSMC algorithm. The net-to-gross ratio of the CEM method was only considered for a single trade, where it is equal to 1, and the stochastic potential-future-exposure multiplier in the SA-CCR method was upper bounded from the same LSMC moneyness indicators. Although the implementation was only considered for interest rate swaps, its generic setup allows for feasible extension to a broader class of (interest rate) products. A bank with an existing XVA engine should be able to do so.

The KVA numbers for interest rate swaps demonstrated to be significant in size, but well reducible via collateral and netting effects. As the numbers were of greater magnitude under the forthcoming SA-CCR regulations, KVA appears to become one of the most imperative XVA’s. In order to deal with future regulations, a switch was considered at the appropriate date, but so called regulatory risk cannot be perfectly hedged.

Lastly, a few directions of further research can be appointed. The presented implementation must be extended to a broader class of products, as regulatory capital covers all asset classes. Further study on how to estimate the NGR for a portfolio of trades would be useful, although the CEM will be discontinued soon. A more relevant direction of research is KVA hedging and optimisation. As the numbers turn out to be imperative, banks will have to reassess the capital costs of each business line, and optimise their balance sheet to reduce inefficient RWA exposure.
Populaire Samenvatting

Een bank verkoopt financiële derivaten aan klanten. Een derivaat is een beleggingsinstrument dat zijn waarde ontleent aan een ander, onderliggend goed. Zulke producten worden voornamelijk gebruikt om ongewenste risico’s af te dekken. De risico’s die een bank zelf loopt wanneer zij een derivaat verkoopt, probeert zij af te dekken door producten met een vergelijkbaar risicoprofiel te kopen. Zo blijft de bank risico neutraal, en verdient zij geld met het faciliteren van derivaten.

De prijs die een bank rekent voor een derivaat is een risico-neutrale prijs, die voortvloeit uit een wiskundig raamwerk dat zo opgezet is, dat niemand in de markt kan arbitreren. Dit raamwerk stelt de bank - en andere marktdeelnemers - bovendien in staat om risico’s in te schatten en correct af te dekken. Echter, omdat banken hun risico’s zelden perfect kunnen afdekken, moeten zij kapitaal vasthouden om eventuele onverwachte verliezen op te kunnen vangen. Dit is vastgelegd in de akkoorden van het Basel comité, dat richtlijnen voor de regulatie van banken ontwikkelt.

In de financiële crisis van 2008, is gebleken dat banken niet altijd de risico’s van hun derivaten op de juiste manier konden inschatten. Enerzijds is dit een gevolg van toegenomen complexiteit in de producten, maar anderzijds een gevolg van de wiskundige opzet van de risico-neutrale prijs. De conclusie die na de crisis werd getrokken, is dat het wiskundige raamwerk op te eenvoudige assumpties was gebouwd: bijvoorbeeld krediet- en liquiditeitsrisico’s werden onvoldoende belicht. Bovendien, bleken de kapitaalbuffers van banken te klein om de verliezen ten gevolge hiervan volledig te dekken, waardoor overheden moesten ingrijpen.

In de nasleep van de crisis, zijn de kapitaaleisen voor banken flink aangescherpt. Het derde Basel akkoord, Basel III, heeft ertoe geleid dat banken nu ook bewuster bezig zijn met kapitaal. Daarnaast ondergaat het wiskundige raamwerk nu een metamorfose, waarbij diverse waardeaanpassingen aan de risico-neutrale prijs worden toegevoegd om de onderliggende risico’s beter te reflecteren. Mede dankzij toenemende aandacht voor kapitaal, is recentelijk een ook waardeaanpassing ontwikkeld die de kapitaal kosten onderliggend aan een derivaat kwantificeert: KVA.

In deze scriptie bestuderen en vergelijken we twee modellen voor KVA. Hoewel de modellen verschillende uitgangspunten hebben en andere wiskundige technieken vereisen, zijn de resulterende formules voor de waardeaanpassing nagenoeg gelijk. In beide gevallen wordt deze gegeven door de verwachtingswaarde van de kapitaaleisen die een derivaat over zijn looptijd genereert. In lijn met de andere waardeaanpassingen, is de verwachtingswaarde te berekenen door de kapitaaleisen in het bestaande wiskundige raamwerk te simuleren en de uitkomsten daarvan te middelen. Uit dit scriptieonderzoek is gebleken dat de KVA waardeaanpassing significante impact heeft op de prijs van derivaten. Het is dus goed dat banken kapitaalkosten op deze manier gaan inprijzen.

---

1 Een luchtvaartmaatschappij kan zich bijvoorbeeld indekken tegen het risico dat de olieprijs stijgt. Het bedrijf koopt een opzione bij de bank, die hen het recht (maar niet de plicht) geeft om over een jaar olie te kopen voor 100 euro per vat, ongeacht de olieprijs op dat moment. Zo betaalt het bedrijf nooit meer dan 100 euro per vat voordat de prijzen in het bestaande wiskundige raamwerk te simuleren en de uitkomsten daarvan te middelen. Uit dit scriptieonderzoek is gebleken dat de KVA waardeaanpassing significante impact heeft op de prijs van derivaten. Het is dus goed dat banken kapitaalkosten op deze manier gaan inprijzen.

2 Arbitreren is winst maken, zonder risico te hoeven nemen. Gratis geld verdienen, in feite.

3 Als een bank een derivaat verkoopt en hiervoor risicokapitaal achterhoudt, zijn daar kosten aan verbonden vanwege het feit dat er onderhouden kapitaal geen rentabiliteit wordt behaald.
Appendix A: Terms and Acronyms

Introduction

**Basel** (The Basel framework) The set of Basel Accords (I, II and III) issued by the BCBS, providing recommendations for regulations in the banking industry. 7

**CVA** (Credit Valuation Adjustment) The price add-on to account for counterparty credit risk in a derivative exit price, calculated via the market price of default risk. 6

**DVA** (Debt Valuation Adjustment) The price add-on to account for the issuer’s credit risk in a derivative exit price, calculated via the market price of default risk. 6

**FVA** (Funding Valuation Adjustment) The price add-on to account for the issuer’s fundings costs in a derivative exit price. 6

**KVA** (Capital Valuation Adjustment) The price add-on to account for the cost of regulatory capital in a derivative exit price, so that an issuer makes its return on equity. 7

**XVA** (X Valuation Adjustments) The generic term referring collectively to a number of different valuation adjustments made to “risk-neutral” prices of derivatives. 7

Chapter 1

**AIRB** (Advanced Internal Ratings Based) The full-advanced calculation method for CCR capital risk-weights, where banks supply internal PD and LGD estimates. 21

**ASRF** (Asymptotic Single Risk Factor) A credit risk model underlying the Basel II regulations, that assumes a single global risk factor and a well-diversified portfolio. 10, 15

**BCBS** (Basel Committee on Banking Supervision) A committee of banking supervisory authorities that provides a forum for regular cooperation on banking supervisory matters. 19

**C** (Commitments) The amount of exposure in a loan that could be drawn in the future. 13

**CEM** (Current Exposure Method) A calculation method for regulatory exposure-at-default, based on the sum of the (current) exposure and a potential future exposure add-on. 21

**CR** (Credit Risk) The risk of a financial loss due to a counterparty not being able to fulfill its financial obligations on a (derivative) contract. Under the Basel III regulatory regime, Credit Risk is divided into Counterparty Credit Risk (CCR) and CVA Risk. 20
**EAD** (Exposure-at-default) The gross exposure to an obligor upon default, that under Basel III is often calculated as the sum of a replacement cost and a potential future exposure. 11

**EC** (Economic Capital) The amount of risk capital held by a financial institution, in accordance with its risk profile, in order to withstand potential difficulties. 9

**ECB** (European Central Bank) The central bank for the euro currency and administrator of the monetary policy in the Euro zone. 11

**EL** (Expected loss) The expected loss on a financial portfolio over a predefined risk horizon. 12

**EN** (Effective Notional) The notional of a trade corrected for risk direction and remaining maturity, under the SA-CCR exposure-at-default calculation. 26

**ES** (Expected Shortfall) A risk measure for financial institutions that is calculated on a portfolio as the average of losses greater than VaR; i.e., the expected loss in the worst cases. 12

**FIRB** (Foundation Internal Ratings Based) A semi-advanced calculation method for CCR capital risk-weights, where banks use internal PD estimates but regulatory LGD estimates. 21

**FRTB** (Fundamental Review of the Trading Book) A comprehensive suite of new capital rules and methodologies for banks’ trading activities, to be implemented as part of Basel III. 21

**HS** (Hedging Set) A set of trades amongst which netting of trade exposures is permitted, under the SA-CCR exposure-at-default calculation. 25

**IMM** (Internal Model Method) A collective term describing all regulatory calculation methods where advanced-status banks are allowed to use (partially) proprietary models. 21, 24

**LGD** (Loss-given-default) The percentage of outstanding credit that is lost upon default of an entity. 11

**MB** (Maturity Bucket) A category to which a trade belongs under the SA-CCR exposure-at-default calculation, based on remaining maturity. 26

**MR** (Market Risk) The risk of a loss due to changes in one or more market variables to which a (derivative) contract is exposed. 20

**MTA** (Minimum Transfer Amount) The minimum amount for which a margin call is made. 25

**NGR** (Net-to-gross-ratio) The ratio that provides netting benefits for a portfolio of trades in the potential future exposure calculation under the CEM methodology. 24

**NICA** (Net Independent Collateral Amount) The difference between the amount of received and posted collateral of two trade parties. 25

**O** (Outstandings) The amount of exposure in a loan that is already drawn by the obligor. 13

**OR** (Operational Risk) The risk that a firm’s internal practices, policies and systems are not adequate to prevent a loss being incurred. 20
PD (Probability-of-default) An entity’s probability of going into default over a given risk horizon. 11

PFE (Potential Future Exposure) The component of a regulatory exposure-at-default calculation that accounts for the exposure a bank might run on a counterparty in the future. 24, 25

RaRoC (Risk-adjusted Return on Capital) A risk-based profitability measurement ratio to measure a business’ financial performance in relation to risk and capital consumption. 10

RC (Replacement Cost) The component of a regulatory exposure-at-default calculation that accounts for the exposure that a bank currently runs on a counterparty. 24, 25

RoE (Return on Equity) A ratio to measure the net profitability of a business in relation to its shareholder equity value. 10

RW (Risk-Weight) The weights in the regulatory capital requirement calculations to differentiate between entities of various risk types. 21

RWA (Risk-Weighted Assets) A bank’s assets weighted according to risk, as prescribed by the Basel banking regulator, in order to determine a bank’s risk capital needs. 19

SA-CCR (Standardised Approach Counterparty Credit Risk) An upcoming calculation method for regulatory exposure-at-default, consisting of a replacement cost and an add-on component, designed to replace the standardised and CEM methodologies. 21

SVaR (Stressed Value-at-Risk) A Value-at-Risk calculated under stressed market conditions. 21

TH (Collateral Threshold) The amount of exposure on the counterparty the issuer of a derivative is prepared to accept before demanding collateral. 25

UL (Unexpected loss) A high loss percentile on a financial portfolio, such that it is an upper bound (at a given confidence level) on the portfolio loss. 12

VaR (Value-at-Risk) A statistical measure of riskiness of financial entities, calculated as the maximum amount expected to be lost over a certain period, at a given confidence level. 12

Chapter 2

AC (Albanese-Crépey) The BSDE KVA approach by Albanese and Crépey [1]. 29

BK (Burgard-Kjaer) The XVA semi-replication approach by Burgard and Kjaer [9]. 29

BSDE (Backward Stochastic Differential Equation) A stochastic differential equation with a prescribed value at some terminal time $T$. 39

CA (Contra-Assets) Under the BSDE approach of Albanese and Crépey, contra-assets are the counterparty default losses and funding expenditures of a bank’s trading business. The CA desk is the equivalent of the classical CVA/FVA desk and trades the contra-assets. 44
CL (Clean) The clean price of a product, under the BSDE approach of Albanese and Crépey, is the price after the credit and funding (CA) expenditures are stripped off. The desks dealing with the remaining market risk are Clean desks. 45

CM (Clean Margin) The collateral account that funds the clean desks, under the BSDE approach. 45

FK (Feynman-Kac) The famous formula that connects partial differential equations and stochastic processes and is often applied in mathematical finance. 33

GK (Green-Kenyon) The KVA-extended semi-replication approach by Green and Kenyon [17]. 29

PnL (Profit-and-Loss) The net profit (or loss) value as generated by a business. 44

RC (Reserve Capital) The account where the CVA and FVA (CA) charges are deposited after each trade, and which is used to cover default costs and funding costs. 45

RM (Risk Margin) The account into which the KVA charges are deposited before being released to shareholders. Moreover, the account serves as a buffer for exceptional losses. 45

SCR (Shareholder Capital at Risk) The capital that shareholders provide to the business (and hence put at risk). The amount of capital is equal to the required economic capital, corrected by the capital available through the Risk Margin account. 46

SDE (Stochastic Differential Equation) A differential equation in which one or more of the terms is a stochastic process, resulting in a solution that is also a stochastic process. 30

UC (Usual Conditions) A condition on a filtration, to be right continuous and contain all $\mathcal{F}$ null sets in the initial $\sigma$-algebra $\mathcal{F}_0$. 30

Chapter 3

IRS (Interest Rate Swap) A forward contract in which two parties agree to periodically exchange interest rates on a given notional value. 54

LIBOR (London Interbank Offered Rate) The average interest rate at which a bank can borrow money from another bank, estimated by a set of leading banks in London. 51

LMM (Libor Market Model) An interest rate model in which the LIBOR forward rates, which have the advantage of being directly observable in the market, are modeled. 50

LSMC (Least Squares Monte Carlo) A Monte Carlo simulation algorithm originally developed for American option pricing, that uses (least-squares) regression to estimate relevant quantities. The method was developed to circumvent nested Monte Carlo simulations. 61

MC (Monte Carlo) A computational method that estimates numerical values, e.g. expectations, by random sampling. 59

SR (Swap Rate) The interest rate on the fixed leg of a swap such that its value equals the value of the floating leg. An interest rate swap traded at the swap rate (par) has zero value. 54
Appendix B: The ASRF Model

In this section, some additional technical background material is provided on the Asymptotic Single Risk Factor model described in Section 1.1.2. Although the original paper from 2002 was (Gordy, 2002, [16]), the mathematical setup and proofs are provided in the terminology of (Rutkowski, Tarca, 2016, [39]) as they provide a more transparent framework. After specifying the required mathematical definitions, we give the proofs of the Propositions in Section 1.1.2.

Consider first the definition of a conditional independence model of a credit portfolio, where the single risk factor model is stated in a slightly more general form.

**Definition B.1.** A conditional independence model of a credit portfolio comprising \( n \in \mathbb{N} \) obligors over a given risk horizon \([0, \tau]\), where \( \tau > 0 \), takes the following form.

1. For each obligor \( i, 1 \leq i \leq n \), denote by \( \delta_i \in \mathbb{R} \) the Exposure-at-default and \( \eta_i \in [0, 1] \) the Loss-given-default. The exposure weight of \( w_i \) of obligor \( i \) is given by
   \[
   w_i = \frac{\delta_i}{\sum_{j=1}^{n} \delta_j}.
   \]
   Moreover each obligor has an unconditional default probability \( p_i \in [0, 1] \).

2. There exist latent (random) variables \( W_1, ..., W_n \) with representation
   \[
   W_i = \gamma_i Y + \sqrt{1 - \gamma_i^2} Z_i,
   \] (B.1)
   where \( Z_1, ..., Z_n \) and \( Y \) are mutually independent random variables. \( Y \) represents the (single) systematic risk factor and \( Z_1, ..., Z_n \) the idiosyncratic risk factors per obligor. The variables have cumulative distribution functions \( H \) and \( G_1, ..., G_n \) respectively; the distributions of \( W_1, ..., W_n \) are given by \( F_1, ..., F_n \). The variable \( \gamma_i \in (-1, 1) \) denotes the asset correlation.

3. The event that obligor \( i \) defaults during risk horizon \([0, \tau]\) is defined by
   \[
   D_i = \{ W_i < F_i^{-1}(p_i) \},
   \] (B.2)
   for \( i = 1, ..., n \).

In accordance with the setup of this definition, the portfolio percentage loss \( L_n \) conditional on the state \( y \) of the systematic risk factor \( Y \), can be written as

\[
L_n(y) = \sum_{i=1}^{n} w_i \eta_i p_i(y),
\] (B.3)

where the conditional default probability can be written as

\[
p_i(y) = \mathbb{P}(D_i | Y = y) = G_i \left( \frac{F_i^{-1}(p_i) - \gamma_i y}{\sqrt{1 - \gamma_i^2}} \right),
\] (B.4)

as follows from a general form of the Vasicek equation (1.14).
Remark B.1. Let \( y = H^{-1}(1 - \alpha) \), where \( \alpha \in (0, 1) \). The default probability of obligor \( i \) conditional on \( Y = y \), may be interpreted as “the probability of default of obligor \( i \) is no greater than

\[
P(D_i | Y = H^{-1}(1 - \alpha)) = G_i \left( \frac{F^{-1}_i(p_i) - \gamma_i H^{-1}(1 - \alpha)}{\sqrt{1 - \gamma_i^2}} \right)
\]

(B.5)

in \((\alpha \cdot 100\%)\) of economic scenarios.”

In accordance with Section 1.1.2, we must study portfolios under two conditions:

1. a portfolio with mutual independence conditional on the state of the systematic risk factor \( Y \) for all obligor pairs

2. an asymptotically fine-grained portfolio

The former condition is already incorporated in Definition B.1 through equation (B.1). The latter is formalised in Definition B.2 below.

Definition B.2. Let \( \Delta = \sum_{k=1}^{\infty} \delta_k \) be an infinite series whose terms \( \delta_k \in \mathbb{R}_+ \) represent the Exposure-at-default values assigned to obligors constituting a credit portfolio. Let its partial sums of order \( n \) be defined as \( \Delta_n = \sum_{k=1}^{n} \delta_k \). The portfolio is said to be an asymptotic portfolio if

\[
\sum_{n=1}^{\infty} \left( \Delta_n \right)^2 < \infty.
\]

(B.6)

Notice that the exposure weights of credits in an asymptotic portfolio diminish rapidly as the number of obligors increases, as follows from Kronecker’s lemma:

\[
\sum_{n=1}^{\infty} \left( \Delta_n \right)^2 < \infty \implies \sum_{k=1}^{n} w_k^2 \to 0 \text{ as } n \to \infty.
\]

(B.7)

Remark B.2. As a practical matter, the restriction is quite weak and would be satisfied by any conceivable real-world bank portfolio. For example, if all EAD values lie in an interval \([a, b]\), then

\[
\Delta_n = \sum_{k=1}^{n} \delta_k \geq na \to \infty \text{ as } n \to \infty,
\]

(B.8)

and

\[
\sum_{n=1}^{\infty} \left( \Delta_n \right)^2 \leq \sum_{n=1}^{\infty} \left( \frac{b}{na} \right)^2 < \infty.
\]

(B.9)

Hence in this case, even though the total EAD diverges, the portfolio still is asymptotic.

In what follows, we will prove the three results from Section 1.1.2, using the definitions above.

1. The first result stated in Section 1.1.2, Proposition 1.1, connects the loss variables conditional and unconditional on the systematic factor \( Y \). The result states that for an asymptotic portfolio, dependent on a single risk factor, the portfolio percentage loss converges almost surely to its condition expectation, as the portfolio approaches asymptotic granularity.

Proposition B.1. Assume a conditional independence model for an asymptotic credit portfolio. Then,

\[
\lim_{n \to \infty} \left( L_n - \sum_{i=1}^{n} w_i \eta_i p_i(Y) \right) = 0, \text{ P - a.s.}
\]

(B.10)
Proof. Let $y \in \mathbb{R}$ be a realisedisation of systematic risk factor $Y$. The random variables

$$\mathbb{1}_{\{Z_i < \zeta(y)\}}$$

for $i = 1, ..., n$ are independent by definition of the model and hence Bernoulli distributed with probability $p_i(y)$. Hence it follows, using the asymptotic credit portfolio assumption, that

$$\sum_{i=1}^{\infty} \text{Var} \left( \frac{\delta_i}{\Delta_i} \mathbb{1}_{\{Z_i < \zeta(y)\}} \right) = \sum_{i=1}^{\infty} \left( \frac{\delta_i}{\Delta_i} \right)^2 \eta_i^2 p_i(y)(1 - p_i(y)) \leq \sum_{i=1}^{\infty} \left( \frac{\delta_i}{\Delta_i} \right)^2 < \infty. \quad (B.11)$$

Moreover, conditional on $Y = y$, the random variables

$$\left( \frac{\delta_k}{\Delta_k} \eta_k \mathbb{1}_{\{Z_k < \zeta(y)\}} \right)_{k \in \mathbb{N}} \quad (B.12)$$

are independent each with expectation $p_k(y)$. The Kolmogorov convergence criterion (cf. (Gut, 2013, [20])) then yields, from (B.11), that

$$\sum_{k=1}^{\infty} \frac{\delta_k}{\Delta_k} \eta_k \left( \mathbb{1}_{\{Z_k < \zeta(y)\}} - p_k(y) \right) \quad (B.13)$$

converges a.s. w.r.t. $P_y$. Applying stochastic Kronecker’s lemma (see Gut, 2005, lemma 6.5.1) yields

$$\frac{1}{\Delta_n} \sum_{k=1}^{n} \delta_k \eta_k \left( \mathbb{1}_{\{Z_k < \zeta(y)\}} - p_k(y) \right) \rightarrow 0, \; P_y - a.s. \quad (B.14)$$

which implies that $\forall y \in \mathbb{R}$

$$\left( L_n(y) - \sum_{i=1}^{n} w_i \eta_i p_i(y) \right) \rightarrow 0, \; P_y - a.s. \quad (B.15)$$

We conclude that

$$\lim_{n \rightarrow \infty} \left( L_n - \sum_{i=1}^{n} w_i \eta_i p_i(Y) \right) = 0, \; P - a.s. \quad (B.16)$$

Remark B.3. The sequence $\{L_n\}_{n \in \mathbb{N}}$ is bounded, but it is not monotone: the exposure weights depend on the number of obligors $n$. Consequently $L_n$, and hence $\sum_{i=1}^{n} w_i \eta_i p_i(y)$, may not converge as $n \rightarrow \infty$.

2. The second result, Proposition 1.2, connects the quantiles of the conditional and unconditional loss distributions: quantiles of the distribution of conditional expectation of portfolio percentage loss, may be substituted for quantiles of the portfolio loss distribution.

Proposition B.2. Consider a credit portfolio comprising $n$ obligors, and denote by $L_n$ the portfolio percentage loss. Let $Y$ be a random variable with continuous and strictly increasing distribution function $H$. Denote by $\psi_n(Y)$ the conditional expectation of portfolio percentage loss $\mathbb{E}[L_n|Y]$. Assume the following conditions hold:

1. $\lim_{n \rightarrow \infty} \left( L_n - \sum_{i=1}^{n} \psi_i(Y) \right) = 0, \; P - a.s. \quad (B.17)$

91
2. There exists an open interval \( I \subseteq \mathbb{R} \) containing \( H^{-1}(1-\alpha) \) and an \( N_1 \in \mathbb{N} \) s.t. \( \forall n > N_1 \), \( \psi_n(y) := \mathbb{E}[L_n | Y = y] \) is strictly decreasing and differentiable on \( I \).

3. There exists an \( N_2 \in \mathbb{N} \) s.t. \( \forall n > N_2 \) the derivative of the conditional expectation function satisfies \( -\infty < -m < \psi'_n(y) < -M < 0 \) for all \( y \in I \), with \( m, M > 0 \) independent of \( n \).

Then it follows that
\[
\lim_{n \to \infty} \mathbb{P}(L_n \leq \psi_n(H^{-1}(1-\alpha))) = \alpha,
\] (B.18)
and moreover
\[
\lim_{n \to \infty} |\text{VaR}_\alpha(L_n) - \psi_n(H^{-1}(1-\alpha))| = 0.
\] (B.19)

Remark B.4. Regarding the technical conditions, the following can be said:

1. Condition 1. is, in addition to the two assumptions of Gordy, already satisfied by asymptotic conditional independent credit models, as follows from proposition (1.1). Moreover, for the proof of proposition (B.2) convergence in probability is sufficient. Hence, it might apply more generally to credit models that satisfy condition 1. in probability-convergence terms.

2. Condition 2. in fact requires smoothness of the conditional expectation of portfolio percentage losses, in states of the economy associated with the tail of the portfolio distribution.

3. Condition 3. holds if \( p_i \) and \( \gamma_i \) are bounded away from zero and one respectively, for \( 1 \leq i \leq n \). In practical situations this will always hold, as capital requirements could be set to either 0 or \( PD \cdot LGD \) if not.

Proof. The proof can be found in (Gordy, 2002, [16]), Proposition 5.

3. The last result of Section 1.1.2, Proposition 1.3, proofs the asymptotic limit for regulatory capital.

**Proposition B.3.** Assume an asymptotic conditional independence model of a credit portfolio. Then the credit risk capital is of the asymptotic form
\[
\lim_{n \to \infty} K_\alpha(L_n) = \lim_{n \to \infty} \sum_{i=1}^{n} w_i \eta_i G_i \left( \frac{F_i^{-1}(p_i) - \gamma_i H^{-1}(1-\alpha)}{\sqrt{1-\gamma_i^2}} \right) - \lim_{n \to \infty} \sum_{i=1}^{n} w_i \eta_i p_i.
\] (B.20)
assuming the latter limits exist.

Proof. It follows from Proposition B.2 that
\[
\lim_{n \to \infty} K_\alpha(L_n) = \lim_{n \to \infty} \left[ \text{VaR}_\alpha(L_n) - \mathbb{E}[L_n] \right]
\] (B.21)
\[
= \lim_{n \to \infty} \left[ \text{VaR}(L_n) - \mathbb{E}[L_n | Y = H^{-1}(1-\alpha)] \right] + \lim_{n \to \infty} \left[ \mathbb{E}[L_n | Y = H^{-1}(1-\alpha)] - \mathbb{E}[L_n] \right]
\] (B.22)
\[
= \lim_{n \to \infty} \left[ \mathbb{E}[L_n | Y = H^{-1}(1-\alpha)] - \mathbb{E}[L_n] \right] - \lim_{n \to \infty} \sum_{i=1}^{n} w_i \eta_i G_i \left( \frac{F_i^{-1}(p_i) - \gamma_i H^{-1}(1-\alpha)}{\sqrt{1-\gamma_i^2}} \right) - \lim_{n \to \infty} \sum_{i=1}^{n} w_i \eta_i p_i.
\]
Bibliography


